Supramolecular amorphous systems: Analysis and control of non-crystalline pharmaceutical systems

A thesis presented in fulfilment of the requirements for the degree of Doctor of Philosophy in the Faculty of Science of the University of Strathclyde

by

Michael Devlin

Strathclyde Institute of Pharmacy and Biomedical Sciences

September 2021

Declaration of Author's Rights

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by the University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Signed: Michael Devlin **Date: 24/09/2021**

Acknowledgements

Firstly, I would like to express my gratitude to Professor Alastair Florence for giving me the opportunity to complete my PhD research in his group, and for his guidance and knowledge throughout the project.

I also wish to thank colleagues from Bayer Healthcare, whom contributed extensively to Chapter 4, including Drs Wolfgang Beckmann, Gabrielle Winter, Jan-Olav Henck and Tia Jacobs. Thanks are also due to Drs Stefan Norberg, Helen Blade, Helen Wheatcroft and Emma Eriksson from Astra Zeneca for providing AZD5718 materials and molecular dynamics simulations shown in Chapter 5.

My gratitude goes to Dr Iyke Onyemelukwe for sample preparation, MATLAB coding and useful discussions surrounding the work in Chapter 4, Dr Michael Chrubasik for help with PCA, Dr Humera Siddique for providing precipitated AZD5718 samples for Chapter 5 and Dr Dean Keeble from Diamond Light Source for assistance in operating the I15-1 beam line and support in data analysis.

Special mention to those who attended beam time visits at Diamond and survived the shifts – the gaffer, Ecaterina Bordos, Dr Maggie Chong, Dr Kenny Smith, Dr Iyke Onyemelukwe, Dr John Robertson and Dr Vijay Srirambhatla. Another special mention to the Ayrshire boys – Drs Alan Martin, John Robertson and Ian Houson for training, useful discussions and facilitating industry interactions, and putting up with me at lunch.

I express gratitude to the CMAC National Facility, and particularly Dr Thomas McGlone and Dr Deborah Bowering, for sample analysis, facilitating training and access to labs. Many others in CMAC also deserve thanks for useful inputs or for general blethering, including Morell Kerr and Drs Martin Ward, Chantal Mustoe, Tony Vasilleiou, Murray Robertson, Iain

Oswald, Carlota Mendez Torrecillas, Carlos Moreno Leon and Andrea Johnston among many others.

Finally, thanks go to my family: my parents, brother and sister and above all, my wee lassies Layla and Allie who gave me the drive to do this.

Abstract

Recent years have seen an increase in interest in amorphous pharmaceutical solids due to their solubility benefits when compared to their crystalline counterparts. This thesis reports approaches for reliable collection and analysis of X-ray pair distribution function (PDF) data to enable structural insights to amorphous pharmaceuticals to improve understanding of structure and properties in these systems. A number of factors are considered and assessed for enabling high-quality laboratory PDF, including sample handling and data collection methodologies, before applying to the amorphous systems indomethacin-polyvinylpyrrolidone (IND-PVP) and AZD5718, a small-molecule active pharmaceutical ingredient (API) currently under development at Astra Zeneca. High-throughput synchrotron X-ray PDF is applied to amorphous paracetamol (PCM) to understand structure and transformations in the amorphous solid.

Optimisation of data collection, sample handling and data treatment procedures enabled laboratory PDF analysis of amorphous pharmaceuticals which compared favourably to benchmark synchrotron PDFs. This enabled in-house structural investigations including limit of detection of crystallinity and the impact of elevated humidity on the amorphous structure. PDF was found to be sensitive to both low levels of crystalline material and structural changes induced by moisture absorption.

Investigations of amorphous AZD5718 revealed subtle structural changes as a result of preparation method, which were largely due to small differences in the water content in the two preparations, demonstrating again the sensitivity of PDF analysis to subtle structural changes. MD simulations revealed significant disruption of hydrogen bonds in the amorphous matrix by water. By extracting sample conformations from the MD simulations, the

dominance of the intramolecular structure to features in the low-r region of the PDF was demonstrated.

High-throughput synchrotron PDF analysis enabled tracking of structural transformations in amorphous PCM, where subtle differences were detected statistically through the β- and glass transitions, and upon crystallisation. MD simulations coupled with PDF allowed production of a structural model of amorphous PCM, and provided a possible explanation for the observed crystallisation behaviour of amorphous PCM.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1. Background

In the pharmaceutical industry, defining and controlling active pharmaceutical ingredient (API) physical form is a crucial part of drug development (Brittain, 2016). For reasons of stability and ease of characterisation, the crystalline phase is typically favoured in pharmaceutical formulations (S. J. Billinge, 2015), however, due to concerns about poor aqueous solubility impacting oral bioavailability of an increasing number of new chemical entities (NCEs) in the pharmaceutical pipeline, the amorphous phase is becoming of increasing interest (Hancock & Parks, 2000). This stems from the fact that amorphous solids possess higher Gibbs free energy and therefore display elevated solubility compares with their crystalline counterparts. Although understanding of the kinetic and thermodynamic behaviors of amorphous drugs has improved (Zografi & Newman, 2017), understanding of the structural features which underpin these properties is still lacking. Thus, structural characterisation of amorphous materials is of significant importance for developing understanding of these increasingly valuable materials.

For crystalline drugs, structural determination is a crucial step in the development process, with single crystal X-ray diffraction (SXD) being the gold standard for structure solution (David et al., 2006) and is required by regulatory authorities where available to enable confirmation of identity. The use of SXD, however, is dependent on crystals of suitable size and quality for this method to be used. Thus, in the absence of crystals that allow for complete single crystal data to be collected, X-ray powder diffraction (XRPD) may be used. These approaches have been developed and exploited increasingly over the last two decades to give an alternative route to access the structure of polycrystalline samples (David, Shankland, McCusker, & Baerlocher, 2010a; David et al., 2006; Alastair J. Florence et al., 2003, 2005). Subsequently, as the use of SXD is dependent on there being reasonable size and quality crystals, the use of

XRPD is dependent on there being crystallites that give rise to Bragg diffraction and the subsequent application of global optimisation methods. Nanocrystals also present a challenge given the impact on line broadening and ability to resolve individual reflections in a diffraction pattern. A further extension of this challenge is encountered when considering amorphous materials: with no long-range periodic order, standard diffraction techniques cannot be used to gain understanding of their structure as no Bragg diffraction of X-ray is produced and so the standard toolbox of crystallographic methods cannot be applied. Therefore, we must turn to other methods to obtain knowledge of the structure of amorphous molecular solids. Recent years have seen an increase in interest in the use of the pair distribution function (PDF) method, produced by Sine Fourier transform of X-ray (or neutron) scattering data, for assessing "local" structure in disordered and amorphous APIs (S. J. Billinge, 2015). In the absence of long-range order, this method essentially provides information on the structure as generally defined by the first or second coordination shell around a given molecule. Despite being widely implemented for the study of nanostructure in inorganic materials (Young & Goodwin, 2011), there remains a significant gap in the routine implementation and analysis of PDF data for molecular organics. As such, the work herein assesses approaches for data collection, analysis and extraction of structural information from PDF data and seeks to move beyond the use of PDF as simply a fingerprinting technique, to enabling structural understanding of amorphous molecular solids.

1.2. Pharmaceutical solid state

Given that up to 80% of all pharmaceutical products are marketed as oral solid dose (OSD) forms (Couillaud, Espeau, Mignet, & Corvis, 2019), the importance of the solid state of

pharmaceuticals is clear. The process of identifying the most suitable phase for manufacturing is crucial in order to access the required drug properties, such as solubility, stability, and manufacturability, amongst others (Brittain, 2016). The range of solid forms available for a given drug can include polymorphs, co-crystals, salts, hydrates, solvates and the amorphous phase (Hilfiker & Raumer, 2018), briefly summarised in Table 1.1.

Table 1.1: Summary of pharmaceutical solid forms.

Solid form	Structural features
Polymorph	Same molecular constituents with the ability to exist in different 3-dimensional periodic or conformational arrangements
Co-crystal	Single phase crystalline material consisting of two or more different chemical constituents in a stoichiometric ratio
Salt	Combination of an ionisable drug with a counter-ion, forming a neutral crystalline complex
Hydrate	Crystal form combinating a drug with water in a stoichiometric ratio
Solvate	As with a hydrate, only the drug is combined with an organic solvent
Amorphous	Drug formulated such that there is no long-range periodic order, may require additional components to stabilise

1.3. Crystalline materials

Crystalline materials are characterised by the presence of 3-dimensional translational and orientational order giving rise to a crystalline lattice (Clegg, 1998). The crystal structure consists of a basic building block, the unit cell, describing the lattice structure of the material. The basic component of a crystal is the asymmetric unit, which in itself can be used to describe the full periodic structure (Figure 1.1), by a series of symmetry operators. As a result of the repetitive nature of the crystalline structure, the conditions for Bragg diffraction are met (Section 1.7. onwards), meaning sharp intensities at distinct angles are observed, allowing structure determination/refinement to be performed (Wilson, 2000).

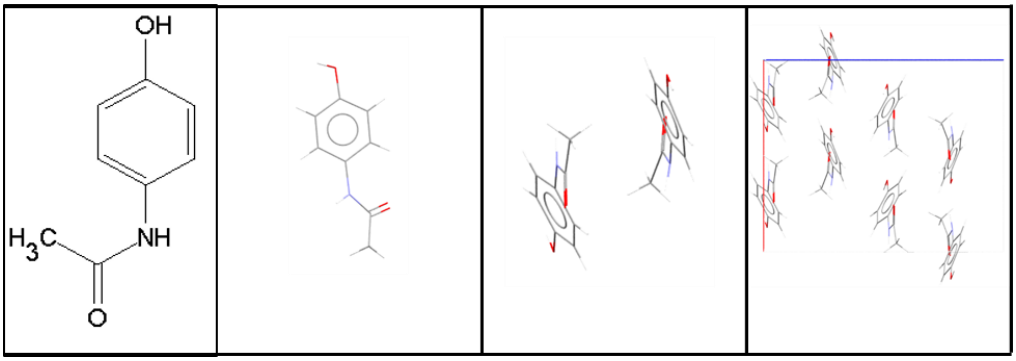

Figure 1.1: 2D chemical structure of paracetamol (left), Single paracetamol molecule (second left), asymmetric unit for paracetamol Form III (second right), unit cell of paracetamol Form III (right). Coloured lines indicate unit cell axes a and c.

1.4. The amorphous phase

The amorphous phase is described as lacking in the 3-dimensional periodic order characteristic of crystalline materials (Craig, Royall, Kett, & Hopton, 1999) (Figure 1.2). Despite the obvious characterisation issue facing these materials, amorphous pharmaceuticals have seen a recent surge in interest in academic and industrial communities (Beyer et al., 2016; Hancock, Shamblin, & Zografi, 1995; Hilden & Morris, 2004; Murdande,

Pikal, Shanker, & Bogner, 2011). This is largely due to the higher aqueous solubility exhibited by amorphous solids (Hancock & Parks, 2000). As solubility is of paramount importance to pharmaceutical product performance, these materials are obviously advantageous in increasing the amount of API available in molecular form for absorption from the gastro-intestinal tract for example. However, the reduced thermodynamic favorability that leads to enhanced solubility can also lead to poor physical stability of amorphous APIs (Kratochvíl, 2011). This, coupled with the poor understanding of their structure and associated material properties has hindered the implementation of rational, predictive approaches to assess risk, performance and stability within the context of the development and production of medicinal products.

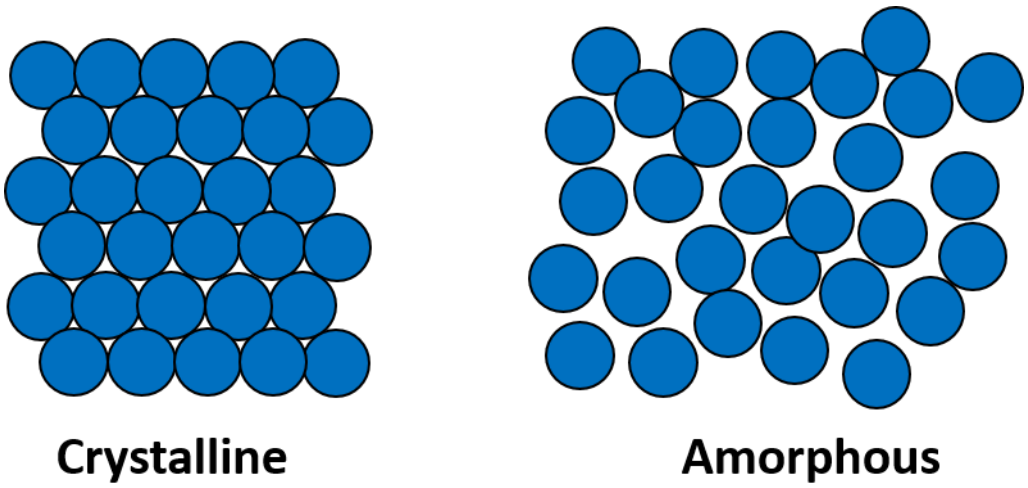

Figure 1.2: Schematic illustration of a crystalline (left) and amorphous (right) structure showing regular repeating packing (crystalline) and the disordered packing in an amorphous solid.

Due to the lack of 3D periodic order in amorphous solids (Stachurski, 2011) amorphous materials lack the regular lattice planes present within crystalline materials, which in X-ray diffraction give rise to constructive interference when Bragg's Law is satisfied that results in sharp diffraction peaks at specific angles. As such X-rays are scattered without

giving rise to X-ray diffraction resulting in a classic broad “halo” being observed during XRPD experiments, as shown in Figure 1.2.

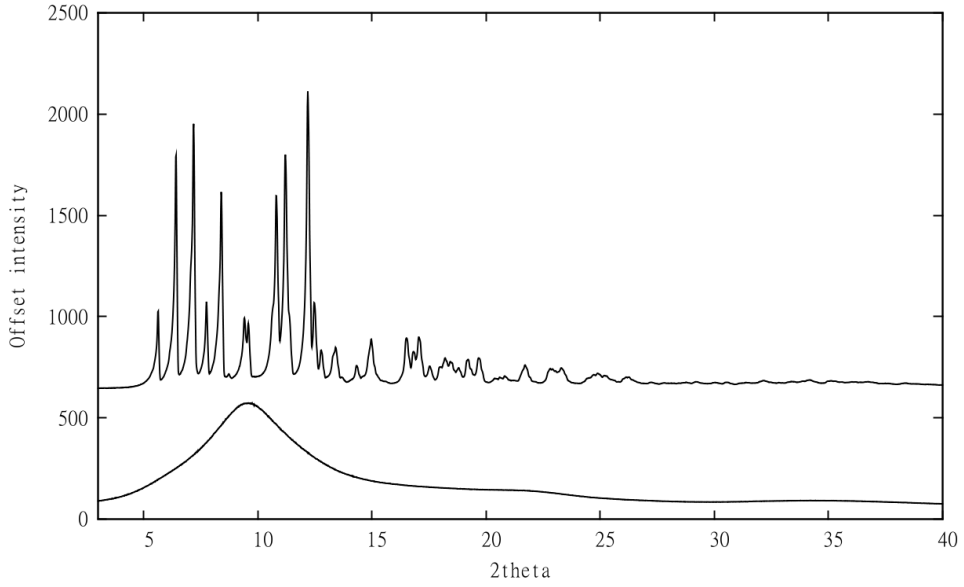

Figure 1.3: In-house molybdenum XRPD data for crystalline Form I (top) and amorphous (bottom) paracetamol.

1.4.1. Stability of amorphous pharmaceuticals

The implementation of amorphous pharmaceuticals has been hindered, in part, due to their inherently poor physical stability (Y. Sun et al., 2012), where their higher Gibbs’ free energy, in comparison to their crystalline counterparts, provides a thermodynamic driving force towards crystallisation. For example, commercially available amorphous tacrolimus was found to crystallise within 4 weeks upon exposure to elevated temperature and humidity (Trasi, Purohit, & Taylor, 2017). This presents a clear challenge for the implementation of amorphous solids in medicinal products, from both a quality perspective, and in the shelf-life of the resulting product.

1.4.2. Stabilisation of amorphous pharmaceuticals

Given their inherent lack of stability against crystallisation, it is often favourable to formulate amorphous APIs with additional components, in a molecular dispersion, which serves to improve physical stability, whilst retaining their solubility benefits (Laitinen, Lobmann, Strachan, Grohganz, & Rades, 2013). These multicomponent systems often rely on polymers within which APIs are molecularly dispersed, though can also include co-amorphous systems where an amorphous solid form is stabilized by mixing with another small molecule component (Moinuddin et al., 2017). Whilst the use of amorphous drugs remains relatively limited , a small number of APIs are currently marketed in the amorphous phase, both as pure drug product, and formulated with polymeric excipients for physical stability enhancement, with a few examples given in Table 1.2 (Wyttenbach & Kuentz, 2017).

Table 1.2: Examples of APIs currently marketed in the amorphous form.

Compound	Trade name	Manufacturer	Carrier	Preparation method
Etravirine	Intelence®	Janssen	HPMC[1]	Spray drying
Griseofulvin	Gris-PEG®	Novartis/Pedinol	PEG[2]	Hot -melt extrusion
Lopinavir & Ritonavir	Kaletra®	AbbVie	PVP VA 64	Hot-melt extrusion
Nifedipine	Afeditab®	Elan/Watson	Poloxamer or PVP[3]	Melt/absorb on carrier

[1] Hydoxypropyl methylcellulose
[2] Polyethylene glycol
[3] Polyvinylpyrrolidone

Telaprevir	Incivek®/Incivo®	Vertex/Janssen	HPMCAS[4]	Spray drying
Troglitazone	Rezulin®	Pfizer	PVP	Hot-melt extrusion
Verapimil hydrochloride	Isoptin® SR-E 240	AbbVie	HPC[5]/HPMC	Hot-melt extrusion
Cefuroxime axetil	Ceftin®	GlaxoSmithKline	N/A	N/A
Quinapril hydrochloride	Accupril®	Pfizer	N/A	N/A

1.5. Structure of liquids and amorphous solids

The development of techniques for studying amorphous structure have been largely pioneered by the inorganic community concerned with water, ceramics and functional materials (Biswas et al., 2018; Lamparter & Kniep, 1997; Treacy & Borisenko, 2012). For example the structures of liquid and solid amorphous water has been extensively studied (Liu, He, & Zhang, 2017; Narten, Venkatesh, & Rice, 1976; Petkov, Ren, & Suchomel, 2012; Skinner et al., 2013; Starr, Bellissent-Funel, & Stanley, 1998), where a tetrahedral orientation of water molecules has been determined as the most likely local arrangement of water molecules. Subsequently, the phenomena of polyamorphism, that is, the ability of an amorphous material to exist in two or more distinct thermodynamic phases, has been described, as in the case of high- and low-density amorphous ices. Examples of organic materials which have been reported to exhibit polyamorphism include D-mannitol (Zhu, Wang, Perepezko, & Yu, 2015; Zhu & Yu, 2017) and triphenyl phosphite (TPP) (Walton et al., 2020). An amorphous-amorphous phase transition has been observed for TPP using infrared

[4] Hydroxypropyl methylcellulose acetate succinate
[5] Hydroxypropylcellulose

spectroscopy and microscopy, where each amorphous phase results in crystallisation of a separate crystal form, which are similar in terms of molecular conformation to the amorphous precursors. Furthermore, comparisons of the liquid and solid forms of amorphous water seem to support the theory that the amorphous solid is merely a continuation of the liquid phase (Mallamace, 2009), in that the structure does not change, however the liquid is significantly more mobile and less viscous.

The assumption that a glass is merely a continuation of the liquid phase is supported by the lack of first order phase transition typically observed between the liquid and amorphous phase (Berthier & Ediger, 2016), that is, a discontinuous change in entropy, corresponding to latent heat (Papon, Leblond, & Meijer, 2006). Instead the glass transition temperature, T_g, is present. The T_g is a second order, or "kinetic" transition, and as such there is a continuous change in entropy, and instead of a re-ordering of the molecular packing, as would be observed in a first order transition, the material undergoes a change in heat capacity and viscosity typically referred to as moving from a glassy (below Tg) to a rubbery state (above Tg, see Figure 1.3).

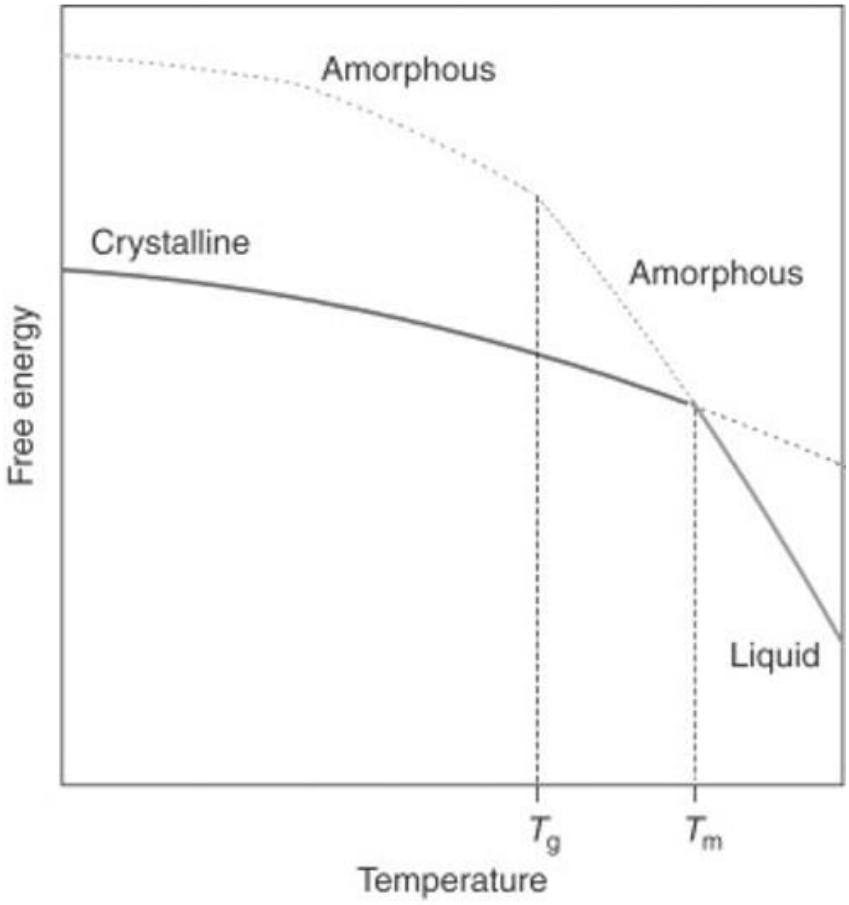

Figure 1.4: Example free energy – temperature phase diagram (Zhang & Zhou, 2009). T_g represents the glass transition temperature, T_m is the melting temperature. Crystal phase is demonstrated to be the lowest free-energy phase below T_m, above which the the liquid is the stable form.

Pharmaceutically relevant studies of amorphous and liquid structure include reports of molecular dynamics (MD) simulations of amorphous indomethacin (IND) (Xiang & Anderson, 2013), wherein it was determined that the structure consisted of both chains of IND molecules, found in the alpha polymorph, and carboxylic acid dimers, found in the gamma polymorph (Figure 1.4). These results compared favourably with spectroscopic analyses of amorphous IND, where similar hydrogen bonding patterns were revealed. MD simulations of glycogen in paracetamol-water solutions have demonstrated that introduction of a single polymer chain can result in ordering of drug molecules around the polymer (Lim,

Feng, & Liu, 2005), by calculating an order parameter and comparing this to the pure paracetamol-water solution. Refinements to neutron and X-ray total scattering data using empirical potential structure refinement (EPSR, see section 1.8.3.5) revealed significant hydration of imidazole molecules in aqueous solutions, where hydrogen bonding was dominated by imidazole-water interactions. Imidazole-imidazole interactions were predominantly due to $\pi - \pi$ stacking of molecules in solution (Al-Madhagi, Callear, & Schroeder, 2020). All studies demonstrate that structure in amorphous and liquid materials is not necessarily a random arrangement of molecules.

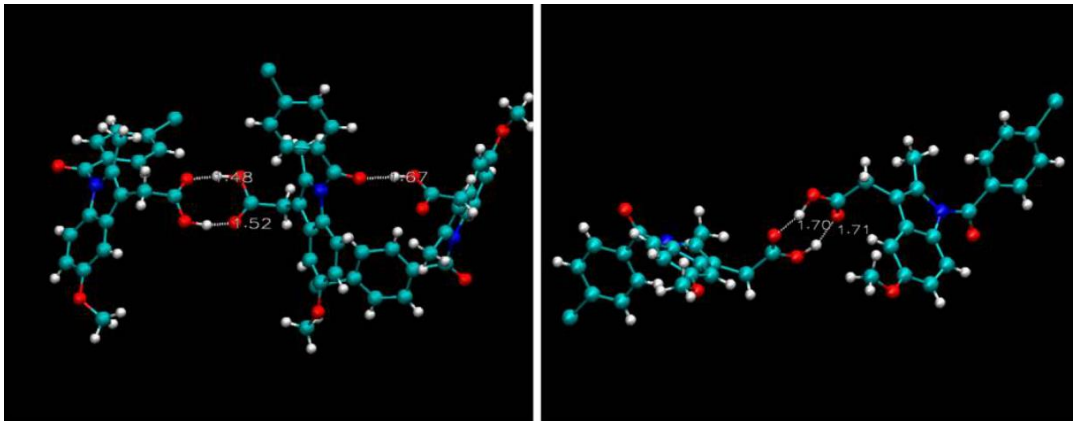

Figure 1.5: Example interactions in proposed model of the structure of amorphous indomethacin, left showing trimer similar to that found in alpha polymorph, right shows dimer similar to that present in gamma polymorph.

1.6. Structure solution and refinement using X-ray diffraction

1.6.1. Principles of X-ray diffraction

X-ray diffraction is a key technique in crystal structure solution. The ability for routine structure determination is essential within the pharmaceutical industry, given the occurrence of diverse polymorphism in a number of molecular solids (Brittain, 2016). The principle of structure solution using X-ray diffraction (XRD) can be described using Bragg's Law (Bragg & Bragg, 1913), describing the conditions under which diffraction occurs when a beam of X-

rays is incident upon a plane of atoms that is scattered by the electrons of the constituent atoms (Figure 1.4). Diffraction occurs due to constructive interference between X-rays which have reflected off a set of repeating interplane distances within a crystal. The occurrence of constructive interference is dependent on the path-length difference of the reflected X-rays, where due to the wave-like behaviour of light, it is required that the difference in path-length between the X-ray is an integer multiple of the wavelength, λ. Due to this requirement for the Bragg condition to be met, sharp peaks are observed at distinct angles that are dictated by the d-spacing between planes in the crystal. Thus, the order of the reflection, $n\lambda$, is related to the interatomic distance, d (Å), and scattering angle, θ (° 2θ), by

$$n\lambda = 2d \sin\theta \qquad (1.1)$$

A short wavelength incident radiation is required for diffraction from molecular crystals as the wavelength needs to be comparable to the interatomic distances within the crystal in order for constructive interference to occur. Figure 1.4 shows a simplified 2-dimensional lattice. In 3-dimensions, these planes are described by a set of Miller indices (hkl), which describe the intersection of the planes with the three unit cell axes (Waseda, Matsubara, & Shinoda, 2011) in reciprocal space. Thus, in the X-ray diffraction experiment, these reflections are detected as a regular series of high-intensity spots, related to distances within the reciprocal lattice, which can then be used to determine the spacings and symmetry in real-space.

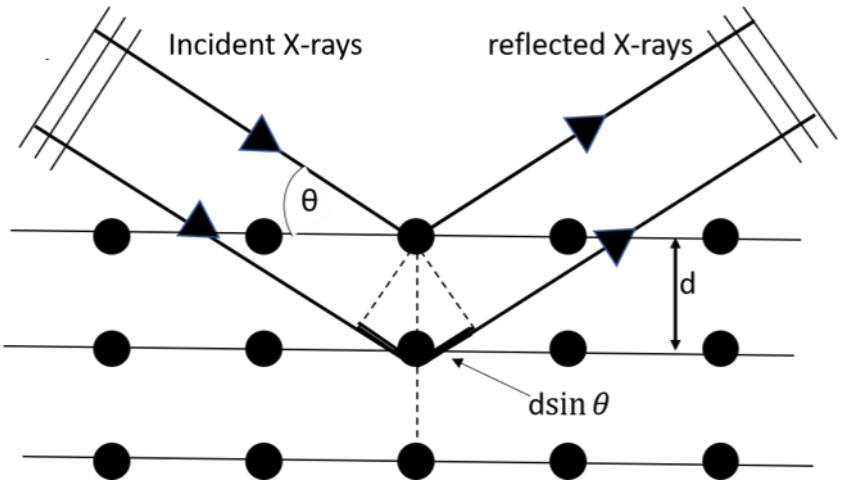

Figure 1.6: Graphic illustrating the principle of Bragg's Law, where θ is the angle of the incident beam relative to the crystal plane and d is the interatomic spacing. Thick line highlighted as $d\ sin\,\theta$ represents the extra distance travelled by the lower diffracted beam.

The principle of Bragg's Law can be understood geometrically using the Ewald sphere (Clegg et al., 2009), shown in Figure 1.5., wherein the crystal is positioned at the centre of a sphere with radius $\frac{1}{\lambda}$. The incident wave vector, s_0, passes through the centre of the sphere and makes contact with the surface of the Ewald sphere. This corresponds to the 0 0 0 reflection and is the origin of reciprocal space. The Bragg condition is only met, however, when the scattered wave vector, s, makes contact with another point which lies on the surface of the sphere. The distance between the origin of reciprocal space and the lattice point touched by the scattered wave is equal to $\frac{1}{d}$, where d is the interplanar distance. However, it is apparent from Figure 1.5. that from a single crystal orientation, not all lattice points are accessible experimentally. Thus, it is necessary to rotate the crystal, and therefore the reciprocal lattice, to enable as many points to be measured as possible. However, where it is not possible to collect data from all points, it may be necessary to use a shorter wavelength to probe more points in the reciprocal lattice.

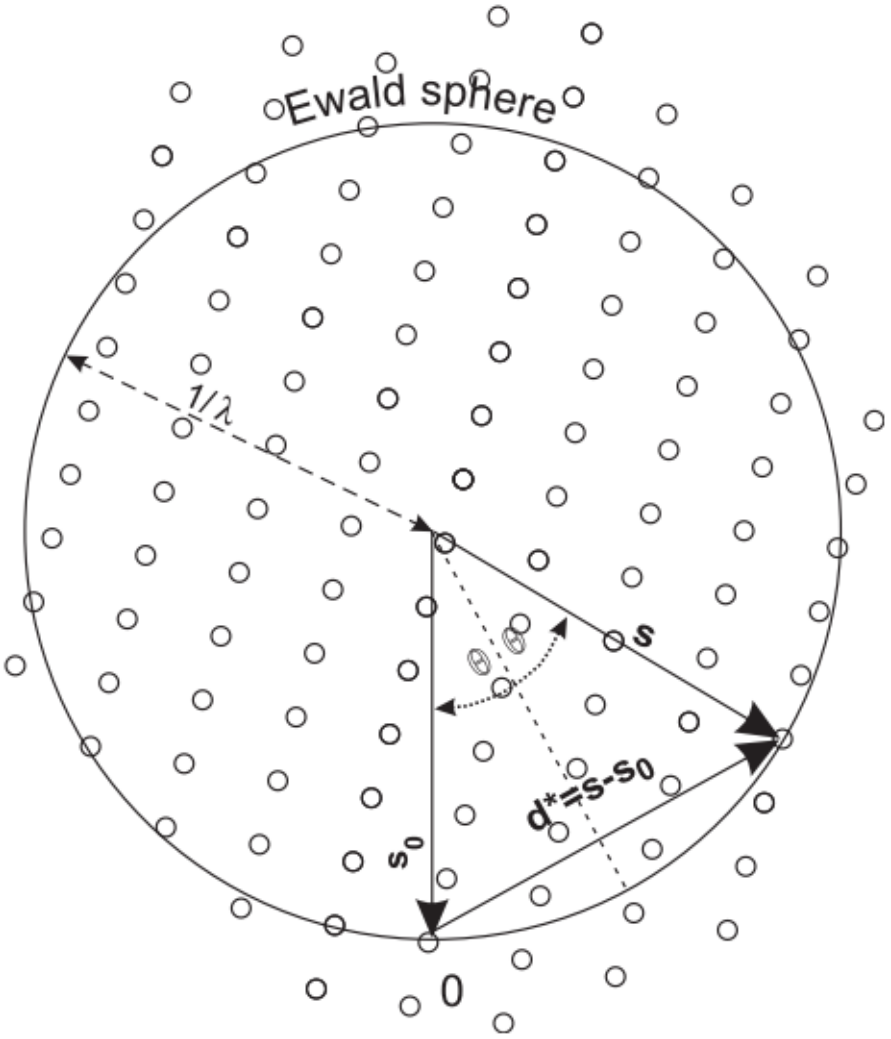

Figure 1.7: Construction of the Ewald sphere, wherein 0 indicates the origin of reciprocal space, s_0 is the incident wave vector, s is the scattered wave vector and d^* is the distance between the two reciprocal lattice points at the tips of the two wave vectors (Dinnebier & Billinge, 2008).

Due to the relationship between real-space and the reciprocal lattice, the reflected intensities contain information regarding distances within the crystal lattice, and the symmetry within the unit cell. The electron density within the crystal is related to the diffracted intensities by Fourier transform, which is described as the structure factor, F_{hkl}, at each reciprocal lattice point hkl. The structure factor for each reflection consists of an

amplitude, $|F_{hkl}|$, and phase, $\emptyset_{hkl}$, and can be determined by Fourier transform of the electron density, by,

$$F_{hkl} = \int_{cell} p(xyz) e^{[2\pi i(hx+ky+lz)]}\, dV \qquad (1.2)$$

where $p(xyz)$ is the electron density and $(hx + ky + lz)$ are the 3-dimensional coordinates of a given atom on a given plane, where the integral is performed for the entire unit cell volume. Conversely, the diffraction pattern may be obtained by performing an inverse Fourier transform on the electron density, by,

$$p(xyz) = \frac{1}{V} \sum_{h,k,l} F_{hkl} e^{[2\pi i(hx+ky+lz)]} \qquad (1.3)$$

where V is the volume of the unit cell. However, only the amplitudes of the scattered intensities can be measured experimentally, and the phase information is lost (Hauptman, 1997). Thus, it is not possible to perform the inverse Fourier transform shown in Equation 1.3. This is known as the *phase problem*, where the structure factors cannot be directly determined from the experiment. A number of methods exist for solving the phase problem, including Direct Methods (Woolfson, 1971), Patterson Methods (A. L. Patterson, 1934) and Dual Space/Charge Flipping (Palatinus, 2013), among others.

1.6.2. Structure determination from X-ray powder diffraction (SDPD)

1.6.2.1. Global optimisation methods

While single crystal XRD is the optimal method for crystal structure solution, in cases where it is not possible to obtain a high-quality single crystal, XRPD may be used for structure determination and refinement, as shown in Chapters 5 and 6. Direct methods and the Patterson method can be used in SDPD where sufficient data exist (Barbas, Kumar, Vallcorba, Prohens, & Frontera, 2020), the lack of spatial resolution accessible and the resulting overlap

of diffraction peaks makes it unfeasible to extract 1000s of independent accurate reflections in most cases concerning molecular solids. This is in part a consequence of their molecular complexity, and low crystal symmetry (David, Shankland, McCusker, & Baerlocher, 2010b). As such, global optimisation methods are typically the chosen technique to solve crystal structures from XRPD data. Global optimisation differs from the standard methods used for structure determination using SCXRD in that they do not attempt to define the electron density from first principles, instead seeking to match the electron density calculated using atomic coordinates in a trial structure and adapting the structural model iteratively until the correct structure is identified.

A significant issue in the use of powder diffraction for structure determination, in comparison to SCXRD, is the loss of information in the powder pattern with respect to a single crystal measurement (Bernstein & Reutzel-Edens, 2019). The number of measurable reflections are often insufficient to fully describe the crystal structure. Thus, structure determination is often not possible without some prior knowledge of the system, such as the known molecular structure of the molecule which reduces the number of independent parameters that need to determined. As such, in global optimisation methods, a trial structure is generated and adjusted over many iterations to improve the fit of the calculated to the experimental data. As such, the structure determination process ends once convergence has been achieved and no better fit to the observed data can be found. Global optimisation seeks to find the global minimum in the search space by trialing a series of structures. This process will be described in the context of simulated annealing (SA), a commonly used global optimisation method implemented within software packages such as DASH (David et al., 2006) and TOPAS (A. A. Coelho, 2003). A number of steps are involved in the SDPD process, which are outlined from Section 1.6.2.2 onwards. Once the structure has

been determined using global optimisation, subsequent structure refinement can be carried out using well established Rietveld methods (A. J. Florence et al., 2008).

1.6.2.2. XRPD pattern indexing

Indexing is the process of determining the unit cell parameters ($a, b, c, \alpha, \beta, \gamma$ and space group) from the peaks in the diffraction pattern (David et al., 2010b). These are calculated by assigning hkl indices to the observed peaks in the diffraction pattern, from which the lattice parameters and space group symmetry may be determined. At this stage in the SDPD process, only accurate peak positions are important, as they relate to d-spacings within the material, whereas the intensities are related to the electron density and are required for structure determination. Examples of software packages which allow indexing of XRPD data include TREOR (Werner, Eriksson, & Westdahl, 1985), X-CELL (Neumann, 2003), DICVOL (Boultif & Louër, 2004) and singular-value decomposition (A. A. Coelho, 2003).

1.6.2.3. Pawley refinement

Once the space group and lattice parameters have been determined, it is necessary to confirm the indexed cell parameters and to extract reflection intensities by performing a Pawley fit (Pawley, 1981). Le Bail fitting (Le Bail, Duroy, & Fourquet, 1988) may also be used, however, it has been found that the Pawley method is more robust mathematically, due to the tendency of the Le Bail method to calculate negative intensities where there has been some over approximation of the background (David et al., 2010b). Pawley refinement methods are model independent, that is, there is no contribution of atomic sites to the calculated intensities, which are freely refined as extracted intensities associated with each

reflection. Therefore, it is only the indexed lattice parameters which are given as input to the fit, which are refined against the experimental data, along with the peak profile, background and zero point. As such, this fitting method gives an indication of the best possible fit to the experimental data, refined unit cell parameters and confirms all observed peaks are described by the unit cell e.g. confirming phase purity.

1.6.2.4. Simulated annealing structure solution

As SA is dependent on chemical information to construct structure models, it is first necessary to define a realistic model of the molecular structure. This can be done using a previously solved structure of the same compound. If this is not available, then models can be constructed via other means, for example using density functional theory (DFT) calculations to optimise the geometry of an input molecule (Altomare, Ciriaco, Cuocci, Falcicchio, & Fanelli, 2017). Once a suitable input has been created, attempts to solve the structure can be made. Due to the loss of data in the 1-dimensional XRPD pattern, it is necessary to limit the number of degrees of freedom (DOF) to improve the data: parameter ratio. Thus, a molecular z-matrix may be constructed to reduce the number of refined parameters (Dinnebier, 2013). This consists of defining rigid constraints between atoms in a molecule, which limits the number of refined parameters from four per atom (x, y and z coordinates, and isotropic temperature factor, b) to six per rigid body (x, y and z translation and rotation), with the additional b-factor for each atom. The molecular z-matrix has the additional benefit of providing necessary constraints on bond distances, angles and torsions, thus ensuring a sensible molecular geometry.

SA consists of generating random crystal structures, which are changed iteratively, and the calculated diffraction compared to the experimental data (Pecharsky & Zavalij, 2005).

The process involves taking each model through a simulated temperature cycle, wherein at higher temperatures, changes which improve and decrease the fit may both be accepted, in order to avoid the structure being trapped in a local minimum before subsequent refinement and elimination of poorly fitting structures at lower temperatures. The exact temperature regime, that is, the cooling rate and starting temperature, are largely dependent on the complexity of the input molecule. For example, a large molecule with a high number of flexible rotatable bonds may require a higher starting temperature or slower cooling rate, or both, to allow a larger search of the space, and increase the chances of finding the global minimum, as more complex structures have a higher probability of getting stuck in a local minimum if an insufficient search is performed.

There are a number of related statistical evaluations to assess the suitability of the model structure. Ultimately, the goal is to minimise the ratio $X^2_{profile}$: X^2_{Pawley}, wherein a value less than 10 is typically considered sufficient (Alastair J. Florence et al., 2005). This ratio evaluates the goodness of fit of the SA run by determining whether the attempted structure solution has resulted in a realistic structure model. Given that X^2_{Pawley} is the lowest value expected for the solution, as the Pawley fit gives an indication of the best possible fit to the data, this ratio should not be lower than one. If values < 1 are obtained, it is necessary to evaluate the structural model and consider that the data have been "over-fitted". Similarly, large values of this ratio should also be cause for concern, as reasonably similar values for $X^2_{profile}$ and X^2_{Pawley} should be obtained, assuming the global minimum has been found and there is minimal disorder or preferred orientation in the sample.

1.6.2.5. Rietveld refinement

Rietveld refinement consists of using an existing structural model to replicate the observed diffraction data, by varying structural and instrument profile factors by a least-squares refinement process (Rietveld, 1969). The step is performed after successful structure solution, using a least squares minimization of the differences between the observed and calculated structures. As with SA, it is important to minimise the number of refined parameters, in the case of covalent solids, by the use of a molecular Z-matrix. While additional information may be obtained via Rietveld refinement, such as grain size or surface roughness (Dinnebier, Leineweber, & Evans, 2019), it is typically only the background and scale factors, and structural parameters (translation and rotation, temperature factors, and flexible torsions) that are refined. Structural refinements to PDF data are also possible, as described in Section 1.8.3.3.

1.6.3. XRPD and amorphous solids

While XRD and XRPD may be used for structure solution and refinement when the material of interest is a crystalline solid, amorphous solids and liquids present a considerable challenge to diffraction and XRPD, due to the nature of Bragg's law (Bragg & Bragg, 1913). As such, they require more advanced methods to study their molecular packing. Nanocrystalline solids also pose a significant challenge when applying standard diffraction methods (S. J. L. Billinge & Kanatzidis, 2004), where the small coherent scattering domain size results in significant line-broadening in the resulting XRPD pattern (Figure 1.1). Such effects in experimental data significantly hinders the ability to resolve individual reflections, and reliably determine peak positions and areas or reflection intensities. Therefore, the standard methods for extracting structural information fail, and we must turn to other techniques to

explore structure. XRPD is used extensively in this work for collection of raw data before transformation to the pair distribution function (see Section 1.8 onwards).

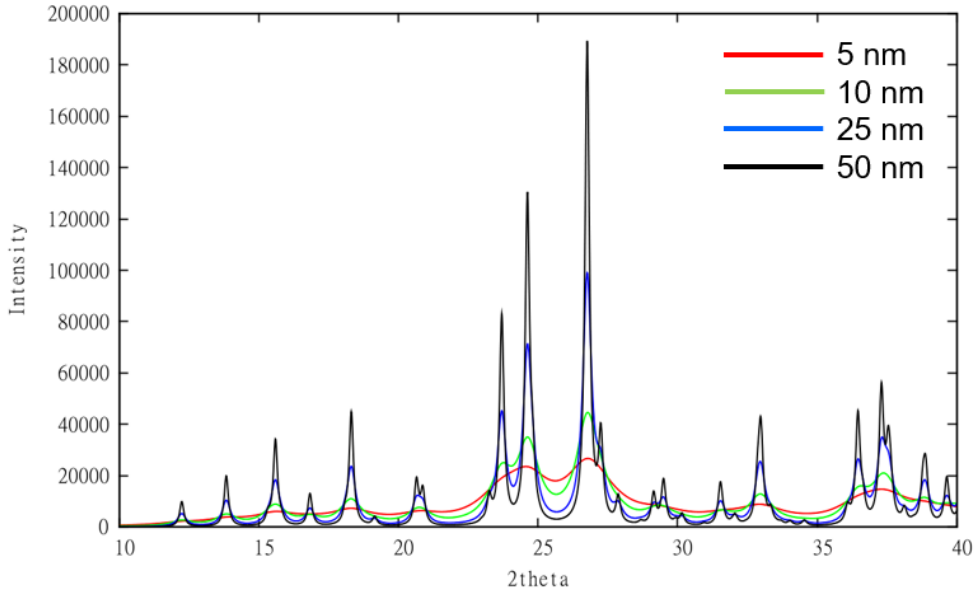

Figure 1.8: Simulated XRPD of paracetamol Form I with different domain sizes. XRPD patterns calculated using TOPAS academic V6 (A. A. Coelho, 2003; Alan A. Coelho, 2018).

Structural analysis is of particular importance, given the occurrence of process-induced amorphisation and intentional nano-sizing that can occur in a number of pharmaceutical processes such as spray drying, milling or compaction (P. A. Priemel, Grohganz, & Rades, 2016). In particular, the wide implementation of size-reduction methods to reduce particle size, such as ball milling (Patel, Baria, & Patel, 2008), have the potential to produce nanocrystalline and amorphous forms of APIs (Bøtker et al., 2011). Various other methods exist for preparation of the amorphous phase, including melt-quenching, spray drying and lyophilisation (Karmwar, Graeser, Gordon, Strachan, & Rades, 2011; Terban, Cheung, Krolikowski, & Billinge, 2016).

1.7. Characterisation of amorphous materials

As described in Section 1.6.3, the standard methods for assessing structure in amorphous solids and liquids, such as XRD/ XRPD, fail, due to the lack of 3-dimensional periodicity. As such, a range of analytical techniques have been used to understand the kinetic and thermodynamic properties of amorphous pharmaceuticals. These include thermal and spectroscopic analyses of amorphous solids, which will be briefly described below.

1.7.1. Thermal analysis of amorphous solids

Differential scanning calorimetry (DSC) and thermogravimetric analysis (TGA) are two thermal analysis methods typically used for studying amorphous pharmaceuticals. Both techniques are applied in Chapter 4 for detecting phase separation and water uptake in formulated drug-polymer systems. DSC consists of applying a controlled temperature profile to a sample, to observe thermally induced transitions within the material (Müllertz, Perrie, & Rades, 2016). These changes are typically observed by measuring the input energy required to maintain the sample at the same temperature, in comparison to a reference, for example an empty sample holder. Important information regarding the physical state of the material can be obtained therefore from the exothermic or endothermic events that can be observed, and as such, the kinetics and thermodynamics of the phase transitions can be studied.

In the context of amorphous pharmaceutical solids, thermal events can give an indication of the stability of the amorphous phase (Figure 1.6). At the simplest level, the observation of a crystallisation exotherm gives a clear indication that the sample is not physically stable against crystallisation. However, the low temperature inflection shown in Figure 1.6 shows the glass transition temperature, T_g. This relates to an increase in molecular mobility, taking

the amorphous solid from a glassy material to a rubbery state (Papon et al., 2006). The glass transition is referred to as being a kinetic transition due to the lack of a first order phase transition between the glassy and supercooled state. This is apparent in Figure 1.6 wherein a subtle shift in the DSC trace is observed, in contrast to the crystallisation (exothermic) and melting (endothermic) peaks shown, which allow the calculation of the Heats of crystallisation and fusion respectively.

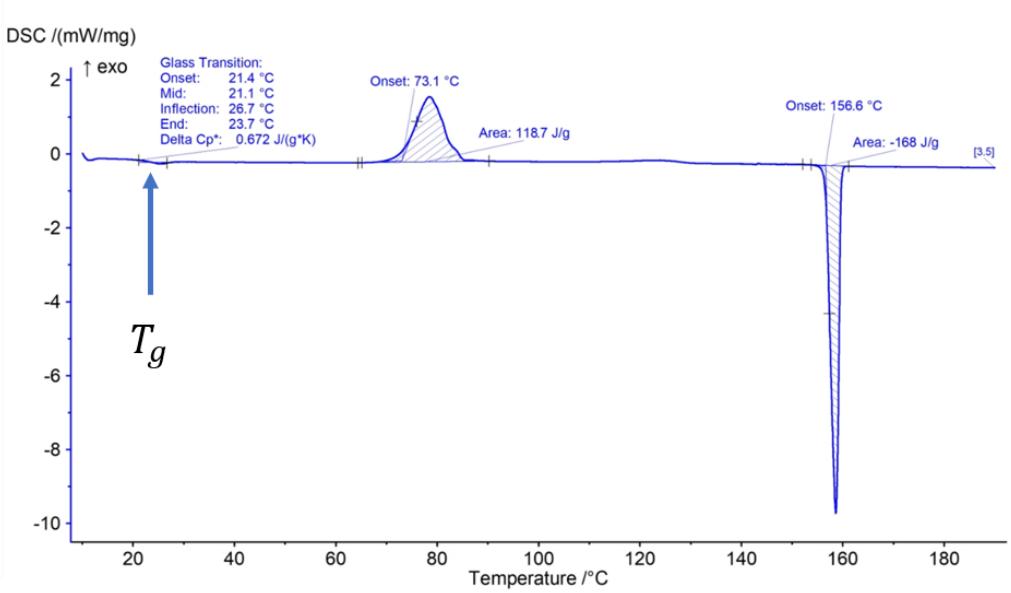

Figure 1.9: Example DSC plot showing heating of crash-cooled amorphous paracetamol. Low temperature inflection at ~ 21 °C shows T_g, exothermic event at ~ 73 °C shows crystallisation, and endothermic event at ~ 157 °C shows melting of the crystalline material.

TGA is similar to DSC in that a sample is taken through a temperature regime and its response in comparison to some reference (again, an empty sample holder) is measured. Contrary to DSC, it is the change in the mass of the sample which is measured (L'vov, 2007) (Figure 1.10), instead of the energy given out or taken up by the sample. However, some instruments are capable of measuring both simultaneously. TGA has a range of applications,

from assessing water absorption (Pyramides, Robinson, & William Zito, 1995) to degradation of materials at high temperature (De Oliveira, De Menezes, & Catharino, 2015).

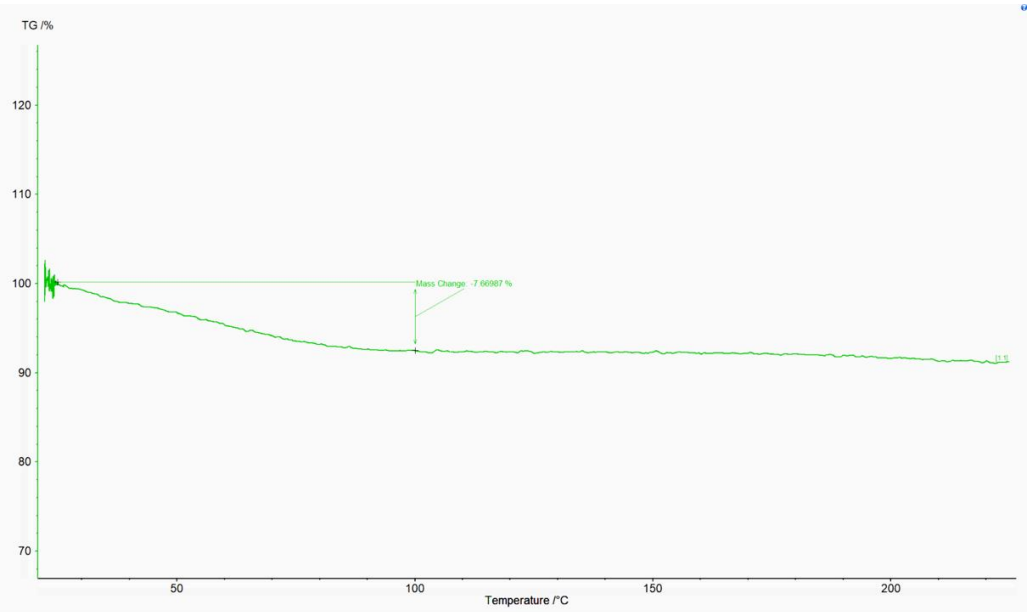

Figure 1.10: Example TGA plot showing heating of PVP polymer. Mass loss between 25 °C and 100 °C corresponds to evaporation of absorbed water.

1.7.2. Spectroscopic analysis

Spectroscopic techniques such as solid-state nuclear magnetic resonance (ssNMR), FT-IR and Raman spectroscopy have been extensively applied in the context of pharmaceuticals to understand structure in amorphous and crystalline solids (Hédoux, 2016; Zimmermann & Baranović, 2011). In the context of his work, FT-IR is used for confirming sample consistency and detecting moisture in formulated materials in Chapter 4. Non-bonded interactions may be probed to compare structure in crystalline and amorphous pharmaceuticals (Zimmermann & Baranović, 2011) and interactions in amorphous drugs prepared via different methods, where it has been demonstrated that the structure of amorphous indomethacin is dependent on the preparation technique, detected using Raman

spectroscopy, among other techniques (Karmwar et al., 2011). Spectroscopic methods have been applied extensively to characterise the structures of amorphous APIs dispersed in polymeric matrices (Fini, Cavallari, & Ospitali, 2008; Meng, Trivino, Prasad, & Chauhan, 2015; Rumondor, Ivanisevic, Bates, Alonzo, & Taylor, 2009; Zhao, Barker, Belton, McGregor, & Craig, 2012), to understand stability and molecular interactions in formulated systems.

1.8. Pair distribution function

The pair distribution function (PDF) describes the probability of two atoms being separated by a given distance (Takeshi & Billinge, 2012b). The PDF is essentially a histogram of interatomic distances within a sample of interest (Figure 1.11). As such, the power of the PDF method is that it can be applied regardless of physical state, such that structural information can be collected from crystalline solids, amorphous materials, liquids and gases. PDF analysis is a key analytical technique used extensively in this work to probe the structure of amorphous pharmaceutical solids.

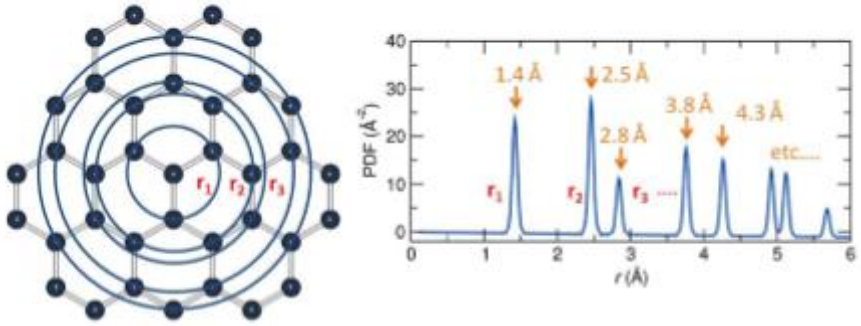

Figure 1.11: Diagram showing radial distances from central atom in graphene model. PDF shows peaks representative of interatomic distances in the graphene structure.

1.8.1. Definitions of the pair distribution function

There are a number of definitions of the PDF, which have been described extensively elsewhere, and which are all described as the PDF, and as $G(r)$ (Keen, 2001). For the purposes of this work, the three most commonly used will be described. As all are commonly referred to as the PDF, they shall at this stage be defined individually as $G(r)$, $T(r)$ and $D(r)$, with the individual derivations and definitions described herein.

Although there are a number of definitions of the PDF, for the purposes of this work, we define the PDF as the function identified as the differential correlation function, $D(r)$, (Keen, 2001), defined as:

$$D(r) = 4\pi r p_0 G(r) \qquad (1.4)$$

Where p_0 is the average number density of the material, r is the interatomic distance in Å and $G(r)$ is the total radial distribution function, which is equal to

$$G(r) = \sum_{i,j=1}^{n} c_i c_j f(Q)_i f(Q)_j \left[\frac{n_{ij}(r)}{4\pi r^2 dr\, p_j} - 1\right] \qquad (1.5)$$

Where the sum goes over all atoms i and j in the material. The proportions of atoms i and j are represented by c_i and c_j, respectively. Additionally, $f(Q)_i$ and $f(Q)_j$ are the atomic form factors of atoms i and j, respectively. The number of atoms of type j that are between distances of r and $r + dr$ from atoms of type i is given by $n_{ij}(r)$, and p_j is equal to the proportion of atom j multiplied by the average number density. There is a third correlation function which is less commonly used in crystallography: the total correlation function, $T(r)$, which is defined as

$$T(r) = D(r) + T^0(r) \qquad (1.6)$$

where

$$T^0(r) = 4\pi r p_0(\sum_{i=1}^{n} c_i f(Q)_i) \quad (1.7)$$

All definitions of the PDF contain the same structural information, however, as will be demonstrated, they change by multiplicative or additive constants (usually related to the atomic number density, p_0), which can have the effect of changing the background or scaling in the resulting dataset. For the purposes of the data and analysis herein, the function $D(r)$ is the most suitable. The $D(r)$ function is preferred over $T(r)$ and $G(r)$ for a number of reasons. Firstly, $D(r)$ itself oscillates around zero, which is related to the average number density of the material (essentially, the function starts at zero and ends at zero), this makes the data easy to visually analyse, and makes the visual assessment of any fitting of the data more simple. The same can also be said for $G(r)$, however, the use of $D(r)$ is preferred due to the multiplication of $G(r)$ by $4\pi r p_0$ (equation 1.1), which emphasises the high-r region of the PDF, which is particularly useful for the study of amorphous materials, where the signal at higher-r values is weak, and attenuates to zero at low r values, due to the lack of long-range order. This behavior differs from that of the total correlation function, $T(r)$, which, due to the addition of $T^0(r)$, as per equation 1.4, adds an r-dependent background to the data, which increases almost linearly with r. Thus, the data starts at zero, but ends at ∞, which makes it difficult to simultaneously observe the full range in $T(r)$.

An additional benefit of the use of $D(r)$ is again due to the multiplication by $4\pi r p_0$, as below the shortest atom-atom distance in the data, the function behaves like -$4\pi r p_0$, which provides a negative-sloping background in the data (Takeshi & Billinge, 2012b). This has the benefit of providing extra confidence in any potential fit to the data, as it will allow determination of whether the correct number density has been used in the experimental PDF transformations, as the p_0 value has an effect on the low-r slope in the PDF, and subsequently, the amplitudes of observed features in the data. This can easily be determined

by visually comparing the experimental and refined datasets. Examples of $G(r), D(r) and\ T(r)$, calculated from experimental total scattering data from PCM Form I, are shown in Figure 1.4, where the key differences between each correlation function are apparent.

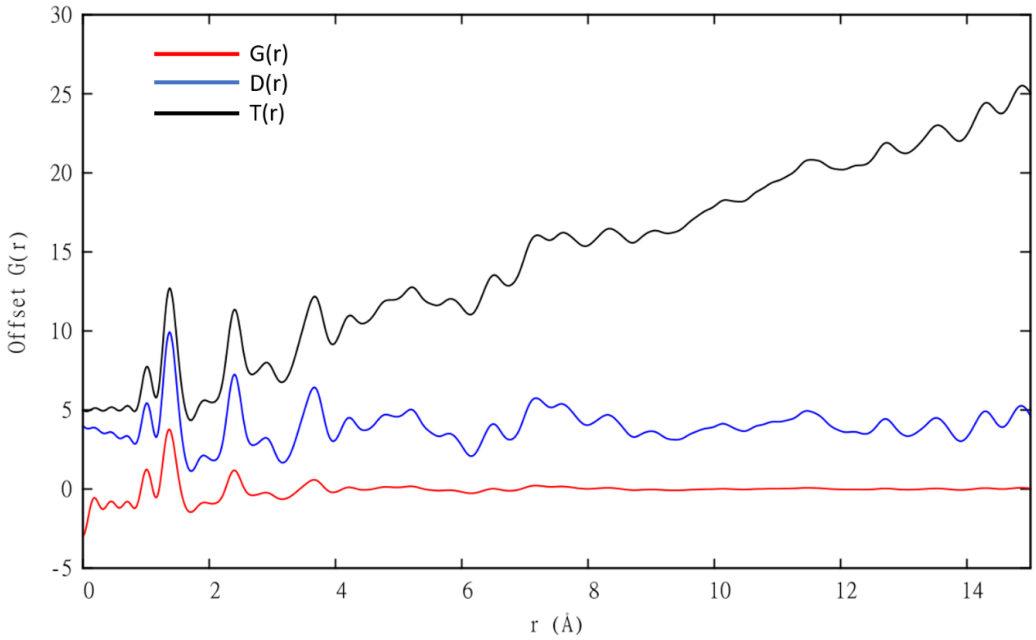

Figure 1.12: Stack of G(r) (red), D(r) (blue), and T(r) (black) for PCM Form I, calculated in GudrunX (A K Soper, 2011).

1.8.2. Calculation of the pair distribution function from total scattering data

As described in Section 1.1, the PDF is produced by the Sine Fourier transform of total scattering data. A number of software packages exist for the calculation of PDFs from X-ray total scattering data, including GudrunX (Alan K Soper, 2011) and PDFgetX3 (Juhás, Davis, Farrow, & Billinge, 2013). As this work focusses on the implementation of X-ray PDF, henceforth all uses of the term PDF will imply the use of X-rays, unless otherwise stated. The term "total" scattering implies that all intensities collected during the XRPD experiment are used. This applies to all elastic scattering from the material of interest, that is, the Bragg and

diffuse intensities. This differs from standard XRPD in that the diffuse scattering in the experiment are usually disregarded, and are considered to be the "background" and are fitted as such (Young & Goodwin, 2011). However the diffuse scattering contains important information describing deviations from the average structure, when studying crystalline materials, and describes an "average" local structure in amorphous solids.

In standard experiments, scattered intensities are described as a function of scattering angle, 2θ. As "intensities" in the PDF are described as a function of interatomic distance, r, it is necessary to convert the 1-dimensional XRPD pattern to an absolute scale. Thus, after subtracting the contribution to the scattering from the background (sample holder, air scatter) the first step in obtaining the PDF is in converting 2θ to Q, which describes the magnitude of the scattering vector, dictating the real-space experimental resolution, by

$$Q = \frac{4\pi \sin\theta}{\lambda} \qquad (1.8)$$

where θ is half the scattering angle, and λ is the wavelength of the X-ray source. The next step is to correct for the atomic form factor, or scattering power of the atomic constituents of the sample, to determine the total scattering structure function, $S(Q)$, by

$$S(Q) = \frac{I(Q)}{\langle f(Q)\rangle^2} \qquad (1.9)$$

where $I(Q)$ represents the intensities from the original diffraction experiment, corrected for background, and $f(Q)$ is the atomic form factor, with $\langle \ldots \rangle$ indicating a compositional average across the sample. Subsequently, a number of corrections are applied to remove the contributions due to Compton scattering, sample polarisation, multiple scattering and absorption, amongst others. Finally, the PDF is obtained by,

$$G(r) = \frac{2}{\pi}\int_{Q_{min}}^{Q_{max}} F(Q)\sin(Qr)\,dQ \qquad (1.10)$$

Where the Fourier transform is restricted by the upper and lower limits in Q. The result of the Fourier transform is essentially a histogram of interatomic distances. Termination of the Fourier transform at some finite value has the unfortunate effect of introducing spurious ripples throughout the real-space data (Masadeh, 2016), which can affect interpretation of the PDF. This will be explained in greater detail in the next section. A simplified graphic demonstrating the data collection and treatment process is shown in Figure 1.5.

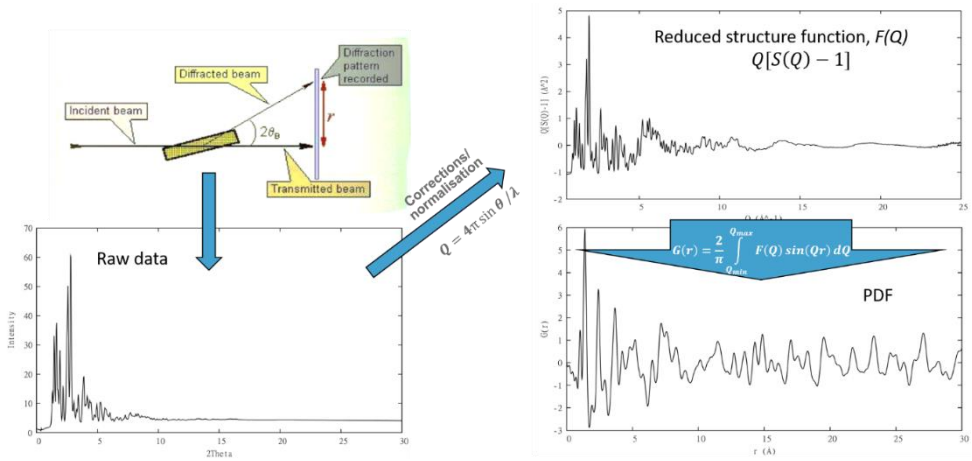

Figure 1.13: Graphic showing data collection and normalisation/treatment process to obtain the PDF. Data shown are for paracetamol Form I.

1.8.2.1. Q_{max} and real-space resolution

The importance of Q_{max} cannot be understated for a number of reasons. First and foremost, Q_{max} dictates the resolution in the PDF, Δr, which is approximately equal to $\frac{2\pi}{Q_{max}}$ (Takeshi & Billinge, 2012b). This differs from typical XRPD experiments in that it is the step size in 2θ that determines the reciprocal-space resolution, which is ideally kept as small as possible, giving a smaller full width at half maximum (FWHM) in the Bragg peaks , which aids in resolving peak positions and areas. As such, in XRPD it is appropriate to use an X-ray source with a relatively long wavelength, such as a copper anode, the X-rays of which scatter the

same reflection to a higher angle than a source with a lower wavelength. For PDF analysis, as the resolution is dependent on Q_{max}, reciprocal-space resolution is often sacrificed to extend the Q-range of the experiment, by having a small sample-detector distance (Jacques et al., 2013), and larger step-size in reciprocal space. This has the effect of maximising the angular range which can be measured, and simultaneously increasing the collection time per step. This in turn maximises Q_{max}, as shown in equation 1.1., and enables optimisation of counting statistics across the full angular range. However, the poor reciprocal-space resolution, although not overly detrimental to the PDF, does result in a damping of features within the real-space data, as shown in Figure 1.4., where this is particularly noticeable at high-r values.

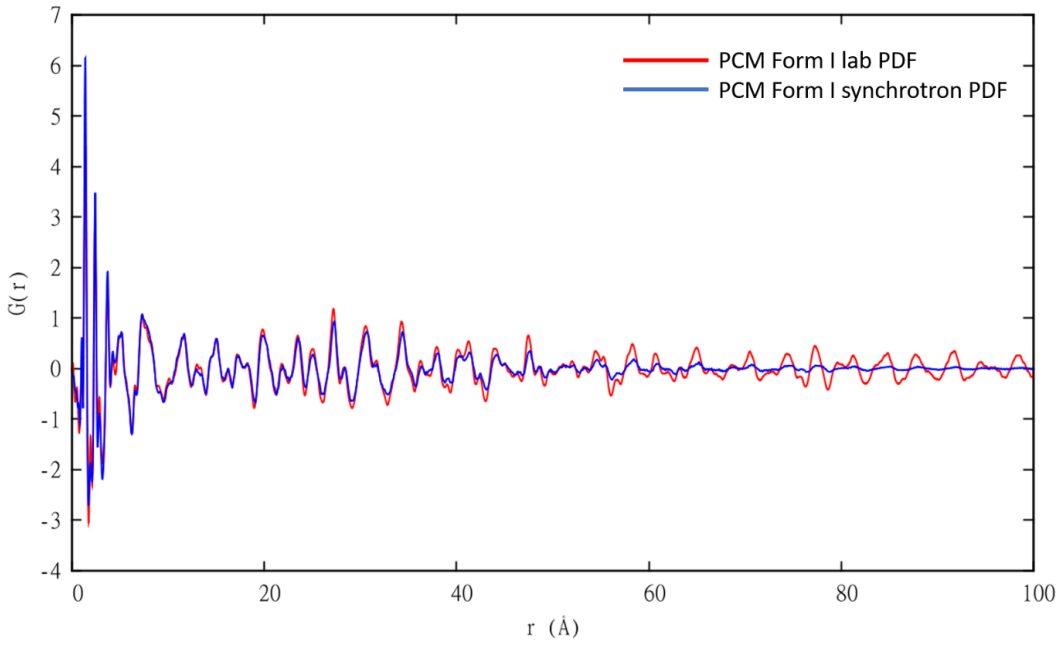

Figure 1.14: Comparison of experimental PDF data calculated from total scattering measurements of PCM Form I from molybdenum laboratory (red) and synchrotron (blue) sources. Poorer reciprocal space resolution in the synchrotron data translates to a damping of features in the PDF, particularly at high-r.

Also demonstrated in equation 1.1, is that Q_{max} is heavily dependent on the wavelength of radiation used in the measurement. As such, the copper anode typically used for XRPD is not suitable for PDF analysis, as the resolution offered by the copper source is not sufficient for real-space analysis. Figure 1.11 demonstrates this, where the PDFs of amorphous paracetamol and indomethacin have been calculated with a Q_{max} comparable to that accessible using a copper anode (8 Å^{-1}), the two plots are very similar, and extraction of any meaningful information from this data is highly unlikely. This is in stark contrast to Figure 1.12, where a Q_{max} of 25 Å^{-1} has been used on the same dataset in the Fourier Transform. The higher real-space resolution results in two distinct PDFs, which is clearly better for accurate fingerprinting, and subsequent analysis of the data.

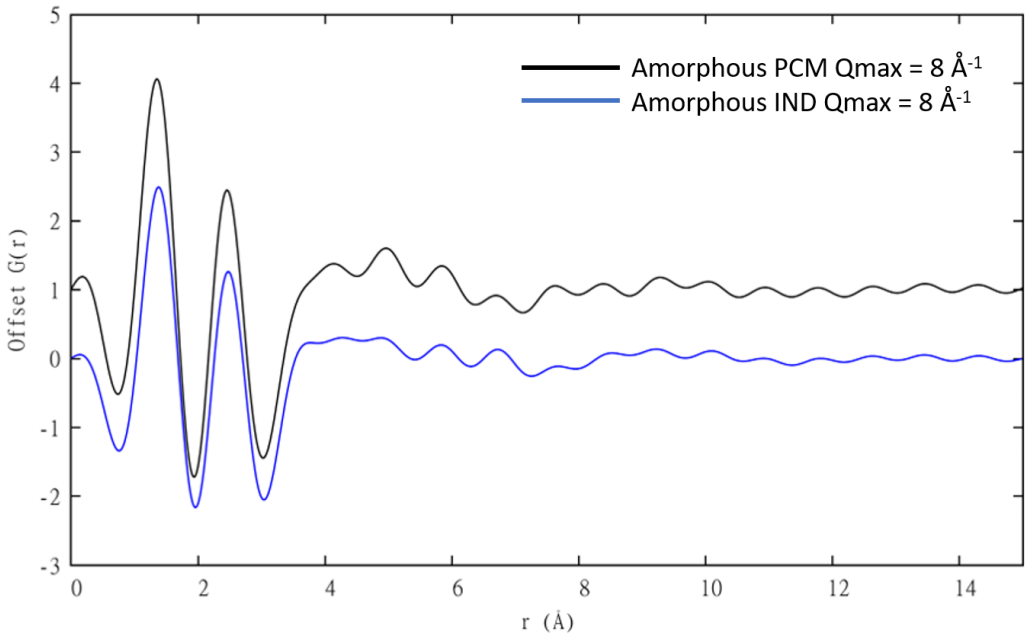

Figure 1.15: PDFs of amorphous paracetamol and indomethacin, calculated with a Q_{max} of 8 Å^{-1}, from experimental synchrotron total scattering data, using GudrunX (Alan K Soper, 2011).

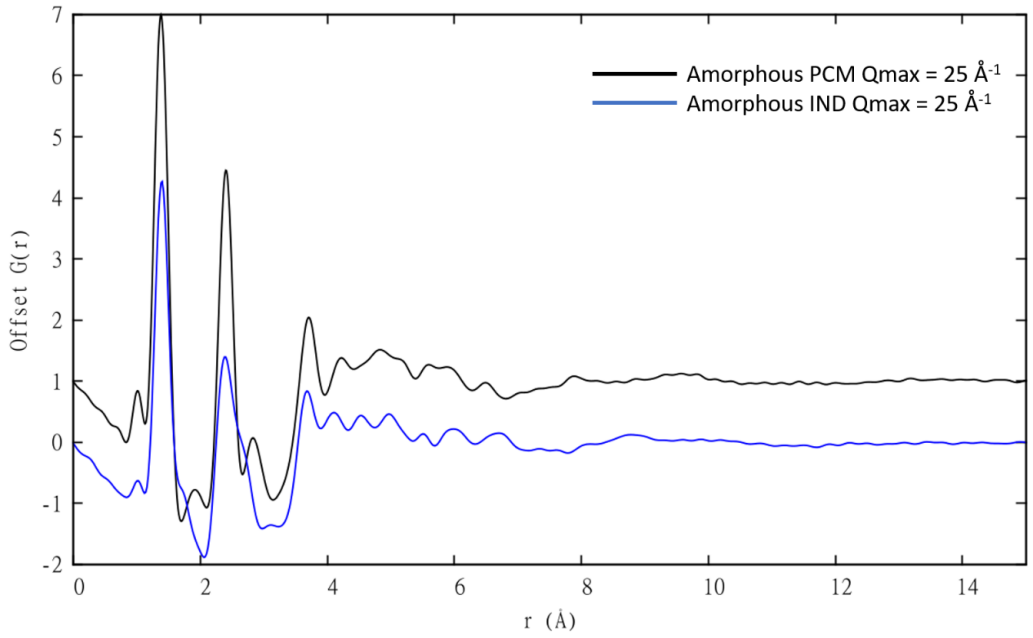

Figure 1.16: PDFs of amorphous paracetamol and indomethacin, calculated with a $Q_{max}$of 25 Å^{-1}, from experimental synchrotron total scattering data, using GudrunX (Alan K Soper, 2011).

It has been recommended that for a unique PDF, a Q_{max} of at least 12.5 Å^{-1} is required (Dykhne, Taylor, Florence, & Billinge, 2011), meaning that, as a bare minimum, a molybdenum anode is necessary for generating PDF-quality data on a laboratory instrument, as shown in Table 1.2.

Table 1.3: Wavelengths and accessible Q_{max} of common laboratory X-ray sources. Q_{max} values assume a $2\theta_{max}$ of 180°.

X-ray source	$K\alpha 1$ wavelength (Å)	Approximate Q_{max} (Å^{-1})
Copper	1.5406	8
Molybdenum	0.7093	17
Silver	0.5594	22

Aside from the effect of Q_{max} on the resolution of the PDF, extending the Q-range of the measurement also reduces the impact of the ripples described in Section 1.9.2., that are generated by the termination of the Fourier integral at finite value. In the ideal experiment, the integral in equation 1.3. has upper and lower limits of ∞ and $-\infty$, respectively. As Q approaches ∞, the scattered intensities drop to zero due to form factor fall-off and the Debye-Waller factor, thus ensuring a smooth drop-off in intensity at Q_{max}. However, in reality, this is impossible due to restrictions in the instrument geometry and wavelength of the source. Thus, termination at some finite value introduces a step-change in the Fourier transform. This ultimately results in the integration of a box-like function, the Fourier transform of which is a *sinc* function (Latry & Rouge, 2000) (Figure 1.2). These ripples are of the form $\frac{\sin(Q_{max}r)}{r}$ (Proffen & Billinge, 1999), and as shown, are particularly troublesome at low-r. The best method to reduce the impact of termination ripples in the PDF is to collect to as high a Q_{max} as possible, and with good statistics on the data. However, as it is impossible to fully eliminate these ripples, a popular method that has emerged is the Soper-Lorch convolution (Alan K. Soper & Barney, 2012), which introduces a broadening in the real-space data, ΔM, which thus minimises the detrimental effect of termination effects. An obvious drawback of this method is of course that all features in the resulting $G(r)$ data are broadened, so care must be taken to ensure a suitable trade-off between minimisation of termination ripples, and resolution of true structural features in the PDF.

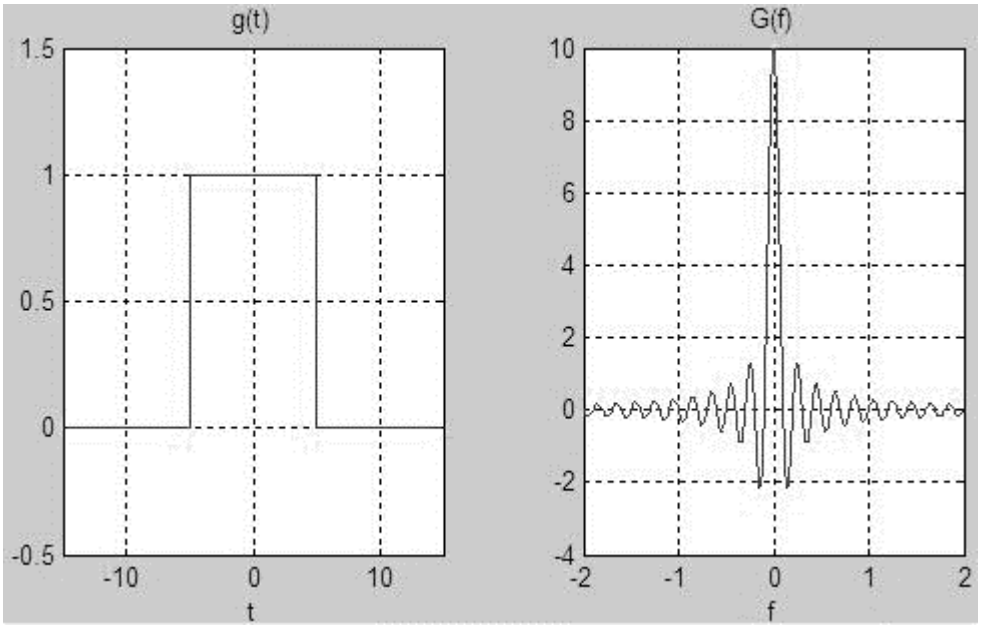

Figure 1.17: Example of a Fourier transform of a box function (Wolberg, 1994).

1.8.3. Analysis of PDF data

While structure solution from PDF data has been attempted to some degree (Juhás et al., 2010; Prill, Juhás, Billinge, & Schmidt, 2016), it is not something that has been routinely implemented, and the same issues facing SCXRD and XRPD exist in structural understanding of amorphous solids using PDF. There are a number of methods which can be used to extract structural information from PDF data. The chosen method largely depends on the physical state of the material, and the length-scale over which the features of interest is observed. As with XRPD, the PDF can be used purely as a fingerprint of the material of interest (Dykhne et al., 2011), but more in-depth analysis may also be performed, from statistical analysis of the PDF (Chapman, Lapidus, & Chupas, 2015), to fitting the data to refine a structural model (Prill et al., 2016). Fitting of PDF data can largely be split into two types: small-box and big-box modelling, which shall be explained in more detail in Sections 1.9.3.4. and 1.9.3.5., respectively.

1.8.3.1. Fingerprinting and statistical analysis of PDF data

While the application of the PDF technique has been implemented to a great extent within the inorganic community (Bell et al., 2008; Jiang, Grande, & Selbach, 2017; White, Provis, Proffen, Riley, & Van Deventer, 2010), the use of the method in the structural assessment of organic solids has been limited. This is probably due in no small part to the relative complexity of even simple organic molecules in comparison to many inorganic systems. As such, in studies of pharmaceuticals, the PDF has served as a fingerprint of the material, with minimal structural information extracted from the data (Dykhne et al., 2011)

While the fingerprinting method seems minimal in terms of its information content, correlations between experimental datasets can be determined to study trends in data or to assess similarity between amorphous and crystalline forms of a material, and to track changes as a function of time, temperature or some other stress method. Principal component analysis (PCA) has been applied to PDF data of inorganic materials to evaluate structural differences between metal chloride, nitrate and sulphate solutions, changes during chemical reactions and structural transformations (Chapman et al., 2015), and in pharmaceutical systems, to infer information regarding the physical state of polymeric dispersions of amorphous drugs (Chieng et al., 2013) and to assess the extent of structural disorder in compressed and milled crystalline forms of APIs (S. Chen, Sheikh, & Ho, 2014).

The software package PolySNAP (Barr, Dong, & Gilmore, 2004b) offers a comprehensive suite of statistical tools for analysis of multiple data formats, including PDF data, Raman, FT-IR and XRPD, among others. The software performs full profile matching by generating a correlation matrix and determining Pearson correlation coefficients (Szczepańska, 2011) between each dataset, which is then used to cluster the data for

identification of trends. Among the tools available in PolySNAP are PCA, Pearson correlation analysis and multidimensional scaling.

1.8.3.2. Calculating the PDF from a structural model

The PDF can be calculated from a structural model by (S. J. L. Billinge, 1998):

$$G_{calc}(r) = \frac{1}{r}\sum_i \sum_j \frac{f(Q)_i f(Q)_j}{\langle f(Q)\rangle^2} \delta\left(r - r_{ij}\right) - 4\pi r p_0 \qquad (1.11)$$

where the sum goes over all atoms i and j, and r_{ij} is the distance between the ith and jth atoms. A number of programs now exist which allow this functionality, such as TOPAS V6 (A. A. Coelho, Chater, & Kern, 2015; Alan A. Coelho, 2018) and PDFfit/ PDFfit2/ PDFgui (Farrow et al., 2007; Proffen & Billinge, 1999). This calculation allows a quick visual comparison of the experimental PDF, and can be used when applying the PDF as a fingerprinting tool (S. J. L. Billinge et al., 2010). This is particularly useful when dealing with a high throughput of data, and can be complemented with software packages such as polySNAP (Barr et al., 2004b), which can automate comparisons and statistical analyses of the data.

1.8.3.3. Refinement of structural models to experimental PDF data

Further analysis based on the calculation of a PDF from a structural model is possible. The same software packages mentioned in Section 1.9.3.1. also allow a Rietveld refinement (Rietveld, 1969b) style of fit to experimental PDF data, wherein the goal is to minimise the difference between the observed, G_{obs}, and calculated PDF, G_{calc}, by changing structural or instrumental parameters to improve the fit to the experimental data. One such

commonly implemented method for quantifying the quality of a given fit is to determine the weighted-profile residuals factor, r_{wp}, by (Peterson, Božin, Proffen, & Billinge, 2003),

$$r_{wp} = \left[\frac{\sum_{i=1}^{n} w_i(G_{obs}-G_{calc})^2}{\sum_{i=1}^{n} wi(G_{obs})^2}\right]^{1/2} x\ 100\% \qquad (1.12)$$

where w_i is the weighting factor. The smaller the r_{wp}, the more closely the refined data matches the experimental. The parameters which may be refined in PDF refinements are in most instances identical to a Rietveld refinement (Dinnebier et al., 2019), and include atomic coordinates (or rigid body translation and rotation for molecular/ covalent materials), lattice parameters, isotropic displacement parameters (anisotropic displacement parameters in the case of PDFfit), scale, domain size and zero-error. As with Rietveld refinements of XRPD data, it is important to minimise the number of refined parameters in order to increase confidence in the resulting fit (Alastair J. Florence et al., 2005), and minimise the computational expense of the refinement. For PDF fits, the latter is significantly more of an issue than with XRPD, due to the number of interactions being modelled, as the number of pair interactions increases as the square of the number of atoms (A. A. Coelho et al., 2015).

There are instrumental factors which also affect the PDF peak shape and amplitude. In practice, as these are instrumental parameters, these should not be freely refined in a given PDF fit, and as such should be refined against a measurement standard, as with XRPD (Cline, Mendenhall, Black, Windover, & Henins, 2015). These factors are related to the reciprocal-space resolution, where damping and broadening (Q_{damp} and Q_{broad}, respectively) of features are observed in real-space, due to the typically poorer Q-space resolution used in PDF measurements (Qiu, Božin, Juhas, Proffen, & Billinge, 2004).

1.8.3.4. Small-box modelling

The small-box modelling method involves fitting the PDF based on a "small" box (typically a single unit cell) containing a small number (up to a few-hundred) of atoms (S. J. L. Billinge, 1998). All PDF refinements shown herein are performed using the small-box method. This is the approach employed in software packages such as TOPAS and PDFgui. As these refinements typically fit a small unit cell under periodic boundary conditions (PBCs), these methods are better suited to crystalline materials. While the refinement capability of this method is inherently limited for understanding disorder in large systems, due to the small box size, it has been implemented extensively for understanding local deviations from the average crystallographic structure in a number of studies (Jurkiewicz et al., 2018; Kumara et al., 2016; White et al., 2010).

1.8.3.5. Big-box modelling

As the name suggests, the "big-box" modelling approach consists of refinement to a supercell of a known structure, again under PBCs. This specific refinement method is not used in this work. Big-box modelling originated with the development of the Reverse Monte-Carlo (RMC) method (McGreevy & Pusztai, 1988), wherein a starting configuration (e.g. the supercell of a known phase) is provided, which is then iteratively changed by movement of a single atom/molecule. Whether a change to the structure is accepted or not is dependent on the change to the resulting fit. The process is repeated until the refinement can no longer be improved.

Several packages have been developed which implement the RMC method, including empirical potential structure refinement (EPSR) (A.K. Soper, 1996; Alan K. Soper, 2011), RMCprofile (Tucker, Keen, Dove, Goodwin, & Hui, 2007) and fullrmc (Aoun, 2016). The RMC

method is particularly useful for local structural studies, as the large cell used can assist in understanding correlated disorder across length-scales. Additionally, it is common for these programs to perform the refinement process against the corrected total scattering data, and the Fourier transform to the PDF is subsequently performed for the calculated and experimental data, allowing comparison in both reciprocal and real-space. EPSR and more recent versions of RMCprofile have implemented Lennard-Jones potentials (see Section 1.9) to minimise the energy as part of the refinement procedure. It is common for these software packages to exclude the contribution of thermal motion in the calculation of the PDF, as the large size of the cell allows for the contribution of correlated motion to peak widths to be replicated in the model. An example of cells used in small and big-box refinements is shown in Figure 1.14.

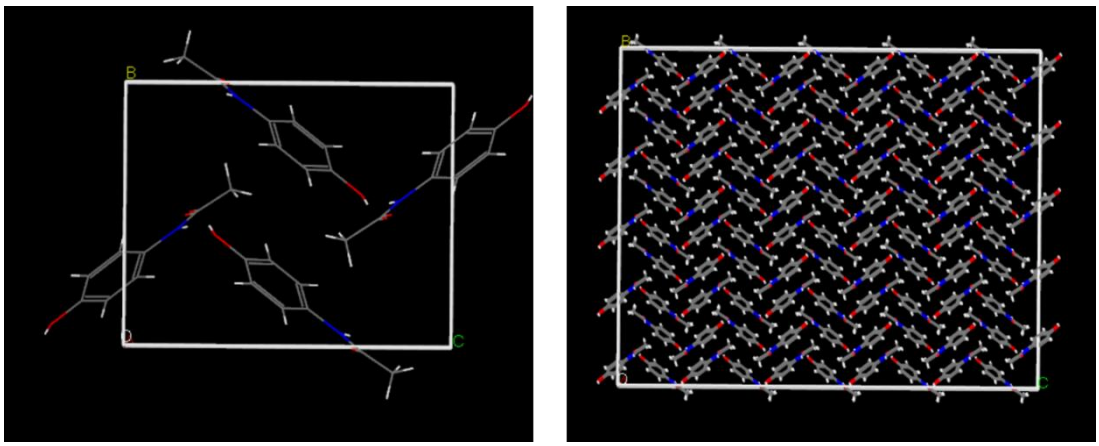

Figure 1.18: Comparison of single unit cell used in small-box refinement methods (left), and a 5 x 5 x 5 supercell used in large-box refinements (right). Both images are of the crystal structure of PCM Form I.

1.9. Molecular dynamics simulations

Molecular dynamics (MD) simulations are a computational tool for analysing the movements and interactions of atoms or molecules by solving Newtons equations of motion

for a system of moving particles (Kamberaj, 2020). MD simulations are used in Chapters 5 and 6 to understand structure in amorphous pharmaceutical solids. Fundamentally, MD packages calculate the forces acting upon particles as a function of some external perturbation, which for a single particle equates to Newton's second law of motion, wherein the force, F, exerted upon some particle, i, is equal to its mass, m multiplied by acceleration, a, (Rapaport, 2004),

$$F_i = m_i a_i \qquad (1.13.)$$

The force acting between two particles, F_{ji}, is a function of the potential energy, V, and the distance, r, between particles i and j, by

$$F_{ji} = -\nabla_{ji} V r_{ji} \qquad (1.14.)$$

where $-\nabla_{ji}$ is the negative gradient of the potential energy. Note that according to Newton's third law that an object exerting a force upon another results in the second objects exerting an opposing force which is equal in its magnitude, however, acts in the opposite direction from the first object. Thus the forces acting upon a given particle j due to any other particle i, F_{ij}, is,

$$F_{ij} = -F_{ji} \qquad (1.15.)$$

Thus, for a system of N particles, a given particle's momentum, p_i, is a function of the external forces acting upon the single particle, f_i^e, and the internal forces, F_{ji}, on every particle in the system due to any given particle, j. Thus, by extension this is related to Newton's second law by,

$$p_i = \sum_{j=1\neq i}^{N} \left(F_i^e F_{ji} \right) \qquad (1.16.)$$

A trajectory consisting of particle positions, forces and velocities at small time intervals, δt, is thus calculated by integrating the differential equation,

$$\frac{dV}{dr_i} = \frac{d^2 r_i}{dt^2} \qquad (1.17.)$$

The standard method for calculating Van der Walls (VdW) attractive and repulsive forces is the Lennard-Jones (LJ) potential (Kamberaj, 2020). LJ potentials have been shown to successfully replicate the physical properties of simple monoatomic systems (Verlet, 1967). While perhaps not best suited to more complex molecular systems, it is still the most commonly used method for approximating the VdW contribution to interatomic interactions. The LJ potential is thus applied by,

$$U(r_{ij}) = 4 \in_{ij} \left(\left(\frac{r_{ij}}{\sigma_{ij}} \right)^{-12} - \left(\frac{r_{ij}}{\sigma_{ij}} \right)^{-6} \right) \qquad (1.18.)$$

where $U(r_{ij})$ is the potential energy, $\in_{ij}$ is the well-depth of the potential energy function and σ_{ij} is the interparticle distance where the potential energy is zero. The individual values for the well depth and optimal interparticle distance are dependent on the VdW radius of a given atom. Additionally, multiple implementations of LJ-type potentials exist, however the LJ 12-6 potential shown in Equation 1.15. is the most commonly implemented within molecular simulations (X. Wang, Ramírez-Hinestrosa, Dobnikar, & Frenkel, 2020). While it is shown here that the potential energy is purely a function of the LJ potential, in reality there are a number of additional components which contribute to this (see Section 1.11.). The LJ potential gives a reasonable physical approximation of simple long-range interactions, which are repulsive at very short distances and attractive at longer distances. In practice, it is necessary to define some cut-off distance, above which the interactions are ignored, for the purpose of reducing computational expense. The exact cut-

off value is of course dependent on the computational power available, as well as the size of the simulation box. A graphical representation of the LJ potential is shown in Figure 1.14.

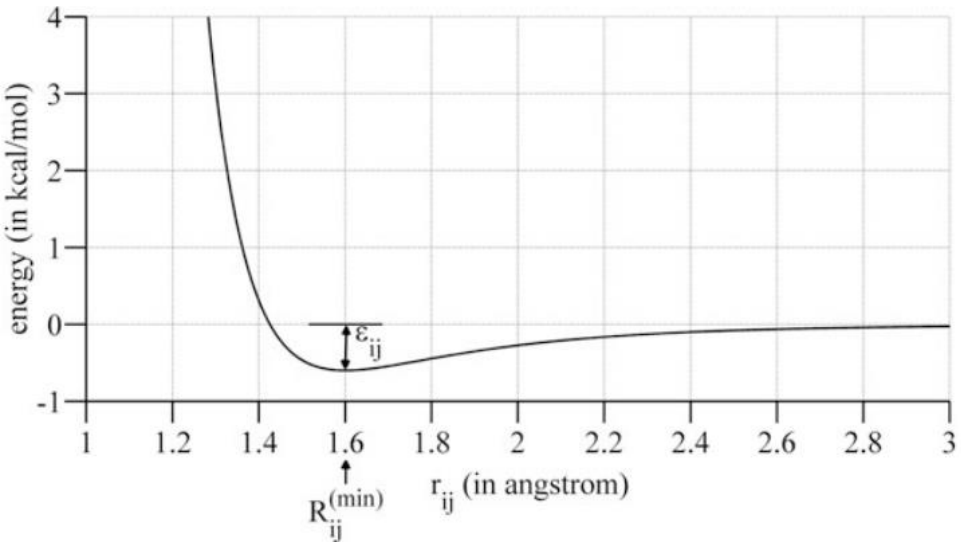

Figure 1.19: Plot of the LJ 12-6 potential, showing the well-depth at ~ 1.6 Å. The optimal separation, at the potential energy minimum, is given as R_{ij}^{min}.

Electrostatic interactions are treated using a Coulombic function, wherein for neutral systems the contribution to the total potential energy is expressed as the sum of all partial charges between all particles pairs in the system. For two interacting atoms, i and j, the partial charge, q_i, is represented as the sum of bond increments, δ_{ij}, which represent the separation of charges between the covalently bonded atoms, by,

$$q_i = \sum_j \delta_{ij} \qquad (1.19.)$$

Thus, the energetic contribution, U_q, from all electrostatic interactions is summed as (Sharma, 2019),

$$U_q = \sum_{ij} \frac{q_i q_j}{r_{ij}} \qquad (1.20.)$$

where the sum goes over all atoms i and j, q_j is the partial charge on atom j and r_{ij} is the distance between atoms i and j.

1.9.1. Force fields

Force fields are used in MD simulations to calculate the forces exerted between atoms or molecules (Papon et al., 2006). Individual contributions to the potential energy are force field dependent, where the total potential energy is calculated as

$$U(r) = U_{bond} + U_{angle} + U_{torsion} + U_{non-bonded} \quad (1.21.)$$

where U_{bond} is the contribution from pairs of bonded atoms, U_{angle} is the sum of all angles between three bonded atoms, $U_{torsion}$ is the contribution relating to the energy change as a function of bond rotation and $U_{non-bonded}$ is the energy from non-bonded interactions between particles, including VdW and electrostatic contributions. The work presented herein used two force fields for MD simulations, the optimised potential for liquid simulations all atom (OPLS-aa) force field (Jorgensen, Maxwell, & Tirado-Rives, 1996) and the condensed-phase optimised molecular potentials for atomistic simulation studies II (COMPASS II) force field (Huai Sun et al., 2016a). As simulations using the OPLS-aa force field were performed externally, the focus from here is on the COMPASS II force field.

The COMPASS force field (H Sun, 1998) was developed as an improvement to the consistent force field (CFF) and polymer consistent force fields (PCFF) (Huai Sun, Mumby, Maple, & Hagler, 1994). Both of these are *ab initio* force fields, where the forcefield parameters are fitted to the potential energy surface of a test functional group obtained via *ab initio* calculations, in order to obtain reliable geometries and intra/intermolecular interactions (Maple, Dinur, & Hagler, 1988). Additionally, the use of Lennard-Jones 9-6 (LJ 9-

6) potentials are taken from these force fields for the description of VdW interactions, which is another commonly used implementation of the standard LJ potential, however one which has been shown to be more effective at short distances, as the standard method has been demonstrated to be "too repulsive" at these short distances (Halgren, 1992).

A key objective in the development of the COMPASS force field was to correct the problems typically encountered by the PCFF and CFF force fields in the accurate calculation of density, and other temperature/pressure and volume dependent properties of systems under study. These problems arose due to the calculations which were used to develop the parameters of the force field. These were calculated at 0 K, wherein the molecules are static, however, the experimental data used to refine the parameters were measured at finite temperature. Thus, the resulting parameters include factors such as thermal expansion, which introduces problems due to the differing properties at the different temperatures. Therefore, these are only valid in cases where the calculations used to develop the parameters are performed under similar conditions to which the experimental data are collected.

To overcome these issues, development of the COMPASS force field involved a complete re-parameterisation of the non-bonded interactions, and modification of the valence parameters, using a hybrid *ab initio*/empirical method, which is necessary for accurate prediction of thermodynamic properties. This is due to the typically small number of molecules used in *ab initio* calculations, which accurately calculate the intramolecular and short-range intermolecular interactions. However, in terms of paramaterisation of MD force fields, this means that these are only suitable for gas-phase applications, as they do not consider the longer-range interactions which are important in condensed-phase materials. Thus, the empirical part of force field development is crucial. This approach is commonly

used to determine non-bonded parameters (Sagarik & Ahlrichs, 1987), and has been found to be capable of accurately replicating experimental properties, for even simplistic Lennard-Jones simulations of molecularly complex solids (Widmann, Laso, & Suter, 1995). Essentially, *ab initio* calculations are performed to generate the initial parameters, and MD simulations of condensed-phase systems are then performed, wherein the parameters are refined against experimental data (e.g. diffusion coefficient, molecular volume, dielectric constant etc.) such that the true physical properties can be replicated. A workflow of the paramaterisation of the COMPASS force field is shown in Figure 1.15. The COMPASS force field was recently updated to expand the number of polymers and molecules which can be accurately simulated (Huai Sun et al., 2016b).

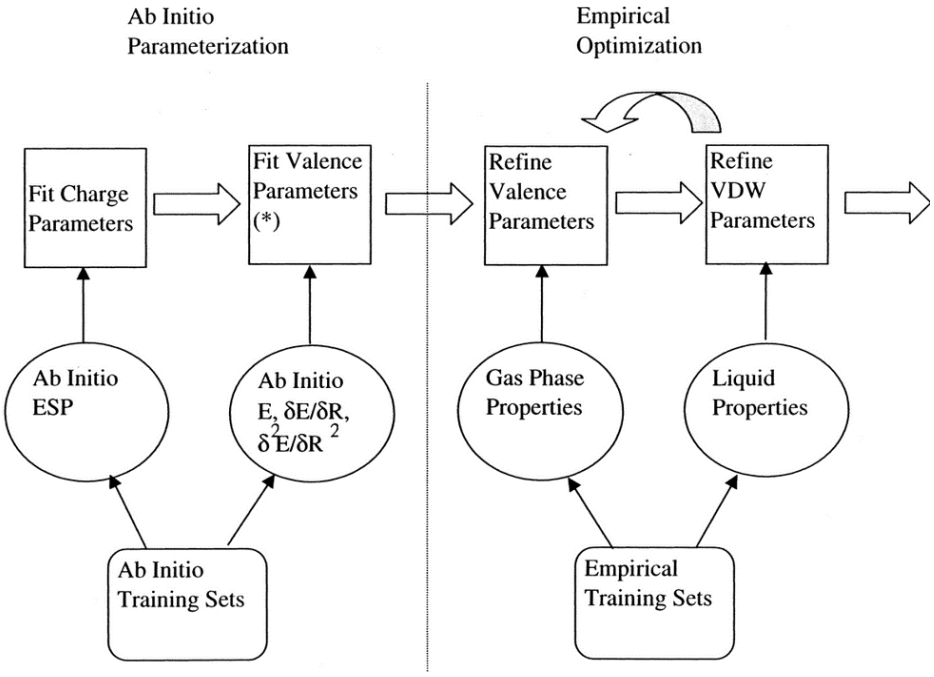

Figure 1.20: Workflow for paramaterisation of the COMPASS force field (H Sun, 1998). Left: Electrostatic potential energies (ESP) are calculated from *ab initio* methods to determine partial charges and valence parameters. VdW parameters are fixed at initial best-guess values. Right: Experimental data from systems in condensed and gas-phases are used to optimise valence and derive VdW parameters.

1.9.2. Thermodynamic ensembles

1.9.2.1. Microcanonical ensemble

The microcanonical ensemble, also known as the NVE (constant number of particles, N, constant volume, V, constant energy, E) ensemble is used when Newton's equations are solved with no temperature or pressure control (Kamberaj, 2020). Thus, a system using the NVE ensemble is completely isolated, and therefore the energy is conserved throughout the dynamics calculations. Fluctuations in the kinetic and potential energy components do occur, however, the total energy is constant.

1.9.2.2. Canonical ensemble

The NVT ensemble, as with NVE, has a fixed number of particles and volume, however the system is coupled to a heat bath/thermostat for temperature (T) control. Thus, due to interaction between the simulation box and the heat bath, energy is able to transfer to and from the simulation box. The obvious benefit of the NVT ensemble is that the dynamic properties of the system can be studied as a function of temperature, which is not possible using the standard microcanonical ensemble.

A variation of the canonical ensemble exists, which is the NPH (constant number of particles, constant pressure, P, and constant enthalpy, H) ensemble. This allows variation of the size, and if required, shape, of the simulation box, by coupling to a barostat, such that pressure-induced changes in a sample material may be examined, without the additional perturbation of temperature affecting the dynamics.

1.9.2.3. Isothermal-isobaric ensemble

The final ensemble described here is the NPT (constant number of particles, constant pressure and constant temperature) ensemble. The temperature and pressure are fixed by coupling to a thermostat and barostat, respectively, allowing the volume, and the shape if required, of the simulation box to change. The NPT ensemble is arguably the best method for simulating "real-life" properties of materials, as the density can be equilibrated to find the true experimental values, assuming the force field parameters are accurate. As such, it is common to run the equilibration stage of the dynamics calculation using the NPT ensemble, before changing to NVE or NVT for the analysis run.

1.9.2.4. Temperature control

In the NVT and NPT ensembles, the simulation box is coupled to a heat bath/thermostat for the purposes of regulating the temperature (Yong & Zhang, 2013). A number of thermostats are available for dynamics calculations, the most commonly used within the Materials studio package Forcite being the Berendsen (Berendsen, Postma, Van Gunsteren, Dinola, & Haak, 1984) and Nosé (Nosé, 1984) methods. Despite being widely implemented in MD simulations, recent studies have demonstrated that the Berendsen thermostat cannot generate reliable trajectories, due to the suppression of fluctuations in the kinetic energy (Braun, Moosavi, & Smit, 2018). While this is beneficial for quickly reaching equilibrium in the dynamics calculations, this results in unrealistic structural and dynamic properties. Thus the Berendsen method is suitable for bringing the system to equilibrium, before changing to another thermostat during the data collection phase. Typically, the Nosé

thermostat is preferred for sampling the equilibrium properties of the simulation, which does not artificially suppress the kinetic energy, resulting in more reliable dynamics.

1.9.2.5. Pressure control

The NPT and NPH ensembles use a barostat to regulate pressure. This is of course important where achieving the correct density of the material is of interest. In the Materials Studio software package a number of barostats are available. For the purposes of this work, two shall be briefly described, the Berendsen (Berendsen et al., 1984) and Parrinello-Rahman (Parrinello & Rahman, 1980) barostats. The Berendsen thermostat is widely implemented, due to it's relative simplicity, however, suffers from the same issues that the Berendsen thermostat introduces (Braun et al., 2019), in that fluctuations in the kinetic energy are suppressed. The main advantage of this method is that it is capable of rapidly equilibrating the pressure, before switching to a more reliable barostat, or an ensemble wherein pressure is not controlled. This pressure regulation method is commonly implemented for liquid or amorphous simulations, as the cell is allowed to change volume, but not its shape. The Parinello-Rahman barostat is commonly used for simulations of periodic sytems, as the volume and shape of the cell are able to change, potenitally allowing the study of strain or phase transformations in crystalline systems. The additional benefit of the Parinello-Rahman method is that is does not suffer from the same issues as the Berendsen method.

1.9.3. Amorphous Cell module in Materials Studio

The Amorphous Cell (AC) module in Materials Studio is used to generate chemically sensible starting configurations of molecules. The method used to construct the simulation

cell is based upon the configurational bias Monte Carlo technique (Akkermans, Spenley, & Robertson, 2013), wherein individual molecules are constructed fragment by fragment. The positioning of fragments are chosen randomly, taking into consideration the positions of surrounding molecules and fragments, and their Boltzmann probabilities (Ilja Siepmann & Frenkel, 1992). As such, the accepted conformation is considered from the probabilities of the previously generated geometries, aiming to construct a box of "random" configurations based on a probability distribution, ultimately building a simulation cell consisting of "likely" conformations.

During the construction stage of the AC, there are a number of additional considerations available to increase the likelihood of producing a cell that is in some minimum on the potential energy surface, and that is of course chemically sensible. The number of steps considered when loading fragments into the cell can be varied. Increasing the number of steps of course increases the likelihood of finding representative conformers from the Boltzmann distribution, however, at a computational cost. Additionally, the total number of steps used to load all molecules into the cell can affect this as it allows more options to be explored in the construction phase. The program is able to test for ring-spearing, for example a covalent bond passing through an aromatic ring, and avoid such fragment placements, and can further limit the possibility of this occurring by restraining rings such that their surface areas are minimised. Imposing restraint on the rings themselves is also beneficial for retaining their geometries and ensuring they make chemical sense.

Sensible close-contacts can be ensured by eliminating the possibility of two atoms becoming too close to each other, by setting a threshold value for the lower distance limit. This is done by defining a distance as a fraction of an atom's VdW radius. Additional checks can be performed for the energies whilst fragments are being loaded into the cells, and by

increasing the number of sampled starting positions for a molecule's construction, and the number of torsions sampled for each fragment, a larger distribution of conformations may be explored for a given molecule, thus, increasing the likelihood of the minimum being found.

When constructing the AC, a molecule of interest is selected and the quantity of that given molecule is specified, wherein the software attempts to load a cell, which is typically cubic, but may also be tetragonal or orthorhombic, at a target density. There is an option to initially build the box at a density lower than the target, which further reduces the chances of unphysical interactions and positionings, such as ring spearing or atom overlap, before ramping to the required density. Multiple cells can be generated and geometry optimised, in order to increase the chances of finding the global minimum.

1.10. Summary

The following experimental chapters will explore the potential of PDF analysis in extracting structural information from amorphous pharmaceuticals. A set of recommendations are described for obtaining high quality PDF data using a laboratory instrument, before being applied to a case study of an amorphous drug substance currently in development at Astra Zeneca. Finally, the use of synchrotron sources for high throughput PDF data, and the information that can be accessed accordingly, are assessed. Complementary to PDF analysis and key to extracting structural information from amorphous APIs is the use of MD simulations, which is also demonstrated in the last two experimental chapters.

Chapter 2: Aims and objectives

2.1. Aims

The principal aim of this work is to investigate structure in amorphous pharmaceuticals using pair distribution function (PDF) analysis. Methods for data collection and treatment will be investigated for optimising a laboratory diffractometer to enable in-house PDF analysis. This includes considerations for sample handling and preparation, using indomethacin (IND) and formulated IND-polyvinylpyrrolidone (PVP) as model crystalline and amorphous pharmaceutical systems. Structure in an amorphous API (AZD5718) will be investigated using PDF analysis and molecular dynamics (MD) simulations, to identify potential structural changes related to preparation method. Similarly, structure and transformations in amorphous paracetamol (PCM) will be investigated using variable temperature synchrotron PDF data, coupled with MD simulations. A range of analytical tools are implemented to complement PDF data and to build confidence in any structural interpretations drawn from this work.

2.2. Objectives

To achieve the stated aims of this work, the following were investigated:

1. Optimisation of data collection and sample handling procedures to enable PDF analysis using a lab instrument.
    i. Explore the impact of using variable count time (VCT) and constant count time (CCT) data collection procedures, and different sample holder sizes on PDF data quality.
    ii. Investigate data treatment procedures for optimising data quality.
    iii. Implement a consistent sample preparation method to build confidence in any potential structural changes detected in PDF data.

iv. Determine the ability of PDF to detect the impact of humidity on samples and to implement suitable controls to minimise uncontrolled exposure of samples under ambient lab conditions.

v. Comparison of the sensitivity of PDF analysis to complementary data from FT-IR and DSC to changes in solid state to understand technique limitations.

2. Investigating the use of PDF to identify of structural changes in samples of an amorphous API (AZD5718) as a consequence of different preparation method.

   i. Evaluate variability in samples produced by the same method to build confidence that PDF can detect structural changes as a function of preparation method.

   ii. Comparison of experimental PDF data to PDFs calculated from theoretical structures derived from MD simulations AZD5718 samples with varying water contents to investigate the influence of water on the structure of amorphous AZD5718.

   iii. Use tools for mapping intermolecular interactions in the Crystal Structure Database (CSD) to assess differences between theoretical amorphous AZD5718 structures.

3. Investigate structure and transformations in amorphous paracetamol (PCM).

   i. Collect variable temperature synchrotron PDF data of amorphous PCM.

   ii. Investigate structural transformations in amorphous PCM by statistical analysis of PDF data that include comparison with known structures, principal component analysis and area under curve calculations for VT PDF data.

   iii. Attempt to determine a thermodynamically feasible structural model for amorphous PCM using MD simulations coupled with PDF.

iv. Apply interfacial MD simulations to investigate observed crystallisation behaviour of amorphous PCM.

Chapter 3: Materials and methods

3.1. Materials

The materials used in the project are listed in Table 3.1. All materials were purchased from Sigma Aldrich, with the exception of AZD5718, which was received from Astra Zeneca, Gothenburg, Sweden.

Table 3.1: Compound names and chemical formula of materials used in the project.

Compound	Chemical formula
Indomethacin (98.5 – 100.5%)	$C_{19}H_{16}ClNO_4$
Polyvinyl pyrrolidone K25 (PVP)	$(C_6H_9NO)_n$
Paracetamol (≥ 97%)	$C_8H_9NO_2$
AZD5718	$C_{24}H_{26}N_6O_3$

3.2. Methods

3.2.1. Sample characterisation and analysis

3.2.1.1. Laboratory XRPD

High-resolution XRPD data for phase identification purposes were collected on a Bruker D8 Advance, equipped with a copper source (wavelength 1.541Å) in Debye-Scherrer transmission and a LynxEye 1D Detector. Samples were lightly ground before loading into 0.7mm borosilicate glass capillaries. Capillaries were rotated during data collection to ensure good powder averaging. Data collection parameters varied between scans.

All lab data for PDF measurements were collected using a Malvern Panalytical EMPYREAN diffractometer, using a molybdenum source (wavelength 0.7093Å) unless

otherwise stated. A GaliPIX3D detector was used for detection of reflections. Measurements were collected at ambient temperatures. Samples were lightly ground before being loaded into either 2 mm or 3.5 mm borosilicate glass capillaries. The samples were rotated during data collection to ensure good powder averaging. A variable count time (VCT) data collection procedure was implemented to optimise signal-to-noise across the range in Q, unless otherwise stated. Instrumental factors affecting PDF data (zero-error, Q-damping and Q-broadening) were refined against NIST standard reference material 640c Si and fixed during any PDF calculations or refinements. Specific data collection parameters varied between scans.

3.2.1.2. Synchrotron XRPD

All synchrotron PDF data were collected at Diamond Light Source beam line I15-1, proposal number E17779. Data were collected for 300 seconds per sample using X-rays with wavelength 0.161669 Å and a PerkinElmer XRD 1611 CP3 area detector. Samples were lightly ground and loaded into 1.5 mm (OD) borosilicate glass capillaries. Capillaries were held in a 12-slot motorised sample changer and rotated during data collection to ensure good powder statistics. Temperature control was achieved using a CryoJet with liquid nitrogen.

3.2.2. PDF calculations and refinements

3.2.2.1. Calculation of PDF from X-ray total scattering data

3.2.2.1.1. xPDFsuite

xPDFsuite (Yang, Juhas, Farrow, & Billinge, 2014) was used in Chapter 4 for calculation of PDF data from corrected and normalised X-ray total scattering data. This software

package was initially selected due to its ease of use and capability of handling a large number of data sets simultaneously for PDF calculations. A scan of an empty borosilicate glass capillary (either 1.5 mm or 2 mm, for synchrotron or laboratory, respectively) was used for background subtraction, before automated normalisation and corrections are applied for Fourier transformation to the PDF. An R_{poly} of 0.9 Å was used, which defines the shortest atom-atom distance in the sample, and a Q_{min} of 0.1 Å^{-1}. The values used for Q_{max} varied between samples.

3.2.2.1.2. GudrunX

GudrunX (Alan K Soper, 2011) was used from Chapters 5 and 6 to obtain the PDF from X-ray total scattering data. A scan of an empty borosilicate glass capillary (either 1.5 mm or 3.5 mm, for synchrotron or laboratory, respectively) was used for background subtraction, before normalisation and corrections are applied for Fourier transformation to the PDF. Data were treated with a Soper-Lorch convolution (Alan K. Soper & Barney, 2012) to minimise termination effects from the finite Fourier transform. The Q_{min} and Q_{max} values selected changed depending on sample and data collection parameters. The minimum radius for Fourier transform, defining the shortest atom-atom distance was 0.9 Å.

3.2.2.2. Calculation of PDFs from structural models

3.2.2.2.1. PDFfit

PDFfit (Proffen & Billinge, 1999) was used for automated batch calculation of PDFs from MD simulation frames in Chapter 5. PDFfit was used in this instance due to its flexibility and capability of handling large volumes of PDF calculations. Anisotropic displacement

parameters (ADPs) were fixed at 0.0025 Å, and PDFs were calculated between 0.1 and 30 Å, with a step size of 0.02 Å.

3.2.2.2.2. TOPAS Academic V6

TOPAS academic V6 (A. A. Coelho, 2003; Alan A. Coelho, 2018) was used to calculate PDFs from MD simulation frames in Chapter 6, and refinement of the crystal structure to the experimental PDF data in Chapter 5. TOPAS was selected in these instances as it was only a single MD frame that was used for comparison to experimental data in Chapter 6, and the use of rigid body refinement (Chapter 1), which is not available in PDFfit, made fitting of the PDFs of molecular compounds possible in Chapter 5. The use of TOPAS is generally more applicable than other software packages for PDF calculations and refinements, as the macros have been modified to account for small thermal parameters at low-r, due to correlated atomic motion, which increase gradually at higher-r to account for decreasing correlated motion due to intermolecular contacts. In packages such as PDFfit, ADPs are fixed across the full r-range, thus it is difficult to obtain suitable parameters.

3.2.3. Sample handling and moisture control

To minimise the contribution of environmental or process-induced sample variations to the PDF data, all materials were handled within a polyethylene glove bag (Aldrich® AtmosBag) under a dry nitrogen flow at ~25°C and <5% RH. This includes grinding of glass into powder, loading of powders into capillaries and long-term storage. Sample drying (where necessary) was carried out by loading powders or glass into a petri dish over P_2O_5 in a vacuum oven at room temperature. Samples were dried for up to 48 hours before being

transferred to the polyethylene glove bag for further handling/storage. Where samples were deliberately exposed to elevated humidity (Chapter 4), samples were transferred to a polyethylene glove bag containing a supersaturated sodium chloride – water solution (67-71% RH) for up to 48 hours.

3.2.4. Statistical analysis

The statistical analysis package PolySNAP (Barr et al., 2004b) was used to perform pattern-matching and quantify the similarity between PDF data sets. PolySNAP quantifies similarity by calculating Pearson correlation coefficients (CCs) (Myers, Well, & Lorch, 2013) between two given patterns, where a CC of 1 suggests complete correlation between two patterns (e.g. superimposable) and a CC of -1 suggests a complete negative correlation. PolySNAP was chosen as it is capable of handling a number of data formats, allowing it to process diffraction, thermal and spectroscopic data if necessary and provides a rapid means of assessing the similarity between multiple data sets.

3.2.5. Molecular dynamics simulations

3.2.5.1. Amorphous cell

The Materials Studio (San Diego, USA) Amorphous Cell (AC) package was used to generate starting configurations of molecules in a large box. The COMPASS II (Huai Sun et al., 2016a) force field was used to calculate interactions between particles. The Ewald summation method was used to calculate long-range electrostatic interactions, with an Ewald accuracy of 1.0×10^{-5}. The atom-based method of summation was used for van der

Waals with a cubic spline truncation at 13.5 Å, and a spline width of 1 Å. Partial charges were calculated as assigned by the COMPASS II force field.

In the construction of the Amorphous Cell, the temperature used was 298 K, and a segment lookahead of 10 steps was selected, with a total number of loading steps of 1000, to appropriately explore the conformational space. Checks were performed for ring-spearing and aromatic ring geometries were fixed. 10 starting configurations were sampled for the initial input position of each molecule and each torsion angle. The cells were constructed with an initial density of 0.2 g/cm^3, before being ramped to the target density. Following construction of the cell, the geometry was optimised using the Smart method, with convergence criteria for energy, force and displacement of 2x10^{-5} kcal/mol, 0.001 kcal/mol/Å and 1x10^{-5} Å, respectively. Standard pressure (1.013x10^{-4} GPa) was applied during geometry optimisation.

3.2.5.2. Dynamics calculations

The Forcite module in Materials Studio was used for molecular dynamics simulations. The same force field parameters and summation methods as outlined in Section 1.2.4.1 were applied for the dynamics calculations. In all cases, the initial velocities were randomly assigned. Regardless of total simulation time, the time step selected was 0.5 fs, and a frame output of 1/ps was chosen for analysis. In the NVT and NPT ensembles, the Nose (Nosé, 1984) thermostat was used. The Berendsen (Berendsen et al., 1984) barostat was used for pressure control in the NPT ensemble, with external pressure of 1.013x10^{-4} GPa. The Berendsen barostat was used, despite its drawbacks, due to its ability to rapidly equilibrate the pressure. This was an acceptable method as the volume of the cell was fixed during the analysis runs and as such did not affect the dynamics.

3.2.6. Searching, visualisation and analysis of structures

All structures were visualised using the *Mercury* software within the CSD (Groom, Bruno, Lightfoot, Ward, & IUCr, 2016). Structure analysis was performed using the Aromatic Analyser and hydrogen Bond Coordination Quick-View modules within *Mercury*. The *Mogul* module in *Mercury* was used for assessing intramolecular geometries from MD simulations of amorphous PCM. Structures used for refinements to PDF and XRPD data are given in Table 3.2.

The Aromatic Analyser module assesses the strength of aromatic interactions using a neural network which has been trained using quantum mechanical calculations of phenyl rings in differing geometries. Aromatic interactions are scored between 0 and 10 and ranked as weak (0 – 3), moderate (3 – 7) or strong (7 – 10). Scores are calculated by selecting a molecule of interest. The software then automatically calculates the interaction score of the surrounding molecules.

The Hydrogen Bonding Coordination Quick-View module allows a rapid assessment of the intermolecular geometries in a given structure. Hydrogen bonding is assessed by calculating the probability for the observed coordination numbers of hydrogen bonds within a structure, based on the data available within the CSD. Thus, the most feasible bonding arrangement in a series of structures can be determined by comparison of the probabilities calculated within the module, where a low probability suggests that there is a more feasible arrangement of intermolecular contacts.

Table 3.2: List of structures used for refinements to PDF and XRPD data, including refcodes where applicable.

Compound	Refcode	Polymorphic form	Source
PCM	HXACAN04	I	CSD
PCM	HXACAN08	II	CSD
PCM	HXACAN40	III	CSD
PCM	HXACAN39	III-m	CSD
PCM	HXACAN47	VII	CSD
AZD5718	N/A	I	AstraZeneca

ConQuest (Bruno et al., 2002a) was used to search the CSD and the custom in-house database created using CSD Editor. Typically, an input structure was drawn within the software with a structural feature of interest identified for the search. *ConQuest* then searches for all structures within the CSD with the matching structural feature. Individual search criteria are specified in Chapter 6.

Chapter 4: Practical considerations for reliable pair distribution function of amorphous pharmaceuticals using laboratory sources

4.1. Introduction

A major challenge facing the pharmaceutical industry is the poor aqueous solubility of a high percentage of new chemical entities (NCEs) emerging from discovery programmes (Kanaujia et al., 2015). As such, there has been a surge in interest in amorphous solids, to exploit the solubility advantages exhibited in contrast to their crystalline counterparts (Hancock & Parks, 2000). The drawback of amorphous pharmaceutical solids, is the risk of spontaneous conversion to a more thermodynamically favoured form that can occur during processing, storage or use (Qian, Huang, & Hussain, 2010). The unstable character of amorphous solids can be countered by the use of a polymeric matrix, which can improve stability through solid-state interactions with the active pharmaceutical ingredient (API) and by kinetically hindering mobility of the drug (Baghel et al., 2016).

As such, a key research objective has been to understand the thermodynamic phase behaviour of formulated drug-polymer systems (Bordos et al., 2019; Rumondor et al., 2009; Yuan et al., 2015). To fully achieve this, it is useful to have the capability to unambiguously characterise the structures of amorphous solids and their polymeric formulations and to detect structural changes indicative of any change in solid state. Structural characterisation of crystalline materials is a routine procedure in drug development. However, given the lack of long-range order in amorphous solids, the requirements for Bragg's law cannot be met, resulting in the amorphous "halo" observed in X-ray powder diffraction (XRPD) experiments (Stachurski, 2011). Despite the lack of long-range ordering, amorphous solids do exhibit a "local" structure, which can be probed to understand nano-scale local ordering, by Fourier transformation of XRPD data to the pair distribution function (PDF) (Takeshi & Billinge, 2012).

The growing interest in the commercial exploitation of amorphous pharmaceuticals has driven interest in the applications of the PDF method in understanding amorphous APIs.

Historically employed in the study of inorganic liquids and glasses (Kirkwood & Boggs, 1942; Lorch, 1969; Thijsse, 1984; Wagner, 1980), PDF analysis has found utility in the analysis of nanocrystalline materials and disordered crystals (Cliffe et al., 2010; Jiang et al., 2017; Masadeh et al., 2007; Young & Goodwin, 2011). The last decade has seen these applications extend to molecular organic materials, with an emphasis on pharmaceutical solids (de Araujo et al., 2017; Shi et al., 2017; Terban et al., 2016, 2020) in an effort to understand what is becoming an increasingly important class of functional materials.

Most PDF studies are performed at synchrotron facilities to benefit from the high flux and detector efficiencies for good signal to noise, rapid data acquisition to high Q (Chupas et al., 2003, 2007; Sutter et al., 2016). As such, a number of challenges arise for generating reliable PDFs using a laboratory diffractometer, particularly in the characterisation of low-atomic number materials, such as organic pharmaceuticals and polymers (Petkov et al., 2013). While there are a few literature examples of PDF studies of pharmaceuticals using a laboratory system (Chieng et al., 2013; Geddes et al., 2019; Nollenberger et al., 2009), there has been little effort to address the specific data collection and sample handling procedures required for reliable PDF data. Firstly, due to the wavelength-dependence of the accessible Q_{max} in the experiment (Chapter 1; Section 1.8.2.1), it has been recommended that a molybdenum anode ($K\alpha 1$ = 0.709 Å, Q_{max} ~ 17 Å^{-1}) is necessary to meet the minimum resolution requirements for a unique PDF (Dykhne et al., 2011). A further significant hurdle in data collection is in altering the diffraction experiment to address the drastic fall-off of elastic scattering at high Q, which is also met with an increase in inelastic scattering. One such approach to optimise counting statistics at high Q is to implement a variable count time (VCT) collection procedure (Dinnebier & Billinge, 2008; Florence et al., 2005; te Nijenhuis et al., 2009), wherein 2θ is split into sections, and corresponding collection times increased with each angular range to minimise noise where measured intensities are weakest. Sample

holders should minimise background contributions and allow for a wide angular range of measurement in 2θ. A transmission geometry instrument is well suited to this (David et al., 2010), in addition to avoiding effects of sample movement during measurement as a result of phase transitions for example. Additionally, containing material within a capillary can also minimise exposure to environmental factors, such as elevated humidity.

Beyond instrument and data quality considerations, there are, of course, sample-specific requirements for reliable PDF data. Common factors that can impact on the PDF include: homogeneity (uniformity of the amorphous dispersion); composition (API and polymer concentration); amorphous content (completely free of crystalline API); residual humidity (minimal moisture content); article size (consistent particle size distribution) and degradation of API/polymer. Indeed, these factors are related, and it is thus crucial to minimise any variables or processes which may have a uncontrolled effect on any of the above. Central to achieving this is the preparation method, and the repeatability and reproducibility of any process implemented to produce the amorphous form. A number of properties of formulated amorphous solids can vary as a consequence of the preparation method, as detailed in Table 4.1.

Table 4.1: Thermodynamic status and destabilisation driving force for solid dispersions/solutions (Qian et al., 2010).

Thermodynamic status	ASD structure	Destabilization driving force
Thermodynamically stable glass	Amorphous solid solution (ASS)/ molecular dispersion of API and polymer	None
Thermodynamically stable liquid	Amorphous solution (AS)/ molecular dispersion of API and polymer	None
Supersaturated glass	Amorphous-amorphous phase separated/ Crystalline-amorphous phase separated (AAPS/CAPS)	Amorphous phase separation, crystallisation of supersaturated API
Supersaturated liquid	Amorphous-amorphous phase separated/ Crystalline-amorphous phase separated (AAPS/CAPS)	Amorphous phase separation, crystallisation of supersaturated API
Supersaturated and immiscible glass	Amorphous-amorphous phase separated (AAPS)	Amorphous phase separation
Supersaturated and immiscible liquid	Amorphous-amorphous phase separated (AAPS)	Amorphous phase separation

The indomethacin-polyvinylpyrrolidone (IND-PVP) solid dispersion is a well-studied miscible amorphous system (Yoshioka et al., 1994; Yuan et al., 2015; Tian et al., 2016) used for the stabilisation of the BCS class II API indomethacin. Due to the presence of hydroxyl group (see

Figure 1.1b), IND is amenable to hydrogen bond formation with PVP (which has a hydrogen bond acceptor group, see Figure 1.1a) and together IND and PVP form an amorphous solid dispersion. PDF analysis was applied to study this amorphous system with a view to achieving a consistent approach to producing high quality lab derived PDFs that can be used to characterise the local structure and stability in this model system as a function of composition, manufacturing route and storage.

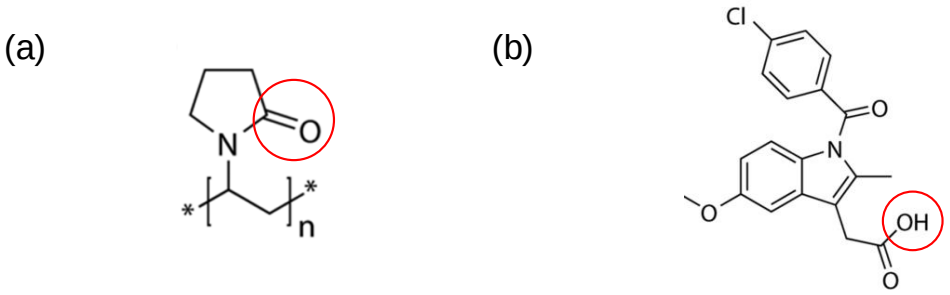

Figure 4.1: Chemical structures of (a) PVP monomer and (b) IND, highlighting the hydrogen bond donor and acceptor groups (red circles) respectively.

4.2. Experimental methods

4.2.1. Materials

4.2.1.1. Reference materials

Indomethacin and PVP Kollidon® 25 were purchased from Sigma Aldrich. Reference data were collected for the as-received materials (as-received indomethacin was purchased in the gamma polymorphic form).

4.2.1.2. Processed materials

Table 4.2 lists the processed ASD materials that were prepared from the reference materials and used throughout this work in the physical states listed. All processed materials were analysed within 30 minutes of preparation.

Table 4.2: Nomenclature used to describe ASD preparations.

ASD composition	Shorthand	Physical states	Description
IND-PVP 85:15 wt%	EXTD-C85	Extrudate C85 powder	Triturated extrudate powder
	EXTD-C85	Extrudate C85 glass	Extrudate glass
	SD-C85	Spray dried C85	Spray dried powder
	-	Melt cooled C85 glass	Melt cooled glass
	-	Melt cooled C85 powder	Triturated melt cooled powder
IND-PVP 70:30 wt%	EXTD-C70	Extrudate C70 powder	Triturated extrudate powder
	EXTD-C70	Extrudate C70 glass	Extrudate glass
	SD-C50	Spray dried C50	Spray dried powder
	SD-C70	Spray dried C70	Spray dried powder
	SD-C85	Spray dried C85	Spray dried powder
IND-PVP 70:30 wt%	EXTD-C50	Extrudate C50 powder	Triturated extrudate powder
	EXTD-C50	Extrudate C50 glass	Extrudate glass
	SD-C50	Spray dried C50	Spray dried powder

IND-PVP 20:80 wt%	EXTD-C20	Extrudate C20 powder	Triturated extrudate powder
	EXTD-C20	Extrudate C20 glass	Extrudate glass

4.2.1.3. Amorphous IND

Amorphous IND was prepared by weighing 10g of as-received IND into a stainless-steel beaker that was submerged into a pre-heated (175°C) silicon oil bath. Temperature control was achieved using a hot-plate and temperature probe. Continuous stirring was performed for 5-15 minutes using an overhead 2-blade pitched impeller at ~400rpm. When all sample was melted, the beaker was transferred to a polyethylene glove bag (Aldrich® AtmosBag) under a dry nitrogen atmosphere at ~25°C and <5% RH to cool for 15 min. Use of the glove bag ensured minimal exposure to environmental humidity. Following this, the amorphous glass was removed and ground for 4 minutes inside the glove bag.

4.2.1.4. Alpha IND

Alpha indomethacin was prepared by antisolvent crystallisation from hot ethanol (80 °C), as described in the literature (Chen et al., 2002). Room temperature filtered water was used as antisolvent. The re-crystallised sample was dried in a vacuum oven at room temperature for 24 hours before recovery.

4.2.1.5. Melt cooled C85

8.5 g of as-received IND and 1.5 g of PVP (dried for 19 hours at 110 °C to minimise water content) were weighed into a 100 ml stainless-steel beaker. A light pre-mixing of the mixture was performed within the beaker using a spatula to assist the melting process. The

beaker was immersed in a silicon oil bath as described in Section 2.1.2.1. Sample was heated for an initial period of 10 minutes with no mixing as this allowed the powders to melt without deposition of powder on the impeller blades and beaker walls, affecting the composition of the final product. After the initial melting period, mixing was achieved using the 2-blade impeller (Section 2.1.2.1.). After 5 minutes of mixing, the sample was transferred to the glove bag and ground for four minutes.

4.2.1.6. Extruded C20, C50, C70 and C85

All extruded materials were prepared using a Thermo Scientific® Process 11 11mm twin-screw hot-melt extruder. All compositions were weighed out as dried powders into 25 ml sample bottles and shaken vigorously to produce a homogeneous blend. Powders were fed into the extruder using a spatula. The screw speed was set to 100 rpm. Temperature was fixed at 50 °C for initial processing, before being ramped to 170 °C before the die zone (die diameter = 1mm), to minimise thermal degradation. A Terwin Instruments 2000 was used to monitor pressure.

4.3. X-ray powder diffraction instrumentation

4.3.1.1. Laboratory silver X-ray powder diffraction

All silver X-ray powder diffraction data were collected using a Malvern Panalytical EMPYREAN diffractometer (molybdenum anode, 0.5594 Å) in Debye-Scherrer transmission geometry, equipped with a GaliPIX3D detector. Data were collected in the angular range 3 – 147 ° 2Θ, with step-size of 0.014 °, and a constant count time of 1 second per step.

All powder samples were lightly triturated to minimise any preferred orientation effects. Samples were loaded into 1 mm borosilicate glass capillaries (Capillary Tubes Supplies Ltd.). A GaliPIX3D detector was used in 1D mode for all scans. Measurements were performed at ambient temperature unless otherwise stated.

4.3.2. Laboratory molybdenum X-ray powder diffraction

All molybdenum X-ray powder diffraction (0.7093 Å) data were collected as described in 2.2.1. Data were collected in the angular range 3 – 147 ° 2θ, employing a VCT procedure as shown in Table 4.3, with a constant step-size of 0.014 ° 2θ. All powder samples were lightly triturated to minimise any preferred orientation effects, or in the case of amorphous samples, to ensure optimal packing density within the capillaries. Samples were loaded into 2 mm borosilicate glass capillaries (Capillary Tubes Supplies Ltd.). GaliPIX3D detector was used in 1D mode for all scans. Measurements were performed at ambient temperature unless otherwise stated.

Table 4.3: XRPD VCT data collection parameters for laboratory PDF analysis

Angular range (° 2θ)	Time per step (s)
3 – 15	0.5
15 – 32	1
32 – 44	1.5
44 – 61	3.5
61 - 147	8

4.3.3. Synchrotron X-ray powder diffraction

All synchrotron X-ray total scattering data was collected on the Diamond Beamline I15-1 (proposal EE17779), using a wavelength of 0.161669 Å and a PerkinElmer XRD 1611 CP3 area detector. Powder samples were lightly triturated, sealed in 1.5 mm borosilicate capillaries, and mounted onto a 12-slot motorized sample changer. The sample changer was used to align each capillary with the X-ray beam for measurements. A cryojet was used for temperature control to achieve and maintain temperatures in the range –173 °C – 175 °C. All programming was carried out using an in-house Python® software.

4.3.4. Pair distribution function

Although a number of software packages exist for the automatic transformation of total scattering data to the PDF, xPDFsuite (Yang et al., 2014) was used in this study enabling its assessment for ease of use, minimal user input for calculations and reliable PDF generation A scan of an empty 2mm borosilicate glass capillary was used for background subtraction, unless otherwise stated. An R_{poly} of 0.9 Å was used, which defines the shortest atom-atom distance in the sample, along with a Q_{min} of 0.1 Å^{-1}. The values used for Q_{max} varied between samples, as termination effects are partly dependent on the amplitude of $F(Q)$, as will be demonstrated.

4.4. Measuring effect of drying and elevated humidity on the PDF

4.4.1. C20

To assess the effect of moisture absorption on the PDF, C20 extrudate was ground in the glove bag for 4 minutes, before being transferred to a petri dish, under P_2O_5, in a vacuum

oven at room temperature for 24 hours. After sampling for analysis, the material was transferred to a glove bag with a supersaturated sodium chloride solution, which provided elevated humidity in the range 67 – 71% RH, for 48 hours. After exposure to humidity, the material was once again placed in the vacuum oven for 48 hours, to remove moisture.

4.4.2. PVP

Before any initial drying of PVP was performed, some pre-processing of the as-received powder was carried out. 5g of as-received PVP was weighed into the stainless steel beaker, and heated as described in 2.1.2.1, only the heating process was allowed to continue for 25 minutes. After cooling, the powder was dried and exposed to elevated humidity, before re-drying, as described in the Section 2.3.1.

4.5. Statistical analysis

The software package PolySNAP (Barr, Dong, & Gilmore, 2004a) was used to quantify similarity between PDFs, by calculating Pearson correlation coefficients (CC) between each data set. A CC of 1 indicates a complete correlation between two PDFs, whereas a CC of -1 indicates a complete negative correlation. Various analysis windows were used for PDF data, as indicated in each section.

4.6. Results and discussion

4.6.1. Reproducibility of the PDF

Reliable PDF analysis of amorphous materials requires reproducible total scattering measurements. It is necessary to optimise the data collection procedure when using a

laboratory X-ray diffractometer (particularly capillary size, sample presentation, source anode, and counting time), due to the limitations imposed by the instrument geometry and flux, in contrast with synchrotron sources. The procedure taken in this work is discussed from Section 3.1.1. onwards.

4.6.2. Capillary size and sample presentation

1 mm and 2 mm borosilicate glass capillaries were tested for data collection for PDF analysis, using a molybdenum source. These contrast with the 0.7 mm capillary typically used for copper XRPD, which, due to the narrow diameter, provide greater angular resolution but at the expense of signal-noise ratio (SNR) due to the lower irradiated sample volume. The data collected from samples held in the 1 mm capillary are shown to have higher levels of noise than the 2mm capillary data. This is illustrated in Figure 4.2, with a comparison of the $F(Q)$ profiles showing higher levels of noise at high-Q for the 1 mm capillary. The improved SNR from the 2mm capillaries also was shown to provide reproducible $G(r)$ curves (Figure 4.3). The effect of noise is visible in the resulting PDF generated from the 1 mm capillary data, as higher noise in $F(Q)$ (Figure 4.2) introduces significant variations in the $G(r)$ despite having a generally similar $F(Q)$ profile and identical Q_{max}. On closer inspection of the G(r) curves, the ripples resulting from higher noise in the 1 mm capillary also have higher amplitudes.

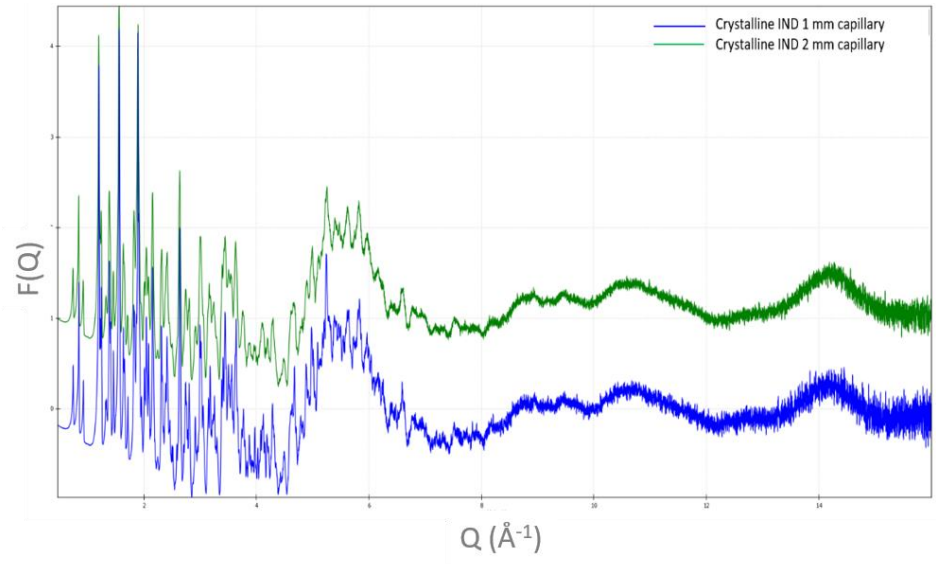

Figure 4.2: Stack of $F(Q)$ profiles of crystalline indomethacin using a 1mm (blue) and 2mm (green) capillary.

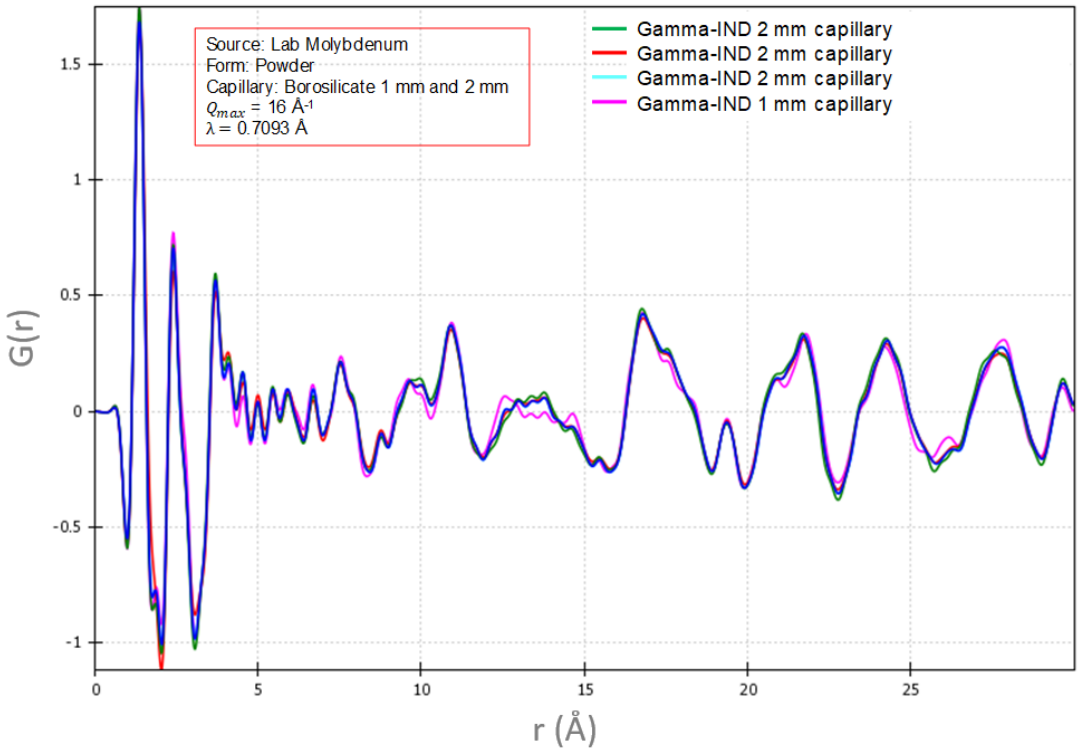

Figure 4.3: $G(r)$ curves generated from 1 mm and 2 mm borosilicate capillaries for Gamma-IND.

The high similarity between the PDFs calculated from crystalline IND in 2mm capillaries is reflected in the PolySNAP analysis (Table 4.4), where PDFs were compared in the region 3.5Å to 30Å. The CCs between all samples are close to 0.99. However, in comparison to the 1 mm measurement, the poor reproducibility due to the limited sample volume is shown, where the highest CC is 0.749. While only a single measurement was taken using the molybdenum source and 1 mm capillary, the poor correlation between the 2mm and 1mm PDFs is striking. This demonstrates that the wider diameter capillary is far more suitable for PDF analysis using a laboratory instrument.

Table 4.4: Pearson correlation coefficients between 1 mm and 2mm PDFs of crystalline IND. Values lower than 0.9 are highlighted bold.

	2 mm (1)	2 mm (2)	2 mm (3)	1 mm
2 mm (1)	1	0.992	0.993	**0.742**
2 mm (2)		1	0.985	**0.749**
2 mm (3)			1	**0.729**
1 mm				1

4.6.3. Effect of noise using a silver laboratory source

The effect of optimising data collection procedures and the detrimental impact of noise on PDF reproducibility was investigated with a low-wavelength laboratory Ag source ($\lambda = 0.5594$ Å), using a constant counting time (CCT) and a 1 mm borosilicate capillary (Section 2.2.1.). Figure 4.4 shows the variability in $G(r)$ curves produced by very noisy signals at high-Q. In $F(Q)$, high levels of noise mask any features in the true signal, particularly at

high-Q. The random noise translated into poor reproducibility in the resultant $G(r)$ curves with significant small-amplitude ripples evident. This renders the PDFs unsuitable for reliable analysis. Whilst the low wavelength of the silver anode has a higher accessible Q_{max}, the results show the impact of data collection where adequate scattering across the Q – range is not implemented. As such, for lab PDF, there is a trade-off between the useable Q_{max}, the quality of the data, and collection time required for suitable data to be obtained.

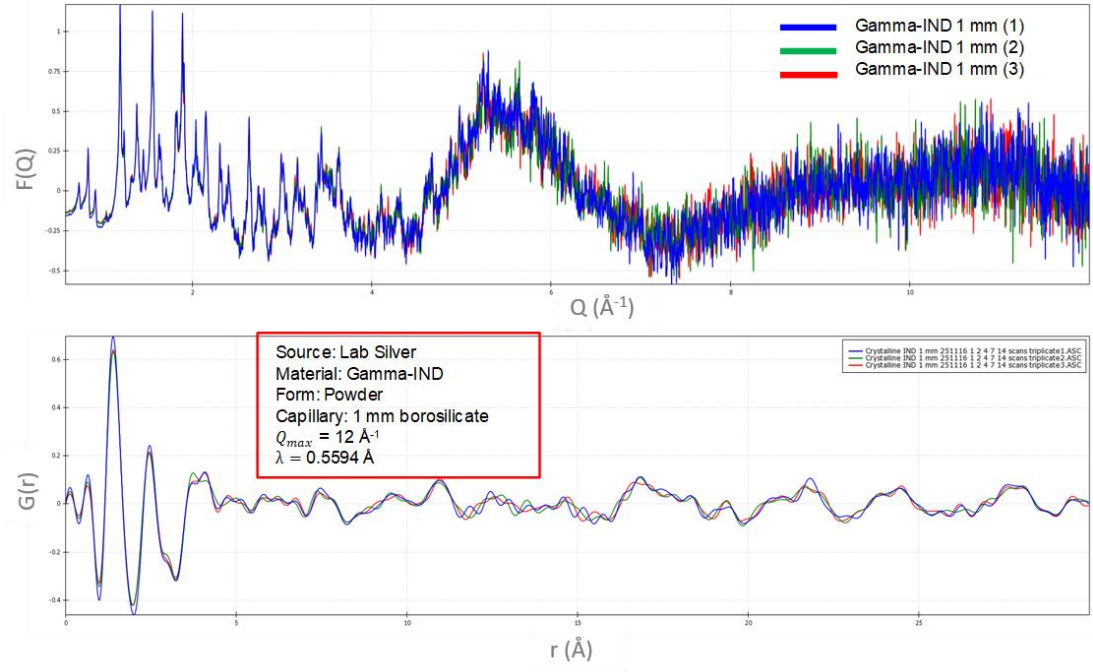

Figure 4.4: $F(Q)$ and corresponding $G(r)$ curves of three crystalline indomethacin samples, collected using a silver X-ray source with CCT data collection procedure.

The poor reproducibility of the Ag measurements is also clearly evident from the PolySNAP analysis (Table 4.5), where all values are significantly lower than 0.9. By contrast to the molybdenum VCT scans using the 2 mm capillary, which were all above 0.9.

Table 4.5: Pearson correlation coefficients between Ag PDFs of crystalline IND. Values lower than 0.9 are highlighted bold.

	Ag (1)	Ag (2)	Ag (3)
Ag (1)	1	**0.818**	**0.718**
Ag (2)		1	**0.686**
Ag (3)			1

4.7. PDF data treatment and improvement

Obtaining atomic PDFs that enable and facilitate accurate characterisation of amorphous APIs and solid dispersions requires carefully collected XRPD data and potential application of data treatment procedures that minimise systematic errors in PDF analysis. In this section, options are explored for improving the quality of the $G(r)$ curves of amorphous and crystalline materials used in this work. These procedures address (i) artefacts introduced from experimental noise, and (ii) systematic errors introduced in the processing of the raw experimental data.

4.7.1. Addressing ripples from Fourier transformation

Ripples caused by termination of the Fourier transform at $Q_{max} < \infty$ can be modelled as $\sin Q_{max}r/r$, and have a wavelength of $\sim 2\pi/Q_{max}$ (Takeshi and Billinge, 2012). This is shown by performing the Fourier transform of $F(Q)$ simulated as a flat line and comparing to the calculated curves of $\sin Q_{max}r/r$ to compare ripple periodicity (Appendix 1: Section 7.). For $F(Q)$ simulated as a flat line, its value remains constant at unity throughout the entire Q-range. This means that the wavelength of ripples is dependent on Q_{max}, and the higher Q_{max} is, the shorter the ripple wavelength. A plot of ripple wavelength versus

Q_{max} demonstrates this (Figure 4.5). Also, the effect of Q_{max} on $G(r)$ profiles generated from a flat line $F(Q)$ is shown in Figure 4.6. Since no structural information is contained in these $G(r)$ profiles, what is observed is purely the effect of Fourier transforming the termination function at a finite Q value. As expected, the wavelengths of the $G(r)$ profiles decrease with increasing Q_{max}.

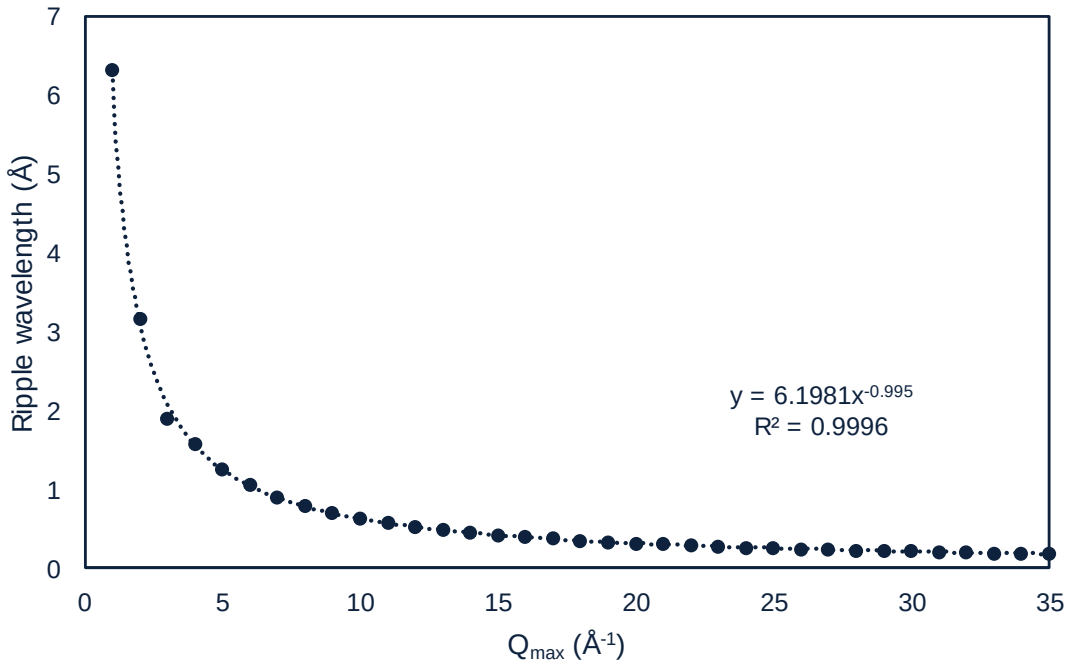

Figure 4.5: Ripple wavelength as a function of Q_{max}. $G(r)$ generated from $F(Q) = 1$ (shown in Figure 4.6).

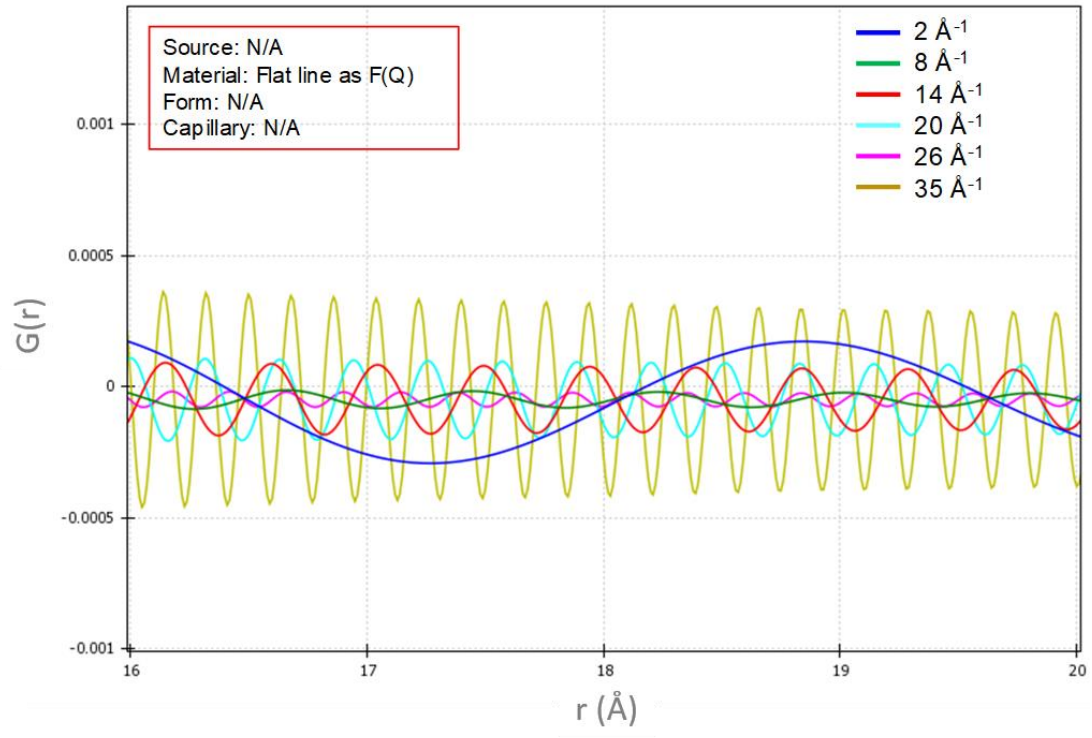

Figure 4.6: $G(r)$ profiles for different values of Q_{max}. $G(r)$ generated from $F(Q) = 1$.

Secondly, the amplitude of termination ripples is shown to be dependent on the value of $F(Q)$ at the termination point i.e. $F(Q)$ at Q_{max}. This is demonstrated in Figure 4.7, where a flat line is input as the $I(Q)$. Termination of the Fourier transform at Q_{max} where the amplitude of $F(Q) \neq 0$ results in higher ripple amplitudes. Where $F(Q) = 0$, ripples are significantly dampened; therefore, the impact of ripples can be further reduced by terminating Q_{max} where $F(Q)$ is close to zero. While these of course represent idealised cases wherein there are no statistical errors in the $I(Q)$ or $F(Q)$ input data, this can be translated to experimental data where noise is present. It is also important to note that for experimental data, the resulting PDF is convoluted with both the Fourier transform of the termination function, and the Fourier transform of the noise (Takeshi & Billinge, 2012b).

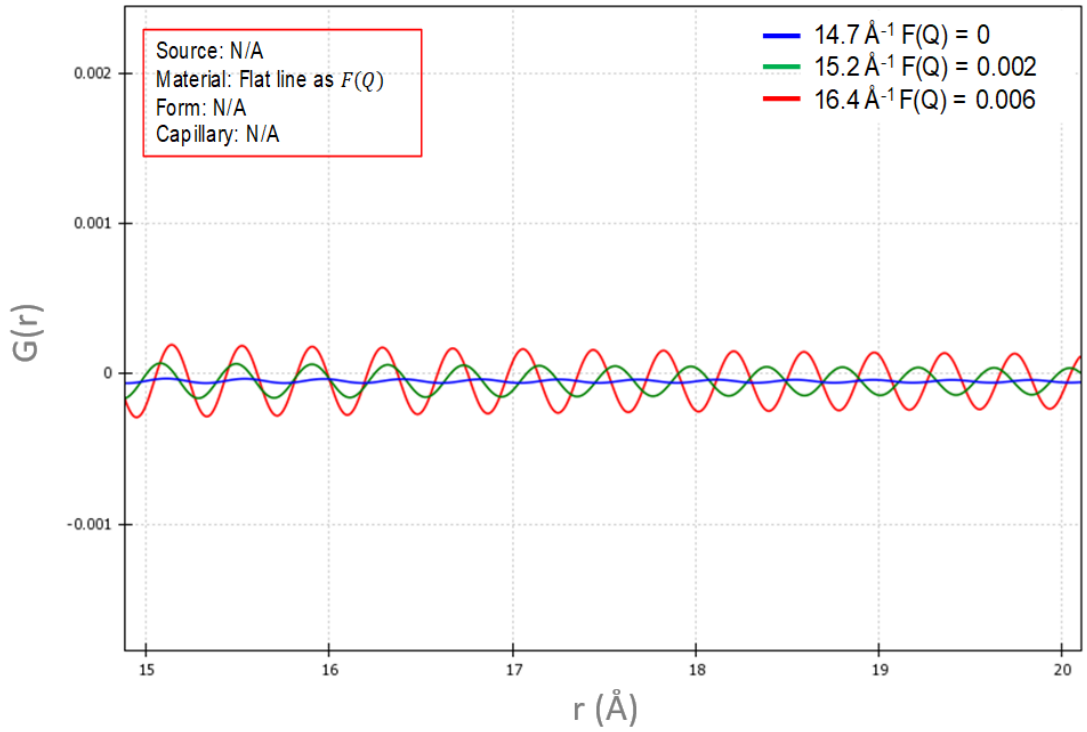

Figure 4.7: Ripple amplitude as a function of $F(Q)$ value. $G(r)$ generated from flat line as $I(Q)$.

4.7.2. Addressing ripples from experimental counting statistics

As shown above, the wavelength and amplitude of ripples in a measured PDF caused by termination at Q_{max} can be well described and at least their impact minimised. However, ripples caused by low SNR in the raw intensity data. For strongly scattering crystalline materials, the $G(r)$ has intense peaks at high-r, therefore ripples in this range are largely inconsequential. However for amorphous systems where featureless $G(r)$ curves at higher r are generated the ripples present potentially bigger impact on the interpretation of PDFs. It is desirable to address ripples that arise from the raw data and result in noise in $G(r)$ to ensure features are not wrongly assigned, especially where materials have significant amorphous content and emerging structural features (e.g. nanocrystallinity) may be weak.

4.7.3. Approaches to mitigating ripples in $G(r)$ curves

Two approaches to correcting undesirable, spurious ripples in $G(r)$ curves are evaluated in the following sections, namely minimising errors and Savitzy-Golay smoothing (Zimmermann & Kohler, 2013). Assessment of these correction techniques was performed in MATLAB® 2016 to aid flexibility and improve understanding already afforded by commercially available PDF software.

4.7.3.1. Error minimisation

One possible approach to reducing the amplitude of termination ripples in $G(r)$ is to terminate $F(Q)$ at a Q_{max} where the value of $F(\mathrm{Q}) \sim 0$, as demonstrated in Section 4.1. In Figure 4.8, the $F(Q)$ of Amorph-IND was terminated at Q_{max} of 16 Å^{-1} and 20 Å^{-1} to observe the effect of the termination point ($F(Q)$ value) on the amplitude of ripples in $G(r)$. An immediate consequence of terminating $F(Q)$ at different Q_{max} is a variation in $G(r)$ features as shown in Figure 4.9. Terminating at lower Q_{max} of 16 Å^{-1} not only resulted in a loss of real-space resolution, evidenced by broadened peaks in the low-r region, but also caused distortions to the low-r features at 2.6 Å, 3.6 Å and 3.9 Å. The most noticeable change in the low-r region occurred at <1 Å where the positions of termination ripples changed alongside reduced amplitudes for Q_{max} of 16 Å^{-1}. In the high-r region, visibly higher amplitudes of termination ripples were observed for Q_{max} of 20 Å^{-1}. A closer view of this range in Figure 4.10 demonstrates this. It is important to note that the increased amplitudes of ripples in the high-r region is also a contribution from a much noisier $F(Q)$ truncation at Q_{max} of 20 Å^{-1}.

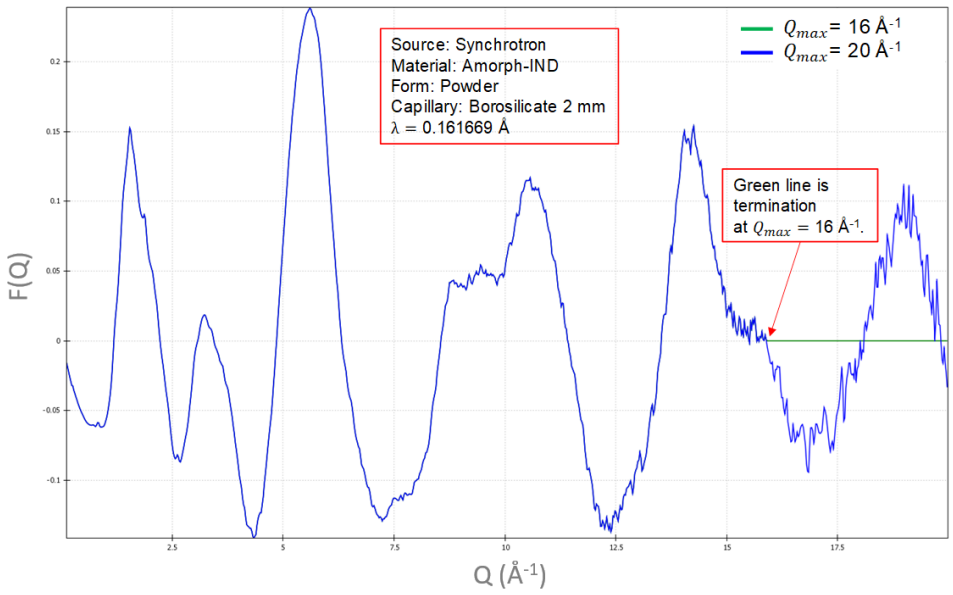

Figure 4.8: Termination of $F(Q)$ of Amorph-IND at different Q_{max} values.

These results illustrate the challenge in terminating at a maximum attainable Q_{max} (where counting statistics are not detrimental) does not necessarily result in better quality PDF where Q_{max} is negatively impacted and therefore the overall quality/information content of the PFD is reduced. The error minimisation method therefore is an optimisation of the termination point which satisfies the requirement of a maximum attainable Q_{max}, and a value of $F(Q)$ that is close to zero, which as seen in Figure 4.9 the location can have significant impact on G(r) quality and interpretability. Thus whilst the error minimisation approach is helpful in mitigating ripples, care is required in implementing the best cut-off positions and in selecting these for each sample compositions/material used.

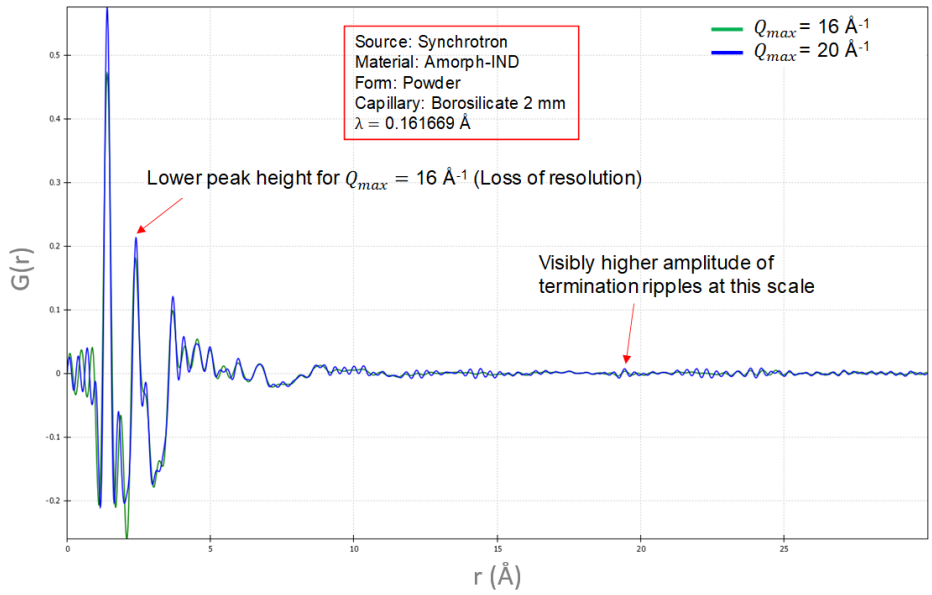

Figure 4.9: $G(r)$ curves for $F(Q)$ terminated at Q_{max} of 16 Å^{-1} and 20 Å^{-1} for Amorph-IND. $r = 0 - 30$ Å.

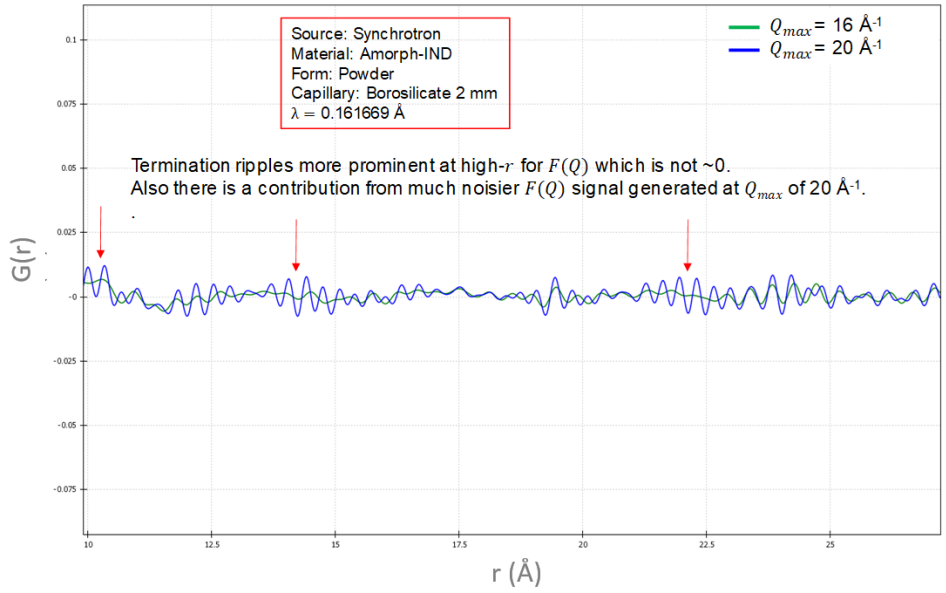

Figure 4.10: $G(r)$ for $F(Q)$ terminated at Q_{max} of 16 Å^{-1} (green) and 20 Å^{-1} (blue) for Amorph-IND. Truncated between $r = 10 - 30$ Å.

4.7.3.2. Smoothing of $F(Q)$

The propagation of random noise from the raw data into the $G(r)$ is a major hindrance to the use of higher Q_{max} values from a laboratory or synchrotron source. Smoothing of the $F(Q)$ is a potential way to de-noise the signal, removing extraneous, random variations in amplitude whilst maintaining the fundamental integrity of the F(Q) signal so that high -Q data can be included in the Fourier transform. This principle is demonstrated in Figure 4.11 for Amorph-IND. Here, the corrected, normalised experimental $F(Q)$ signal was de-noised by smoothing with a Savitsky-Golay algorithm. The Savitsky-Golay filter was selected as it is regularly applied for data smoothing and de-noising in spectrum analysis, signal processing, and image processing fields (Zimmermann & Kohler, 2013). The key advantage of this filter is due to the method in which it processes data points. As any given dataset is split into a series of analysis windows, consisting of only a few points (e.g. 5), the important features in the data (e.g. Bragg peaks or broad diffuse features in the case of XRPD data) are largely unaffected (Savitzky & Golay, 1964). Note that at the selected truncation point of $Q_{max} = 20$ Å^{-1} used in this study, $F(Q) = 0.0039$ (i.e. ~0) the resultant ripples in the $G(r)$ are due to contribution from random noise in $F(Q)$ and finite termination. On inspection of the normal and smoothed F(Q) derived $G(r)$ curves respectively, significantly more ripples are present in the original $G(r)$, with the smoothing reducing ripple amplitudes by ~50% in the high-r region (see bottom pane of Figure 4.11), whilst showing no significant alteration to the underlying shape of the F(Q) profile. Given that the key features in $F(Q)$ remain unchanged, it is apparent that the structural information encoded within the $F(Q)$ passes the filter unaffected.

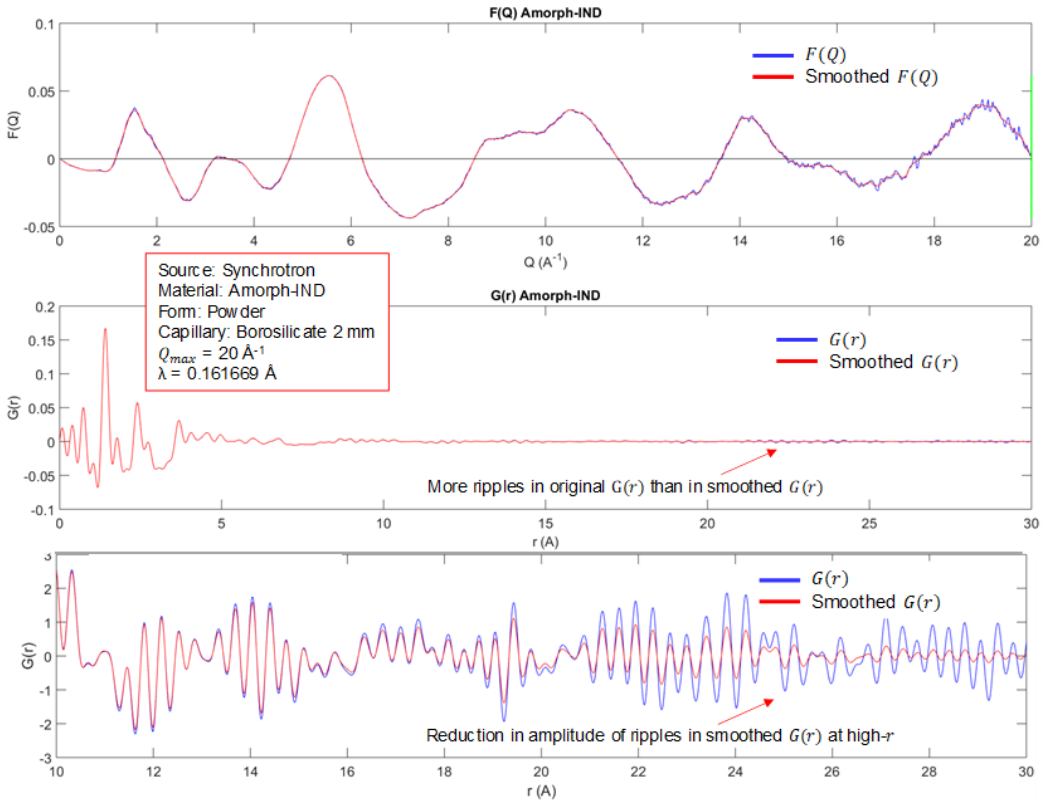

Figure 4.11: (Top) original (blue) and smoothed (red) $F(Q)$, (middle) $G(r)$ calculated from original (blue) and smoothed (red) $F(Q)$ for Amorph-IND. r = 3 – 30Å. (Bottom) truncated view of middle pane showing r = 10 – 30Å.

The effectiveness of smoothing $F(Q)$ to de-noise data shows potential promise though does require further evaluation. Since statistical noise propagated from $S(Q)$ is amplified in $F(Q)$, a potential approach would be to smooth the $S(Q)$ data. Smoothing has been applied as an error mitigation technique in the PDF analysis of nanoparticles (Mullen & Levin, 2011). In their work, an *aws* (adaptive weights smoothing) algorithm was able to handle SNRs as low as 25 and appeared to reconstruct the noise-free signal reasonably well. While smoothing $F(Q)$ is shown to be effective at minimising the effect of noise at high-r, it is apparent from Figure 4.12 that this does not affect the low-r termination ripples to any significant extent. This is due to truncation effects and as such, smoothing will not change the amplitudes of these features, as it is an inherent limitation of the finite integral (Chapter

1). As such, other approaches could be assessed to minimise this, such as the real-space Soper-Lorch convolution function (Alan K. Soper & Barney, 2012), however, these are not discussed or implemented in this chapter.

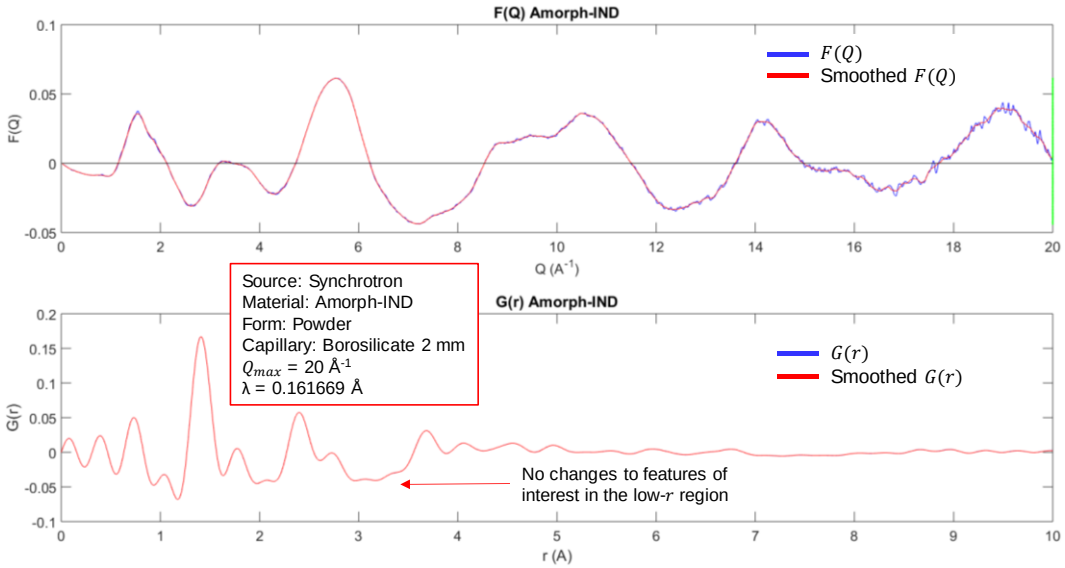

Figure 1.12: Original and smoothed $F(Q)$ and corresponding $G(r)$curves for Amorph-IND. PDF truncated between 0 and 10 Å.

4.8. Lab and synchrotron derived PDFs of reference samples

Having selected a suitable lab data collection and data treatment procedure, $G(r)$ curves of reference samples (Processed PVP, as-received PVP, γ-IND and α-IND) were generated using the PDF method outlined for molybdenum and synchrotron X-ray sources (Section 2.2.4.). Figure 4.13 shows stacked $G(r)$ curves for the reference samples generated from synchrotron X-ray total scattering data. The $G(r)$ curve for γ-IND shows well defined features which extend into the high-r region, beyond the upper-r limit shown reflecting the long-range ordering in the sample, characteristic of a crystalline material. Amorph-IND and all PVP

samples displayed a lack of features in this r-range, reflecting their amorphous character and absence of long range order.

On close observation of the PDF curves of the amorphous samples, the presence of periodic small-amplitude sinusoidal features across $G(r)$ is evident. Commonly overlooked in literature (Peterson et al., 2012; Shi et al., 2017), these graphical features in the calculated PDF G(r) may be misinterpreted as actual structural features. For example, falsely indicating the presence of nanocrystalline domains in an otherwise homogeneous amorphous sample. However, these fluctuations are mainly due to noise and truncation errors (Takeshi and Billinge, 2012). These ripples can be modelled and have a wavelength of $\sim 2\pi/Q_{max}$ (Takeshi & Billinge, 2012a). This is a mathematical consequence of the theoretical $G(r)$ being convoluted with noise and the Fourier transform of the termination function (Chapter 1). This is discussed in more detail in section 4.4. The small-amplitude termination ripples are visible in the high-r region of the $G(r)$ curves of Processed PVP, as-received PVP and Amorph-IND due to the lack of underlying structural features at high-r Though are less apparent in the crystalline $G(r)$ curve due to the relatively strong structural features in the crystalline PDF in this region.

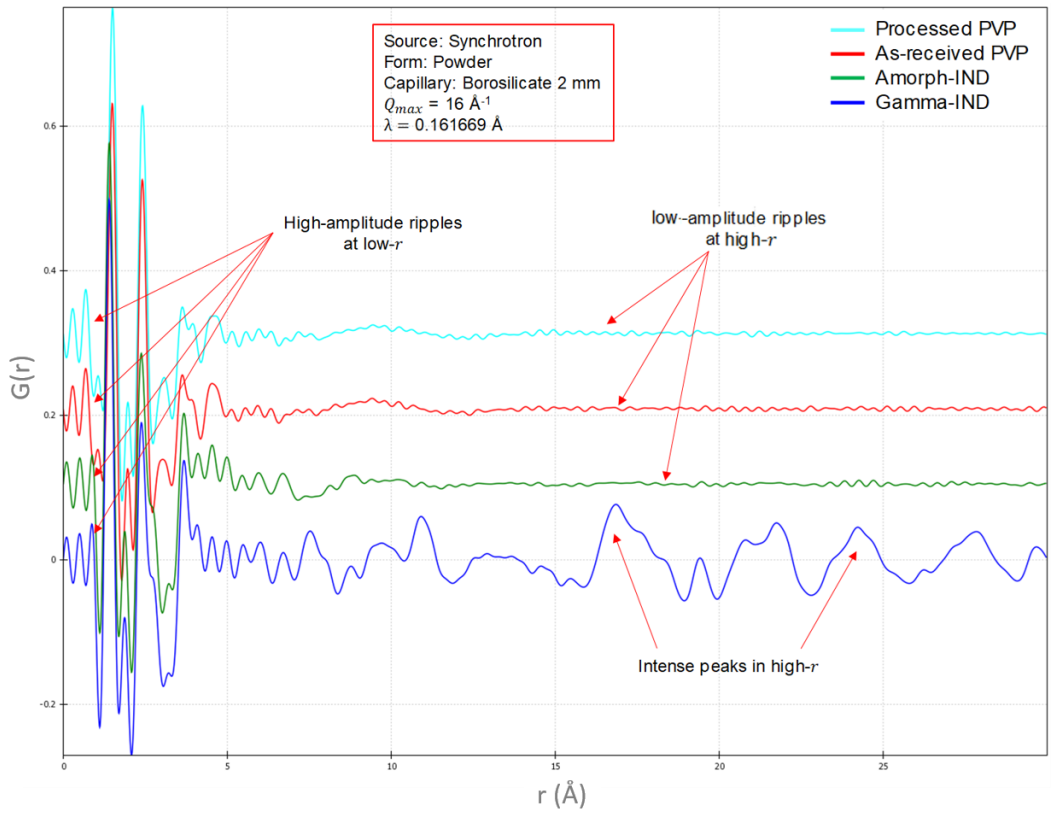

Figure 4.13: Stacked PDF traces of reference materials generated from synchrotron X-ray source. High and low intensity termination ripples are highlighted across the range in r. Structural features at high-r for crystalline IND are also highlighted.

In the low-r region (0 – 1 Å) in Figure 4.13, high-amplitude ripples are also observed and these artefacts may be misinterpreted as real atom-atom correlations in the molecular structure, with the exception of the feature observed at ~ 1 Å, which is characteristic of the shortest C – H and O - H covalent bonds present in IND and PVP.

In Figure 4.14, the $G(r)$ curves obtained from molybdenum X-ray data show similar characteristics for crystalline and amorphous materials to those displayed by the $G(r)$ curves obtained from synchrotron measurements in Figure 4.13. Interestingly, the low-r termination ripples in molybdenum $G(r)$ curves have smaller amplitudes than observed in the synchrotron $G(r)$ curves for the same Q_{max} termination. This was found to be an effect

of the different $Q_{max} - inst$ values between the laboratory molybdenum instrument and synchrotron X-ray source. The $Q_{max} - inst$ value selected in xPDFSuite sets the range over which the data are fitted for the *ad-hoc* corrections implemented in the software package, thus affecting how resulting PDFs are scaled (Juhás et al., 2013; Yang et al., 2014).

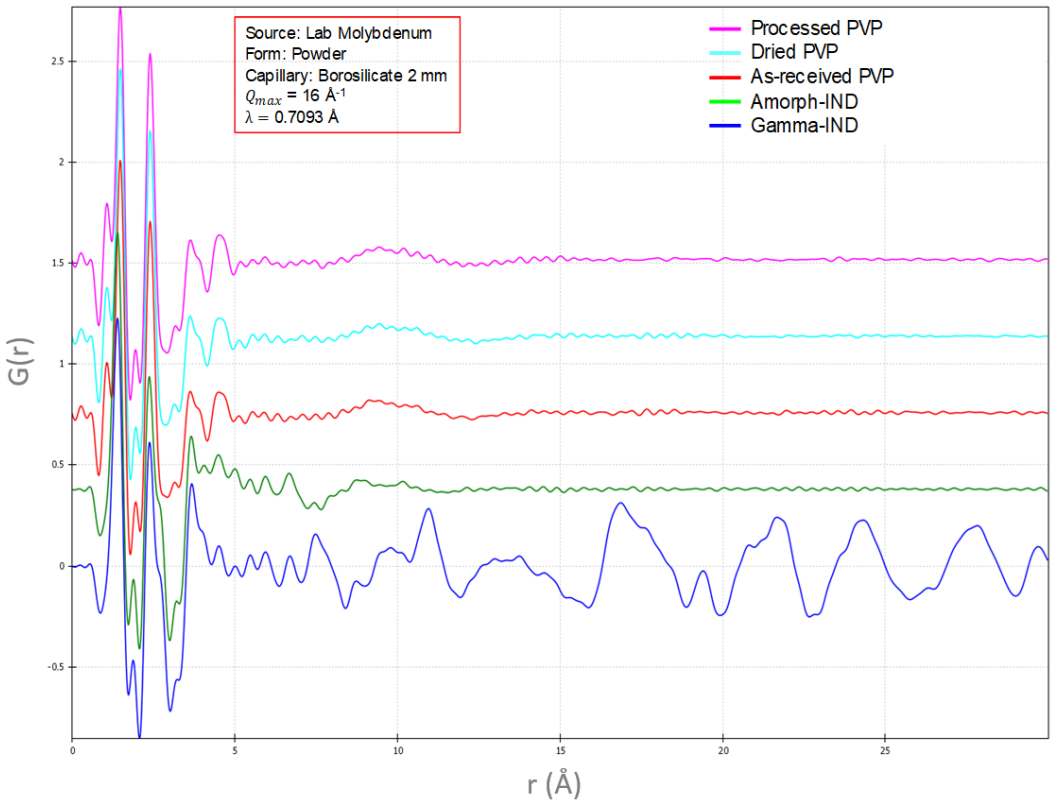

Figure 4.14: Stacked PDF traces of reference materials generated from laboratory molybdenum X-ray source. From top to bottom: Processed PVP, dried PVP, as-received PVP, amorphous IND, and gamma-IND.

To demonstrate the effect of $Q_{max} - inst$ on the resulting PDFs, an overlay of Gamma-IND $F(Q)$ and $G(r)$ curves obtained from synchrotron and molybdenum sources (Figure 4.15) shows the lower synchrotron $F(Q)$ signal resulting from fitting the data over a wider range at $Q_{max} - inst =$ 23.4 Å^{-1}. The resulting $G(r)$ curve has smaller absolute

intensities than the $G(r)$ from laboratory molybdenum radiation. The major peaks and features in both $G(r)$ curves are at the same r-positions, however at <1 Å differences in profiles occur. This highlights the impact that termination ripples, as with the low-r ripples in the synchrotron-derived $G(r)$ which have higher amplitudes, can mask real structural features in this low-r region. This highlights a general issue with the particular implementation of data analysis approaches in the XPDFSuite software package, in that the $G(r)$ is not the Fourier transform of the true differential scattering cross section (Alan K Soper, 2011), which should be normalised such that $S(Q)$ and $G(r)$ are near identical regardless of experimental factors such as the X-ray wavelength and Q_{max}. Of course, other instrumental considerations can affect the resulting profile, such as any error in 2θ values and step-size in the diffraction experiment, among others as discussed in Chapter 1. Thus, a direct statistical comparison of the lab and synchrotron PDFs is not possible when xPDFsuite is used to determine the experimental $G(r)$.

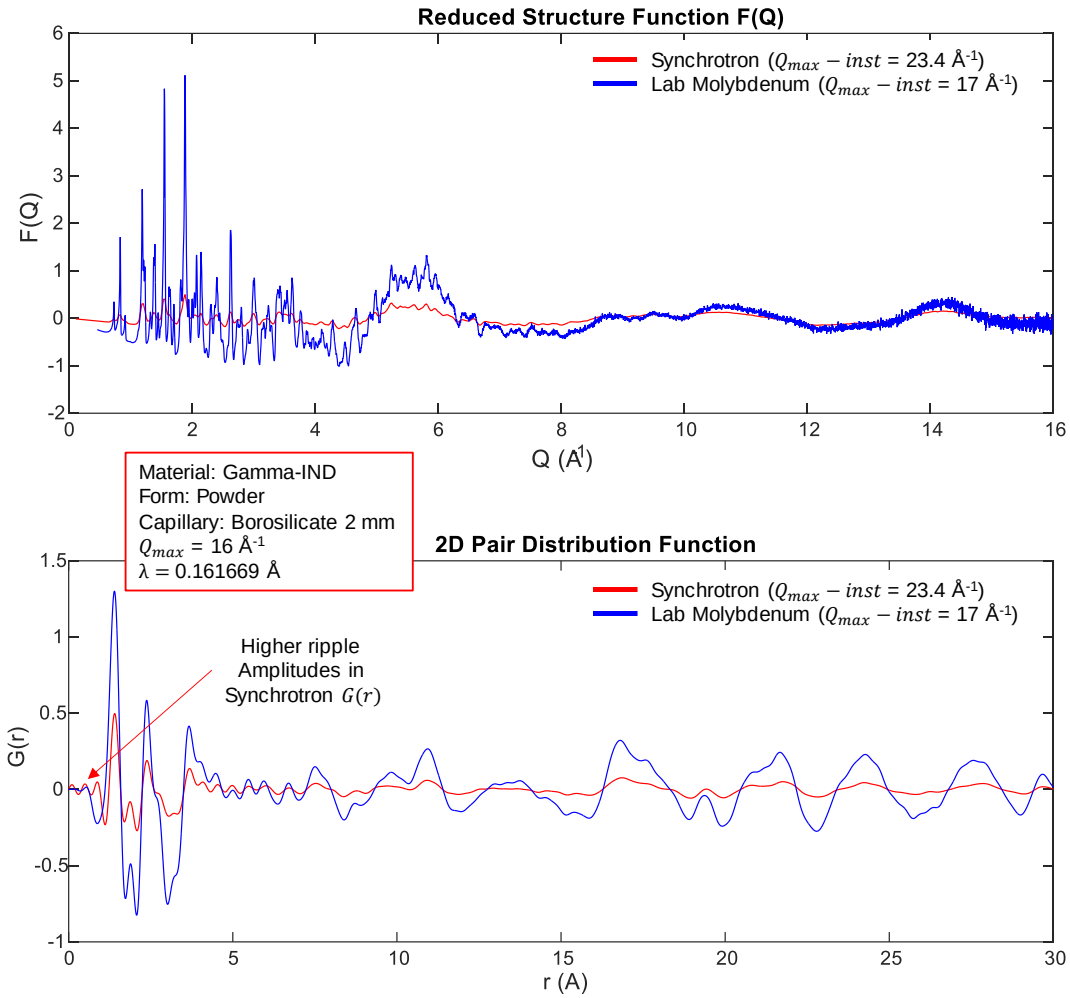

Figure 4.15: Comparison of $F(Q)$ and $G(r)$ from synchrotron and molybdenum sources at different $Q_{max} - inst$.

Generating $F(Q)$ and $G(r)$ at the same $Q_{max} - inst$ of 17 Å^{-1} (Figure 4.16), still well above the limit for reliable PDFs, produces comparable absolute $G(r)$ intensities for both sources, with similar magnitudes in both the F(Q) and $G(r)$ profiles. Also, in the low-r region (<1 Å) the ripple amplitudes are similar. Aside from the lower intensities observed in $F(Q)$, the amplitude differences in the resulting PDFs are due to damping and broadening of features in real-space due to the poorer Q-resolution in the synchrotron data (Takeshi & Billinge, 2012b). The poor reciprocal-space resolution is due to both the angular resolution of the measurement, and shorter wavelength used in the diffraction experiment. These

instrumental factors also complicate any direct absolute statistical comparisons between the data sets, however, it is clear by a visual assessment that the lab PDFs capture all relevant structural information collected in the synchrotron measurements.

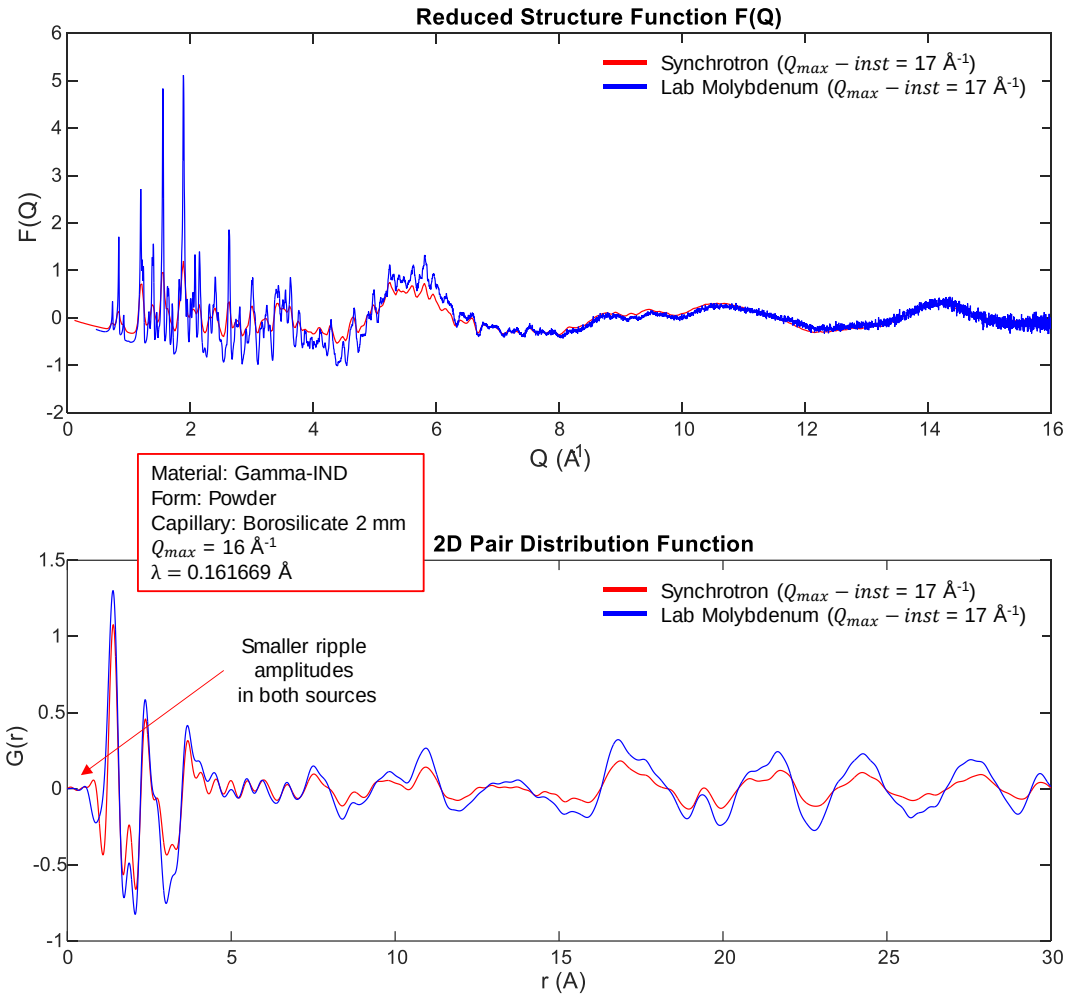

Figure 4.16: Comparison of $F(Q)$ and $G(r)$ from synchrotron and molybdenum sources at same $Q_{max} - inst$.

Overall, both $G(r)$ curves of the reference materials contain unique information that can be used for accurate fingerprinting of a sample. Importantly the lab collected data show excellent reproducibility as well as similarity to the synchrotron derived PDF features confirming their suitability. Aside from the identification of crystalline and amorphous

materials from long-range characteristics, the $G(r)$ curves of Amorph-IND and γ-IND have excellent overlap in the region between 1.5 – 4 Å for both molybdenum and synchrotron PDFs. This suggests the local packing at these distances are similar in the amorphous and crystalline forms of IND, whereas long-range order beyond ca. 10Å is absent in the amorphous material. Given that the molybdenum PDFs compare favourably with the synchrotron data, the value of the laboratory instrument for high-quality PDF generation is demonstrated.

For extrudate ASDs in Figure 4.17 and Figure 4.18, the regions at $r = 3$Å and 3.7Å show a response correspondence with increasing and decreasing IND content. This potentially opens the possibility of using these PDF data quantitatively for identifying unknown IND-PVP ASD compositions.

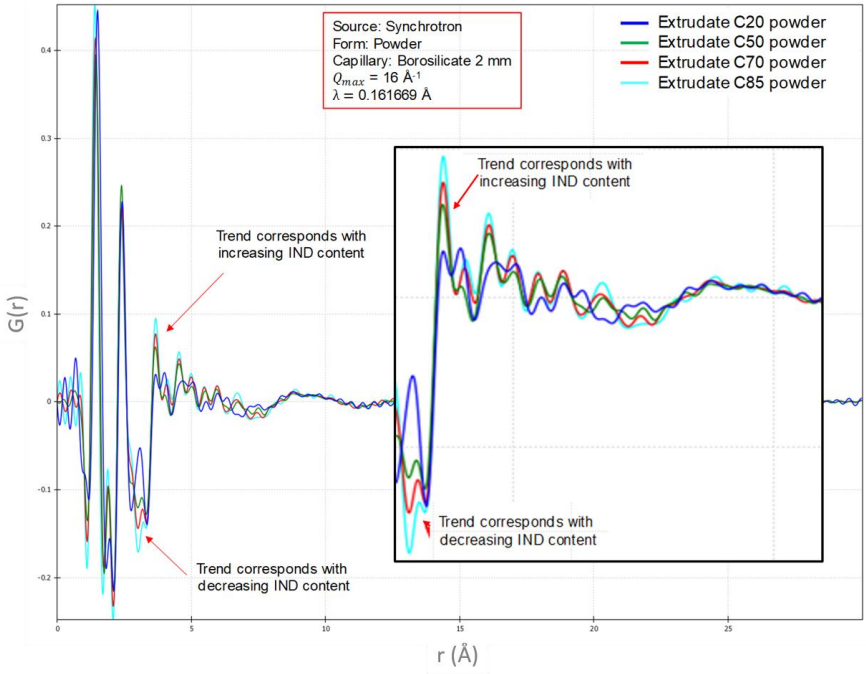

Figure 4.17: Overlaid $G(r)$ curves of extrudate ASD compositions generated from synchrotron X-ray source. Inset shows plot truncated between ~3Å and 10Å for clarity.

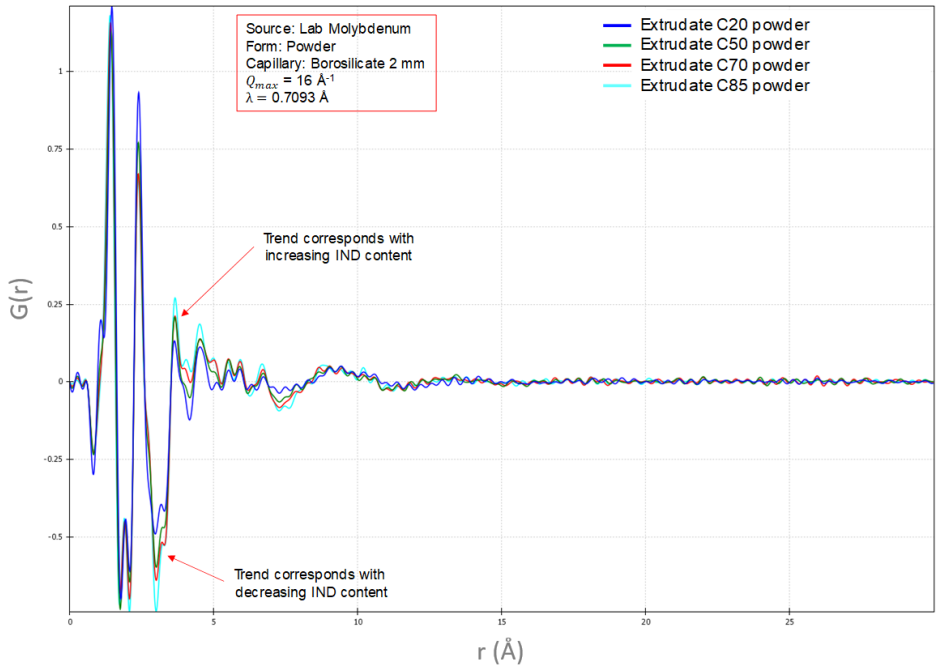

Figure 4.18: Overlaid $G(r)$ curves of extrudate ASD compositions generated from laboratory molybdenum X-ray source.

4.9. Reproducibility of ASD preparation method

To assess the reproducibility of the lab derived data, multiple PDFs were prepared from a fresh ASD sample obtained from a standardised preparation method. Obviously ensuring a consistent ASD sample is an important criterion, avoiding uncontrolled sample variation as a possible contribution to changes observed in $G(r)$ curves. To assess this, three different samples of Melt cooled C85 were prepared on different days following the method in section 2.1.2.3. C85 was used in this instance as it represents a model amorphous pharmaceutical system, and the preparation method had been optimised allowing rapid sample preparation. The freshly prepared Melt cooled C85 samples were analysed using FTIR and DSC as references for control (Appendix 1: Section 3.) as well as using PDF.

FTIR and DSC data (Appendix 1: Section 3) confirmed the consistency of the sample preparation method. The FTIR data of the three samples showed near-superimposable first derivative traces, indicating the high similarity between samples. Similarly, the DSC data were consistent in the detected glass transition temperatures (~46°C) and melting point of recrystallised IND (~154°C), showing that their physical properties are similar. These analyses overall indicate that the chosen preparation method was sufficient to produce reproducible ASD samples.

Comparison of the $G(r)$ curves in Figure 4.19 revealed similar local structures of all samples in the range 1.5 – 6 Å. Given that the FTIR and DSC data indicate that the three samples are similar, this gives further confidence in the preparation method, but also in the reliability of the data collection procedure used for PDF measurements. Based on the results, the standardised preparation method is suitable for the preparation of consistent ASD samples, and the proposed data collection procedure is sufficient for tracking structural changes in ASDs prepared in the lab.

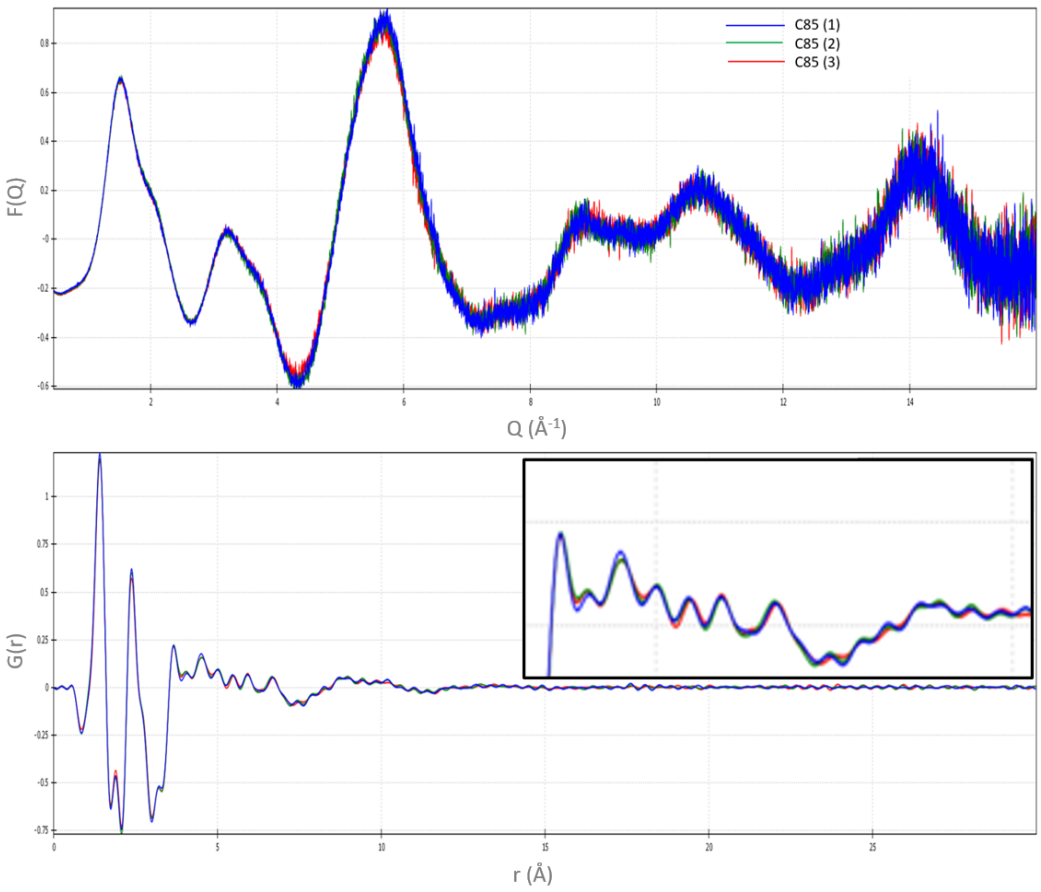

Figure 4.19: F(Q) (top) and G(r) (bottom) curves of day 0 melt cooled C85 powders. Good agreement is shown between low-r features. No crystallinity is apparent from either $F(Q)$ or at high-r in $G(r)$. Inset shows PDF truncated between ~3Å and 10Å for clarity.

The PolySNAP analysis of the G(r) for the fresh C85 samples shows all CCs are below 0.9 (Table 4.6) in contrast with the results from the crystalline material, all of which were above this value. Given the lack of structural information above ~ 9Å, the window of analysis was reduced to 3.5 – 8.5Å (Table 4.7) to minimise the impact of noise on the PDF comparison. As shown, this improves the comparisons between C85 sample PDFs allowing a more informative assessment of the similarity between the PDFs in the region of most interest for amorphous sample comparison. On this basis, it was concluded that the collection procedure used is sufficient to obtain reproducible PDFs of the amorphous material.

Table 4.6: Pearson correlation coefficients between the PDFs generated from molybdenum XRPD data collected from fresh C85 samples, analysed between 3.5 and 30Å. Values lower than 0.9 are highlighted bold.

	C85 (1)	C85 (2)	C85 (3)
C85 (1)	1	**0.83**	**0.886**
C85 (2)		1	**0.879**
C85 (3)			1

Table 4.7: Pearson correlation coefficients between fresh C85 samples, analysed between 3.5 and 8.5Å.

	C85 (1)	C85 (2)	C85 (3)
C85 (1)	1	0.989	0.992
C85 (2)		1	0.991
C85 (3)			1

4.10. Detecting levels of crystallinity with PDF upon aging

With reliable lab derived PDF measurements and sample handling capabilities, it is possible to apply PDF analysis to stability studies and assess the sensitivity of the technique to detecting early signs of crystallinity in ASDs stored for extended periods of time (i.e. aged ASDs). To do this, ASD powders and extrudates were stored for up to 130 days to allow for thermodynamically driven phase separation (AAPS) and recrystallisation to occur. The outcomes of the study are discussed below.

4.10.1. Aged Extrudate C50 glass and powder

Figure 4.20 shows the $F(Q)$ and $G(r)$ curves of Extrudate C50 glass and powder stored for 130 and 125 days, respectively. It was expected that the powder would recrystallise before the glass due to the mechanical stress from the grinding process (Bhugra, Shmeis, & Pikal, 2008). It is apparent from the $F(Q)$ data that there is no detectible crystallinity in any of the samples, due to the absence of Bragg peaks in the diffraction patterns, which are almost superimposable. Similarly, the $G(r)$ curves show no crystallisation, due to the absence of peaks at high-r values. The PDFs remain very similar to the day 0 sample, aside from subtle differences at ~ 3 - 6Å. These small changes are expected to be due to noise/termination ripples. This is supported by the PolySNAP analysis (Table 4.8), wherein all CCs were above 0.9, indicating that the $G(r)$ curves remain highly similar. Thus, the PDF and diffraction data demonstrate that C50 is stable over the storage period. These results are supported by triplicate DSC traces of EXTD-C50 powder D125 (Appendix 1: Section 4.) which also showed no signs of recrystallisation, suggesting stability after prolonged storage.

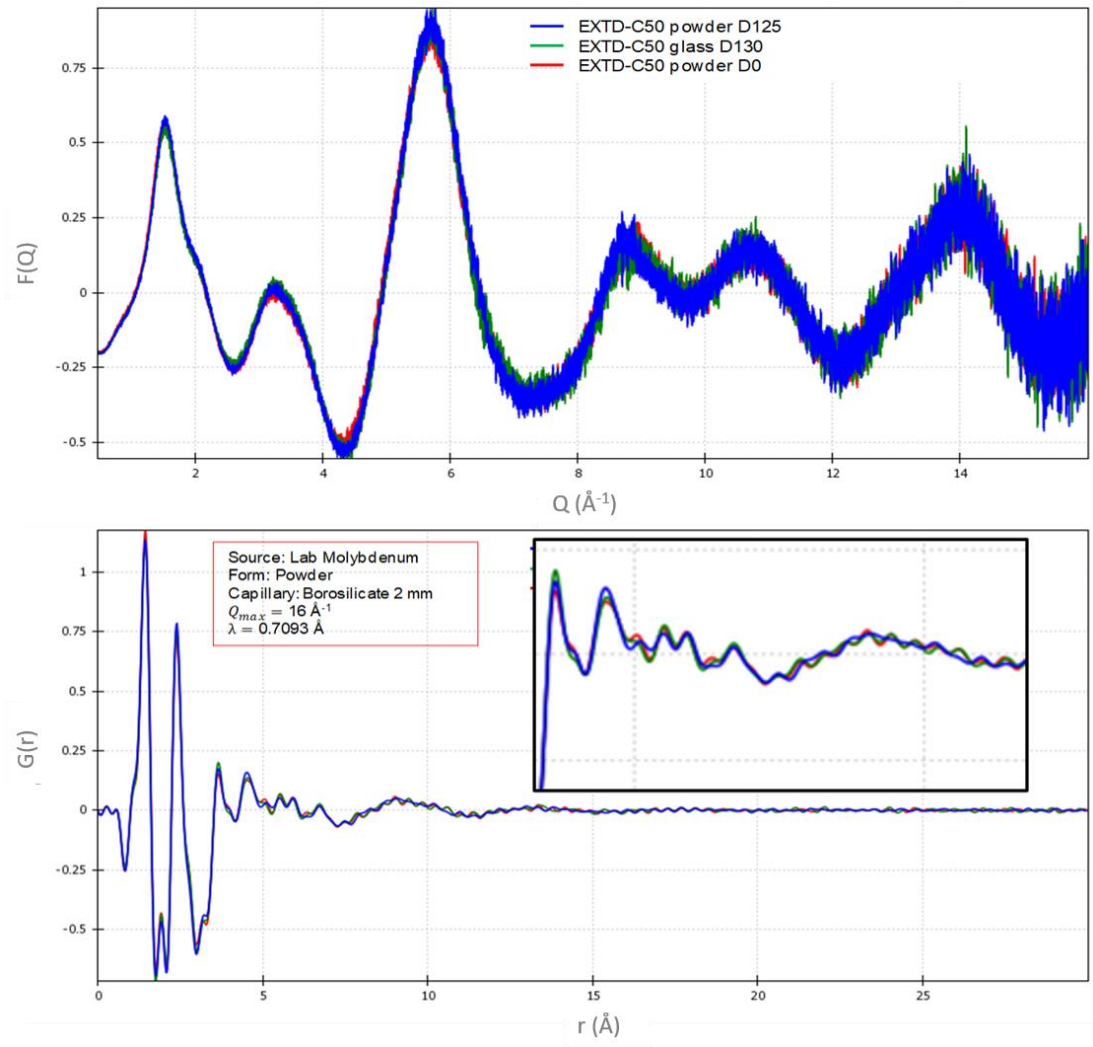

Figure 4.20: $F(Q)$ and $G(r)$ of Extrudate C50 stored in glass and powder form. Inset shows PDF truncated between ~3Å and 10Å for clarity.

Table 4.8: Pearson correlation coefficients between the PDFs generated from molybdenum XRPD data collected from fresh and aged C50 samples, analysed between 3.5 and 30Å.

	C50 powder D125	C50 glass D130	C50 powder D0
C50 powder D125	1	0.963	0.975
C50 glass D130		1	0.98
C50 powder D0			1

4.10.2. Aged Extrudate C70 glass and powder

Figure 4.21 shows the $F(Q)$ and $G(r)$ curves of Extrudate C70 glass and powder stored for 123 and 127 days respectively. As with C50, the absence of Bragg peaks and the good overlap of $F(Q)$ profiles of EXTD-C70 powder D127 and EXTD-C70 glass D123 with the day 0 Extrudate C70 powder (EXTD-C70 powder D0) suggests that both samples were stable for the duration of storage. The PDF data also reflects this, where the curves are visually very similar, and CCs are all above 0.9 (Table 4.9). However, the DSC trace of EXTD-C70 powder D127 detected two overlapping T_gs, revealing the presence of DSC-detectable phase-separated amorphous domains (AAPS) (Appendix 1: Section 4.). This demonstrates the enhanced stability of the C70 glass compared to the powdered form, as would be expected due to the increased surface area of the powder (Zografi & Crowley, 2002). While the unstable nature of this system has been confirmed using DSC, it is apparent that the PDF is not as sensitive to these subtle changes in the structure, based on the available data.

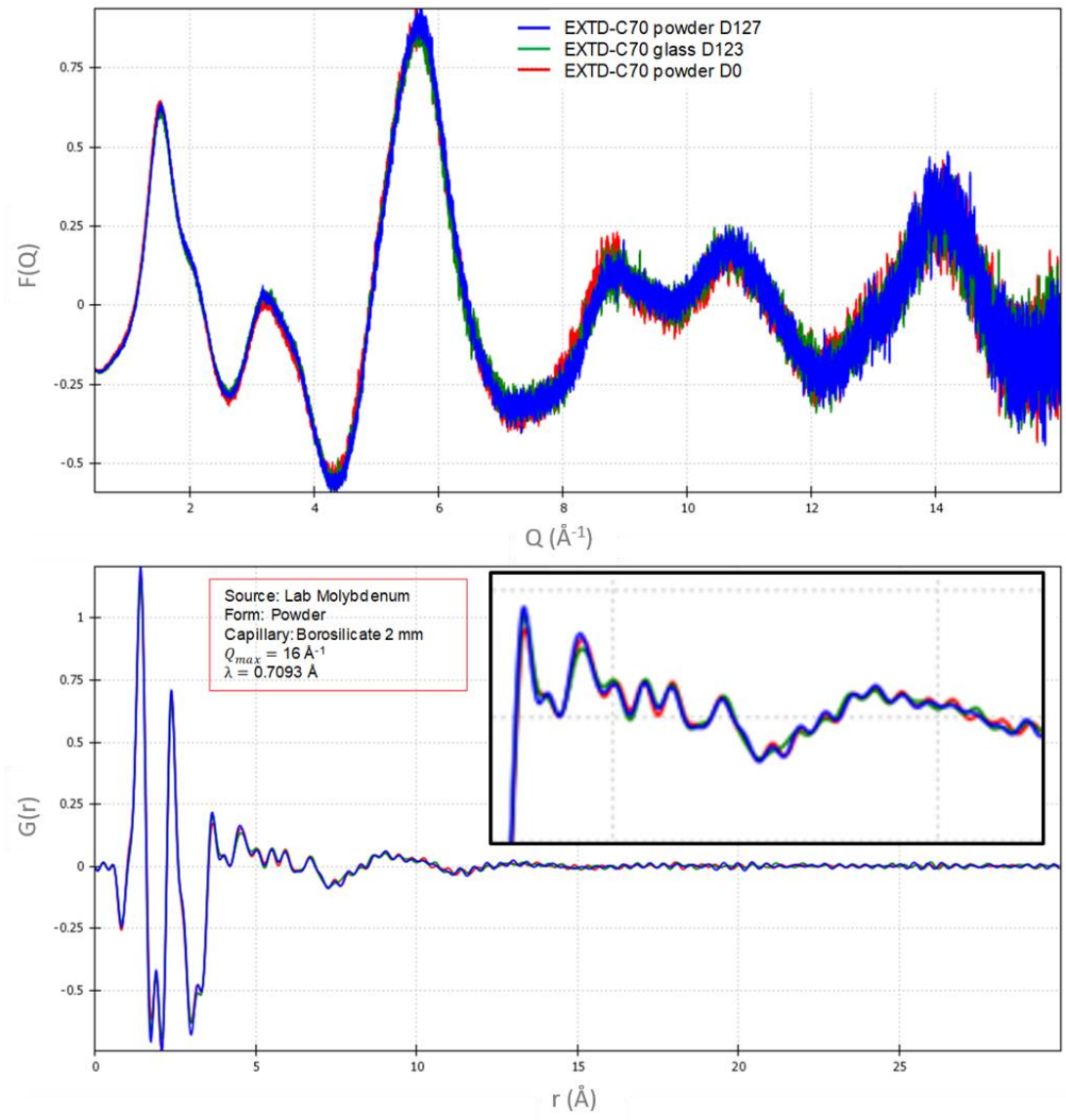

Figure 4.21: $F(Q)$ and $G(r)$ of Extrudate C70 stored in glass and powder form. Inset shows PDFs truncated between ~3Å and 10Å for clarity.

Table 4.9: Pearson correlation coefficients between the PDFs generated from molybdenum XRPD data collected from fresh and aged C50 samples, analysed between 3.5 and 8.5Å.

	C70 glass D123	C70 powder D127	C70 powder D0
C70 glass D123	1	0.984	0.982
C70 powder D127		1	0.984
C70 powder D0			1

4.10.3. Aged Extrudate C85 glass and powder

In Figure 4.22, the $F(Q)$ of EXTD-C85 powder D128 (Extrudate C85 powder aged for 128 days) and EXTD-C85 glass D134 (Extrudate C85 glass aged for 134 days) show similarity in the overall signal profile associated with the amorphous system except for the emergence of sharp peaks between 2Å^{-1} – 7Å^{-1} in the $F(Q)$ data for EXTD-C85 powder D128, indicating emerging crystallinity from the bulk amorphous sample. Similarly, the emerging crystallinity is apparent in the PDF, where subtle increases in the peak amplitudes are evident above 10Å. Given that these features are present beyond the r-values associated with intramolecular contacts in the IND molecule, it can be concluded that these features are consistent with the emergence of long-range order in the otherwise amorphous sample. These data confirm the ability of PDF to detect relatively low levels of crystallinity in ASDs. This is consistent with the DSC data (Appendix 1: Section 4.) where a crystallisation even was observed at ~110°C, with a subsequent melting event. The results suggest that extruded C85 is less stable when stored as a powder than as extrudate and less kinetically stable than extrudate C70 powder. This is in line with theoretical PC-SAFT predictions (Prudic et al., 2015) as the C70 composition is closer to the solubility limit of IND in PVP. By contrast, C85 is supersaturated, far less stable, and therefore expected to phase-separate and recrystallise more readily. The amount of

crystalline IND, X_c, already present in the EXTD-C85 powder D128 sample was determined as 2.61% by DSC (Appendix 1: Section 5.).

Further work would be required to quantify the level of crystallinity directly from the PDF trace. An overlay of the $F(Q)$ profiles of EXTD-C85 powder D128 and γ-IND confirmed that the peaks detected at 2.3 $Å^{-1}$, 3.2 $Å^{-1}$, 4.5 $Å^{-1}$, and 4.9 $Å^{-1}$ (red tick marks in Figure 4.23) were characteristic of a small amount crystalline γ-IND in the aged powder sample.

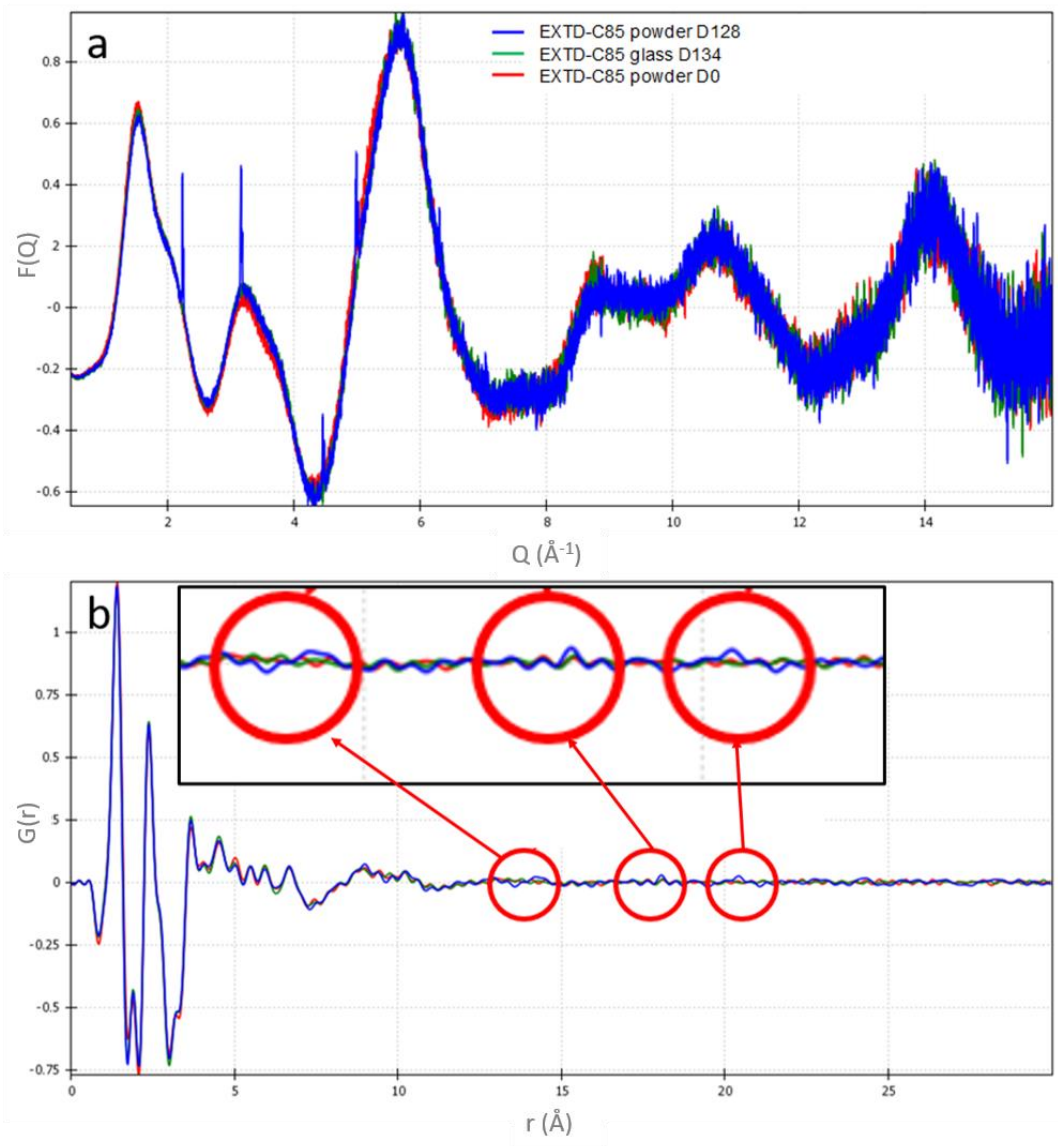

Figure 4.22: a) $F(Q)$ and b) $G(r)$ of Extrudate C85 stored in glass and powder form. Emergence of crystallinity is apparent in the PDF of the aged C85 powder (blue), highlighted in red. Inset shows highlighted regions for clarity. Crystallisation evident in $F(Q)$ by the presence of Bragg peaks in the profile.

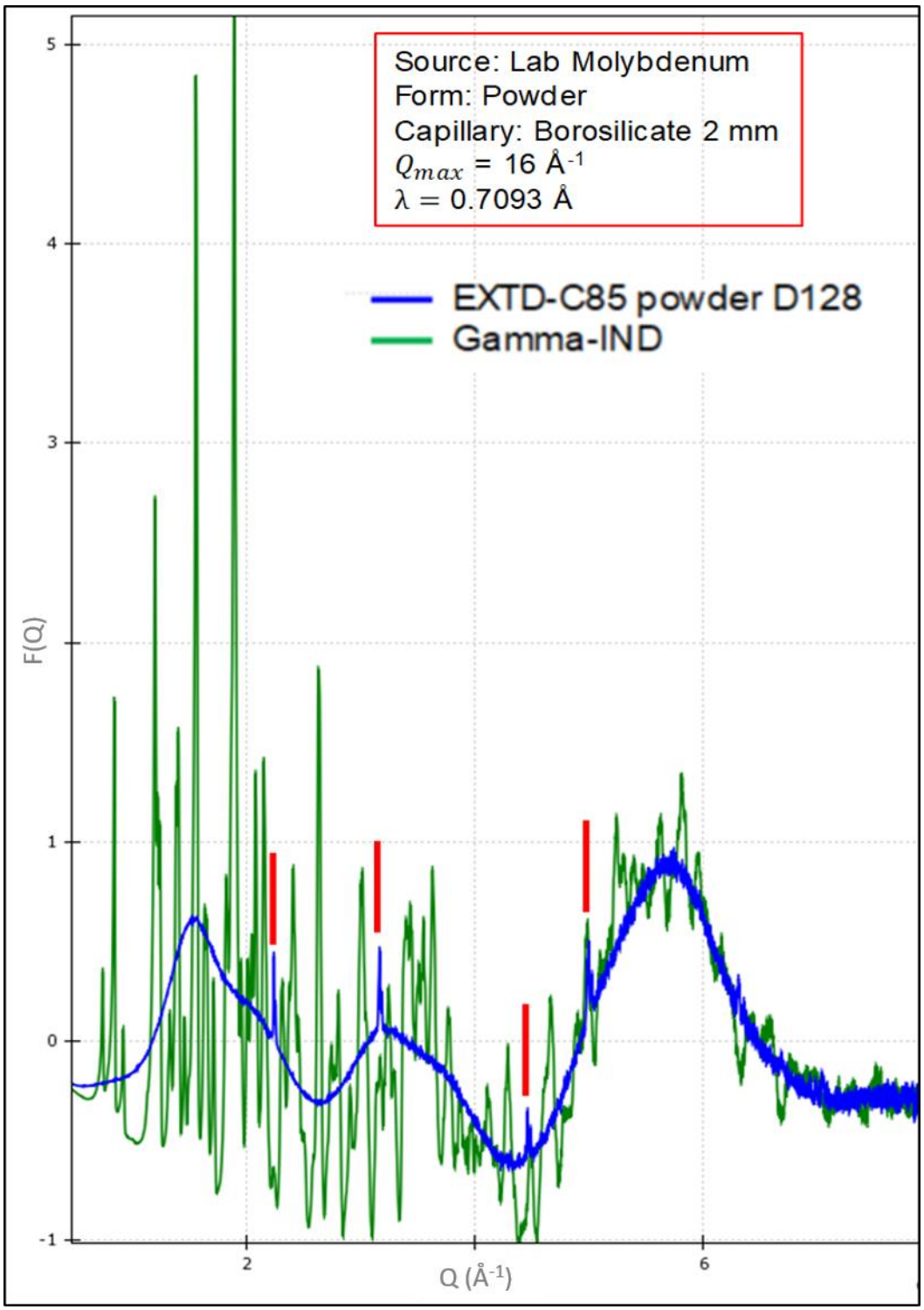

Figure 4.23: Overlay of $F(Q)$s of Gamma-IND and EXTD-C85 powder D128. Red tick marks indicate Bragg diffraction peaks observed in the aged sample.

4.11. Investigating structural changes with PDF: Effect of moisture

The effect of moisture on the local structure of Extrudate C20 and Processed PVP was investigated by drying, humidifying, and re-drying samples of both materials according to the procedure outlined in Section 2.6. $G(r)$ curves were generated from laboratory molybdenum X-ray total scattering measurements as described in Section 2.2.4, and samples were also analysed by FTIR and TGA to correlate moisture content to the observed $G(r)$ curves.

4.11.1. Processed PVP

Figure 4.24 shows the $F(Q)$ and $G(r)$ data for the dried, humidified and re-dried PVP samples. The difference between the wet and dry samples is strongly evident in the $F(Q)$ with significant changes in the 1-8A^{-1} region, where striking changes in the peak positions and intensities are observed. Re-drying largely reversed these changes, with the re-dried $F(Q)$ being almost superimposable with that of the original sample.

This is also observed in the PDF data. Re-drying the humidified sample reversed the effects shown in the PDF, and resulted in closely overlapping $G(r)$ curves for DWD-PVP-DRIED and DWD-PVP-REDRIED with subtle differences indicating the difficulty in controlling residual moisture content in the re-dried sample. The reversible effect of low moisture content on the $G(r)$ confirmed the sensitivity of PDF to changes in the local structure of Processed PVP that arise from the absorption of water altering the local molecular packing. Differences between the wet and dry powder samples were also detected by FTIR analysis with a shift of the non-H-bonded amide C=O stretch (1676 cm^{-1} peak) to the right (Appendix 1: Section 6.).

The $G(r)$ for DWD-PVP-WET showed distinct differences in the low-r region when compared with DWD-PVP-DRIED and DWD-PVP-REDRIED. A relative increase in moisture content in wet PVP of 8.5 wt% (Appendix 1: Section 6.) resulted in reduced peak heights between 1.5 – 3.5

Å, as well as significant distortions to features and shifts in positions of peaks between 3.6 – 10 Å. There is striking variation in the peak amplitudes between 3.5 and 5 Å between the dry and wet sample. The intensities of these features are significantly higher for the wet material, indicating possible interaction of water with the PVP chains, suggesting an increase in molecular contacts at this range. The effect of such structural changes on the PDF have not been extensively studied, and as such, are poorly understood. Above this range, there is a slight drop in the amplitudes of the features when compared to the dry sample. Again, while the specific structural meaning of these observations cannot be definitively explained at this stage, it is possible that this is evidence of a subtle decrease in inter-chain contacts. This is potentially due to the plasticising effect of water, which results in swelling of the material and an increase in the inter-chain distance (Hodge, Bastow, Edward, Simon, & Hill, 1996).

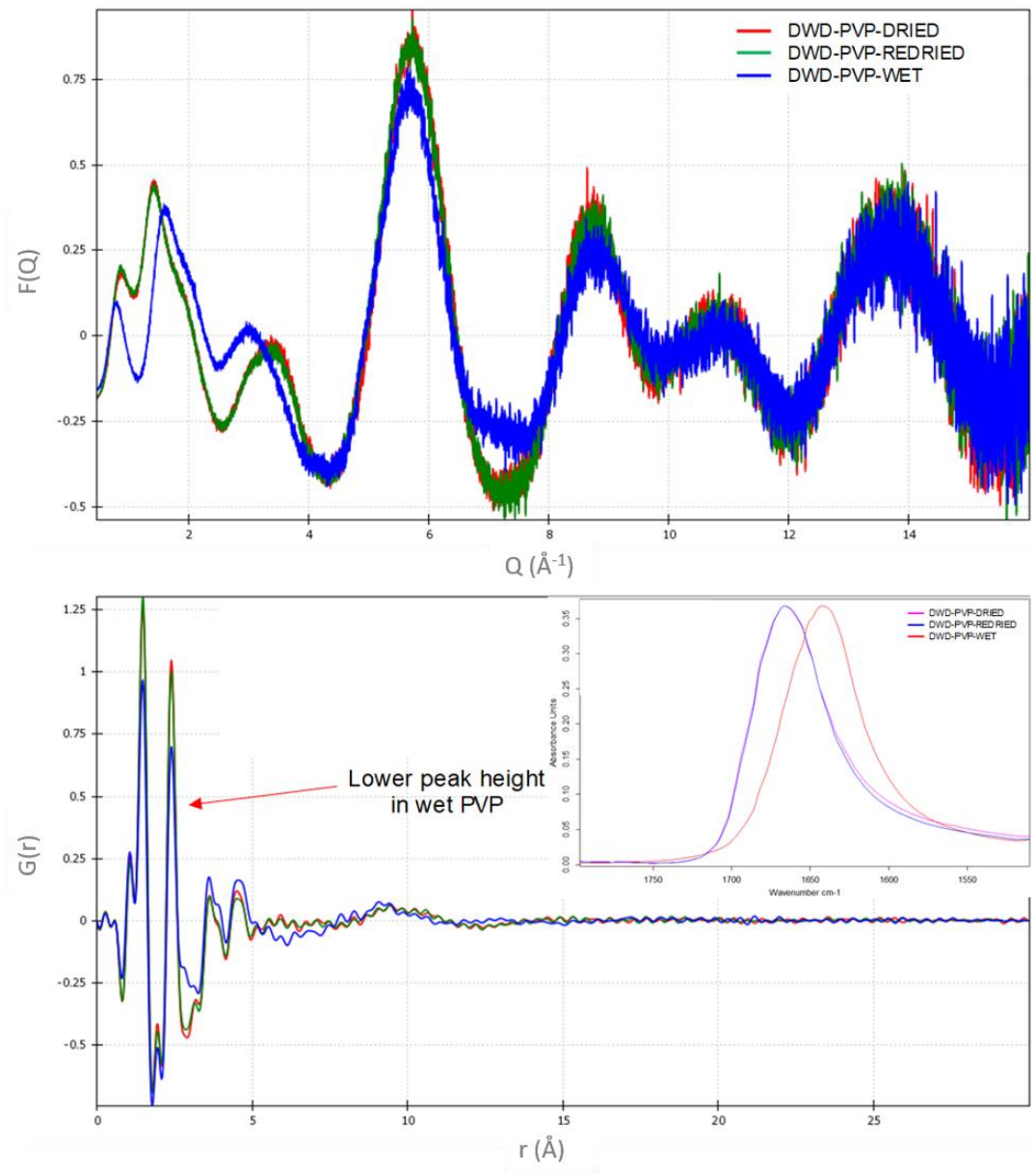

Figure 4.24: $F(Q)$ (top) and $G(r)$ (bottom) curves of dried, wet, and re-dried Processed PVP. Inset shows the corresponding FTIR data. Peak shift from ~1660cm^{-1} to ~1640cm^{-1} indicative of moisture uptake in the wet sample

4.11.2. Extrudate C20

Figure 4.25 shows a similar reversible effect of moisture on the local structure of Extrudate C20 powder. The relative increase in DWD-C20-WET of 8.3 wt% (Appendix 1: Section 6) resulted in both the $F(Q)$ and $G(r)$ showing significantly different profiles with distortions to features and shifts in peak positions between 3.6 – 10 Å. The effect of water on C20 is similar to that of the pure PVP sample, indicating that the interactions of water with the polymer are not disturbed by the presence of the IND molecules. This is an expected finding as C20 is below the solubility limit of the IND-PVP system, allowing both water and IND to interact with the polymer chains. Although there may also be water-drug interactions, this was not evident from the available data.

Drying of C20 appeared to be more complete than for pure PVP evidenced by identical $F(Q)$ and $G(r)$ for the raw and redried samples. The reversible effect indicated that Extrudate C20 powder remained physically stable and did not undergo any permanent structural changes resulting from prolonged exposure (~12 hours) to moisture at 21 °C. This is consistent with the fact Extrudate C20 is a sub-saturated amorphous solid solution at room temperature and would be expected to remain thermodynamically stable. FTIR also detected the presence of moisture in the DWD-C20-WET sample with a corresponding shift of the 1676 cm^{-1} peak to the right (Appendix 1: Section 6), indicating the hydrogen bonding of water to carbonyl functional groups.

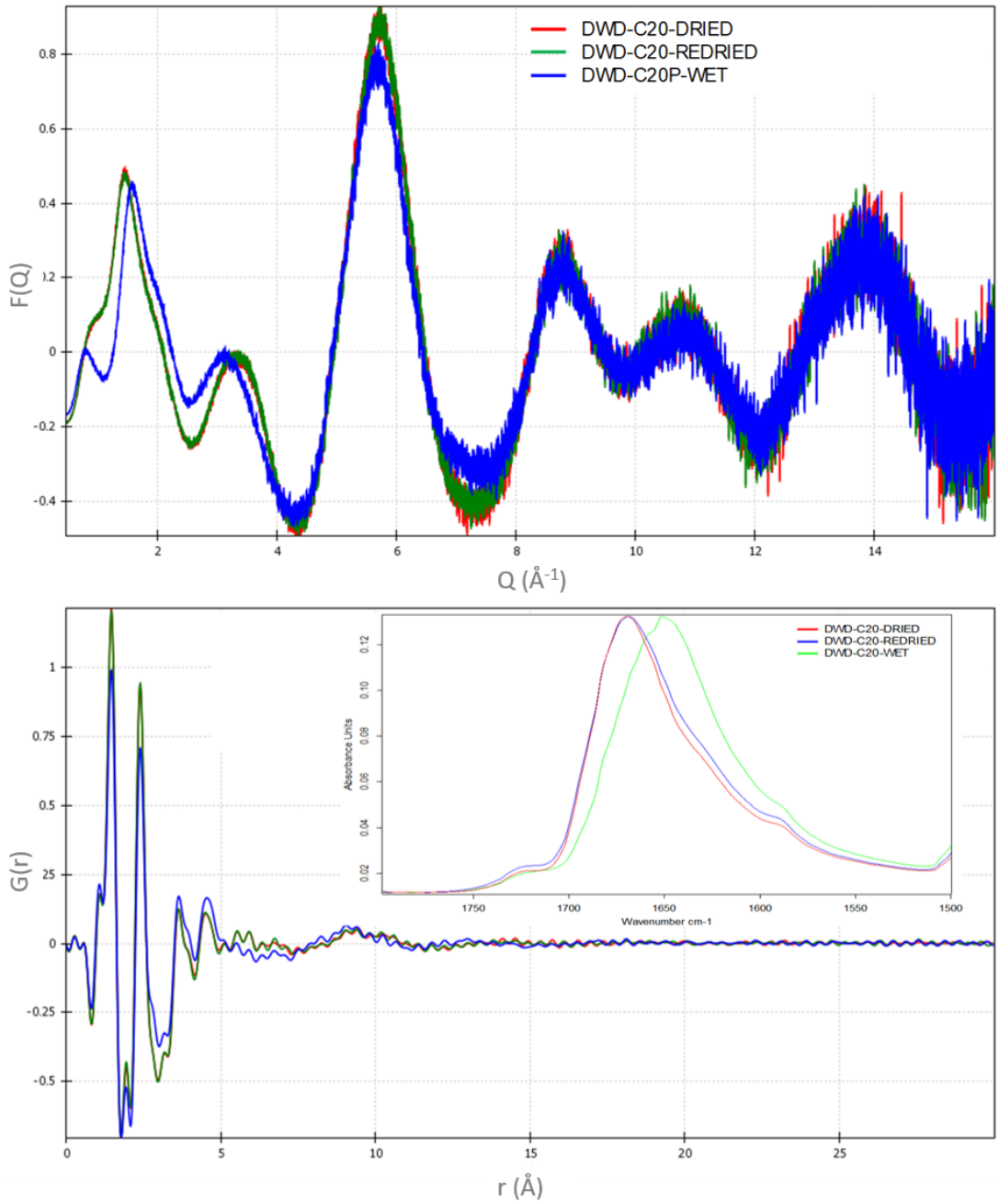

Figure 4.25: $F(Q)$ and $G(r)$ curves of dried, wet, and re-dried Extrudate C20 powders. Inset shows the corresponding FTIR data. Peak shift from ~1660cm^{-1} to ~1640cm^{-1} indicative of moisture uptake in the wet sample

4.12. Recommendations for PDF data collection and treatment

Based on the findings from the different data collection, sample presentation, data treatment and analysis approaches evaluated in this section, the following recommendations are made with the objective of enabling access to data with sufficiently large Q_{max} values with appropriate signal to noise to enable the routine generation of high quality, true $G(r)$ curves using laboratory equipment for the study of amorphous pharmaceuticals.

a. Reproducible sample preparation method to ensure consistent sample composition
    i. Use of complementary techniques to ensure sample consistency
    ii. Minimal exposure to environmental humidity and temperature by use of a glove bag or other dry sample container.
b. Optimised data collection for high quality PDF
    i. Implementation of a short-wavelength source (molybdenum or silver) to meet resolution requirements in real-space.
    ii. Use of VCT data collection procedure to optimise data quality, particularly at high-Q.
    iii. Larger sample volume using ≥ 2mm capillaries to improve scattering signal whilst avoiding X-ray absorption from large samples.
c. Obtaining a reliable PDF from diffraction experiment
    i. Identification and minimisation of termination ripples in $G(r)$ to avoid misinterpretation of features (see below).
    ii. Optimising termination point in Fourier transform such that the value of $F(Q)$ at Q_{max} is close to zero.
    iii. Smoothing of $F(Q)$ where necessary to reduce impact of noise in $G(r)$. This is not implemented within xPDFsuite and as such, needs to be performed in external software. j

d. Validation of PDF reproducibility using appropriate statistical comparison

    i. PolySNAP offers a suite of tools for data analysis. Pearson correlation coefficients offer a rapid method for assessment of similarity.

    ii. Selection of suitable analysis window depending on physical state of material.

        i. For crystalline materials where PDF extends to high r-values, a suitable range includes both short and long-range components of PDF, for example 3.5Å to 30Å, as used here.

        ii. For amorphous materials, analysis window can be reduced to reduce impact of noise on any comparisons, for example 3.5Å to 8.5Å, as used here.

4.13. Conclusions

This work has addressed important practicalities that are often overlooked in PDF literature, particularly concerning the analysis of amorphous pharmaceuticals. Discussions have highlighted factors that should be considered when preparing and handling amorphous samples to ensure accurate and reproducible production of homogeneous and fully amorphous ASDs for fingerprinting. Intricacies surrounding total scattering measurements of poorly scattering pharmaceutical materials on in-house laboratory X-ray sources, including appropriate choice of radiation source and attainable Q_{max} have been discussed. Moreover, collection procedures to ensure good counting statistics, including capillary sizing and VCT implementation have been proposed.

Data treatment procedures were also investigated, where it was found that an optimisation of the termination point in the Fourier transform (Q_{max} where $F(Q)$~0) is able to significantly reduce the amplitudes of termination ripples in the resulting PDF (Section 3.2). It was also

determined that smoothing the $F(Q)$ data using a suitable filter, in this case a Sovitzky Golay filter, was able to reduce the impact of noise on the high-r ripples in the PDF, while retaining the key structural features, thus enabling further investigation with confidence in the data quality.

Our investigation has demonstrated that the application of optimised laboratory total scattering data collection procedures outlined in this work can yield high-resolution, reproducible PDF fingerprints of crystalline and amorphous pharmaceuticals that are comparable to synchrotron-generated PDFs (Section 3.3.). In addition, consistent amorphous APIs and ASDs are obtainable through standardised preparation methods as demonstrated in this work. Importantly, we have shown that optimised total scattering measurements of accurately and consistently produced samples can enable reliable PDF analysis to be carried out on amorphous APIs and ASDs of interest.

Regarding the capabilities of PDF analysis in studying amorphous pharmaceuticals, laboratory generated PDF was able to detect structural differences in amorphous materials subjected to stress conditions (Section 3.6.). For instance, the effect of moisture on the local structure of humidified PVP and C20 samples was easily detected by PDF and corroborated by FTIR data. A visual analysis of the PDF suggests that the introduction of water molecules in the polymer chains increases the number of atom-atom contacts at short distances (~ 3.5 – 5 Å). This is potentially due to water-polymer interactions, while simultaneously decreasing the number of interactions beyond this distance, possibly due to an increase in inter-chain distance caused by swelling due to the plasticising effect of water. This demonstrates the need for careful sample handling and storage when dealing with amorphous pharmaceuticals, particularly hygroscopic materials, such as the PVP used in the study. As

expected, moisture-induced structural changes in the undersaturated C20 sample were reversible owing to its thermodynamically stable nature.

Stability studies performed on ASDs of IND-PVP demonstrated the sensitivity of laboratory generated PDFs in detecting the presence of crystalline domains (Section 3.5.). In particular, laboratory generated PDF was found to have a limit of detection of crystallinity beyond 2.6% in phase-separated ASD. Additionally, the physical stability of IND-PVP ASDs was found to be dependent on the physical state in which samples were stored. For example, a powder sample of supersaturated C85 ASD was found to be less stable than the extrudate glass form as recrystallisation occurred sooner, as detected by PDF and DSC. As expected, powder samples of C50, C70, and C85 behaved as kinetically stable glasses, with varying degrees of AAPS and recrystallisation occurring after finite storage time according to increasing IND wt%. Interestingly, all extrudate glass samples remained stable for the period of storage. The effect of surface area on the physical stability of amorphous indomethacin has been investigated (Zografi & Crowley, 2002), where increased surface area was found to correspond to a decrease in the physical stability of the amorphous material.

Overall, the work shown herein demonstrates that PDF analysis has potential to enable better understanding of amorphous materials alongside other traditional characterisation methods. Further optimisation of the data collection procedures may be necessary to obtain high-quality PDF data at shorter time scales. This will enable routine in-house rapid fingerprinting of amorphous pharmaceutical materials with respect to preparation method, and the study of structural changes in metastable formulations. Additionally, it can be useful to explore other methods, such as molecular modelling/ MD simulations to assist in extracting structural information from the PDF of amorphous pharmaceuticals, as will be demonstrated in the following experimental chapters.

Chapter 5: Structure and stability in amorphous AZD5718 prepared by different manufacturing routes

5.1. Introduction

AZD5718 is a novel active pharmaceutical ingredient (API) currently under development for the treatment of coronary artery disease (Ericsson et al., 2018). The drug shows promise as a potential treatment, however, aqueous solubility of the crystalline phase (AZD5718 Form A), of which there is one known single-component form, is poor. Hence, the amorphous phase is of interest for the purposes of targeting higher solubility (Hancock & Parks, 2000). However, lack of structural understanding of amorphous solids creates a challenge to characterize and explain preparation-dependent differences in properties of the amorphous API. The implementation of pair distribution function (PDF) analysis, coupled with molecular dynamics (MD) simulations is of considerable interest therefore as means to provide a greater understanding of the structure of amorphous systems, the impact of processing conditions on the resultant structure and therefore stability and performance,, and on the structural changes that may accompany phase changes during manufacture, storage and/or during administration (Y. Chen et al., 2016; Karmwar et al., 2011; Karmwar, Graeser, Gordon, Strachan, & Rades, 2012; J. E. Patterson et al., 2005).

The known crystal form of *AZD5718* has a $P2_12_12_1$ orthorhombic unit cell (a = 5.67096, b = 18.00112, c = 21.2832), shown in Figure 5.1. The structure largely forms chains and rings of up to six *AZD5718* molecules. The crystalline conformer contains an intramolecular hydrogen bond, between a carbonyl oxygen and the nitrogen on the carboxamide group, resulting in an “L-shaped” conformer, which has been confirmed as the global energy minimum by density functional theory (DFT) calculations at the B3LYP-D/6-31G level of theory.

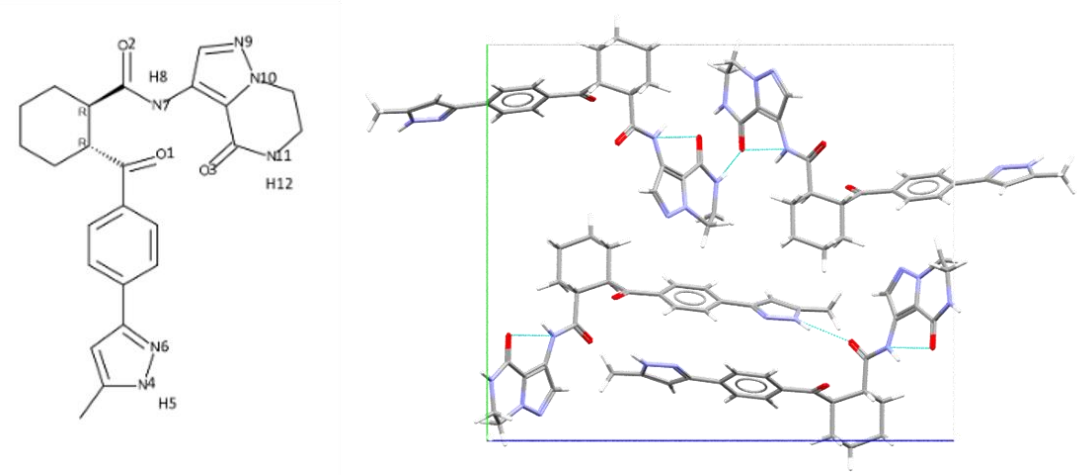

Figure 5.1: Left: molecular structure of AZD5718. Right: Crystal structure of AZD5718, viewed down crystallographic a axis.

As mentioned, the poor solubility of the Form A form has led to the investigation of this drug as an amorphous phase to access the higher solubility and dissolution rates that accompany the increased free energy. Amorphous AZD5718 was originally prepared using spray drying (Norberg, *personal communication,* September 2019), however, due to the poor scalability of the method, an alternative preparation route was sought. A continuous precipitation method was developed for production of the material (Siddique, *personal communication,* September 2019). Both methods have been demonstrated to successfully deliver amorphous AZD5718. Whilst confirmation of structural similarity is straightforward for crystalline materials using Bragg diffraction for example, it is less straightforward when considering amorphous systems (Stachurski, 2011). Hence the interest in amorphous AZD5718 is to investigate whether PDF can confirm that both routes produce equivalent amorphous materials or if any measurable differences can be observed. In principle, changes in the underlying structure of an API in the solid state can lead to differences in the manufacturability, stability or performance of the final product (Karmwar et al., 2011), where it has been found that for amorphous indomethacin, the preparation method (milling, melt-

quenching or spray drying) affected both the local structure of the amorphous solid (assessed using Raman spectroscopy) and the physical stability against recrystallisation. For AZD5718 specifically, whilst both routes result in the formation of an amorphous material by standard XRPD and DSC analysis, differences in particle size/morphology was observed.

The possibility of polyamorphism, that is, the ability of an amorphous material to exist as two or more distinct thermodynamic forms, has been suggested for pharmaceuticals (Hancock, Shalaev, & Shamblin, 2002). Examples include the reports of separate amorphous phases of paracetamol, affecting polymorphic outcome upon recrystallisation (Nguyen Thi, Rademann, & Emmerling, 2015), and the alleged amorphous forms of D-mannitol, which have shown apparent evidence of a first-order phase transition (Zhu et al., 2015; Zhu & Yu, 2017). This is a well-established phenomenon for inorganic materials, particularly with water, where low- and high- density amorphous ices have been extensively examined (Finney et al., 2002; Narten et al., 1976; Tulk et al., 2002). Despite this, there remains limited evidence of thermodynamically distinct amorphous phases of APIs, and no reported examples where a complete assessment of structure, thermodynamics and phase transitions has been performed.

MD simulations have been used to study structure in amorphous pharmaceutical materials (Gupta, Nunes, Vyas, & Jonnalagadda, 2011; Lerbret et al., 2009; Xiang & Anderson, 2013, 2014). However, most have lacked the complementary structural measurement provided by PDF analysis, with a single exception being a study of amorphous felodipine in a polymeric dispersion (Geddes, Blade, McCabe, Hughes, & Goodwin, 2019). As such, the majority of studies have had no way to confirm that the simulated structures were a true reflection of the material itself. Clearly therefore the combination of MD and PDF,

supplemented by traditional thermal, diffraction and spectroscopic techniques, offer the potential to provide true understanding of structure in these complex materials.

Here, samples of AZD5718 produced by both spray drying and continuous precipitation were assessed by PDF analysis. These experimental data were compared to MD simulations to give insight to the structure of the material and provide understanding of the physical properties of amorphous AZD5718 and their process dependence.

5.2. Materials and methods

5.2.1. Preparation of amorphous AZD5718

Samples of spray dried (SD) AZD5718 were received from AstraZeneca. Precipitated (PREC) AZD5718 was prepared at the University of Strathclyde using a continuous precipitation method (Siddique *et al.*, *unpublished,* 2019).

5.2.2. X-Ray powder diffraction

High resolution X-Ray powder diffraction (XRPD) data for AZD5718 Form A was collected on a Bruker D8 Advance II diffractometer, equipped with a copper ($k\alpha 1$ = 1.54 Å) anode and a Lynxeye Series 1 position sensitive detector, for the purposes of phase identification. The sample was loaded into a 0.7 mm borosilicate glass capillary (Capillary Tube Supplies Ltd, UK) and flame sealed. Data were collected between 3° and 40° 2θ, with step-size 0.017 ° 2θ at 10 seconds per step.

For PDF analysis, data were collected on a PANalytical EMPYREAN series 2 diffractometer, equipped with a molybdenum ($k\alpha 1$ = 0.7093 Å) anode, and a GaliPIX^{3D} detector. Samples were loaded into 3.5 mm borosilicate glass capillaries in a glove bag under

a dry nitrogen flow, to minimize moisture absorption from environmental humidity. Capillaries were then removed from the glove bag to be flame sealed. For a reference PDF of water, de-ionised water was injected into the capillary using a syringe and flame sealed. A variable count time (VCT) procedure (Shankland, David, & Sivia, 1997) was used to optimise counting statistics across 2θ, with a constant step-size of 0.3° 2θ. Angular ranges and counting times at each range are shown in Table 5.1.

Table 5.1: XRPD VCT data collection parameters for PDF analysis

Angular range (° 2θ)	Time per step (s)
3 - 32	18
32 – 44	36
44 – 61	72
61 - 147	216

5.2.3. Pair distribution function

For PDF analysis, data were collected as described in Section 2.2. GudrunX (A K Soper, 2011; Alan K. Soper & Barney, 2011) was used to apply corrections and normalisations for background, multiple scattering, polarisation, Compton scattering and atomic form factor to extract the differential scattering cross-section for transformation to the PDF, using a Q_{max} of 16.95 $Å^{-1}$. A Soper-Lorch (Alan K. Soper & Barney, 2011, 2012) broadening of 0.08 Å was applied to minimise the impact from ripples arising as a result of termination of the Fourier transform at finite value.

5.2.4. Molecular dynamics simulations

MD simulations were received from AstraZeneca. Simulations of AZD5718 containing 0, 0.5, 1, 2 and 4 wt% water were provided. AZD5718 – water compositions for each simulation are shown in Table 5.2. The Materials Studio (San Diego, USA) Amorphous Cell module was used to generate starting models for amorphous AZD5718. MD Simulations of AZD5718 were then performed using the GROMACS code (Hess, Kutzner, Van Der Spoel, & Lindahl, 2008), employing the optimised potential for liquid simulations all atom (OPLS-aa) force field (Jorgensen et al., 1996). The NPT ensemble was used during equilibration, which was run for 600 ns. The Berendsen thermostat (Berendsen et al., 1984) was used for temperature control and the Parrinello-Rahman barostat (Parrinello & Rahman, 1980) was used for pressure control during dynamics runs.

Table 5.2: AZD5718-water composition for each simulation

Water content (wt%)	AZD5718 molecules	Water molecules	Total number of atoms
0	128	0	7552
0.5	128	16	7600
1	128	32	7648
2	128	65	7747
4	128	132	7948

5.2.5. Analysis of MD simulations

GROMACS commands *gmx hbond* and *gmx clustsize* were used to assess hydrogen bonding and water clustering in MD simulations, respectively. The last 100 ns were used for all analyses, to ensure the system had equilibrated. For *gmx hbond*, OH and NH groups are always assumed to be donors, whereas O and N are always acceptors. As this is not necessarily the case, the total number of hydrogen bonds were assessed, as opposed to individual donor or acceptor groups. Two molecules are identified as being hydrogen bonded if the donor and acceptor sites are within 3.5 Å of each other.

5.2.6. Calculation of PDFs for MD simulations

PDFfit (Proffen & Billinge, 1999) was used for batch calculation of PDFs from MD simulations. For each frame in every simulation provided, a PDF was calculated, by (Takeshi & Billinge, 2012b)

$$G_{calc}(r) = \left(\frac{1}{r}\right) \sum_i \sum_j \left[\left(\frac{f(Q)_i f(Q)_j}{\langle f(Q)\rangle^2}\right) \delta(r - r_{ij})\right] - 4\pi r p_0 \qquad (5.1)$$

where $f(Q)_i$ and $f(Q)_j$ are the atomic form factors of atoms i and j, respectively, $\langle \dots \rangle$ indicates the compositionally averaged form factor, and p_0 is the average number density. To obtain the time-averaged PDF for comparison to experimental data, the mean PDF was determined for each trajectory. PDFs were also calculated for a range of AZD5718 single molecules with a distribution of conformations, as well as a number of PDFs for two interacting AZD5718 molecules in the simulation box, for which the mean PDF was also

obtained, for comparison to experimental data. A Q_{max} of 100 Å^{-1} was used in PDF calculations to minimise simulated termination ripples in calculated data, to mimic the Soper-Lorch broadening applied to experimental data.

5.2.7. PDF and Pawley refinements for AZD5718 Form A

Pawley and PDF refinements were performed using TOPAS Academic Version 6 (A. A. Coelho, 2003; A. A. Coelho et al., 2015; Alan A. Coelho, 2018). For Pawley refinements (Pawley, 1981), peak profile, zero error and axial divergence were fit against NIST silicone 640c standard reference material. For PDF fits, reciprocal-space resolution-dependent broadening and damping were refined against the same standard material. The molecular Z-matrix was generated automatically using DASH (David et al., 2006). Molecular translation and rotation were refined, along with scale factor, zero-error and isotropic temperature factors for each atom (modified to increase as a function of interatomic distance, r, accounting for correlated motion at low-r).

5.2.8. Statistical analysis

PolySNAP 3 (Barr et al., 2004b) was used to compare experimental PDFs, by generating Pearson correlation coefficients (CC) (Szczepańska, 2011) between each dataset, to quantify the similarity between PDFs. Using this method, a CC of one indicates a complete correlation between two datasets, zero indicates no correlation, and -1 implies a negative correlation. For all datasets, comparisons were performed between 3.5 and 8.5 Å.

5.2.9. Aromatics analyser

The aromatics analyser (AA) functionality within the Mercury software (Bruno et al., 2002b) from the Cambridge Crystallographic Data Centre (CCDC) was used to assess the strengths of aromatic interactions in amorphous and AZD5718 Form A. Aromatic interaction strengths are predicted using a neural network, which is based upon quantum mechanical calculations performed on phenyl rings with differing relative geometries. Interactions are then scored between 0 and 10 and ranked as either weak (0 – 3), moderate (3 – 7) or strong (7 – 10). For AA calculations, 12 molecules were picked at random from the first frame of the last 100 ns of the trajectories corresponding to 2% and 4% water (closest to experimental). The 12 molecules represent approximately 10% of the total number of AZD5718 molecules present within each simulation.

5.2.10. Hydrogen bonding analysis

The H-bond Coordination Quick-View module within Mercury was used to assess feasibility of hydrogen bonding in amorphous and AZD5718 Form A. Bonding is rapidly assessed by calculating probabilities for coordination numbers of hydrogen bonds, based on available structural data in the crystal structure database (CSD). The stability of hydrogen bonding networks can thus be quickly assessed by comparing observed coordination numbers to the optimal values, with a low probability suggesting that a more favourable hydrogen bond network exists for the material.

5.3. Results and discussion

5.3.1. Characterisation of reference materials

While both preparation routes result in the formation of amorphous AZD5718, some differences exist between the materials. For example, small agglomerated particles were produced via the PREC method, whereas SD AZD5718 exhibited larger, spherical particles, observed using scanning electron microscopy (SEM) (Norberg, *personal communication,* September 2019). Additionally, it was determined that PREC had a higher water content than the SD powder, determined by thermal analyses (Siddique, *personal communication,* September 2019). While both amorphous materials exhibit long term physical stability, understanding of the impact of preparation method on the underlying structure is necessary to determine whether a consistent product is produced.

5.3.2. XRPD analysis of AZD5718 reference materials

5.3.2.1. Copper XRPD of AZD5718 Form A

XRPD data for AZD5718 Form A are shown in Figure 5.2. The presence of Bragg peaks gives clear evidence of the long-range order in the sample. For the purposes of phase identification, and a rapid purity assessment, a Pawley fit (Pawley, 1981) was performed on the known single crystal unit cell parameters and space group, against the data obtained for the crystalline material (Figure 5.2). The excellent fit (R_{wp}= 3.99 %) obtained gives confidence in the phase purity of the sample, meaning that any further analysis and comparison to SD and PREC AZD5718 is representative of the AZD5718 molecule, and the known crystalline phase.

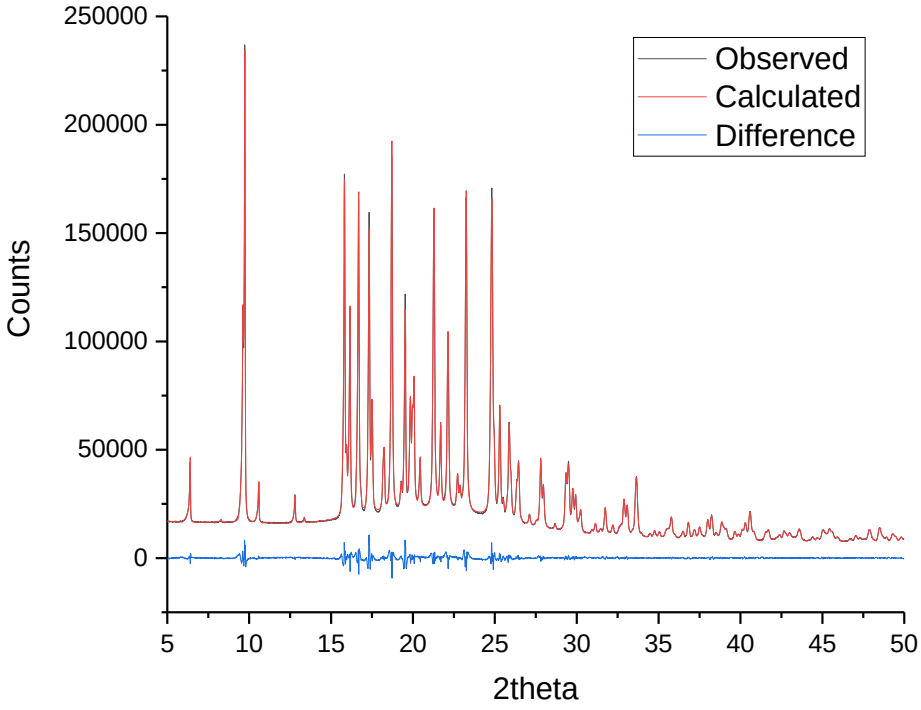

Figure 5.2: Pawley fit to the XRPD data for AZD5718 Form A, with red showing observed profile (y_{obs}), black the calculated profile (y_{calc}), and blue the difference plot ($y_{obs} - y_{calc}$).

5.3.2.2. Molybdenum XRPD of SD and PREC AZD5718

The absence of crystallinity is apparent in the XRPD of SD and PREC materials (Figure 5.3), where a broad, featureless halo is present in both cases. From a visual inspection, there is little noticeable difference between the two, and little information may be inferred based upon this data.

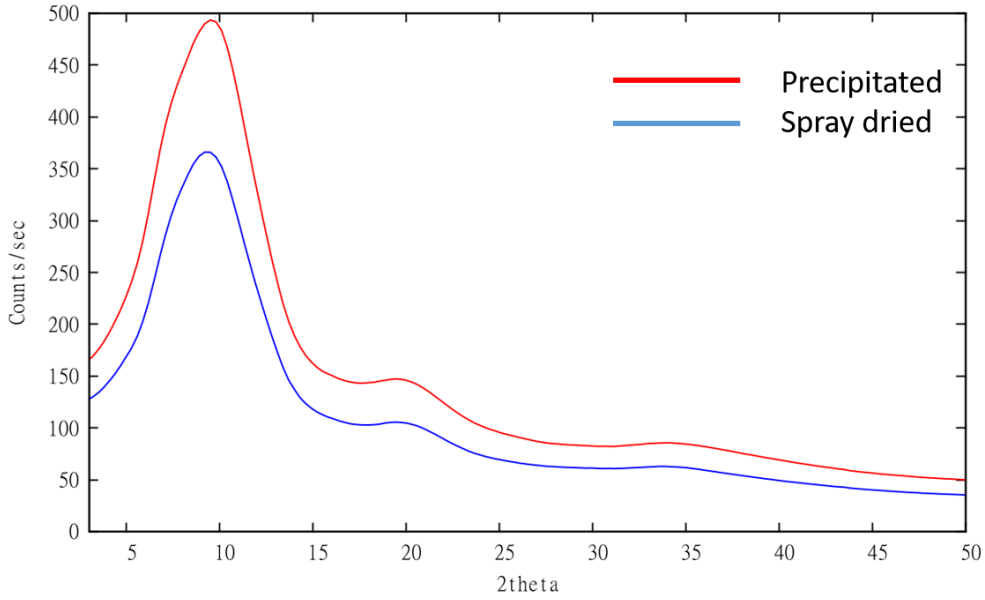

Figure 5.3: Stack of molybdenum XRPD data for SD and PREC AZD5718, truncated between 3 ° and 50 ° 2θ.

5.3.2.3. PDFs of AZD5718 reference materials

The data from Figure 5.3 were corrected, normalised and transformed to PDFs as described in Section 2.3. Figure 5.4 shows the experimentally derived PDF patterns, determined from the molybdenum XRPD data for the AZD5718 reference materials. The PDFs are truncated at 20 Å for clarity. For AZD5718 Form A, the peaks in $G(r)$ extend out to high-r values, well beyond the upper-r limit shown. This is expected and reflects the long-range 3D periodic order within the crystal lattice of AZD5718. Conversely, for PREC and SD AZD5718, the absence of long-range order is apparent and the features in the PDF attenuate to zero at around 10 Å. From a basic visual comparison, it appears that the local ordering in SD and PREC AZD5718 are very similar. However, closer analysis of the SD and PREC AZD5718 PDFs show subtle but significant differences on closer inspection.

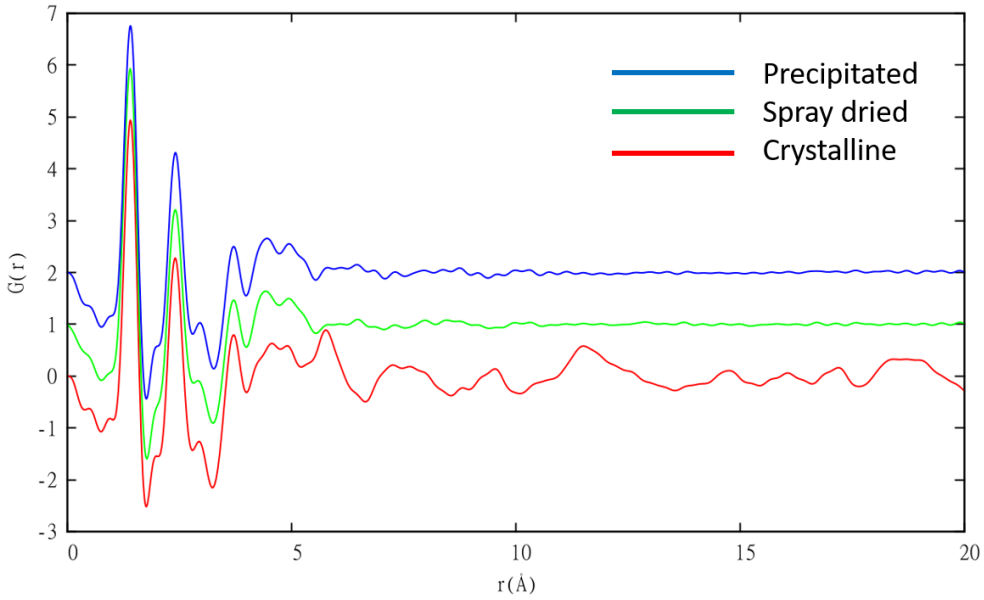

Figure 5.4: Stacked PDFs of AZD5718 reference materials.

In comparison to the crystalline material, local structure in SD and PREC AZD5718 are significantly different. The main PDF peak positions are, however, similar, out to ~ 6 Å (Figure 5.5), with the exception of the broad feature at ~ 4.4 Å in the amorphous material. However, upon an inspection of the features in the crystalline PDF, there are two peaks in the crystalline phase, at 4.25 Å and 4.55 Å, which sit either side of this feature observed in the amorphous material. It is possible that in amorphous AZD5718, these two peaks "merge" due to the significant disorder in the amorphous sample. However, the more obvious differences between the amorphous and crystalline are the amplitudes of the features in the PDFs. These differences may be explained by the long-rang ordering within the crystalline material, with the repeating distances, even at low-r, effectively increasing the amplitudes of features in the data, which are typically higher in the PDF of AZD5718 Form A than amorphous. Similar observations have been made in a PDF study of low and high density amorphous ices, where differences in peak amplitudes have been assigned to differences in coordination of water molecules (Mariedahl et al., 2018).

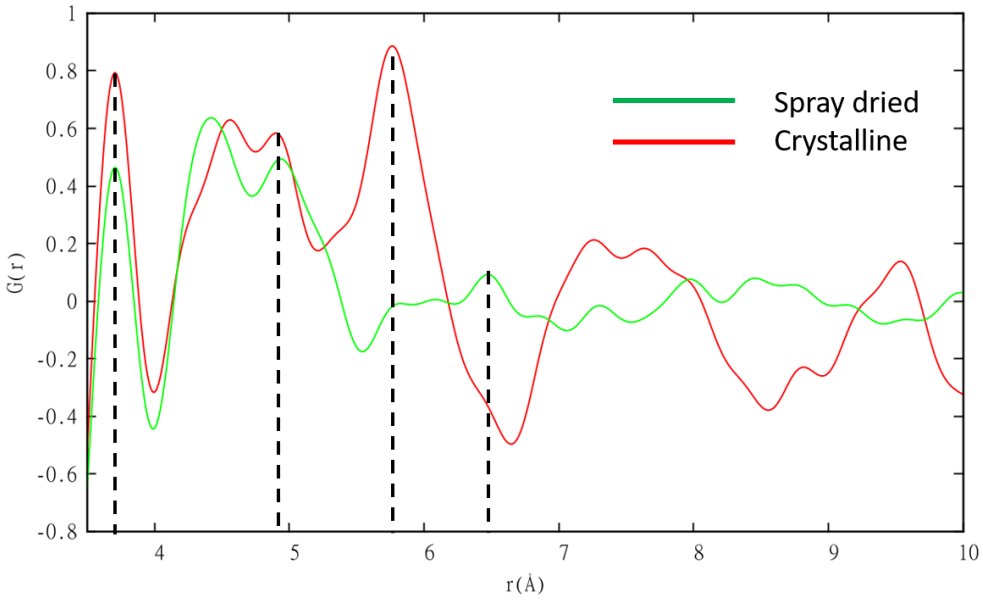

Figure 5.5: Comparison between PDFs of SD and AZD5718 Form A, truncated between 3.5 Å and 10 Å. Dashed lines highlight common peaks in the low-r region.

5.3.2.4. PDF fit of AZD5718 Form A

A fit of the PDF calculated from the AZD5718 crystal structure to the experimentally determined PDF of AZD5718 Form A was performed in which the structure is adjusted to achieve the best fit to the experimental data by minimising the difference between the calculated PDF and the experimental data (Section 2.8.; Figure 5.6). An overlay of the crystal structure and the structure obtained by refinement to the experimental PDF data is shown in Figure 5.7. The quality of the fit (R_{wp} = 9.604%) demonstrates that the VCT procedure used to collect the data is sufficient to achieve an acceptable signal-noise ratio across the angular range, which is particularly important at high-Q. Additionally, the close fit achieved to the experimental provides confidence that the corrections and normalisations performed were reliable. It is also evident from Figure 5.7 that the refined structure is sensible, given the good visual comparison in the overlay of the single crystal structure, and that obtained from the PDF refinement. The low MSD value of 0.0227 further supports this. Despite the good visual agreement between the observed and calculated profiles, there are clear differences

between the two, which are caused by several factors. The propagation of instrumental or sample-derived differences and uncertainties in the raw data through the Fourier transform to generate the PDF is not currently accounted for in available software packages such as TOPAS and PDFgui (Farrow et al., 2007) and may have an impact on the data. Whilst preferred orientation corrections, which have recently been developed for PDF data (Cervellino & Frison, 2020) have not yet been implemented in the TOPAS code for PDF, as opposed to routine XRPD analysis (Dollase, 1986). Finally, local deviations from the average crystallographic structure, which are observable in the PDF if present, cannot be fitted easily using the small-box method (Hou, Zhao, Paterson, Li, & Jones, 2018) applied in TOPAS, which consists of refinement to a single unit cell, and would require a big-box style refinement to understand and fit (Neilson & McQueen, 2015), allowing refinement to a supercell of a known phase, examples of which include RMCprofile (Tucker et al., 2007) and Fullrmc (Aoun, 2016).

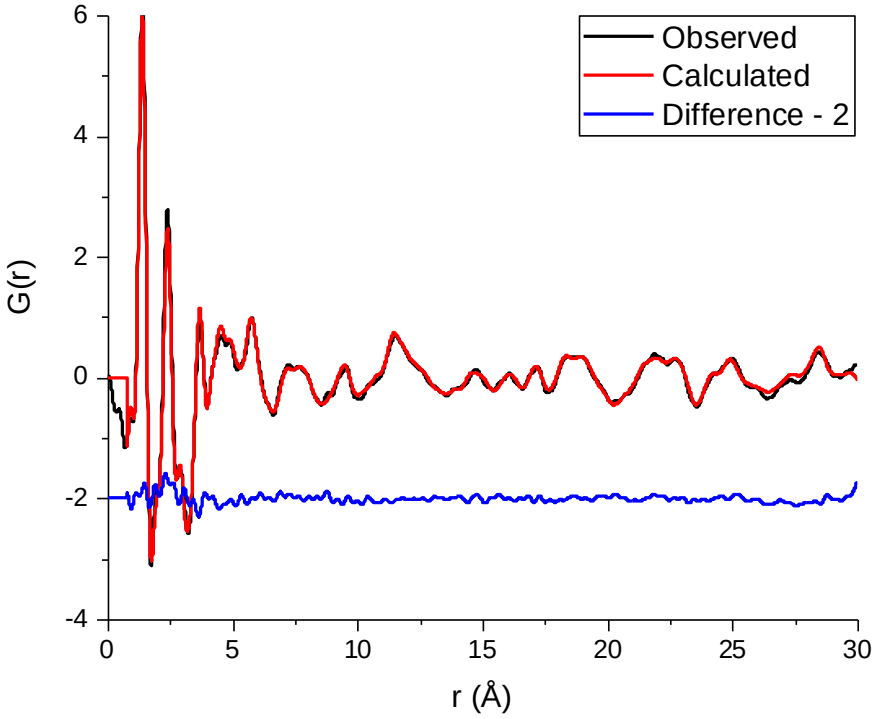

Figure 5.6: PDF fit of AZD5718 Form A, with red showing observed profile (y_{obs}), black the calculated profile (y_{calc}), and blue the difference plot ($y_{obs} - y_{calc}$).

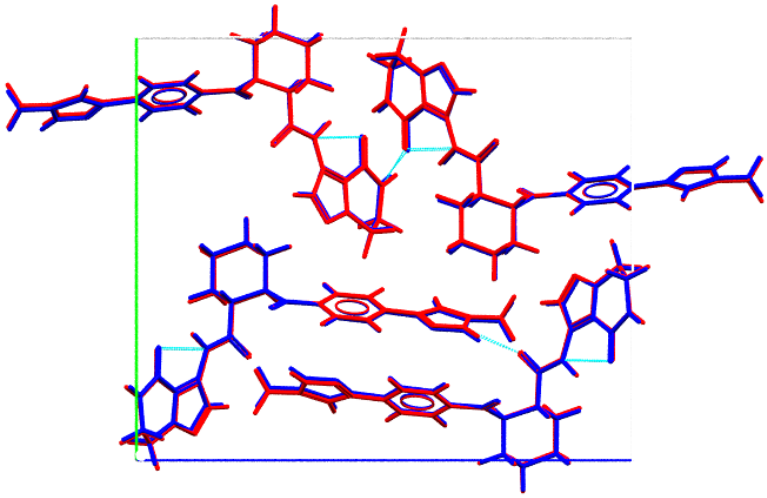

Figure 5.7: Structure overlay of single crystal (blue) AZD5718 crystal structure, and calculated structure obtained from fit to experimental PDF data for AZD5718 Form A (red). RMS = 0.0227. Structure viewed down crystallographic *a* axis.

5.3.3. Comparison between SD and PREC AZD5718

5.3.3.1. Reproducibility of PDF measurements for SD and PREC AZD5718

In order to ensure that any differences observed between the PDFs of SD and PREC AZD5718 are due to structural differences between the materials, as opposed to sample or data analysis variations or artefacts, three samples of each material were measured for PDF analysis. The overlays are shown in Figure 5.8. The multiple scans of each material show a high degree of similarity across $G(r)$. However, some subtle differences are apparent between different samples from the same preparation method. These differences may be explained by several factors. Firstly, although it is possible to approximate the background contribution to the data by subtraction of a reference pattern, in practice this is almost impossible to achieve with 100% accuracy. This is largely due to differences in capillary packing efficiency and variations in the incident beam intensity. Additionally, it is possible that the sample undergoes come constant change/ relaxation (Newman, Hastedt, & Yazdanian, 2017; Vranić, 2004) as a result of mobility in the amorphous phase, as well as similar changes within the material as a result of moisture. Although samples were loaded into capillaries in a glove bag, there is still potential for some exposure to ambient conditions whilst sealing the capillary, and subsequently whilst transferring from the glove bag to the instrument.

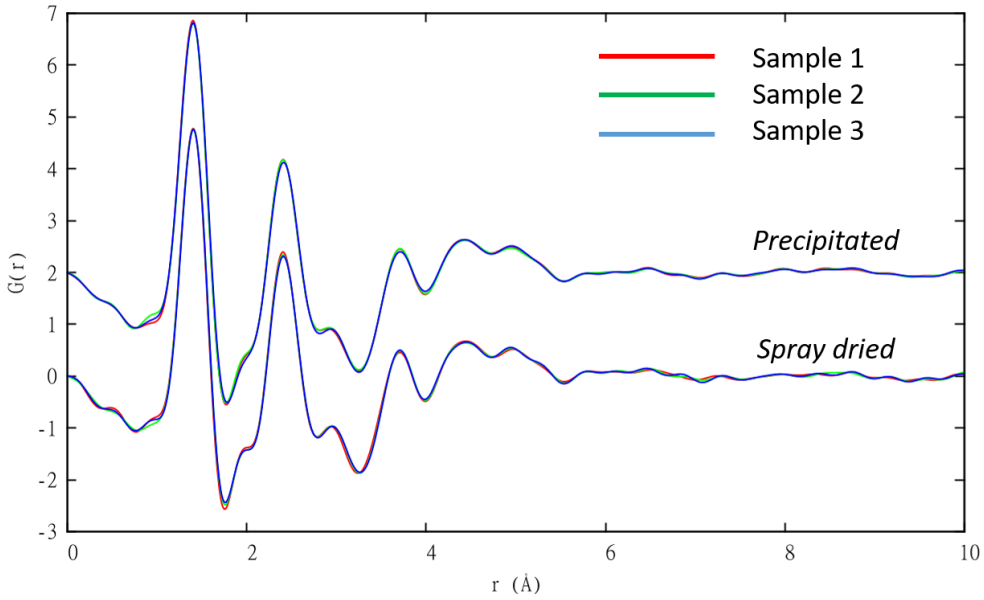

Figure 5.8: Stacked overlays of PDFs of SD and PREC AZD5718.

Table **5.**3 shows the Pearson correlation coefficients (Szczepańska, 2011) (CCs) for all PDFs when compared to each other, as calculated using polySNAP (Barr et al., 2004b). The use of CCs gives a quantitative measure of the similarity between observed data sets, and as such is an ideal tool to identify subtle differences in the PDF traces shown. For all AZD5718 samples prepared by the same method, the CCs are all higher than 0.99, whereas when comparing SD and PREC, they are all below or equal to this value. Therefore, the differences observed in the PDFs between SD and PREC appear significant, and worth further investigation. At this stage, it appears likely that the differences are due to the increased water content present within the PREC material, however, the relation between these differences and the structural features within each material will be investigated further.

Table 5.3: Summary of correlation coefficients between PDFs of AZD5718 samples. Correlations higher than 0.99 are shown in bold

	PREC 1	PREC 2	PREC 3	SD 1	SD 2	SD 3
PREC 1	**1**	**0.996**	**0.996**	0.987	0.98	0.99
PREC 2		**1**	**0.993**	0.985	0.98	0.988
PREC 3			**1**	0.983	0.971	0.986
SD 1				**1**	**0.991**	**0.996**
SD 2					**1**	**0.991**
SD 3						**1**

Figure 5.9 shows the overlays of the PDFs of SD and PREC AZD5718 samples. Although there is good agreement between the data, in terms of the peaks present in the PDFs, there are minor differences in the amplitudes of these features, where SD has generally higher intensities than PREC from around 5 Å onwards. Although issues in background subtraction have been highlighted above, the correlation coefficients are highest for samples prepared by the same method. The correlation coefficient between the two sample types show subtle but consistently larger differences across the PDF range used in the comparisons which can be related to structural differences between the SD and PREC samples.

Potential explanations for the basis of these structural differences in SD and PREC AZD5718 may be that the increase of water molecules present in the PREC sample increases the local density of the material, which in turn decreases the amplitude of features within the PDF in PREC. The baseline of which is dependent on the number density, or number of atoms per sample volume (Figure 5.20). Additionally, the higher water content may decrease the number of AZD5718- AZD5718 contacts in the material, which will be replaced by

AZD5718-water interactions. As the scattering power of water is significantly lower than that of AZD5718, due to the lack of higher atomic number scatterers in the molecule, this could decrease the amplitude of the features observed in the PREC material. Thus, from a visual inspection of the PDF data, it may be inferred that the increase in water content in PREC results in the disruption of AZD5718- AZD5718 interactions in the amorphous solid. However, it appears that the local structure observed within SD is largely preserved in PREC, with the addition of water molecules potentially filling voids in the structure. The causes of these differences are explored further by comparison with MD from section 3.2.3. onwards.

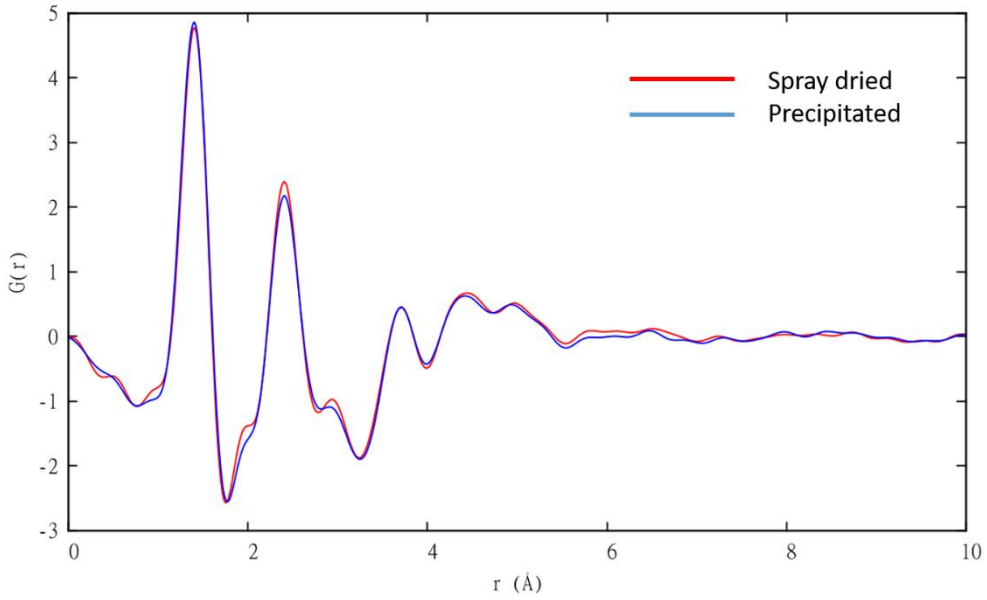

Figure 5.9: Overlay of PDFs of SD and PREC AZD5718.

5.3.4. Comparison of experimental PDFs to MD derived structural models

Table 5.4 shows the CCs for the experimental data when compared to the mean PDFs calculated from MD simulations. In line with experimental observations, SD AZD5718 has the highest CC with the mean PDF of the 2 wt% water MD simulation, whereas, PREC appears to be more similar to those containing 4 wt% water. Although this is expected, due to the higher

water content in PREC than SD, it demonstrates the reliability of the MD simulations, in that at the very least, the subtle impact of water on the structure of amorphous AZD5718 is observed. Overlays of the best fit MD PDFs with experimental data are shown in Figure 5.10. It is clear from comparison of the data in Figure 5.10 that there is a striking similarity between the observed and calculated PDFs for both sample preparation methods. Whilst differences are present between the experimental and MD PDFs, it is important to note that there has been no structural refinement of the models to the experimental data, and as such, the observed differences are not unexpected.

Table 5.4: Summary of correlation coefficients for experimental PDFs compared to simulated PDFs from MD simulations. Highest CCs are shown in bold.

	PREC	SD
0%	0.762	0.780
0.5%	0.814	0.851
1%	0.821	0.890
2%	0.807	**0.904**
4%	**0.868**	0.835

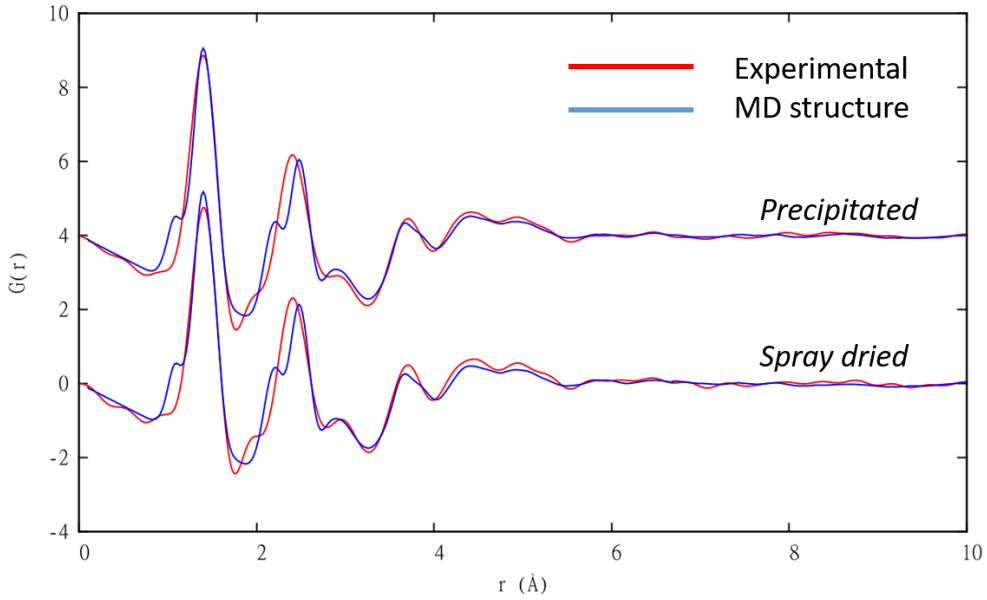

Figure 5.10: Overlay of PDFs of SD and PREC AZD5718 with PDFs calculated from MD simulations of 2 and 4 wt% AZD5718, respectively.

5.3.5. Analysis of MD simulation boxes

5.3.5.1. Water clusters in amorphous AZD5718

Figure 5.11 shows the maximum observed water cluster sizes for each simulation. The clustering of water molecules in amorphous solids is of considerable importance, due to their plasticising effects (Authelin, Mackenzie, Rasmussen, & Shalaev, 2014; Shalaev et al., 2019), and in this case, in understanding the structural differences between SD and PREC. The maximum observed cluster size is dependent on the concentration of water in the material, which has been reported elsewhere (Authelin et al., 2014), where at 0.5 wt% water,

the largest cluster contains 3 water molecules, whereas at 4 wt%, the largest contains 14 water molecules.

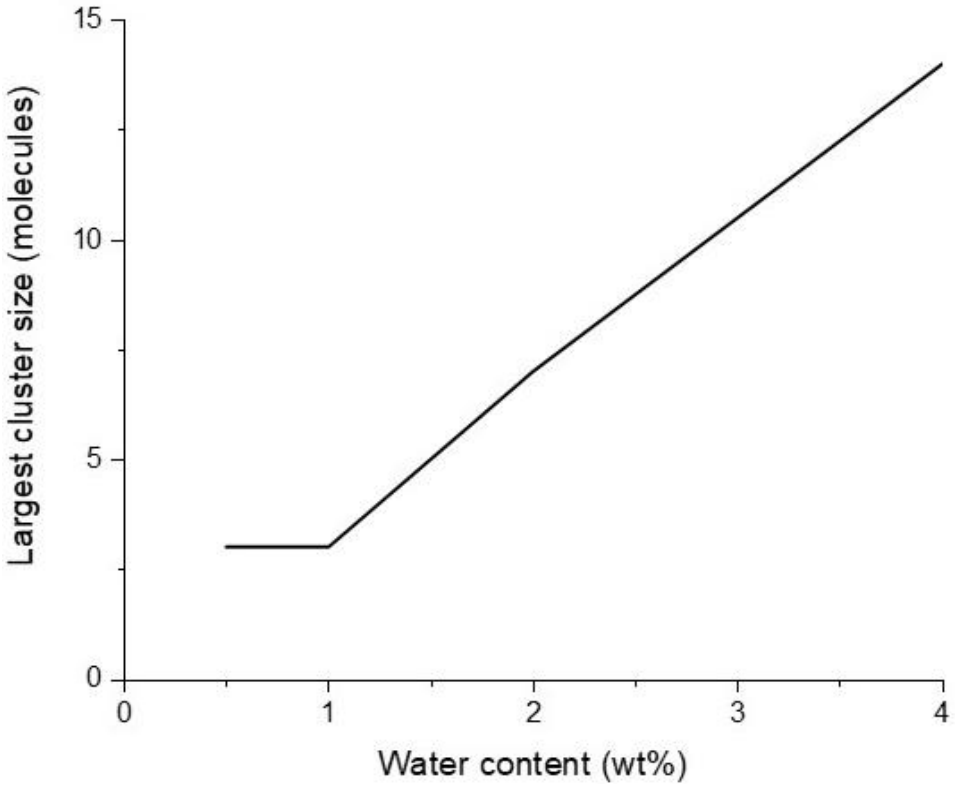

Figure 5.11: Largest water cluster observed for each AZD5718 simulation.

The probabilities of a given water cluster size being observed are shown in Figure 5.12. Note that no description of how probabilities are determined is given in the GROMACS documentation, and as such, these are treated as arbitrary values. To normalise the probabilities for comparison across water contents, they are shown here as a percentage of the sum of all probabilities for each water content. As expected, there is a lower probability of water clustering occurring at low water contents, particularly at 0.5% water, where there are approximately 8 AZD5718 molecules for every water molecule. The probability, therefore, of two water molecules being near each other is low. Subsequently, as water content increases, the probability of clustering increases. It is only at 4 wt% water, that

clusters larger than 6 water molecules are observed. While this analysis shows that water clustering can occur at all water contents, the behaviour of water at larger concentrations may help explain the apparent differences between the PDF data of SD and PREC (Section 3.2.). This is consistent with Figure 5.15, where the PDFs of SD and PREC are overlaid with that of water.

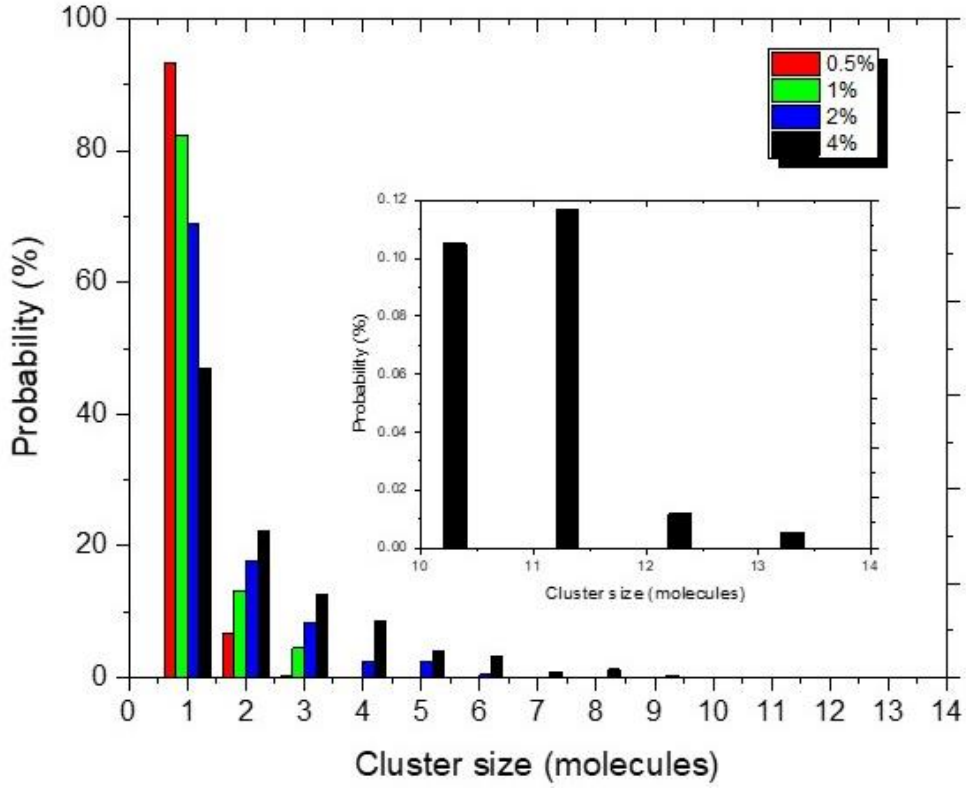

Figure 5.12: Probabilities of water cluster sizes observed in MD simulations. Inset shows probabilities for cluster sizes larger than 10 molecules.

Figure 5.13 shows the first frame of the last 100ns of 2 and 4 wt% water simulations, with all AZD5718 molecules removed, showing the distribution of water in the amorphous solids. As would be expected, there is a significant increase in the number of water molecules at 4 wt%, and as such, a visual inspection reveals that there is a subsequent increase in the number of water-water interactions. Despite the low weight percentage of water in the simulations, it is clear that there is potential for significant hydration of AZD5718 molecules

at both shown water concentrations. In fact, at 4% water, there are more water molecules than AZD5718, with an obvious potential to disturb AZD5718- AZD5718 interactions.

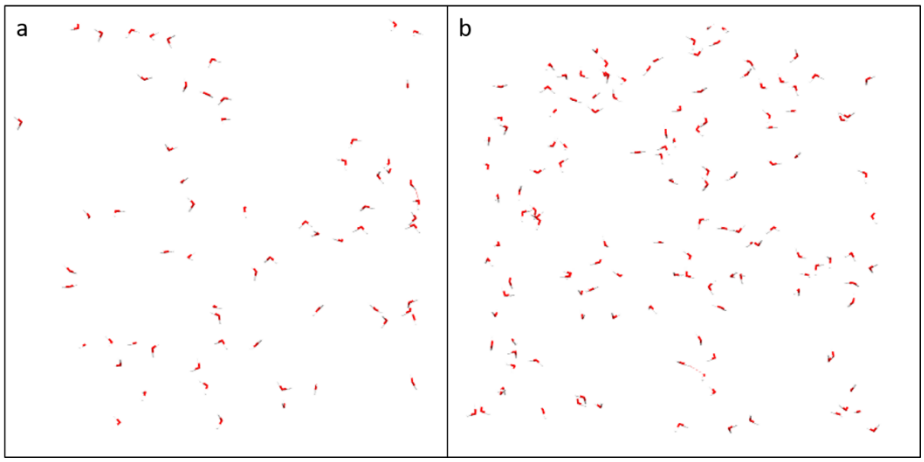

Figure 5.13: Snapshots of MD simulations, showing only water molecules, with a) 2% water and b) 4% water.

Figure 5.14 shows a few typical water clusters observed in the simulations. From these images alone, it is clear that water does not only interact with neighbouring AZD5718 molecules, but also with itself. The presence of phase separated water is observable in the PDF data when compared to a reference PDF of water, where there are changes in the data for AZD5718 that correspond with water-water contacts (Figure 5.15). The clustering of water molecules, demonstrated in Figure 5.11 and Figure 5.12, can explain these changes, in that the larger occurrence of phase separated water at higher water concentrations gives rise to bulk water-like features in the experimental PDF data for AZD5718.

Figure 5.14: Example water clusters found in MD simulations.

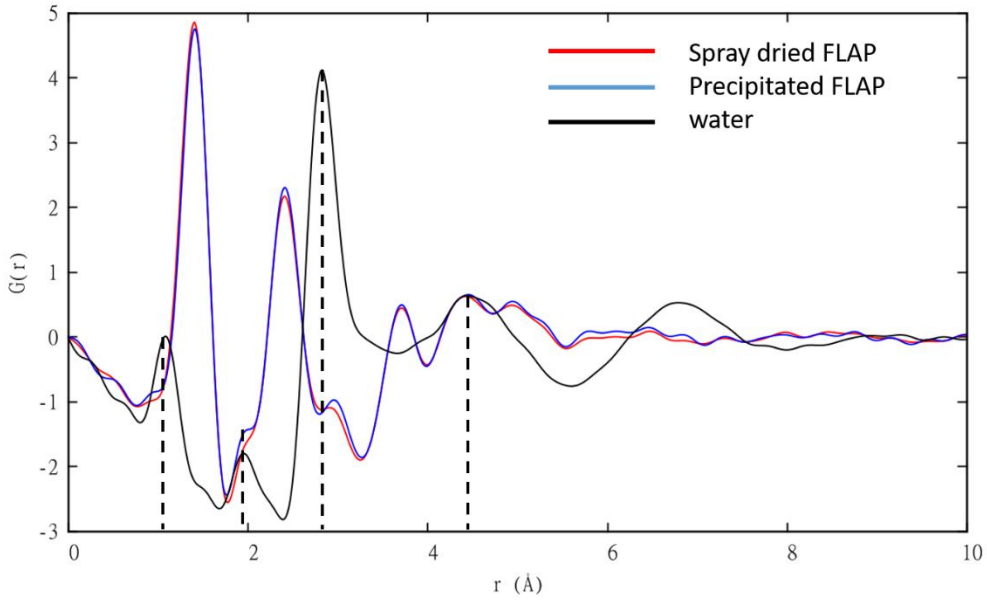

Figure 5.15: Overlay of the PDFs of SD and PREC AZD5718 with water. Dashed lines highlight differences between SD and PREC which align with features in the PDF of water.

5.3.5.2. Hydrogen bonding in amorphous AZD5718

Figure 5.16 shows the mean number of bonds per frame for all trajectories, displayed as a percentage of the total number of bonds. As shown, the percentage of AZD5718-

AZD5718 hydrogen bonds reduces as a function of water content, whilst all interactions involving water drastically increase. By 4 wt%, AZD5718 molecules interact more with water than they do with neighbouring AZD5718 molecules. While the bulk structure of amorphous AZD5718 is similar, regardless of SD or PREC preparation method, as evidenced by the PDF data (Section 3.2.), it is evident that there is some disruption of AZD5718- AZD5718 interactions at higher water contents.

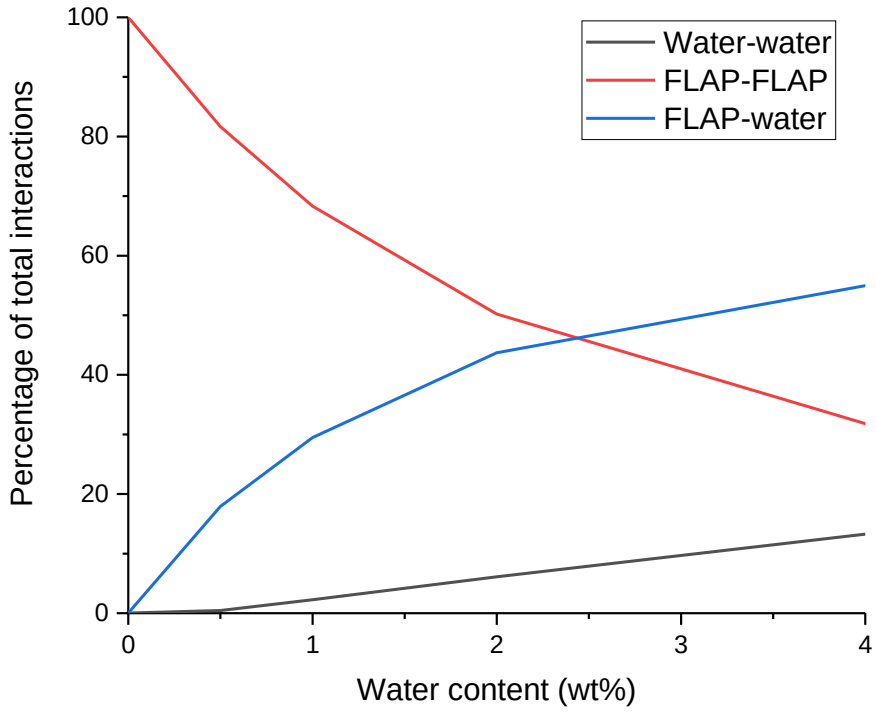

Figure 5.16: Mean number of bonds per frame for each trajectory, calculated for AZD5718-AZD5718, AZD5718-water and water-water contacts as a percentage of the total number of interactions in the system.

5.3.5.3. Assessment of aromatic interactions

A summary of the analysis of the aromatic interactions is shown in Table 5.5. For AZD5718 Form A, the aromatic interactions are the same for every molecule, due to the

periodic nature of the material, and as such, only a single AA calculation was used. For the crystalline material, there are no “strong” interactions, with the majority of them being “moderate”, and only a single “weak” interaction, with a mean score of 5 (moderate strength). As this is the crystalline form, a higher score than the amorphous material is expected, which is the case here. Given that this is the only known pure polymorph of AZD5718 (there is a solvate crystal structure), a higher score may be expected than that shown here, however, it is possible that in the case of AZD5718, the hydrogen bonds play a larger role in stabilising the structure, as opposed to any other attractive force. This would appear to also be the case for the amorphous phase. Despite there being a small number of “strong” interactions, the majority of them are “weak”. However, it is possible that hydrogen bonding is responsible for the stability of the amorphous material.

Table 5.5: Summary of aromatic interactions for 2% and 4% water MD simulations, and AZD5718 Form A.

	2 wt% water	4 wt% water	Form A
Strong interactions	2	3	0
Moderate interactions	10	12	4
Weak interactions	46	46	1
Mean interactions per molecule	4.83	4.62	5
Mean score	1.66	1.95	5
Mean rank	Weak	Weak	Moderate

5.3.5.4. Assessment of hydrogen bonding

Table 5.6 shows a summary of the hydrogen bonding analysis. AZD5718 Form A has the most favourable bonding arrangement out of the three, where all interactions are coordinated according to the most suitable identified by the Mercury software. The high mean probability (0.7249) of hydrogen bond coordination observed for AZD5718 Form A confirms the feasibility of the bonding arrangements. For amorphous AZD5718, a high percentage of all interactions are identified as being unsuitable (> 95% are unsuitable). Given that this is a non-crystalline material, this is not unexpected. However, this analysis is based purely upon coordination numbers of the hydrogen bonds and gives no consideration to the molecular geometries or locations of the hydrogen bonded molecules. Despite this, Table 5.6 demonstrates the structural similarity between SD and PREC, in that both exhibit structural behaviour which contrasts with the crystalline phase.

Table 5.6: Summary of hydrogen bonding coordination analysis for 2% and 4% water MD simulations, and AZD5718 Form A.

	2 wt% water	4 wt% water	Form A
Non-ideal coordinated (% of all interactions)	95.06	95.19	0
Ideal coordinated (% of all interactions)	4.94	4.81	100
Mean probability (all interactions)	0.05	0.05	0.725
Mean probability (all non-ideal coordinated)	0.001	2.49E-05	N/A
Mean probability all (ideal coordinated)	0.98	0.99	0.725

Mean probability (AZD5718 non-ideal coordinated)	0.001	3.18E-05	N/A
Mean probability (AZD5718 ideal coordinated)	0.98	0.99	0.725
Mean probability (water)	3.88E-05	2.14E-09	N/A

Figure 5.17 shows the distribution of hydrogen bond motifs observed in the crystalline and amorphous structures, based upon graph-set analysis in Mercury. Due to differences in the number of independent units in the cells for crystalline and amorphous AZD5718, these are shown as a percentage of the total number of hydrogen bonds. There are significant differences in hydrogen bonding between crystalline and amorphous forms. For both amorphous structures, the motifs present are very similar, reflecting the structural similarity between the two. For amorphous AZD5718, molecules largely form dimers, or other finite arrangements, whereas for the crystalline material, chains and rings are the dominant structural features. Whilst a very small amount of ring motifs exists in the amorphous structure, there are no chains of AZD5718 molecules, which form the majority of hydrogen bonds in the crystalline material. For all of the structures, intramolecular hydrogen bonds are present, which gives an indication of the presence of the L-shape conformer which exists in the crystal structure. In the case of the amorphous structures, there may be other conformers present which also contain this intramolecular interaction, however, this does suggest that some similar conformations contribute to the distribution in the amorphous form, the majority of which appear to be distinct from that observed in the crystalline material. For example, in the case of 2 wt%, a total of 27 intramolecular hydrogen bonds are present in the frame used for this analysis, out of a possible 128 AZD5718 molecules.

Similarly, 25 are observed at 4%. There does not appear to be any significant difference in the conformer distribution between 2% and 4% water, as these are likely to change during the course of the dynamics runs, therefore, no definitive conclusions regarding the conformer distributions can be determined by assessing a single frame. A complete assessment of the conformations in the amorphous material was not performed during the analysis.

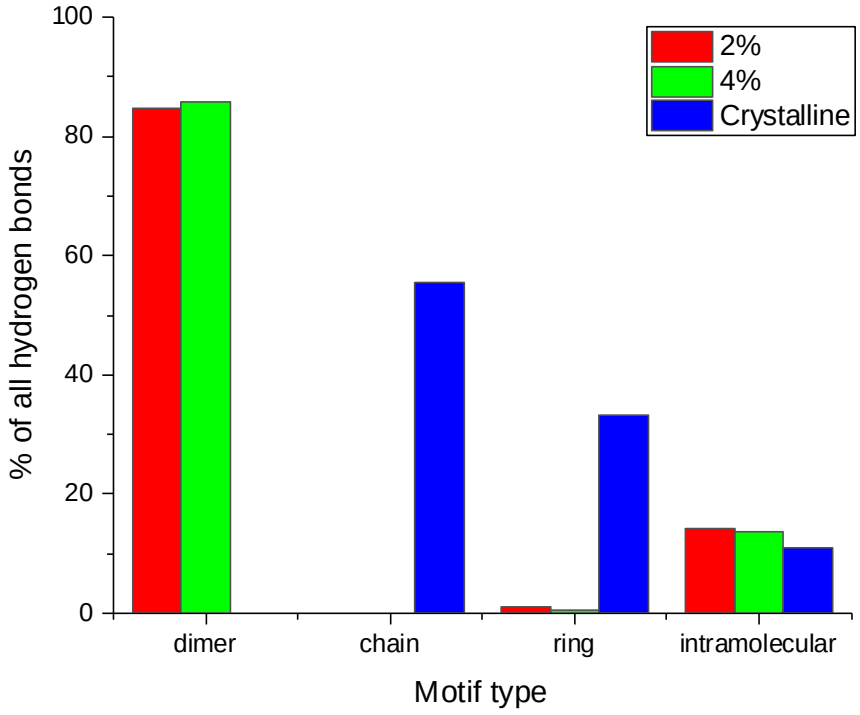

Figure 5.17: Structural motifs present in 2% and 4% water MD simulations, and AZD5718 Form A.

5.3.5.5. Conformation in amorphous AZD5718

A distribution of conformations is observed in the amorphous AZD5718 structures, as shown in the overlay in Figure 5.18. AZD5718 can feasibly adopt a range of internal geometries, owing to the number of flexible rotatable bonds, from “V-shaped” molecules, to fully extended conformers. It has already been established from the previous analyses (Section 3.2.3.2.; Figure 5.17) that the low energy L-shaped conformation observed in the crystalline form is not a dominant conformer in the amorphous phase. The difference in conformation between amorphous and AZD5718 Form A potentially enhances the resistance of the amorphous solid to crystallisation, as conformational rearrangement would require significant energy input to occur.

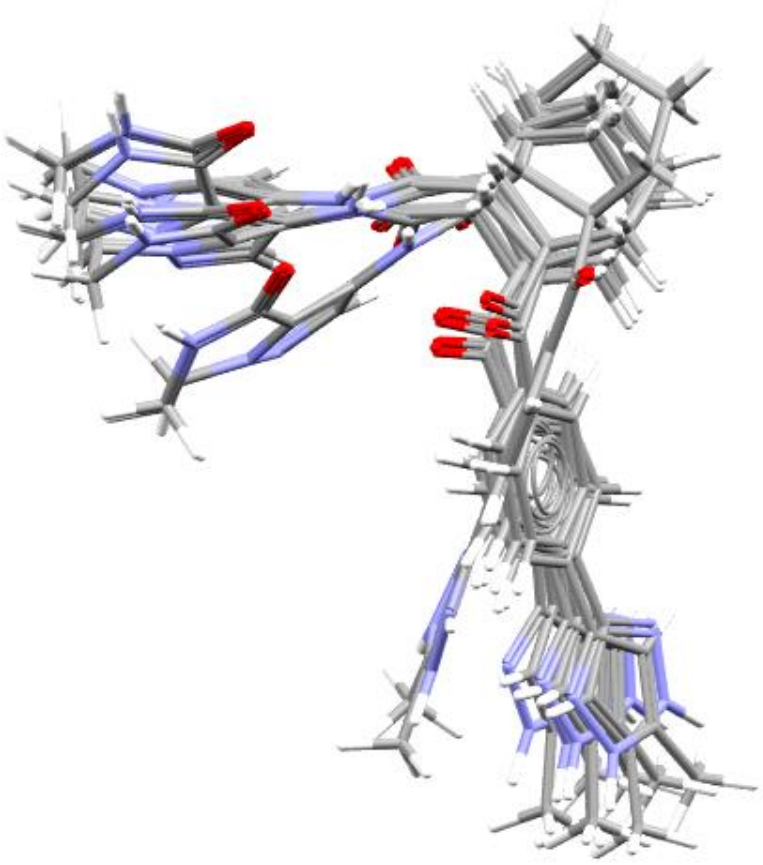

Figure 5.18: Overlay of 100 example AZD5718 conformations adopted in the amorphous phase, extracted from multiple frames of MD trajectories for 2 and 4 wt% water. The range of potential AZD5718 conformations is evident, due to the conformational flexibility of the molecule as a result of the large number of flexible rotatable bonds.

Due to the lack of long-range ordering in the amorphous material, it is likely that the PDF signal is dominated by molecular and short-range intermolecular interactions. Figure 5.19 shows an overlay of SD AZD5718 with the mean PDF calculated from a range of single AZD5718 molecules with a distribution of conformations. A total of 100 single molecules were used to obtain this distribution. Several peak positions in the experimental data are reproduced in the mean single molecule PDF, however, the amplitudes are significantly different, where in the simulated data, the amplitudes are higher. This is because the single molecule calculations are based upon a single AZD5718 molecule in the large MD box. This

drastically reduces the number density of the material, which drops to near zero, as the majority of the cell has no contents. As the baseline in the PDF is dependent on the number density, any atom-atom contacts will result in significant peak amplitudes, as the density of these contacts will be significantly larger than the average across the cell. The effect of box size, and effectively the average number density, on single molecule PDF calculations, is demonstrated in Figure 5.20, where a single molecule was placed inside three different cell sizes, with cell lengths of 15 Å, 25 Å and 50 Å, where it is observed that the larger box sizes have higher amplitudes, attributed to the lower number density of the cell. It is clear, however, that not all of the features in the PDF are reproduced, and it is likely that the features not observed in the calculated PDF are due to intermolecular contacts, related to the first coordination shell around a given AZD5718 molecule

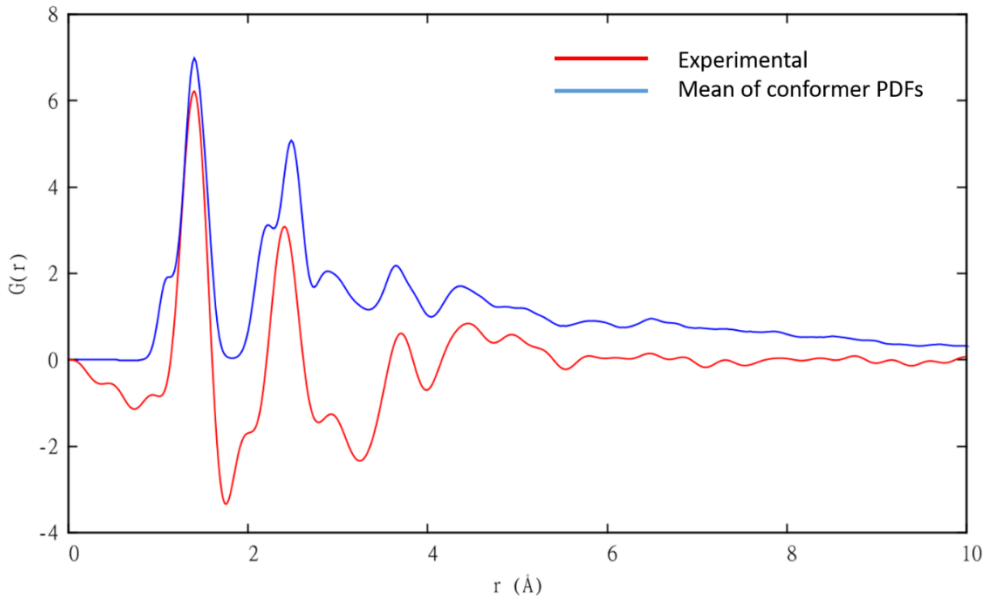

Figure 5.19: Plot of PDF for amorphous AZD5718 and the mean PDF calculated from 100 different AZD5718 conformers.

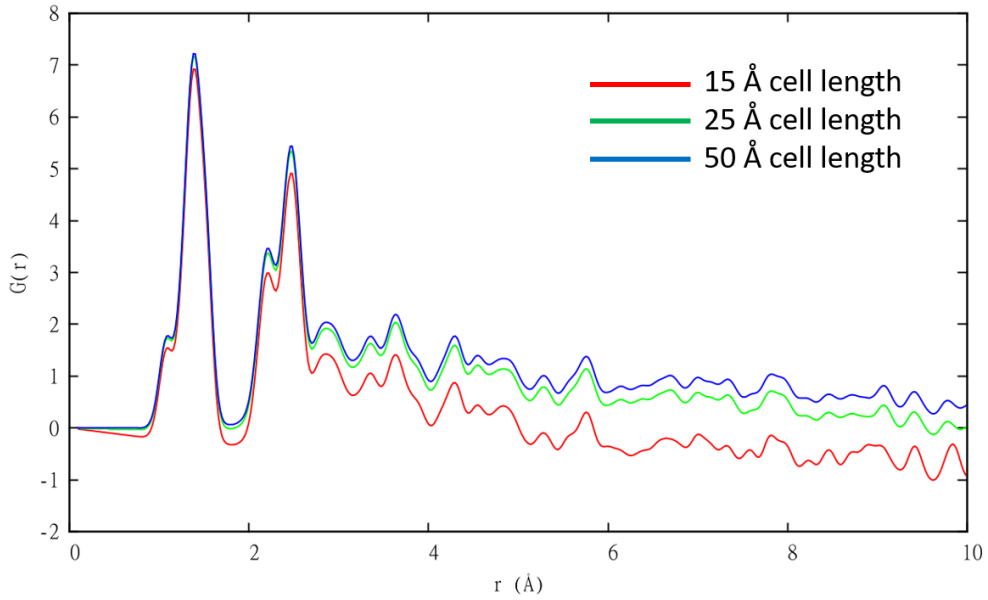

Figure 5.20: PDFs calculated for a single AZD5718 molecule, with varying cell dimensions, with 15 Å cell length (red), 25 Å cell length (green) and 50 Å cell length (blue) shown. PDF is truncated at 10 Å for clarity.

5.3.6. Mapping structural features in PDF of amorphous AZD5718

Some example interactions extracted from the MD simulations, which reproduce distances observed in experimental amorphous PDFs, are shown in Figure 5.21. The identified distances are all less than 6 Å, as these are the dominant features in the PDF, after which, the local structure is less well-defined, as the disordered nature of the material is reflected in the low peak amplitudes. The peak at ~ 3.7 Å in this case has not been investigated, as it is apparent from Figure 5.19 that the main contribution to this feature is from intramolecular interactions, as evidenced in the mean PDF of the single AZD5718 molecules.

As the PDF shows distributions of interatomic distances, centroids were calculated for a functional group or ring of interest, to obtain an average distance for all points, which is manifest as the peak maxima in the experimental data. Good agreement is obtained by measuring distances between functional groups involved in hydrogen bonded interactions, and ring-ring centroids accurately reflect longer distances in the material. The shortest distance compared to the experimental data, at 3 Å, seems to correspond to atoms involved

in hydrogen bonds in AZD5718, with the example showing an NH…N hydrogen bond. It is worth noting that there are several other close contacts which will contribute to these features, including intramolecular distances. However, for simplicity, and ease of identification, the focus here has been on the hydrogen bonding. Beyond this, the distance at ~ 4.4 Å, in this case, corresponds to a ring-ring centroid, with ring positions close to the functional group involved in the hydrogen bond. Given the common occurrence of rings in the chemical structure of AZD5718, these types of interactions are likely to contribute significantly to the data. At ~ 5 Å, this is again related to atoms involved in the hydrogen bonding, however, given the positions and number of atoms near the bond in the carboxamide group, the centroid-centroid distance is larger than that at 3 Å. The final distance shown at ~ 5.3 Å corresponds to a ring-ring centroid between two molecules, at positions further from the hydrogen bond, again showing the large contribution of the rings to the observed peaks. All distances compare favourably to the experimental data. Due to the large size of the AZD5718 molecule, only first-neighbour molecules appeared to contribute to these low-r distances. As such, the types of interactions and ring-ring contacts highlighted are largely representative of the features observed in the data. Of course, the peaks in the PDF represent distributions of interatomic contacts, and therefore, a great number of other distances will exist in the material.

It is apparent from the snapshots shown of interactions in amorphous AZD5718, that the structure differs significantly from the crystalline form, which is also evident in Figure 5.17, where the dominant features are chains and rings of AZD5718 molecules, whereas in the amorphous form, more finite arrangements are formed, such as the dimers observed below. The formation of dimers in the amorphous phase has been reported in simulations and spectroscopic studies of pharmaceuticals (Kamińska et al., 2020; Strachan, Rades, & Gordon, 2007; B. Wang & Pikal, 2010), as opposed to the more complex motifs often

observed in their crystalline counterparts. Additionally, notable structural diversity is present in the amorphous solid, in comparison to the crystalline form. The observation of a larger distribution of motifs is, of course, expected due to the disordered nature of the material. Hydrogen bonds in the crystal are also directed by NH…O=C contacts. From the few images shown below, there is appreciable deviation from this in the amorphous phase, where a significant number of NH…N interactions are formed.

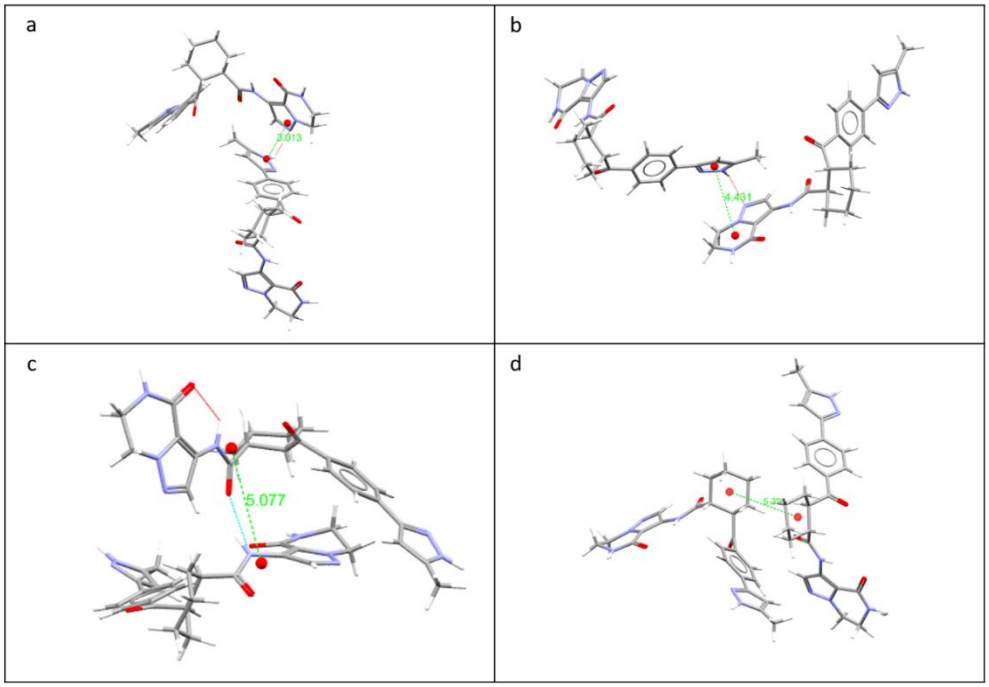

Figure 5.21: Distances identified in MD structures which are comparable to experimental PDF data, with a) identification of distance at ~ 3 Å, b) distance at ~ 4.4 Å, c) distance at ~ 5 Å and d) distance at ~ 5.3 Å. For all images, the hydrogen bond is highlighted by a blue line, and the distance of interest is shown by a green line.

A comparison of the experimental SD reference PDF to the mean PDF calculated for the interactions shown in Figure 5.21 is shown in Figure 5.22. There is good agreement between the peak positions and relative peak amplitudes between the experimental and calculated data. Aside from the differences described in Section 3.2.3.6., related to the

number density, the differences present here can be attributed to the small sample size used for the calculation, as it is likely that not all the “average” interactions in the MD simulations are captured in the four frames used. However, this demonstrates that the small number of interactions selected above are sufficient to describe the majority of the low-r peaks in the PDF.

While it is evident from Figure 5.19 that the intramolecular structure contributes significantly to the amorphous PDF, it is clear that the PDF also contains useful information related to intermolecular interactions, as a number of the features in the experimental data are not replicated using only the single molecule calculations. In particular, the feature at ~ 5 Å, which is not accounted for in the mean PDF for the calculated conformers, is clearly reproduced in the calculated PDF shown in Figure 5.22. Beyond ~ 5 Å, however, there is poor comparison between the observed features, which suggests that the sample interactions shown are not representative of all interactions in the amorphous material. Given the structural diversity observed in the MD simulations and the weak peak intensities, which reflect the loss of structural coherence at these distances, this is expected.

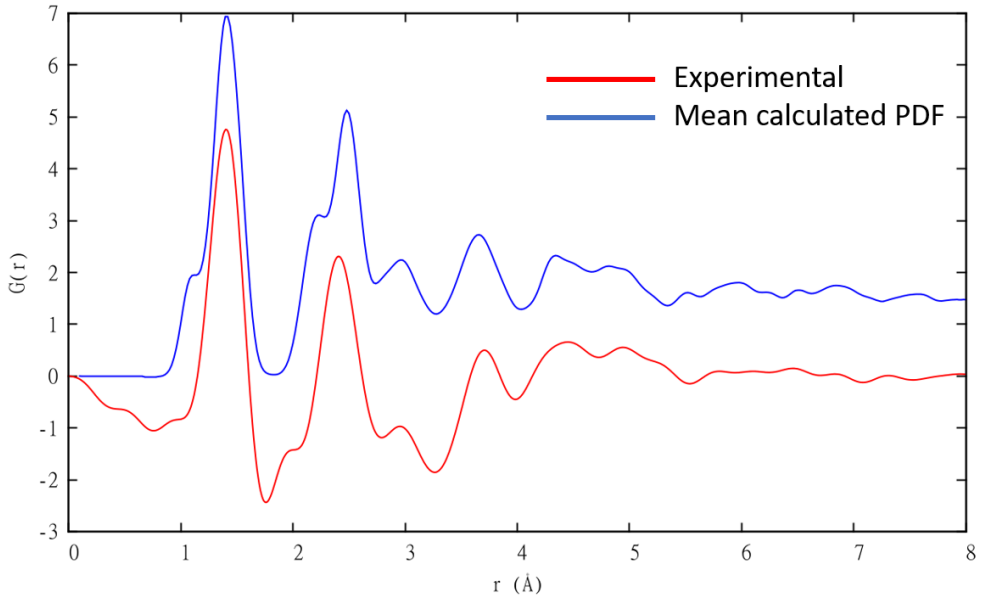

Figure 5.22: Comparison of experimental PDF of SD AZD5718 to the mean PDF calculated for selected interactions.

5.4. Summary

Subtle changes in the local structure of amorphous AZD5718, which are directly attributed to the presence of water as a result of preparation method, were detected using PDF analysis. This demonstrates the sensitivity of the technique in probing the local structure of amorphous materials, where a subtle change in the water content had a significant impact on the data, as evidenced by the change in CCs, where for samples prepared by the same method, CCs were > 0.99, which dropped to ≤ 0.99 when comparing samples prepared by different methods. As such, the technique could be used to unambiguously identify SD and PREC AZD5718, the latter of which has a higher water content.

Comparison of the reference experimental PDFs of SD and PREC AZD5718 to those calculated from MD simulations showed excellent agreement, where PREC showed higher similarity to the MD simulations with higher water content. The opposite was true for SD, in line with experimental observations. An investigation into the structure in MD simulations of

amorphous AZD5718 revealed potential disruption of AZD5718- AZD5718 hydrogen bonds by water at higher water content. The disruption of molecule-molecule hydrogen bonds by a plasticiser has been suggested in an MD study of sucrose (Gupta, Nunes, & Jonnalagadda, 2013), wherein it was suggested that this may contribute to increased mobility at elevated humidity. Similarly, this has been reported in an experimental study of thermoplastic starch (Ma & Yu, 2004), where the use of formamide and glycerol as plasticisers seemed to disrupt interactions between starch chains, evidenced by FT-IR. Water clustering was identified in the MD simulations for AZD5718 (Section 3.2.3.1; Figure 5.12), particularly at high water contents, which contributes to the differences between the PDF data for SD and PREC (Section 3.2.3.1.).

As expected, the intramolecular structure of AZD5718 contributes significantly to the strongest PDF features. This has been confirmed using single molecule PDF calculations (Section 3.2.3.6.; Figure 5.19), where the feature at ~ 3.7 Å was shown to align with several intramolecular contacts, specifically third and fourth nearest neighbour atoms in the molecular structure. However, a calculated mean PDF obtained from pairs of interacting AZD5718 molecules (Section 3.2.4.; Figure 5.22) reproduced the key features in the experimental PDF highlighting contributions from intermolecular interactions out to ~ 6 Å. Key distances observed in the reference experimental PDF for SD AZD5718 were mapped to selected structural features from the MD simulations at 2% water (Section 3.2.4.; Figure 5.21), showing good agreement. The identified features were related to hydrogen bonded interactions from AZD5718- AZD5718 contacts.

The need for careful sample handling and processing, particularly for non-crystalline materials, is an important aspect of the experimental investigation of amorphous materials by PDF. The long analysis times (~ 22 hours for a PDF measurement), and sample preparation do mean samples can undergo relaxation or dehydration which could impact on the results,

due to the sensitivity of PDF analysis in this context. However, aside from the sensitivity of the PDF, the technique allowed the differences resulting from the two manufacturing methods to be confirmed, and through correlation with the MD simulations, assign this difference to changes in the water content of the two samples. Given the structural changes that were identified as a direct result of water interaction with AZD5718, the need for careful sample handling is reinforced, and it is crucial to control environmental variables, such as temperature and humidity, during the development stage of any drug product.

Chapter 6: Structure and dynamics in amorphous paracetamol

6.1. Introduction

The painkiller paracetamol (PCM) is a model pharmaceutical system whose crystalline solid-state behaviour has been widely investigated and characterised (Di Martino, Conflant, Drache, Huvenne, & Guyot-Hermann, 1997; Nichols & Frampton, 1998b; Perlovich, Volkova, & Bauer-Brandl, 2007; Reiss, Van Mechelen, Goubitz, & Peschar, 2018a; Shtukenberg et al., 2019; Zimmermann & Baranović, 2011). PCM exhibits extensive polymorphism, with five single-phase forms being structurally characterised to date (Nichols & Frampton, 1998b; Reiss et al., 2018a; Shtukenberg et al., 2019) and several additional polymorphs reported although full crystal structures have not yet been determined for them all (Shtukenberg et al., 2019; Smith, Bishop, Montgomery, Hamilton, & Vohra, 2014). The known crystal structures of PCM and lattice parameters are shown in Figure 6.1 and Table 6.1, respectively.

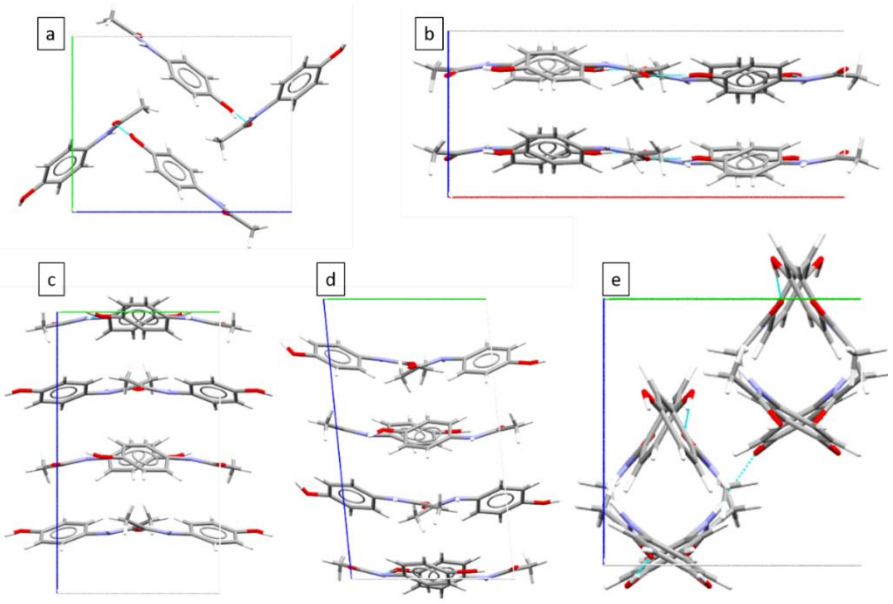

Figure 6.1: Known structures of PCM, showing (a) Form I, (b) Form II, (c) Form III, (d) Form III-m, and (e) Form VII.

Table 6.1: Lattice parameters for Forms I (Naumov, Vasilchenko, & Howard, 1998), II (Nichols & Frampton, 1998a), III and III-m (Reiss, Van Mechelen, Goubitz, & Peschar, 2018b) and VII (Shtukenberg et al., 2019).

Form	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)	Z’
I	7.094	9.263	11.657	90	97.672	90	1
II	17.166	11.777	7.212	90	90	90	1
III	11.838	8.569	14.818	90	90	90	2
III-m	11.755	8.572	14.516	90	90	90	4
VII	16.844	9.482	9.142	90	90	90	2

While there is significant structural diversity in solid forms of PCM, hydrogen bonds form exclusively between the OH in the phenyl group and the C=O in the N-acetyl group; and the NH in the N-acetyl group and the phenyl OH. No deviation from these motifs have been reported for the crystalline phases. PCM molecules are hydrogen bonded to four other PCM molecules, in all reported structures except Form VII, in which one of the molecules in the asymmetric unit forms only two hydrogen bonds (OH…O=C) in a chain-like motif. The structural similarity between Forms II, III and III-m is apparent in Figure 6.2, wherein sheets of hydrogen-bonded PCM molecules form chains and rings, which differ from the herringbone-type structures shown for Forms I and VII. The differences in packing between Forms II, III and III-m are due to subtle changes in the torsion angles in the N-acetyl and hydroxyl groups, which change the direction of the hydrogen bonds. The structural similarity between these forms is also reflected in their lattice parameters, where for III and III-m these are almost identical, and when comparing III and III-m to Form II, the parameters are closely related, with $a \approx b$, $b \approx \frac{1}{2}a$ and $c \approx 2c$.

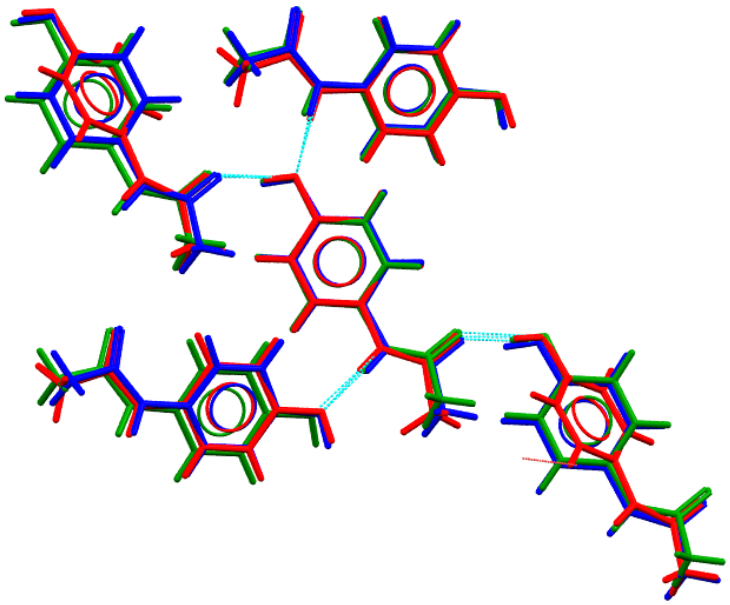

Figure 6.2: Structure overlay of Forms II (red), III (blue) and III-m (green) highlighting the common structural motif in these forms.

Structural characterisation of Form III eluded researchers for a number of years due to the poor physical stability of the polymorph and difficulty in isolating the form as a stable sample to allow measurement although a number of studies have reported the detection of the phase using thermal and spectroscopic analyses (Gaisford, Buanz, & Jethwa, 2010; Zimmermann & Baranović, 2011). The first report of the crystal structure of Form III was in 2009 (Perrin, Neumann, Elmaleh, & Zaske, 2009), where the metastable phase was isolated following crystallisation from the melt, and a Rietveld refinement performed to a predicted crystal structure. Whilst this was an achievement, given the lack of success in isolating Form III, the fit to the experimental data was suboptimal, with a reported r_{wp} of 13.8%. The authors attributed this poor fit to disorder in the material that was not taken account of in the structural model. Following this, a reassessment of the structure was performed (Reiss et al., 2018a), wherein it was possible to solve the structure from the XRPD data, with a more

reliable r_{wp} of 4.2%. This yielded a similar result to the initial structural report, which differed by the positioning of the methyl group, and distance of the OH…O=C hydrogen bond.

It has been reported that crystallisation of Form III from the bulk amorphous solid is dependent on the material being confined, for example, under a glass cover slip or within a glass capillary (Nanubolu & Burley, 2012; Perrin et al., 2009). Under these conditions Form III crystallises within the bulk material whereas Form II crystallisation begins on the surface. It has been suggested that surface mobility can be a significant factor in the crystallisation of molecular glasses (Yoshioka, Hancock, & Zografi, 1994; Yu, 2016). Where molecular mobility is reduced at the surface of a pharmaceutical solid this may reduce the mass transfer or reorientation processes involved in their re-crystallisation, and thus stabilise the amorphous solid. This is supported for example by experimental observations that nano-confinement of paracetamol in porous activated carbon can produce particles loaded with amorphous solid (Miriyala, Ouyang, Perrie, Lowry, & Kirby, 2017), supporting the theory that suppressing surface mobility can hinder crystallisation. Similarly, adhesion of particles of amorphous indomethacin to polymeric surfaces (Petra A. Priemel et al., 2013) has been demonstrated to improve the stability of the amorphous form due to a reduction in surface molecular mobility induced by drug-polymer interactions at the interface.

The potential for the amorphous form to template crystallisation of pharmaceutical polymorphs has been raised for PCM, where a study of crystallisation from organic solvents reported two distinct amorphous precursors to crystallisation, resulting in either the formation of Form I or II, detected by PDF analysis and Raman spectroscopy (Nguyen Thi et al., 2015). Whilst this suggests a possible templating effect of the amorphous structure, the work does not address the potential impact of residual solvent in the PDF attributed to the reported amorphous form. As such, the contribution from residual solvent to the total PDF

(see Chapters 4 and 5) would be difficult to deconvolute and so a definitive description of local ordering in the amorphous precursors remains lacking. Therefore careful control of sample environment is essential during preparation and measurement to avoid any contribution from environmental or process-induced variations in the sample or experimental data.

Here, amorphous PCM was produced by melt-quenching the liquid phase, and transformations in the amorphous material and crystalline forms tracked using synchrotron derived PDF analysis. Comparison to known polymorphs of PCM and MD simulations of bulk amorphous PCM are used to provide insight to local ordering in the amorphous phase. MD simulations of amorphous PCM at vacuum and SiO_2 interfaces are also performed to understand the role of surface molecular mobility in directing the crystallisation of metastable forms.

6.2. Materials and methods

6.2.1. Materials

Paracetamol was purchased from Sigma Aldrich and used as received.

6.2.2. Synchrotron X-ray powder diffraction

Synchrotron XRPD data were collected at Diamond Light Source (DLS) beam line I15-1 (proposal EE17779), using a wavelength of 0.161669Å and a PerkinElmer XRD 1611 CP3 area detector. Samples were loaded into 1.5mm (outer) diameter borosilicate glass capillaries. Data were collected for 300 seconds per sample. A camera positioned near the sample stage allowed images of the sample to be taken at each temperature.

Variable temperature (VT) data were collected on the sample, with temperature control achieved using a CryoJet. Following a room temperature (RT) reference scan of as-received crystalline Form I, the sample was heated to 180°C, to induce melting, before crash cooling to -155°C. This was followed by a temperature ramp, returning to the melting point in 5°C increments, with a measurement at each temperature interval.

6.2.3. Pawley fitting of crystalline PCM

TOPAS Academic V6 (A. A. Coelho, 2003; A. A. Coelho et al., 2015) was used to perform Pawley fits (Pawley, 1981) to as-received and recrystallised PCM samples for the purposes of phase identification. Lattice parameters used in Pawley fits were the same as those shown in Table 6.1.

A peak shape function, describing the reciprocal-space resolution dependence of the peak shapes and amplitudes, was refined against a silicone reference material. These values were then fixed in subsequent Pawley refinements to the experimental data.

6.2.4. Pair distribution function

GudrunX (Soper, 2011; Soper & Barney, 2011) was used to apply corrections and normalisation for background, multiple scattering, polarisation, Compton scattering and atomic form factor to extract the differential scattering cross-section before transformation to the PDF, using a Q_{max} of 25 Å^{-1}. A Soper-Lorch (Soper & Barney, 2012) broadening of 0.08Å was applied to minimise the impact from ripples arising as a result of termination of the Fourier transform at finite value.

6.2.5. Molecular dynamics simulations

6.2.5.1. Production of bulk amorphous cell

The Amorphous Cell module in the Materials Studio (BIOVIA, San Diego, USA) package was used to generate 10 cubic sample cells of 100 paracetamol molecules, at a density of 1.25g/cm^3. The initial density was chosen as this is lower than, but close to, the density of crystalline forms of PCM (~1.3g/cm^3) (Naumov et al., 1998). Each cell was generated with an initial density of 0.2g/cm^3, to avoid atomic overlap and ring-spearing (see Chapter 1), before ramping to the target density. The Condensed-phase optimised molecular potentials for atomistic studies (COMPASS II) (Huai Sun et al., 2016b) force field was used to describe interactions between particles. The Ewald summation method was used to calculate long-range interactions. Geometry optimisation (GO) was performed using the Smart method, which is a cascade of the steepest descent, adjusted basis set Newton-Raphson (ABNR), and quasi-Newton methods of geometry optimisation, to accelerate GO. Subsequently, the lowest energy cell was used for further analysis.

6.2.5.2. Construction of PCM-interface boxes

Boxes consisting of PCM-vacuum and PCM-SiO_2 interfaces were produced to study the effect of a silica-air interface on the mobility of amorphous PCM at an air glass interface. The PCM-silica box was produced using a starting model for amorphous SiO_2 (included with Materials Studio software), with cell dimensions a = b = 28.51Å, c = 40Å. The c-axis is expanded such that approximately 60% of the volume of the cell is unoccupied. This allowed the Amorphous Cell packing task to be used to fill the void space with PCM molecules, at a density of 1.22g/cm^3, in line with the equilibrated density of bulk amorphous PCM (see

Section 3.5.3). Following packing of the cell, the geometry was optimised using the Smart method, before dynamics simulations were run.

The PCM-air cell was produced using the bulk amorphous PCM cell after equilibration under the NPT ensemble, as a starting configuration. The *b*-axis was expanded to 55Å, allowing approximately half of the cell to be occupied with PCM molecules. The remaining volume was left unoccupied, to reflect the PCM-air boundary. No input atoms were used due to the low density of air. A geometry optimisation was performed using the Smart algorithm before further analysis.

6.2.5.3. Molecular dynamics simulations of amorphous PCM boxes

The Forcite module within the Materials Studio package was used for molecular dynamics (MD) simulations. The lowest energy cells were used for dynamics simulations in all cases. Simulations of the bulk amorphous PCM box was performed for 2ns initially using the NPT ensemble to equilibrate the density, using the Nose thermostat (Nosé, 1984) and the Berendsen thermostat (Berendsen et al., 1984) for temperature and pressure control, respectively. After an initial equilibration, calculations were performed for a further 2ns using the NVT ensemble, with the Nose thermostat for temperature control. The final density of the NPT simulations was 1.22g/cm^3, which was used as the target density of the PCM populated volume of the cell in the PCM-SiO_2 box. The interfacial amorphous cells were equilibrated for 2ns, using the NVT ensemble, with a further 2ns performed for analysis of the trajectories. Further simulations at 60°C were performed using the same settings as above, to compare mobility in the amorphous solid at high and low temperature.

6.2.5.4. Calculation of PDFs from MD-derived structural models

TOPAS Academic V6 (A. A. Coelho, 2003; A. A. Coelho et al., 2015) was used to calculate PDFs from individual MD frames using Equation 6.1 (Takeshi & Billinge, 2012b)

$$G_{calc}(r) = \left(\frac{1}{r}\right) \sum_i \sum_j \left[\left(\frac{f(Q)_i f(Q)_j}{\langle f(Q)\rangle^2}\right) \delta(r - r_{ij})\right] - 4\pi r p_0 \qquad (6.1)$$

where $f(Q)_i$ and $f(Q)_j$ are the atomic form factors of atoms i and j, respectively, p_0 is the average number density, and $\langle f(Q)\rangle$ indicates the compositionally averaged form factor. Calculated PDFs were convoluted with a $sinc$ function to replicate the Q_{max}-dependent termination ripples in the experimental data. A further convolution was applied to account for the Soper-Lorch broadening (Alan K. Soper & Barney, 2012) applied to the experimental data. When calculating the PDF from the MD-derived structural model of amorphous PCM, only the scale and zero-error were refined.

6.2.5.5. Calculation of diffusion coefficients in interfacial MD simulations

Diffusion coefficients (DCs) were calculated for the MD simulations of bulk amorphous PCM, PCM-vacuum and PCM-SiO_2 by fitting a straight line of the form $Y = mx + c$ to the plot of the mean squared displacement (MSD) to simulation time, where $m/6$ is equal to the diffusion coefficient (Mauritz, Storey, & George, 1990). Standard deviations (SDs) were obtained by splitting the plot into four sections and calculating the DC independently for each. SDs are thus calculated as the standard deviation of the 4 DC values.

6.2.6. PDF refinements of known PCM structures to experimental amorphous PCM

TOPAS Academic V6 was used to perform PDF-refinements of the MD-derived structures of known PCM polymorphs to the experimental data for amorphous PCM. A damping parameter was used to simulate the known forms as 7.5nm nanocrystals, such that the attenuation of the PDF signal was comparable to the experimental data. This was calculated by multiplying the PDF by the autocorrelation function (S. J. L. Billinge et al., 2010):

$$f(r;d) = \left[1 - \frac{3r}{2d} + \frac{1}{2}\left(\frac{r}{d}\right)^3\right]\Phi(d-r) \qquad (6.2)$$

where d is the diameter of the domain, assuming a spherical particle.

Coordinates of known forms were constrained by dummy atoms, with distance restraints applied, to prevent the resulting structure from deviating uncontrollably from the starting positions. Q-related damping and broadening were refined against a silicon reference standard and were fixed during refinement of PCM structures to the experimental amorphous PDF.

6.2.7. Statistical analysis

PolySNAP 3 (Barr et al., 2004b) was used to compare the experimental PDFs across the temperature range to the calculated PDFs of known PCM forms, by generating Pearson correlation coefficients (CC) (Szczepańska, 2011) between each dataset. This coefficient provides a quantitative measure of the similarity between PDFs. Using this method, a CC of one indicates a complete correlation between two datasets, zero indicates no correlation, and -1 implies a negative correlation. For all datasets, comparisons were performed between 3.5 and 8.5Å, as above this value there is minimal structural information in the PDF of amorphous materials.

6.2.8. Principal component analysis

The PCA module in Origin (OriginPro 2018, academic) was used to identify trends in the experimental VT PDF data. The r-range used in the analysis was 3.5 – 8.5Å. Three principal components were used as these were found to account for >96% of the variance in the data. The first principal component (PC1) was able to account for > 90% of the variance in the data, while principal component 2 (PC2) only accounted for ~ 5% of the variance. A third component (PC3) was included in the analysis, however, this was only able to account for ~ 1% of the variance. Score plots were calculated for PC1 and PC2, and PC1 and PC3.

6.2.9. Area under the curve (AUC)

The trapezoidal method was used to integrate the area, I, under PDF curves, by (Liengme, 2014),

$$I \approx \sum_{i=1}^{n} \frac{1}{2}(y_{i+1} + y_i)\Delta x \qquad (6.3)$$

where the area under the curve is divided into small segments based on the step-size in the x-axis, Δx, and is calculated as the height of each segment multiplied by Δx. The PDF data were analysed in the region between 3.5Å and 8.5Å.

6.2.10. **Conquest search of structural features in small-molecule crystal structures**

A search of the Crystal Structure Database (CSD) was performed using ConQuest (Bruno et al., 2002a) to identify the occurrence of C=O dipole interactions in small molecule crystal structures, for comparison to the MD-derived amorphous PCM structure. An input structure was drawn describing the N-acetyl group found in PCM (Figure 6.3). A non-bonded contact defined between the two drawn carbonyl groups, up to a maximum distance of 3.5Å was used as the search criteria.

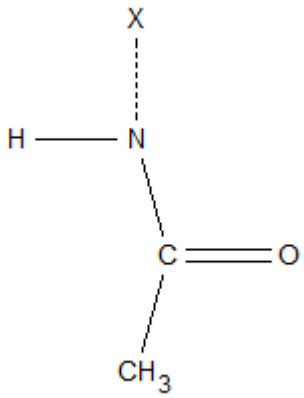

Figure 6.3: Input structure describing N-acetyl functional group found in PCM. Used as search criteria for ConQuest search.

ConQuest was also used to search torsion distributions in known crystalline PCM structures and the MD-derived structural model, using a custom CSD database (Section 2.12). A 2D drawing of a PCM molecule was used as input for the ConQuest search, with two flexible torsions defined as shown in Figure 6.4.

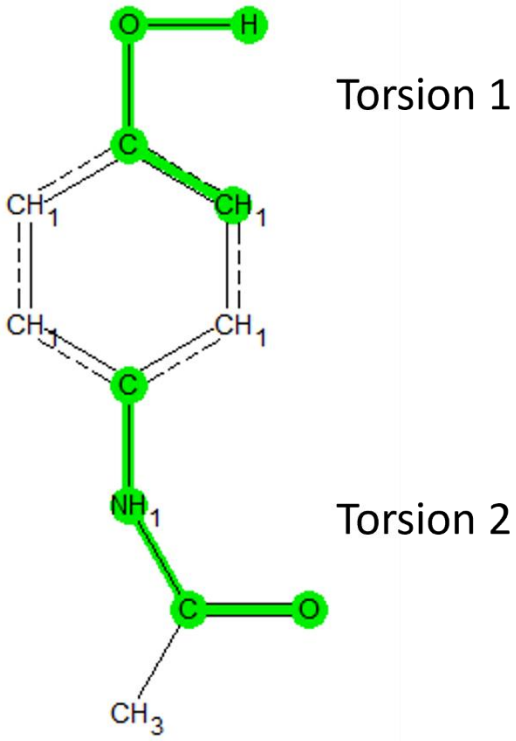

Figure 6.4: Input structure used for ConQuest search of PCM flexible torsions in crystalline forms and MD-derived structural model.

6.2.11. Full Interaction Maps (FIMs)

Full Interaction Maps (FIMs) are a tool available within the *Mercury* software that provide an alternative way to visualise and identify intermolecular interactions. They allow the calculation of contour plots around a given functional group, where the contours give an indication of the preferred hydrogen bonding location, or other non-bonded interaction. Thus, where a donor-acceptor site sits out with the contour, this indicates that there is a more feasible or preferred bonding arrangement. Here, FIMs are used to assess the intermolecular geometries within the MD-derived structural model of amorphous PCM. The contours were calculated with a likelihood level of 2, 4, and 6, meaning that the contour will

be generated around an area wherein the density of interactions is 2, 4 or 6 times that which would be expected at random.

6.2.12. Construction of IsoStar plot from MD frames

A crystal structure database (CSD) searchable database was created using the Industrial CSD editor software (V2.21.0) (Groom, Bruno, Lightfoot, & Ward, 2016). The database consisted of frames from the RT bulk amorphous PCM MD trajectory. The IsoStar (Bruno et al., 1997) plot was calculated by searching the custom database using ConQuest (Bruno et al., 2002a). The corresponding files are then manually exported to IsoGen, where the 3D plot is automatically generated.

6.2.13. Conformational assessment of amorphous PCM

The CSD software package MOGUL (Bruno et al., 2004) was used to assess the intramolecular geometries generated from a geometry optimised frame extracted from the MD simulations, to ensure they were chemically reasonable. The program DSNAP (Barr, Dong, Gilmore, Parkin, & Wilson, 2005) was used to assess conformational similarity between PCM molecules in the MD simulation frames and the known crystal structures of PCM. The comparison was enabled by exporting the ConQuest files described in Section 2.10 to DSNAP.

6.3. Results and discussion

6.3.1. VT-XRPD

Selected VT-XRPD data are shown in Figure 6.6. The scan of as-received Form I is shown for reference. The absence of long-range order where the XRPD pattern is a broad halo due

to melting is evident from the liquid melt at 180°C. Subsequent scans of the crash-cooled glass also reflect this lack of 3D periodicity. From visual inspection of the amorphous XRPD data (Figure 6.5), there is very subtle change in the amorphous "halo" as a function of temperature, reflecting the similarity in the local structure of the glass and liquid samples. There is a subtle shift in the maxima for the melt towards higher scattering angles, which would suggest shorter average distances in the local structure of the liquid. This contrasts with the general expectation of thermal expansion increasing average intermolecular distances and that liquids have lower density than solid materials (Khachan, 2018), with water being the exception (Ball, 2008; Tanaka, Girard, Davis, Peuto, & Bignell, 2001). While this was not investigated further, it is an unusual observation, and at the time of writing there are no known reports of similar behaviours in amorphous pharmaceuticals.

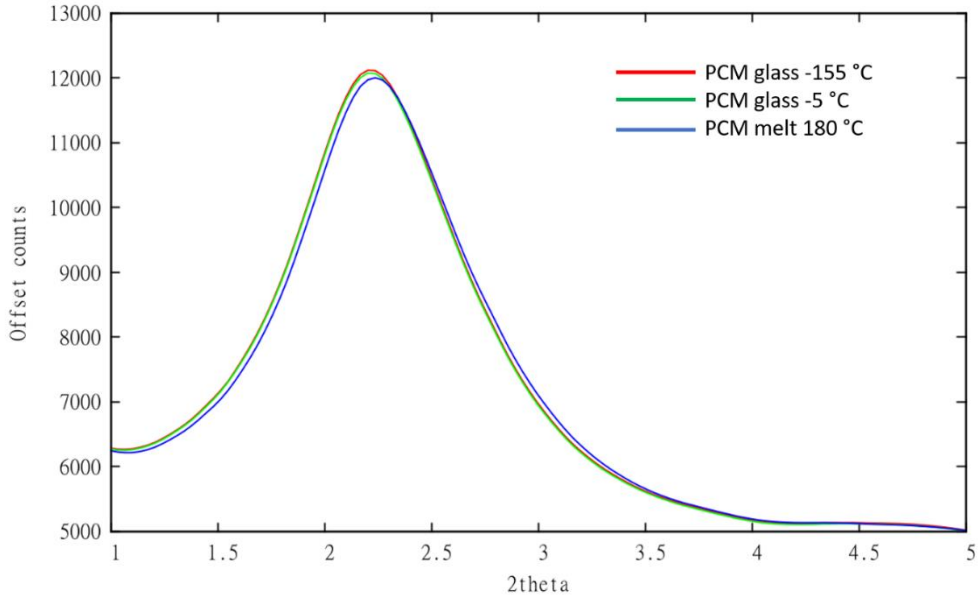

Figure 6.5: Overlay of XRPD data for PCM glass at -155 °C (red), glass at -5 °C (green), and melt at 180 °C (blue). Truncated between 1 and 5° 2θ for clarity.

The sample underwent crystallisation to Form III upon heating the glass to ~60°C, confirmed by Pawley refinement of the lattice parameters to XRPD data (Section 3.2).

Following the initial crystallisation of form III the sample underwent a subsequent solid-solid transformation to Form II at 100°C, and a subsequent conversion to Form I at 155°C. The observed events are close to reported values for PCM polymorph crystallisation from the melt (Burley, Duer, Stein, & Vrcelj, 2007; Di Martino et al., 1997). By comparison of the reference Form I scan to the recrystallised Form I, it is evident that there is significant preferred orientation in the recrystallised sample.

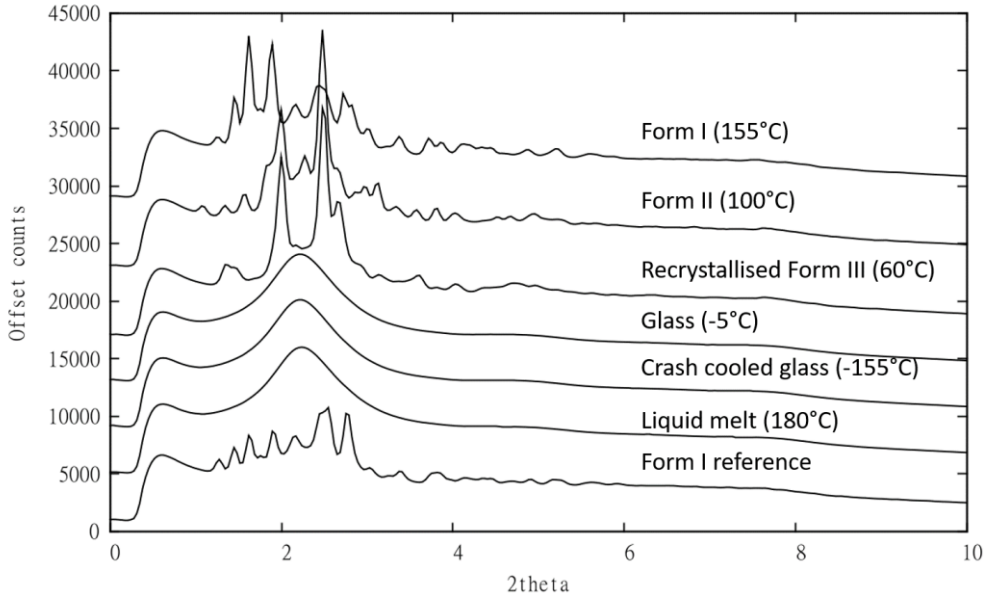

Figure 6.6: Stack of selected plots from VT-XRPD data collection. From bottom to top: RT reference scan of Form I, liquid PCM after melting Form I at 180 °C, crash-cooled glass after cooling the melt to -155 °C, glass after heating from 155 °C to -5 °C, recrystallised Form III after heating glass to 60 °C, transformation to Form II following heating of Form III to 100 °C, transformation to Form I after heating Form II to 155 °C. Data shown is truncated between 0 and 10° 2θ for clarity.

6.3.2. Pawley fits of known lattice parameters to recrystallised samples

For the purposes of phase identification, the XRPD data from samples of as-received and recrystallised sample were used in a Pawley refinement against the known lattice parameters for each polymorph. The resultant fits and refined lattice parameters are shown in Figure 6.7

and Table 6.2, respectively. Whilst the poor reciprocal-space resolution is worth noting, the excellent fits give confidence in the identified phases.

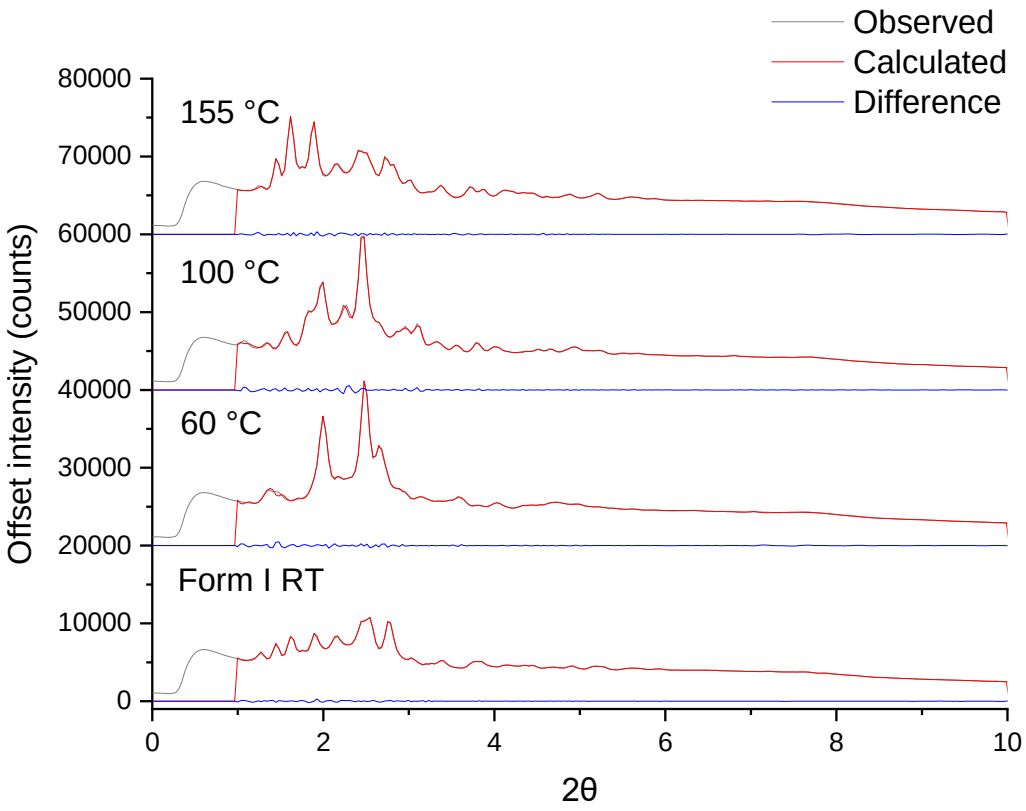

Figure 6.7: Pawley fits of known lattice parameters to as-received and recrystallised PCM. From bottom to top: Fit of Form I lattice parameters to as-received material, fit of Form III lattice parameters to recrystallised sample, after heating the amorphous solid to 60 °C; fit of Form II lattice parameters to sample after heating recrystallised Form III to 100 °C; fit of Form I lattice parameters after heating recrystallised Form II to 155 °C. In all cases grey shows the observed data (Y_{obs}), red shows the calculated (Y_{calc}) and blue shows the difference ($Y_{obs} - Y_{calc}$). R_{wp} values were all <4%.

Table 6.2: Refined lattice parameters from Pawley refinement to experimental VT XRPD data.

Sample	*A* (Å)	*B* (Å)	*C* (Å)	α (°)	β (°)	γ (°)
Form I reference	7.0805	9.3303	11.6903	90	97.4091	90
Re-crystallised Form III	11.7603	8.5304	14.9652	90	90	90

Re-crystallised Form II	17.0084	11.5597	7.2545	90	90	90
Re-crystallised Form I	7.1155	9.5979	11.5510	90	97.7758	90

6.3.3. VT-PDF data

Figure 6.8 shows selected PDF data sets, calculated from appropriately corrected and normalised (Section 2.4) synchrotron total scattering data (see Section 2.5), across the temperature range. From visual inspection, no significant change in the amorphous structure is observed as a function of temperature before the first crystallisation event at 60℃. However, upon closer inspection of the features in the amorphous PDFs (Figure 6.10), there are subtle differences between the PDFs of the liquid melt and subsequent measurements of the amorphous solid.

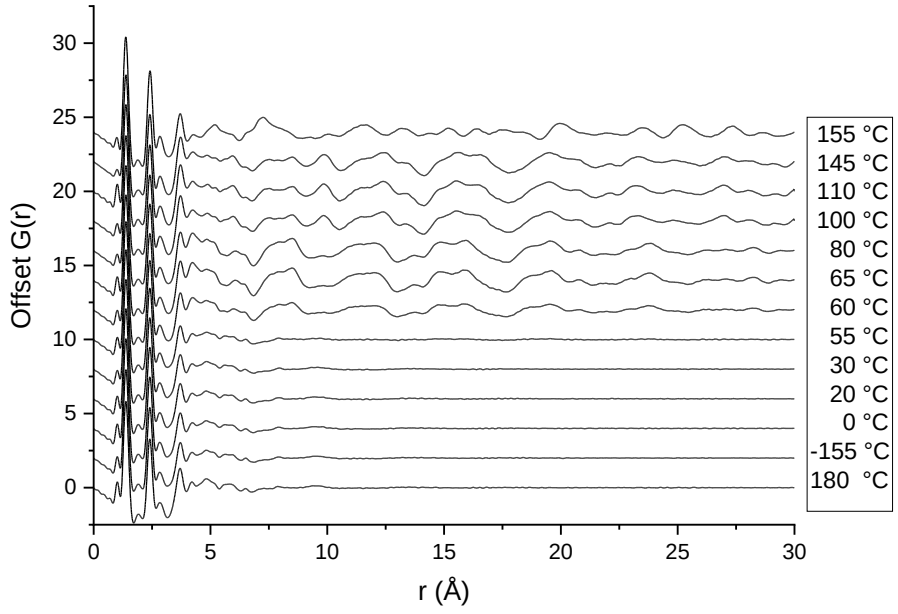

Figure 6.8: Selected VT-PDF data in the range 0-30 Å showing transformations from the amorphous solid, starting with the liquid melt (bottom) produced by melting Form I at 180 °C, followed by the PDF of the amorphous glass following cooling of the melt to -155 °C. Subsequent data sets show transformations induced in the material after heating the glass to the temperatures shown in the plot.

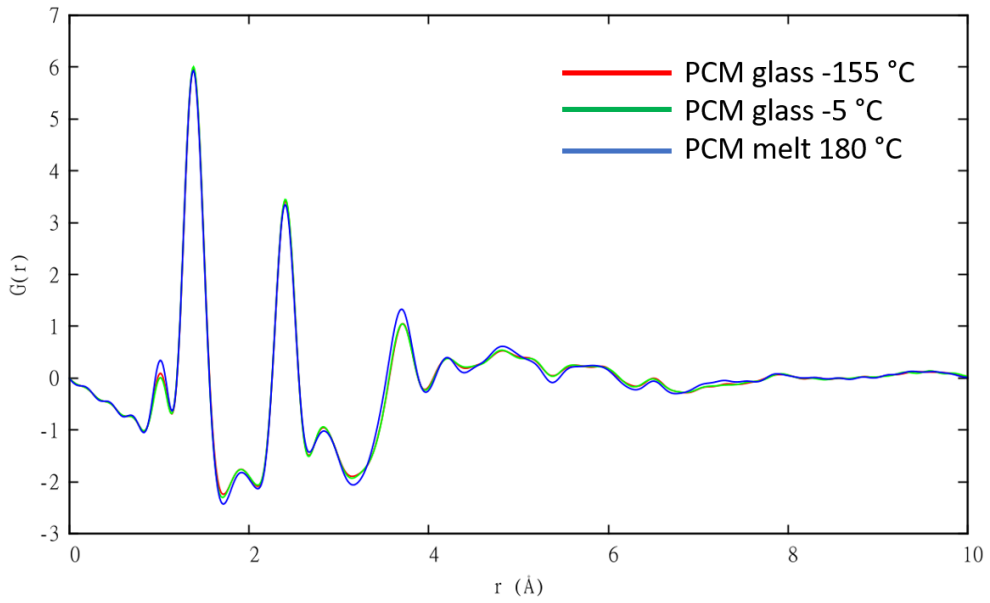

Figure 6.9: Overlay of PDF data in the range 1-10 Å for PCM glass at -155 °C, after cooling from the melt at 180 °C (red), glass at -5 °C following heating of the sample from -155 °C (green), and melt at 180 °C, after melting of as-received Form I (blue).

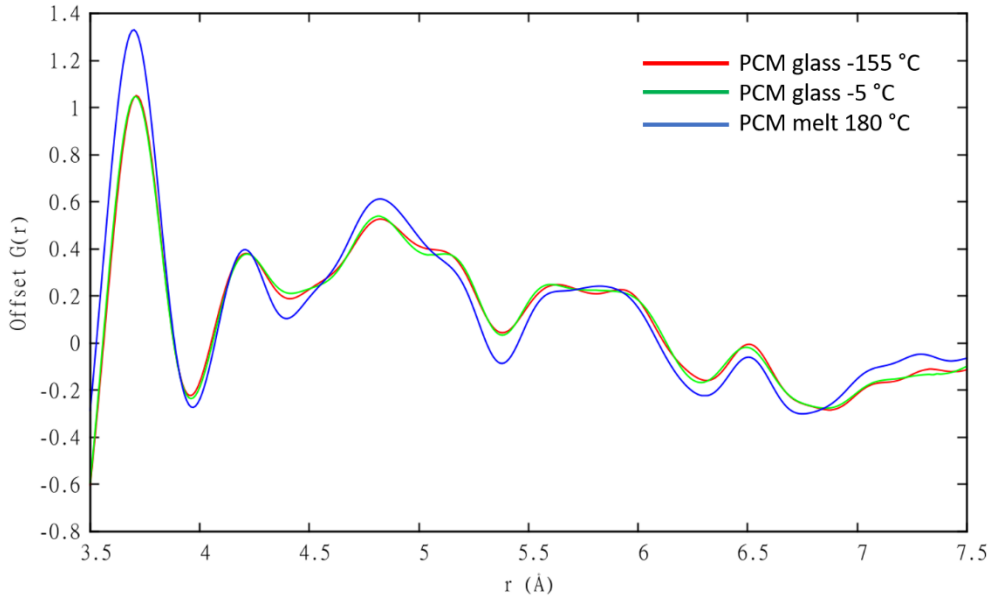

Figure 6.10: Overlay of PDF data in the range 3.5-7.5 Å for PCM glass at -155 °C (red), to the glass at -5 °C (green) and melt at 180 °C after melting of as-received Form I (blue).

Attempts to reproduce the observed intensity differences between the PDFs of solid and liquid PCM using MD simulations performed at 500K failed (Figure 6.11). Similarly, there was no unexpected changes in the density, which was lower for the high temperature MD simulation than those run at lower temperatures. While some structural change is expected between the experimental PDF data for the melt and the amorphous solid, the differences shown in Figure 6.10 differ from what may be expected for a liquid and solid. Notably, it would reasonably be expected that the peaks of the liquid would be broader than that of the solid material, due to increased mobility and thermal vibrations in the high-temperature melt. This is not necessarily the case here as in some cases, the features in the solid are observed to be broader. While thermal parameters do not affect the PDF data of amorphous materials as significantly as that for crystalline due to the significant static disorder in the amorphous material (S. J. L. Billinge, 2011), it is reasonable to assume that the liquid PDF would exhibit broader features than the solid.

The expected effect of thermal vibrations on the PDF on an amorphous material is demonstrated in Figure 6.12, where PDFs are calculated from the bulk amorphous PCM MD box (See sections 2.5.4) using varying thermal parameters (*beq*s). As expected, the increasing thermal parameters introduce a broadening effect on the calculated PDF. This change is minimal initially but becomes more apparent as thermal parameters are increased. This is particularly evident in the low-r region as the features become increasingly broad and dampened as a result, making it difficult to resolve individual features in the data. The significant changes in the low-r region are due to the largely intramolecular nature of these features, where there is significant correlated atomic motion as a result of the covalent bonds between the atoms (Prill, Juhás, Schmidt, & Billinge, 2015). Above ~3.5Å a similar damping and broadening of intermolecular features is observed, demonstrating the expected behaviour of an amorphous material as a function of temperature. The observed effect on

the calculated PDFs contrasts with the experimental data, indicating some structural change between liquid and solid samples of PCM which cannot be explained by increasing thermal vibrations.

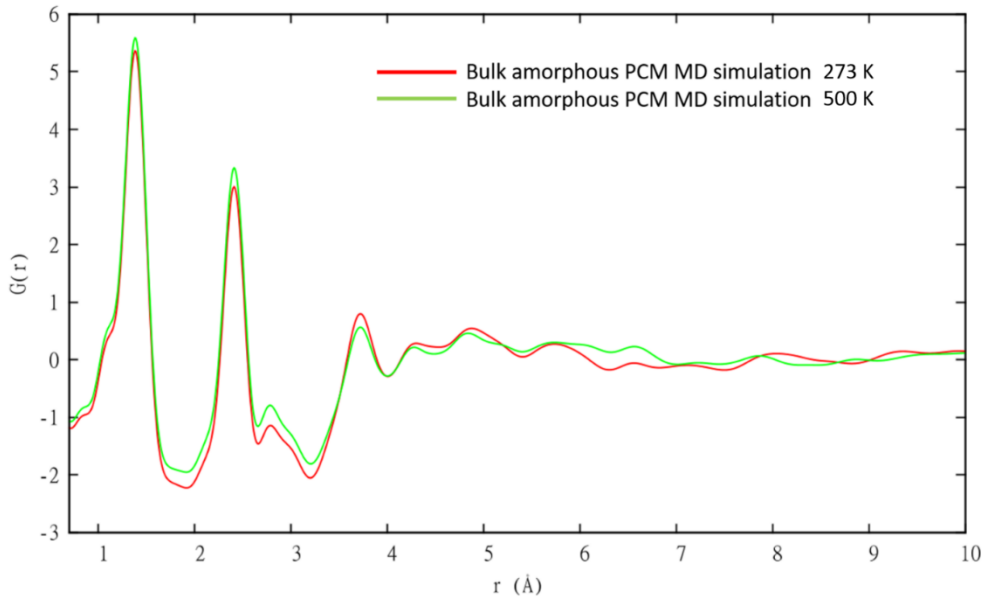

Figure 6.11: Comparison of PDF calculated from MD simulations of bulk amorphous PCM ran at 25°C and 227°C.

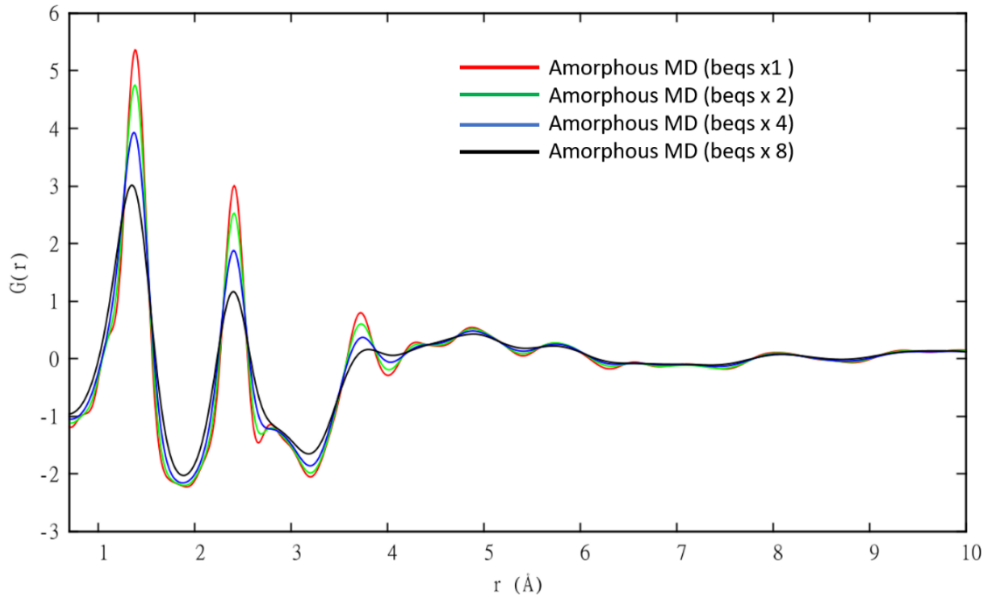

Figure 6.12: Comparison of PDFs calculated from the bulk amorphous PCM MD box using different thermal parameters (beqs).

6.3.4. PDF fits of known PCM polymorphs to experimental amorphous PDF

To understand local packing in the amorphous solid, PDF fits of known PCM polymorphs to the data obtained from a crash-cooled sample of PCM at -155°C were performed (Figure 6.13). The calculated PDFs were modified to resemble a 7.5Å nanoparticle (See section 2.6) such that the PDF attenuates to zero at r-values comparable to an amorphous material. To demonstrate the effect of this modification, Figure 6.14 shows the calculated PDF of PCM Form I before and after modification by the nanoparticle shape function. When the PDF is modified to resemble that of a nanoparticle, the PDF attenuates to zero at much shorter distances. Otherwise the features in the unmodified PDF extend to beyond the upper limit shown. Additionally, the amplitudes of the PDF peaks are smaller than the unmodified calculation. The use of the modification function thus allows comparison of local order in the amorphous solid to crystalline forms while minimising the contribution of long-range order to the PDF.

All fits compare reasonably in the low-r region, which is expected due to the dominance of intramolecular bonding in this region of the PDF (Prill et al., 2015). The PolySNAP (Barr et al., 2004b) analysis of the data (Figure 6.15), reveals that Form III has the closets similarity to the experimental data, where the highest CC is for Form III (CC = 0.957). This value shows a strong similarity and is significantly higher than the comparisons between any of the other refined PDFs, where the next highest value is for Form I (0.917).

The high CC between the calculated PDF for Form III and the experimental amorphous PDF implies there is significant similarity between the local structures of Form III and the amorphous material. This similarity may suggest that the amorphous structure possibly acts as a template for Form III. Amorphous PCM has previously been reported to template polymorphic form in solution crystallisation experiments (Nguyen Thi et al., 2015). This is consistent with Ostwald's rule of stages (Ostwald, 1897), wherein phase transitions are directed by the smallest change in free energy. Structural similarity between amorphous and crystalline forms have been reported for inorganic (Mavračić, Mocanu, Deringer, Csányi, & Elliott, 2018) and molecular (Mariedahl et al., 2018) systems, where the tetrahedral arrangement of water molecules observed in the crystalline forms is retained in the liquid.

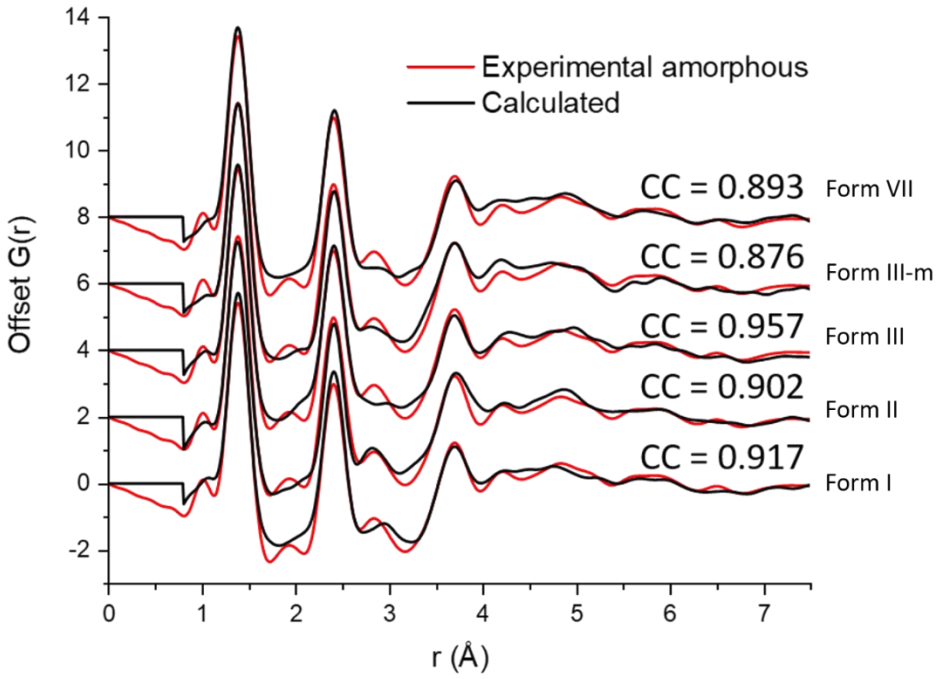

Figure 6.13: Stack of refined PDFs calculated from the known crystals structures to the experimental PDF of amorphous PCM at -155°C. From top to bottom: Form VII, Form III-m, Form III, Form II, and Form I. CCs calculated between refined and experimental data shown above each plot.

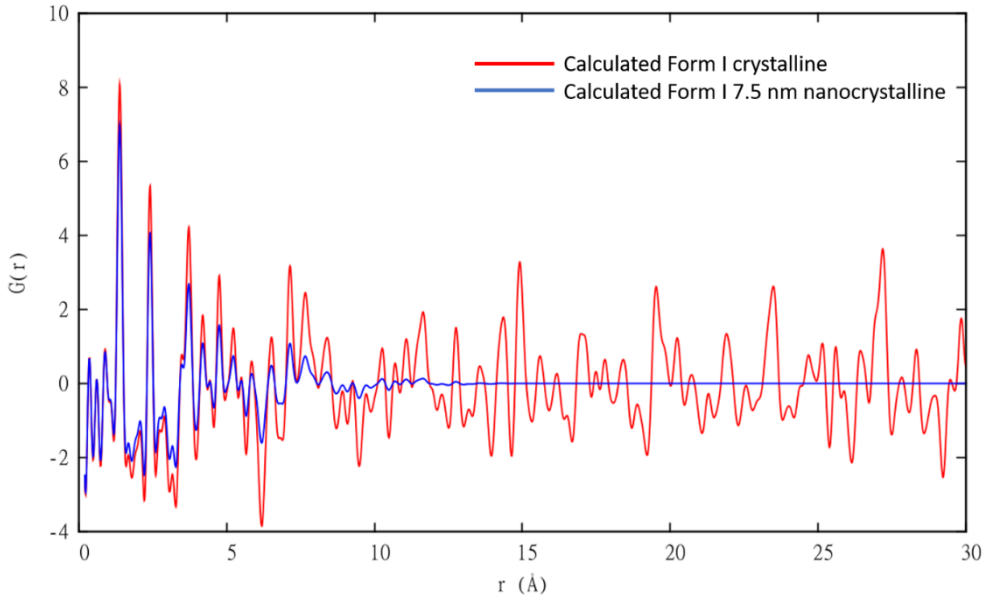

Figure 6.14: Overlay of calculated PDFs of PCM Form I with and without nanoparticle modification function. Red line shows bulk crystalline material, blue is PDF of Form I modified to resemble a 7.5nm nanocrystal. Attenuation of the PDF peaks at low-r is apparent in the nanocrystalline example, whereas for bulk crystalline material, peaks extend to beyond the upper r-limit shown, reflecting long-range order. PDFs calculated using TOPAS Academic V6.

Figure 6.15 shows the CCs for the refined PDFs of known forms when compared to experimental PDF data across the measured temperature range -5-50°C. The crash-cooled amorphous solid displays a very similar local structure to both the liquid and Form III (both CCs > 0.95). As temperature increases, a significant drop in CC for the melt is apparent, where the lowest CC is 0.918 at 50 °C, immediately before crystallisation. The similarity to Form III, however, remains relatively constant, with the lowest CC being 0.949. This value is still considerably higher than the largest CC value for any of the other forms. CCs for the crystalline forms also see a striking reduction in their similarity to the amorphous solid across the temperature range.

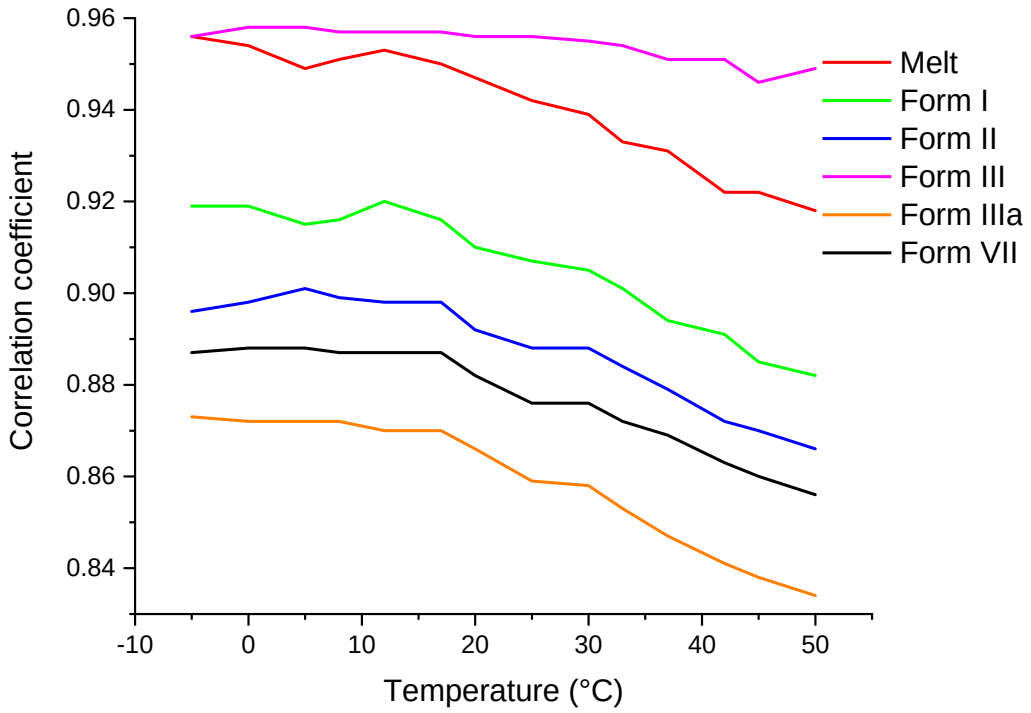

Figure 6.15: Plot of correlation coefficients for refined crystalline PDFs of each known form to experimental VT-PDF data for amorphous PCM.

6.3.5. Area Under Curve (AUC) Analysis for VT-PDF data

It is apparent from Figure 6.15 that the amorphous structure undergoes some subtle change as a function of temperature however the exact cause of these is difficult to interpret. Whether these changes are due to structural re-organisation or purely a temperature effect that reflects the mobility of PCM molecules in the amorphous material is unknown. In an effort to examine the changes in PDF that occur as a function of temperature the area under each PDF curve was determined and plotted across the temperature range (Figure 6.16). Significant variations in the AUC occur as a function of temperature, transformations to and between crystalline forms being readily discernible. Overall, a downward trend is shown to occur with increasing temperature, after the initial crystallisation. This decrease in the AUC is expected to be due to broadening and damping of features in the PDF due to increased thermal motion and expansion.

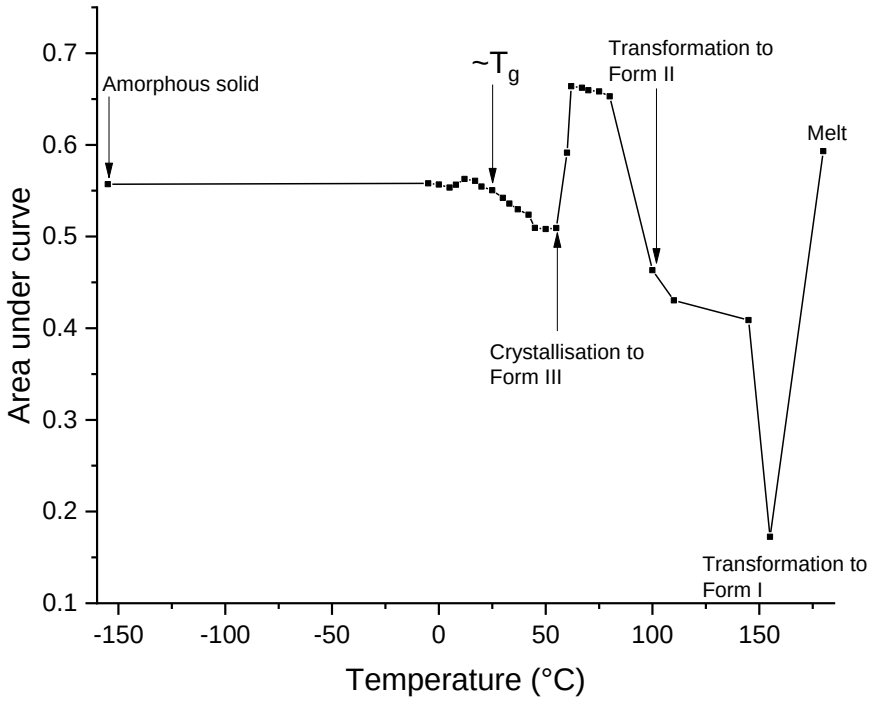

Figure 6.16: Area under PDF curves as a function of temperature. Figure annotated with key transformations.

When focussing on the temperatures below the crystallisation temperature, it is apparent that subtle changes occur in the amorphous phase (Figure 6.17). Changes in the plot shown are approximately consistent with reported transformations in amorphous PCM, where an increase in the AUC is shown from ~ 12°C, directly before a reported β-relaxation in PCM (~17°C) (Johari, Kim, & Shanker, 2005; Sibik et al., 2015), and the reported T_g of PCM of ~23°C (Sibik, Sargent, Franklin, & Zeitler, 2014). The β-relaxation process relates to an intramolecular relaxation and is often viewed as a precursor to the glass transition temperature. Above these events, there is a steady drop in the AUC. This steady drop, as described earlier in this section, is expected to be due to increased mobility in the sample. The AUC then seems to plateau at ~45°C, where it remains relatively stable directly before crystallisation to Form III.

The observed changes in the AUC also align with images taken of the sample on the beamline during data collection. Specifically, the glass changes from transparent to opaque before the reported β-relaxation (Sibik et al., 2015). Following this, the sample changes to a transparent material at the T_g. At ~40°C, before the AUC plateaus, the sample is again transparent. However, the sample was visibly more mobile inside the borosilicate glass capillary. Whilst AUC is not a widely used method for analysis of PDF, this analysis shows the total PDF signal does appear to vary around known physical transformation in PCM. Hence, we propose AUC as a relatively simple and convenient means to analyse temperature series PDF data to identify transition points quickly, allowing individual PDFs of interest to be selected for more detailed analysis.

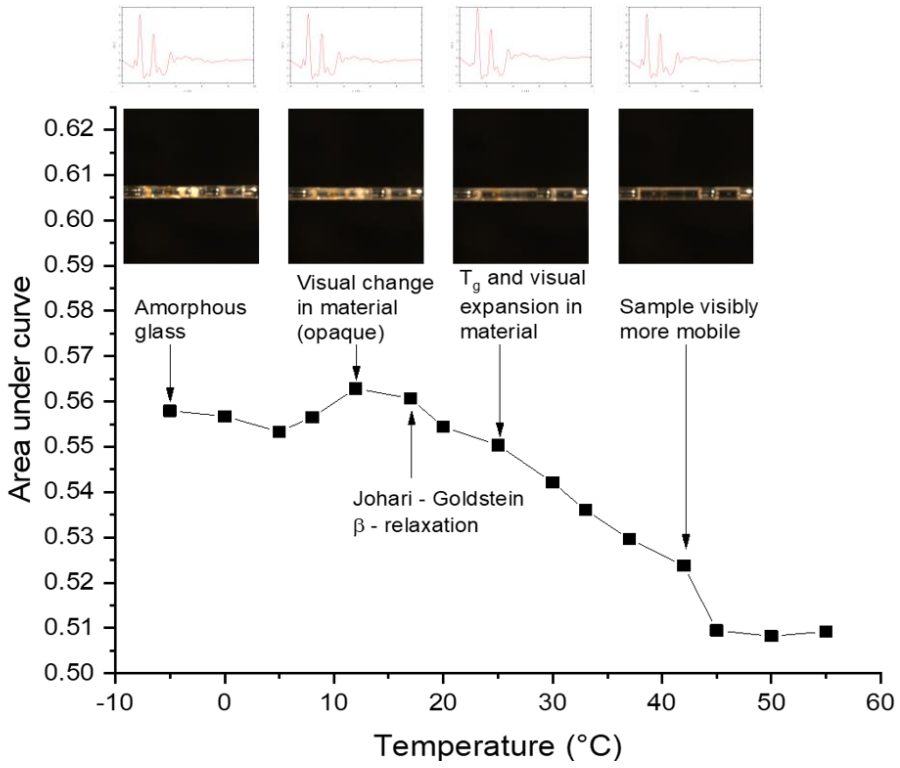

Figure 6.17: AUC of PDFs in the range T = -10-55 °C i.e. up to the crystallisation temperature. Figure is annotated with key physical transformations in the amorphous solid as well as images of the sample at each key temperature. Opaque regions are due to cracks in the glassy material not the presence of crystalline PCM. In all cases, X-ray beam was focussed on right-hand side of each image, ensuring data were collected from the sample volume closest to the Cryojet sample heater. PDF data collected at the same temperature as each image shown.

6.3.6. PCA of VT-PDF data

Score plots (see Section 2.8) for the VT-PDF data are shown in Figure 6.18. A clear temperature dependent trend is present in the data, shown in plot 16 (a) of PC1 against PC2 where PDFs below the T_g are clustered (state T range labels or cluster colour). Above T_g (T>25°C), there is a systematic trend with increasing temperature. The melt (180°C) appears distinct from all other data sets. In Figure 6.18b, showing score plot for PC1 and PC3, a similar trend is apparent. However, of interest is the region highlighted purple in the score plot, where temperatures associated with transitions reported in amorphous PCM (Section 3.6.) are clustered. Specifically, the temperatures in this purple cluster are associated with the T_g and β-relaxation events (12°C – 25°C) and the temperature range at which the sample was observed to be more mobile in the capillary (42°C). It is likely that there is some structural change associated with this cluster of points, whether related to structural or conformational relaxation, or otherwise. These temperatures were also identified using the AUC method and by visual observation of the sample (Section 3.6). It is worth noting that while PC3 describes only a small percentage of the variance (1.3%), the separation of clusters of PDFs corresponding to changes in the material demonstrate the value in including PC3 in the analysis.

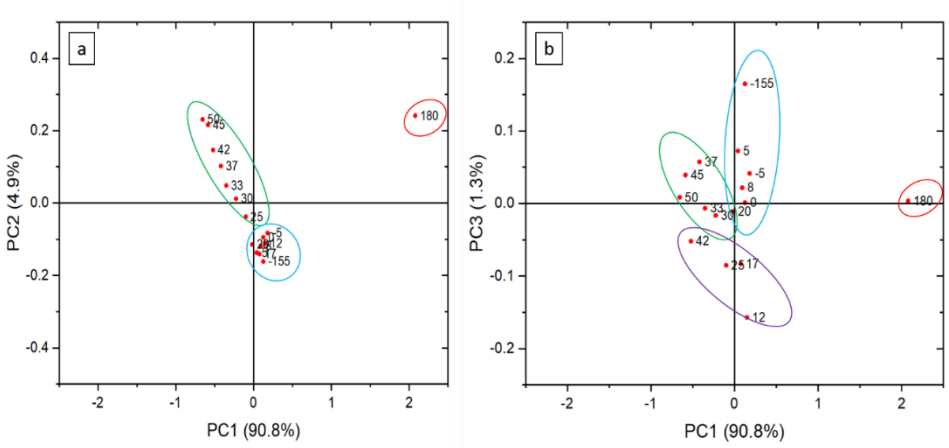

Figure 6.18: PCA score plots for VT-PDF data: (a) score plot of PC1 against PC2 and (b) score plot of PC1 against PC3. Trends are highlighted for ease of interpretation. Data labels are temperatures (°C) that corresponding data set was collected at.

6.3.7. Comparison of experimental amorphous PDF to MD-derived structural model

Figure 6.19 shows the overlay of the experimental PDF of amorphous PCM with the PDF calculated (see Section 2.5.4) from a single geometry optimised frame from the bulk amorphous PCM trajectory (see Section 3.10, Figure 6.28). Although small differences are apparent between the two PDFs there has been no refinement of the structural model to the experimental data. Only the scale factor and zero-error were refined. This initial fit shows all the main features observed in the experimental data appear to have reproduced suggesting that the MD simulation have captured the representative interactions present in the amorphous solid.

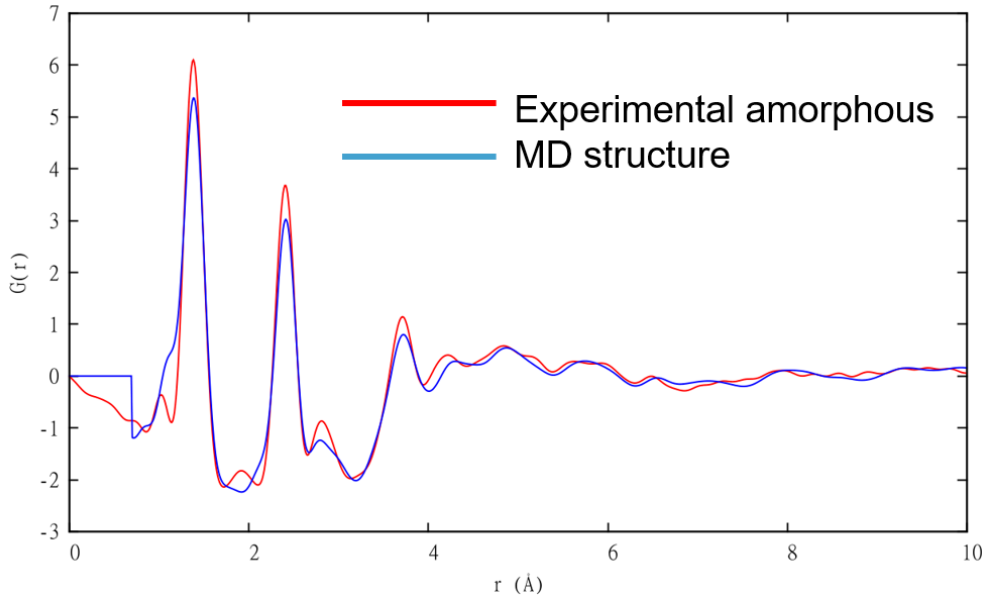

Figure 6.19: Overlay of experimental PDF of crash cooled amorphous PCM with PDF calculated from geometry optimised MD structure.

6.3.8. Assessment of intramolecular geometries in bulk amorphous PCM MD box

6.3.8.1. MOGUL geometry check

An assessment of the intramolecular geometries in the bulk amorphous PCM box was performed using MOGUL (Bruno et al., 2004) to ensure that the resulting structure was chemically reasonable. All distances and most angles were in agreement with similar fragments within the CSD, demonstrating that sensible intramolecular geometries were retained using the COMPASS II force field. The angles identified as unusual were always related to the N-acetyl group and were only slightly out with the threshold identified for PCM molecules. However, a ConQuest search of the N-acetyl group reveals similar values have been found for other molecules in the CSD. Mogul identified 11 molecules out of 100 where torsion values lay out with the normal torsion angle distributions. This demonstrates that most MD-derived geometries are in agreement with the data available within the CSD, suggesting that they are low-energy conformations. While the majority of conformers align

with the CSD, the 11 identified as “unusual” may have a high conformational potential energy, as would be expected due to the conformational distribution within the high-energy amorphous solid, which may be otherwise stabilised by non-bonded intermolecular interactions.

6.3.8.2. Torsion distributions

Figure 6.20 shows the torsion distributions obtained by a search of the CSD using ConQuest (Section 2.10). A distribution of the torsion angles defined in Section 2.10 is apparent in the bulk amorphous MD box, compared to the structurally characterised forms of PCM. This is expected due to the relatively low Z' of all PCM crystal structures, all of which are ≤4. By contrast, the MD-derived structure contains 100 independent units. This is reflected in the torsion distributions, where rotation of the N-acetyl and hydroxyl groups result in a number of distinct conformers. For both torsion angles, there is a large distribution within the MD-derived structure and there is no obvious similarity or preference for any of the known crystalline forms of PCM.

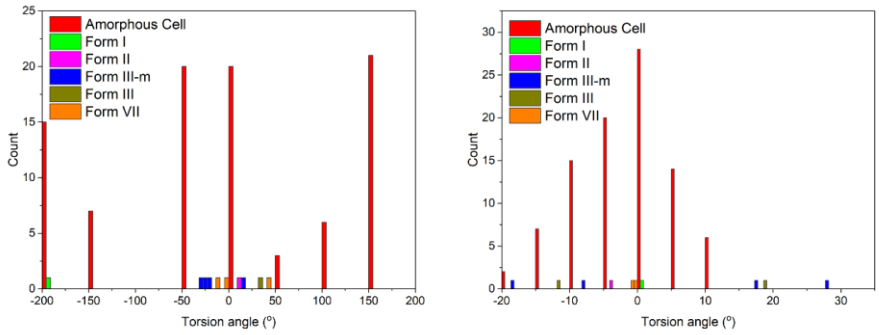

Figure 6.20: Comparison between torsion distribution for bulk amorphous PCM MD box and known crystalline forms for (left) Torsion 1 as defined in Section 2.12 (right) Torsion 2. Despite the small number of rotatable bonds, a significant distribution of conformations is apparent, largely due to rotation of the N-acetyl and hydroxyl groups.

6.3.9. 3D IsoStar plot of interactions in bulk amorphous PCM

Figure 6.21 shows the 3D IsoStar plots corresponding to short-range non-aromatic interactions, calculated from the bulk amorphous PCM MD trajectory file (Section 2.10). For the contact points shown in the plots, red points represent oxygen, blue are nitrogen and grey are carbon. The software used to calculate the 3D plot (IsoGen) assumes that a single conformation is present across all structures provided. Thus, the reference conformer shown corresponds to that described by the first set of coordinates encountered by the software in the calculation. This will impact on the exact geometries of subsequent interactions mapped in the density plots however should not have a significant impact on the overall distributions observed in the disordered material.

Analysis of the intermolecular interaction types in the amorphous PCM model show he N-H···O=C interactions. This interaction is not observed in any of the polymorphic forms of PCM (see Section 1) although has been observed in amorphous PCM samples using FTIR spectroscopy (Trasi & Taylor, 2012), where the peak attributed to the C=O stretch (assigned to the O-H···O=C hydrogen bond) splits. This peak spitting is attributed to both the O-H···O=C interactions, also observed in the crystal forms, but also to N-H···O=C. This N-H···O=C hydrogen bond therefore appears to be exclusive to the amorphous solid. It is also worth noting that this feature emerged at higher wavenumbers than the O-H···O=C peak, indicating that this interaction is weaker than the favoured O-H···O=C. This gives a possible explanation for the absence of this interaction in the crystalline phases, where stronger interactions are formed.

Further differences in the amorphous PCM model from crystalline structures are apparent in the phenyl O-H group interactions, where OH···OH interactions are present (Figure 6.20c). Although OH···OH interactions do not represent a majority of non-bonded contacts involving this functional group, these are absent from all of the single-component polymorphic crystal forms. In line with the known forms, there is an abundance of C=O and

N-H interactions with the OH group. There is also evidence of a small number of C=O…C=O dipole interactions (Figure 6.20a) which have been shown to be comparable in strength to medium-strength hydrogen bonds, and around half that of a strong hydrogen bond, such as C=O…H-O (Allen, Baalham, Lommerse, & Raithby, 1998). Although not observed in the crystalline PCM polymorphs, the NH···NH interaction (Figure 6.20b) also appears in the amorphous solid. While not a favoured interaction for PCM, a search of the CSD using ConQuest (Section 2.10) (Bruno et al., 2002b) reveals this is a relatively common interaction within small molecule crystal structures containing the acetyl or N-acetyl functional groups.

Around the N-H group, there is a narrower distribution of intermolecular contact sites compared with the OH or C=O groups. This is expected to be due to steric effects arising from the aromatic ring, effectively limiting the available space for PCM molecules to feasibly form non-bonded interactions. Crystal forms of PCM exhibit similar behaviour (Figure 6.22). This contrasts with the C=O and O-H groups, which lie at either end of the molecule, increasing the potential volume available for interactions to occur.

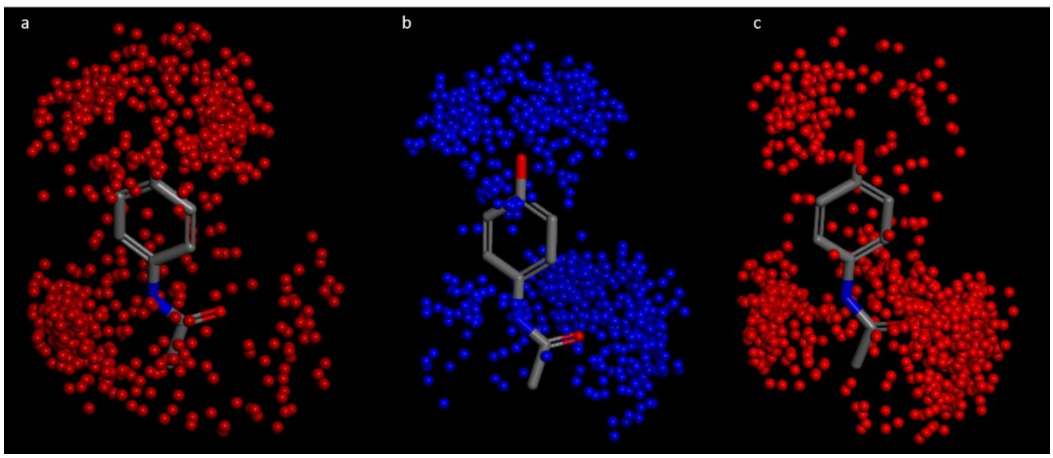

Figure 6.21: IsoStar plot for bulk amorphous PCM MD trajectory. Plot shows interactions from functional groups in PCM molecules, where (a) shows contact points corresponding to carbonyl C=O interactions, (b) shows points representing N-acetyl N-H interactions, and (C)

shows points representing phenyl O-H interactions. Cut-off distance for non-bonded contacts was 4Å.

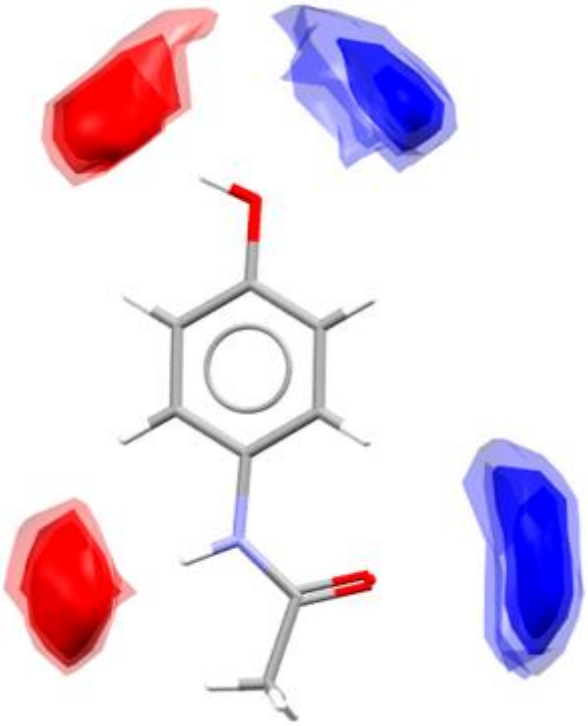

Figure 6.22: Full interaction maps (FIMs) calculated for functional groups for PCM conformer from form I. The smaller contour around the N-H group demonstrates the relatively small volume available for non-bonded interactions to occur due to steric effects from the aromatic ring.

The aromatic interactions also show a large distribution of molecular contacts (Figure 6.23). As aromatic ring centroid-centroid distances were probed in the ConQuest search, the non-bonded cut-off was increased to 5Å. The occurrence of such interactions suggests a significant role of π- π interactions in the structure of the amorphous solid. As with the other non-bonded contacts (Figure 6.20) there is a distribution on $\pi - \pi$ geometries, with parallel and parallel displaced dominating over other geometries such as edge-to-face. This again differs from the crystal forms (Figure 6.24), where a narrower distribution of potential π interaction sites are shown in keeping with the disordered structure.

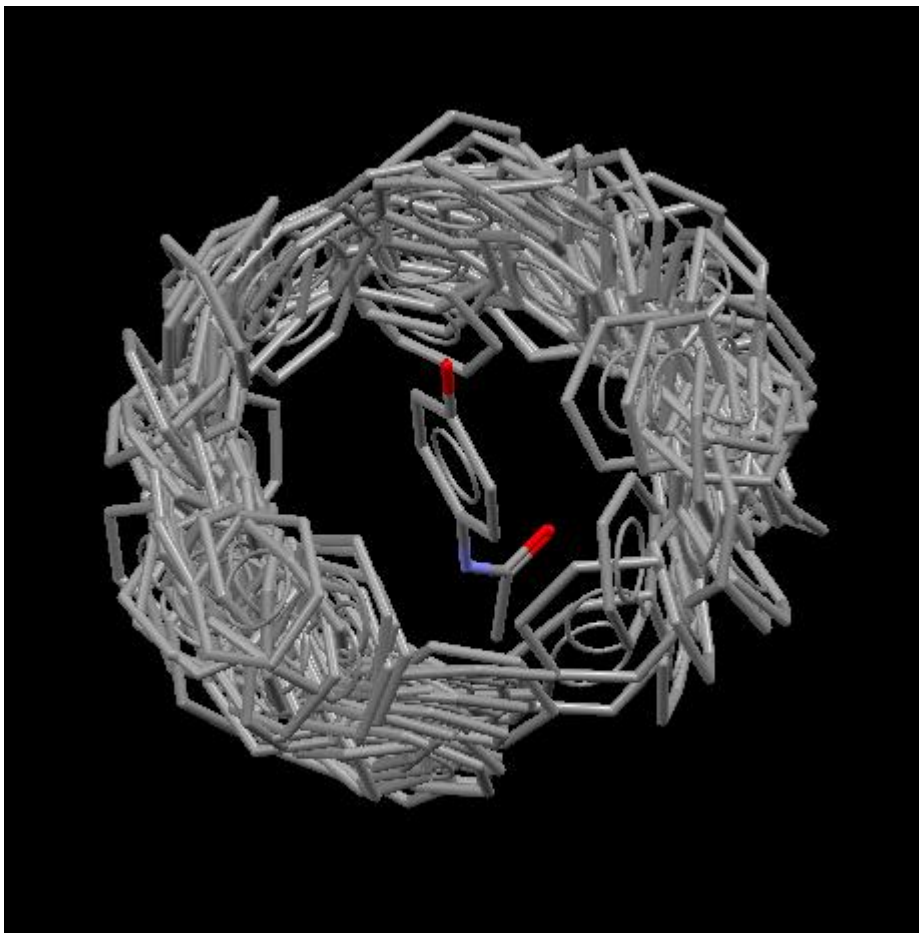

Figure 6.23: IsoStar plot for bulk amorphous PCM MD trajectory. Plot shows pi-pi interactions around the PCM aromatic ring. Cut-off distance for non-bonded contacts was 5 Å.

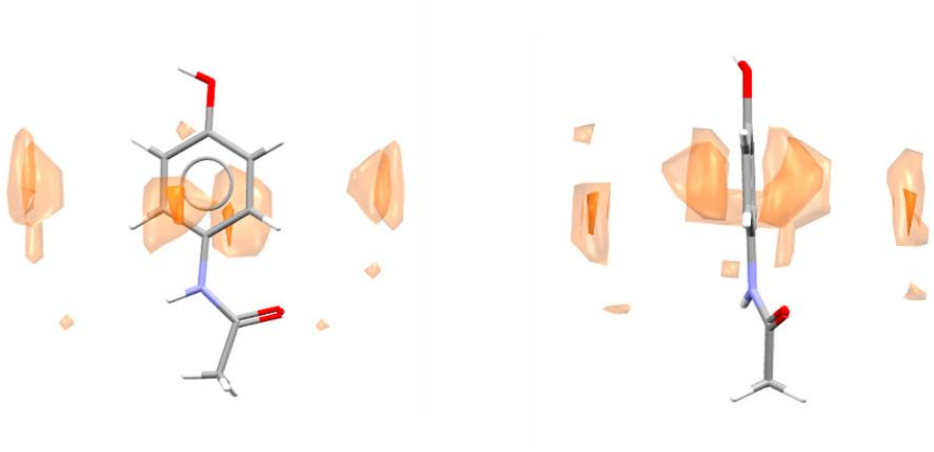

Figure 6.24: FIMs calculated for aromatic ring in PCM conformer from Form I. Contours correspond to aromatic carbons. Contours show a narrower distribution of potential interaction sites for π interactions to occur, contrasting with that observed for amorphous PCM. Two views of calculated FIMs shown for clarity.

6.3.10. Assessment of intermolecular interactions in bulk amorphous PCM MD box using Full Interaction Maps (FIMs)

FIMs (Section 2.11) calculated around sample interactions from the bulk amorphous PCM MD box are shown in Figure 6.25. The majority of interactions fall within the contours calculated around the central PCM molecules, demonstrating the feasibility of the MD-derived structure. The majority of the PCM molecules are coordinated via hydrogen bonds to four other PCM molecules as is observed in all but one of the known PCM polymorphs (Section 1). In the amorphous cell some variation from this is observed with a number of occurrences where the central molecule forms only three hydrogen bonds.

The images shown demonstrate that the majority of hydrogen bond motifs in amorphous PCM are observed in the crystalline forms. However as expected without the constraints of symmetry a wider distribution of geometries and types of intermolecular interactions are observed in the amorphous PCM model (Figure 6.26). For example, two cases are shown here of N-H···O=C hydrogen bonds and an OH···OH bond. Furthermore, where the structure deviates significantly from the crystalline behaviour, the hydrogen bond location sits out with the calculated contour. Deviation from the contour location demonstrates that the hydrogen bond is in a suboptimal location based on the data within the CSD. As this is a high-energy amorphous form, some deviation from the preferred locations may be expected. This is particularly evident at the N-H···O=C interactions. Structural deviations are also demonstrated in the central image in Figure 6.24, where the OH…OH interaction also sits out with the calculated contour and is not observed in the crystalline forms.

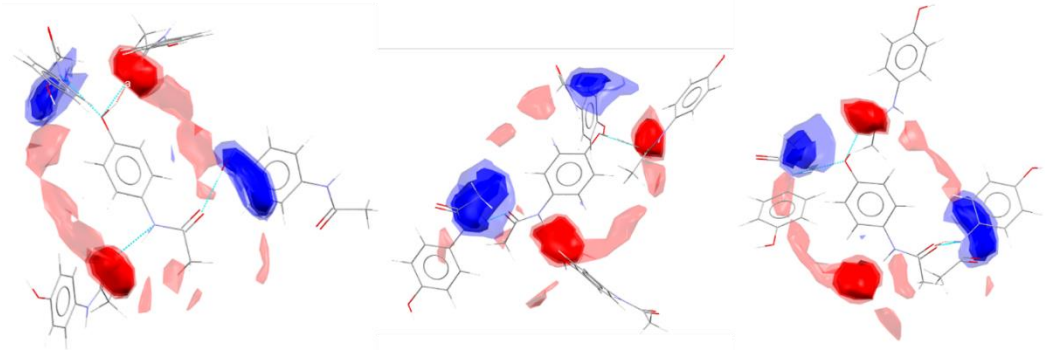

Figure 6.25: FIMs calculated around central PCM molecule for selected interactions found in bulk amorphous PCM MD box.

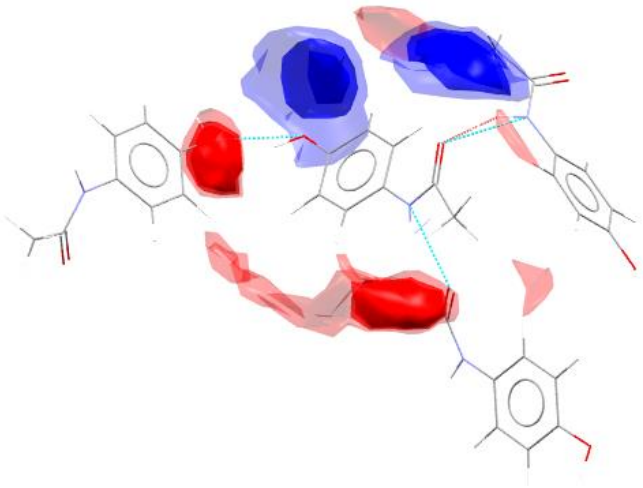

Figure 6.26: Example interaction from bulk amorphous PCM box. Hydrogen bonds are shown in turquoise. The hydrogen bonded motif shown is not observed in the crystalline forms, due to the single hydrogen bond formed by the OH group on the central PCM molecule. In addition, the OH···OH interaction shown is not present in crystalline forms.

6.3.11. Interfacial MD simulations

6.3.11.1. Role of molecular mobility in directing crystallisation

Figure 6.27 shows snapshots of the three MD simulation boxes described in section 2.5. DCs were calculated for each trajectory after equilibration (Figure 6.28; Section 2.5.5), to assess the impact of surface and bulk mobility on crystallisation in the amorphous solid.

All DC values are summarised in Table 6.3. From highest to lowest, the DCs calculated from the RT MD simulations are PCM-vac (1.783×10^{-7}cm/s), PCM at PCM-SiO_2 interface (1.208×10^{-8}cm/s), PCM bulk (9.167×10^{-9}cm/s) and SiO_2 at PCM-SiO_2 interface (3.750×10^{-10}cm/s). Notably, the DC for the PCM-vacuum box is by far the largest (1.78×10^{-7}cm/s), at two orders of magnitude greater than all other PCM values. This indicates significantly higher mobility of PCM molecules at the surface of the amorphous material in comparison to PCM molecules in the bulk or at the PCM-SiO_2 interface (1.97×10^{-9}cm/s and 1.21×10^{-8}cm/s, respectively). This suggests that interactions of PCM with SiO_2 therefore significantly reduce surface mobility. Mobility of PCM molecules at the SiO_2 interface is close to that determined for the bulk material.

Table 6.3: Summary of DC values calculated from interfacial MD trajectories.

MD box	Compound for DC calculation	DC at 25°C (cm/s)	DC at 60°C (cm/s)
PCM-SiO_2	SiO_2	3.750×10^{-10}	3.333×10^{-10}
PCM- SiO_2	PCM	1.208×10^{-8}	1.583×10^{-8}
PCM bulk	PCM	9.167×10^{-9}	5.625×10^{-8}
PCM-vacuum	PCM	1.783×10^{-7}	N/A

The kinetically driven crystallisation of Form II, which occurs when the amorphous solid is left exposed to the environment (Section 1), is hindered by the introduction of the SiO_2 interface, which significantly reduces surface mobility. In the bulk material, crystallisation may instead be driven thermodynamically (relative thermodynamic stability of PCM forms

observed here: III<II<I), as the kinetic drive for recrystallisation has been reduced by the SiO_2 at the surface. Recrystallisation here thus results in Form III before converting to the more thermodynamically favoured Form II as temperature is increased. These findings complement previous reports that Form II crystallisation begins at the surface of the material, whereas Form III emerges from the bulk (Nanubolu & Burley, 2012). The role of interfaces in stabilising the amorphous form has been previously reported (Palomäki, Lipiäinen, Strachan, & Yliruusi, 2020), where trehalose and melibiose were able to stabilise amorphous PCM. Thus, MD simulations offer some insight to the observed crystallisation behaviour when combined with the structural similarity between the amorphous phase and Form III determined by comparison of known crystal forms to the experimental amorphous PDF (Section 3.4).

A different trend is observed for the simulations run at 60°C. From highest to lowest, the DCs calculated from the MD simulations ran at 60°C are PCM bulk ($5.625x10^{-8}$cm/s), PCM at PCM-SiO_2 interface ($1.583x10^{-8}$cm/s) and SiO_2 at PCM-SiO_2 interface ($3.333x10^{-10}$cm/s). However, like the RT simulations, the SiO_2 DC ($3.333x10^{-10}$cm/s) is the smallest for all high temperature calculations. Additionally, the PCM DC at the SiO_2 boundary ($1.583x10^{-8}$cm/s) also remains relatively small. It appears that even at high temperatures, where PCM crystallisation to Form III is observed, SiO_2 is able to hinder molecular mobility of PCM at the glass-glass interface. At this interface, the DC is comparable to the RT simulations. At 60°C, the mobility of the bulk material is almost an order of magnitude greater than the all other values at this temperature ($5.625x10^{-8}$cm/s). Mobility of the bulk material is still significantly lower than the surface at RT. While this demonstrates the significant mobility on the surface on the sample leading to Form II crystallisation, it is apparent that bulk mobility is sufficient to enable reorganisation to another metastable state. Given the structural similarity between

the amorphous phase and Form III (Section 3.4), the emergence of Form III is favoured, as it represents the lowest change in free energy within the system.

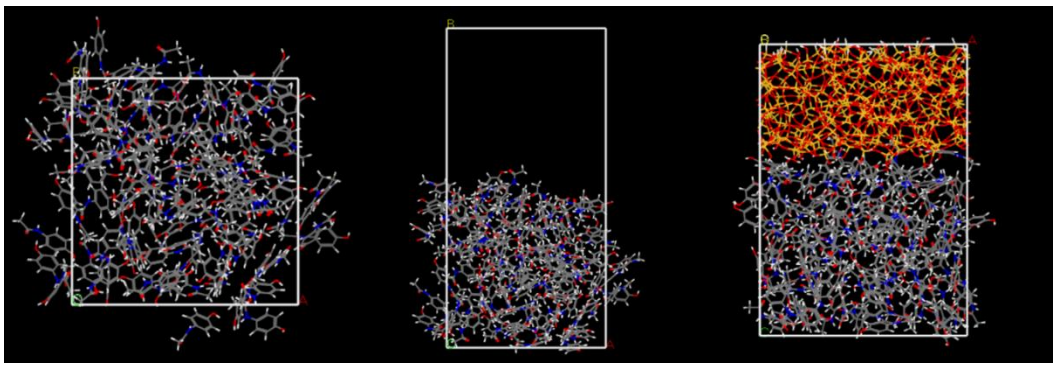

Figure 6.27: Snapshots of MD boxes. From left to right: bulk amorphous PCM, PCM-vacuum interface box, PCM-SiO_2 interface box.

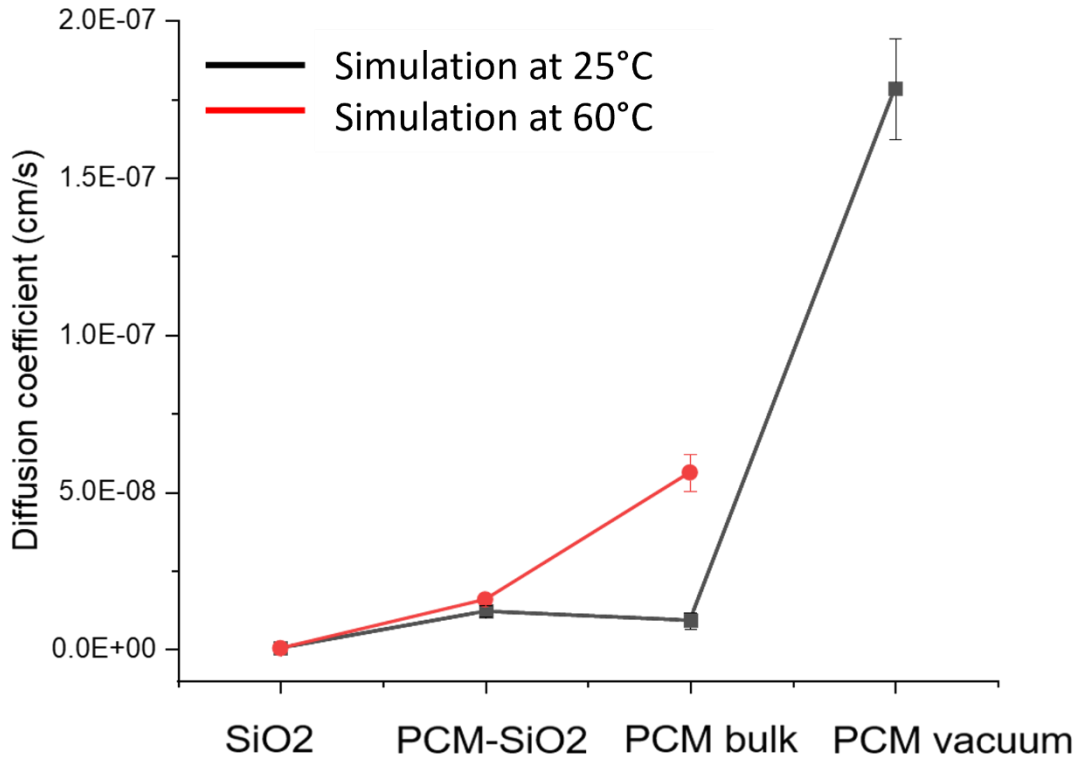

Figure 6.28: Diffusion coefficients for each MD box. Error bars represent 1 standard deviation of the mean DC value. Black points correspond to DCs for simulations ran at 25°C, red points correspond to simulations ran at 60°C.

6.4. Summary

The structure of amorphous PCM has been studied using PDF analysis and MD simulations. Statistical analysis of the PDF data (Sections 3.5 and 3.6) allowed subtle changes in the local structure of amorphous PCM to be identified as a function of temperature. Clustering in PCA and an AUC approach aligned with reported secondary relaxations in the material (T_g and β-relaxation) (Sibik et al., 2015) and crystallisation events. These also aligned with visual changes in the material (Figure 6.17) as shown in the images taken of the sample on the beam line. These results demonstrate the sensitivity of PDF analysis in detecting subtle changes in the local structure of amorphous molecular solids.

Use of MD simulations has enabled the first thermodynamically feasible structure for amorphous PCM to be proposed. Analysis of the intra- and intermolecular structure provided an understanding of the distribution of conformations and intermolecular interactions that accurately accounts for the experimental amorphous PDF (Section 3.7). By fitting the structures of known crystalline forms of PCM it was also demonstrated that structure in amorphous PCM has the closest similarity to Form III (Section 3.4).

Structure in bulk amorphous PCM was found to contain similar hydrogen bonding patterns to all known forms (Section 3.10). In addition, some new hydrogen bonded motifs were observed for example an N-H…O=C interaction, which has previously been described in a spectroscopic study of amorphous PCM (Trasi & Taylor, 2012), and OH···OH hydrogen bonds were also observed. Analysis of the torsion distributions in the amorphous phase (Section 3.9.2) revealed no preference or similarity to the crystalline PCM forms.

MD simulations were also used to provide insight to the preferred crystallisation of PCM Form II when the amorphous solid is left exposed to the environment. Interfacial MD simulations revealed significantly increased mobility of surface molecules in the amorphous

solid (Section 3.11.1), where the DC calculated for the PCM-vacuum box was two orders of magnitude higher than those calculated for the bulk and PCM-SiO_2 boxes. These results are consistent with previous reports of Form II crystallisation occurring on the surface of the material, whereas Form III emerges from the bulk (Nanubolu & Burley, 2012). MD simulations at 60 °C indicate that the bulk DC increases drastically at high temperature. However, this is still an order of magnitude smaller than the RT surface DC, reflecting the high mobility of surface molecules and their role in directing crystallisation to Form II. Thus, due to the SiO_2 interface minimising the kinetic drive for Form II crystallisation, phase transitions in the sample were directed by the thermodynamics of the system and follow Ostwald's rule of stages (Ostwald, 1897).

The sensitivity of PDF analysis in detecting subtle changes in local ordering as a function of temperature has been demonstrated. Subsequent analysis of the PDF, supported by MD simulations, gave insight into structure and transformations in the amorphous solid, and offered an explanation for the confinement-dependent crystallisation behaviour of PCM. It was also shown that molecular mobility on the surface of the amorphous material may play a role in directing the kinetic crystallisation of Form II. The findings from this may be used to inform formulation development of amorphous solids and can assist in targeting metastable phases or stabilising the amorphous form. Thus, the value in employing PDF analysis to understanding the interrelationships between structure and function in amorphous pharmaceuticals is demonstrated.

Chapter 7: Conclusions and future research

7.1. Conclusions

The work presented herein addresses practical approaches for the reliable collection, treatment and interpretation of PDF data from amorphous pharmaceutical solids. A set of recommendations are presented in Chapter 4 for laboratory PDF data collection with approaches for sample handling, data collection procedure and data treatment discussed. These approaches are applied to an amorphous API (AZD5718) in Chapter 5 to understand the impact of manufacturing method on local structure. Further investigation couple experimental PDFs with calculated PDFs from MD simulations to access thermodynamically feasible structures. Finally, approaches for analysis of high-throughput synchrotron PDF data are presented using amorphous paracetamol as a case study. The value of MD simulations coupled with PDF data is again demonstrated. This technique pairing enabled structural understanding and insight into the crystallisation behaviour of the amorphous solid.

The implementation of a laboratory diffractometer for in-house PDF studies of amorphous pharmaceuticals presents a number of challenges. Specifically, the poor scattering power of most amorphous pharmaceutical materials, a consequence of low atomic number components and the absence of constructive interference in the scattering experiment, is a significant hurdle for delivering high-quality data suitable for PDF analysis. The use of appropriate data collection and treatment procedures, as demonstrated herein, can assist in overcoming these challenges. There are a range of available software packages for calculation of the PDF from experimental data. In this work, GudrunX and xPDFsuite were implemented and each has been shown to produce reliable PDF data. As discussed in Chapter 4, whilst xPDFsuite is a useful tool for rapid calculation of the PDF requiring little user input, it does not allow the same extent of customization of data treatment and correction available in other packages such as GudrunX. For example, the automated data correction procedures

can result in a reduced structure function, $F(Q)$, that is not properly normalised. As a consequence given PDFs are not on an absolute scale, using xPDFsuite, the comparison of PDFs generated from data collected using two different instruments requires scaling of PDF data to allow direct comparison. This cannot be done with the current version of xPDFsuite. GudrunX allows this direct comparison due to normalisation of the scattering data to the differential scattering cross section by normalising the experimental intensities to the source flux. Also, the implementation of the Soper-Lorch function in GudrunX, for minimising the impact of noise and termination effects on $G(r)$, provides an additional data treatment procedure when those discussed in Chapter 4 are insufficient for obtaining high-quality PDF.

It has been demonstrated that as long as long data collection times are implemented, laboratory PDF can assist in enabling structural understanding of amorphous pharmaceuticals. While laboratory PDF collection has been previously applied to provide a fingerprint of a given material, structural changes can also be detected as a function of stress, storage or processing. Challenges remain for non-ambient data collection. For example, in variable temperature experiments, the large sample volumes can make reliable temperature control difficult using capillary and Cryostream setups. In such instances or where high-throughput PDF data is required, it is often desirable or even necessary to seek access to synchrotron sources for the higher flux and shorter collection times. The laboratory PDF data shown here have been demonstrated to be comparable to that obtained at a synchrotron in terms of the useful representation of amorphous structure.

Molecular dynamics simulations coupled with PDF have been found to be invaluable in enabling structural understanding of amorphous solids in this work. The Materials Studio software package provides a series of tools for generation of large boxes of molecules (Amorphous Cell), geometry optimisation and dynamics (Forcite), providing

thermodynamically feasible trajectories of molecules. The COMPASS II force field represents the state-of-the-art in MD simulations for organic systems due to extensive parameterisation on representative functional groups. It is thus recommended to implement this force field if available. The OPLS-aa force field applied in Chapter 5 is a popular alternative to COMPASS II and has been found to produce reliable MD simulations for a range of materials (Fluitt & De Pablo, 2015; Moultos, Tsimpanogiannis, Panagiotopoulos, Trusler, & Economou, 2016; Robertson, Qian, Robinson, Tirado-Rives, & Jorgensen, 2019). Hence, selection of a specific force field is often governed by availability and access to a given MD package. Beyond the force field itself, there are a number of other considerations for obtaining reliable dynamics. For example, the use of the Ewald summation method for calculation of long-range interactions and the Nosé thermostat for temperature control have been recommended (Chapter 1, Section 1.9.2.4) as suitable methods for MD calculations. Where either of these methods have been applied, the resulting intra- and intermolecular geometries have been verified using the tools within the CSD. As the use of the Berendsen barostat has been found to produce unreliable dynamics, the Berendsen barostat was only applied to equilibrate the density in this work before switching to an ensemble with no pressure control. Thus, the dynamics are not expected to be adversely affected by the use of this barostat (Chapter 1; Section 1.9.2.5).

Structural analysis of large boxes of simulated molecules necessary to describe an amorphous material also represents a challenge to visualisation and interpretation of structural features. By contrast, in small-molecule crystallography, the structure can be explained by a small unit cell repeating in three dimensions. In this work, a simplification of the large MD structure was generated using the tools within the CSD. Specifically, the CSD Editor, Isogen and IsoStar, were used to create 3D spatial distribution plots where the average locations of non-bonded interactions can be observed. The implementation of these

methods for MD simulations has enabled the first proposed structure of amorphous PCM. Given the hurdles in applying and interpreting PDF data from organic molecular solids, this proposed structure represents a significant step in advancing the use of PDF within the pharmaceutical materials community.

The methods applied here offer useful techniques for extracting information from PDF data and demonstrate that application of PDF to the materials discussed here is not trivial and can be greatly improved using the methodology recommendations discussed. A number of challenges and opportunities for further progress remain in this area. Advances in computing methods and power as well as data and structural analysis methodologies will help to further the applicability of PDF to an important class of functional materials. These advances will allow the application beyond the single-component systems discussed here to formulated systems for example.

7.2. Future research

A number of opportunities remain in the application of PDF to amorphous pharmaceutical solids, where structural understanding can be enabled through high quality PDF data coupled with MD simulations (Anderson & Xiang, 2016; Munjal & Suryanarayanan, 2021). There are various methods which may be used to enhance the value of PDF data in pharmaceutical development, some of which are discussed below.

While the use of a lab instrument for reliable PDF data has been demonstrated, further optimisations of data collection and treatment protocols could enable the use of a silver source for PDF analysis of poorly scattering molecular solids, enabling higher resolution PDF. Advances in detector technology enabling higher sensitivity to low-wavelength X-rays, and lab sources with higher flux may allow the implementation of such sources for PDF analysis.

The use of a silver source has the potential to increase sensitivity to the subtle structural changes indicative of phase separation in formulated drug-polymer systems during stability studies. PDFs generated from these data offer the potential to enable better prediction of stability as well as detecting small structural changes resulting from process history. Alternatively, the current data collection procedure may be tailored to enable high-throughput PDF data in the laboratory, allowing rapid fingerprinting of amorphous APIs to accelerate development programs. This could be achieved via a number of methods, for example by using a larger diameter capillary or by increasing the step-size in the scattering experiment to allow longer counting times.

The value of MD simulations coupled with PDF for providing structural insights to amorphous single-component systems could be extended to formulated drug-polymer or co-amorphous drug-drug systems. Given the large volume of research in understanding stability and solubility in these materials, coupling MD simulations and PDF may assist in identification of stable formulations, with the experimental PDF data used to validate structural models calculated from the MD simulations. This approach would lead to the potential for greater predictive design in amorphous formulations and therefore minimise the requirement for time- and material- consuming experimental efforts.

Developing robust approaches for probing amorphous structure will benefit the from the use of multiple techniques. Techniques which may offer complementary information to PDF include small-angle X-ray scattering (SAXS) (Laggner & Paudel, 2018) and time of flight secondary ion mass spectrometry (ToF-SIMS) (Iuraş et al., 2016). These techniques can allow tracking and correlation of structural changes from atomic (PDF, Å to nm scale) to nano (SAXS, up to few hundred nm scale) to micro (ToF-SIMS, up to few hundred μm) scales, allowing complete analysis across the entire system. As the value of other complementary

techniques (e.g. FTIR, DSC, TGA) has been demonstrated, additional analysis tools such as SAXS and ToF-SIMS may enable a thorough understanding of structure and properties of these industrially important systems.

PDF has been shown to provide useful insights to structure and transformations in amorphous materials. The results have demonstrated the value of lab PDF in development of amorphous pharmaceutical solids and emphasised the potential of MD simulations in providing complementary structural insights. With the ability to apply PDF analysis in the laboratory allowing routine application, future developments as discussed may enable the routine determination of amorphous pharmaceutical structure for the first time enabling structure property relationships to be explored, feeding a new era of rational design of amorphous systems.

References

Akkermans, R. L. C., Spenley, N. A., & Robertson, S. H. (2013). Monte Carlo methods in Materials Studio. *Molecular Simulation*, *39*(14–15), 1153–1164. https://doi.org/10.1080/08927022.2013.843775

Al-Madhagi, L. H., Callear, S. K., & Schroeder, S. L. M. (2020). Hydrophilic and hydrophobic interactions in concentrated aqueous imidazole solutions: a neutron diffraction and total X-ray scattering study. *Physical Chemistry Chemical Physics*. https://doi.org/10.1039/C9CP05993H

Allen, F. H., Baalham, C. A., Lommerse, J. P. M., & Raithby, P. R. (1998). Carbonyl-Carbonyl Interactions can be Competitive with Hydrogen Bonds. *Acta Crystallographica Section B: Structural Science*, *54*(3), 320–329. https://doi.org/10.1107/S0108768198001463

Altomare, A., Ciriaco, F., Cuocci, C., Falcicchio, A., & Fanelli, F. (2017). Combined powder X-ray diffraction data and quantum-chemical calculations in EXPO2014. *Powder Diffraction*, *32*(S1), S123–S128. https://doi.org/10.1017/S088571561700015X

Anderson, B. D., & Xiang, T.-X. (2016). Molecular Dynamics Simulations of Amorphous Systems. In *Computational Pharmaceutical Solid State Chemistry* (pp. 331–373). Hoboken, NJ: John Wiley & Sons, Inc. https://doi.org/10.1002/9781118700686.ch13

Aoun, B. (2016). Fullrmc, a rigid body reverse monte carlo modeling package enabled with machine learning and artificial intelligence. *Journal of Computational Chemistry*, *37*(12), 1102–1111. https://doi.org/10.1002/jcc.24304

Authelin, J. R., Mackenzie, A. P., Rasmussen, D. H., & Shalaev, E. Y. (2014, September 1). Water clusters in amorphous pharmaceuticals. *Journal of Pharmaceutical Sciences*. John Wiley and Sons Inc. https://doi.org/10.1002/jps.24009

Ball, P. (2008, March 20). Water: Water - An enduring mystery. *Nature*. Nature Publishing Group. https://doi.org/10.1038/452291a

Barbas, R., Kumar, V., Vallcorba, O., Prohens, R., & Frontera, A. (2020). Sildenafil–Resorcinol Cocrystal: XRPD Structure and DFT Calculations. *Crystals*, *10*(12), 1126. https://doi.org/10.3390/cryst10121126

Barr, G., Dong, W., & Gilmore, C. J. (2004a). High-throughput powder diffraction. II. Applications of clustering methods and multivariate data analysis. *Journal of Applied Crystallography*, *37*(2), 243–252. https://doi.org/10.1107/S0021889804000391

Barr, G., Dong, W., & Gilmore, C. J. (2004b). PolySNAP: A computer program for analysing high-throughput powder diffraction data. *Journal of Applied Crystallography*, *37*(4), 658–664. https://doi.org/10.1107/S0021889804011173

Barr, G., Dong, W., Gilmore, C. J., Parkin, A., & Wilson, C. C. (2005). dSNAP: A computer program to cluster and classify Cambridge Structural Database searches. *Journal of Applied Crystallography*, *38*(5), 833–841. https://doi.org/10.1107/S0021889805021308

Bell, J. L., Sarin, P., Driemeyer, P. E., Haggerty, R. P., Chupas, P. J., & Kriven, W. M. (2008). X-Ray pair distribution function analysis of a metakaolin-based, KAlSi 2O6·5.5H2O

inorganic polymer (geopolymer). *Journal of Materials Chemistry*, *18*(48), 5974–5981. https://doi.org/10.1039/b808157c

Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, *81*(8), 3684–3690. https://doi.org/10.1063/1.448118

Bernstein, J., & Reutzel-Edens, S. M. (2019). Powder diffraction and pharmaceuticals (pp. 767–781). https://doi.org/10.1107/97809553602060000979

Berthier, L., & Ediger, M. D. (2016). Facets of glass physics. *Physics Today*, *69*(1), 41–46. https://doi.org/10.1063/PT.3.3052

Beyer, A., Radi, L., Grohganz, H., Löbmann, K., Rades, T., & Leopold, C. S. (2016). Preparation and recrystallization behavior of spray-dried co-amorphous naproxen–indomethacin. *European Journal of Pharmaceutics and Biopharmaceutics*, *104*, 72–81. https://doi.org/10.1016/j.ejpb.2016.04.019

Bhugra, C., Shmeis, R., & Pikal, M. J. (2008). Role of mechanical stress in crystallization and relaxation behavior of amorphous indomethacin. *Journal of Pharmaceutical Sciences*, *97*(10), 4446–4458. https://doi.org/10.1002/jps.21291

Billinge, S. J. (2015). Atomic pair distribution function: a revolution in the characterization of nanostructured pharmaceuticals. *Nanomedicine*, *10*(16), 2473–2475. https://doi.org/10.2217/nnm.15.116

Billinge, S. J. L. (1998). Real-space Rietveld: Full profile structural refinement of the atomic pair distribution function. *Plenum, New York*, 137.

Billinge, S. J. L. (2011). Total Scattering Pair Distribution Functions (TSPDF) for Fingerprinting Amorphous Pharmaceuticals. *Pharmaceutical Powder X-Ray Diffraction Symposium (PPXRD)*.

Billinge, S. J. L., Dykhne, T., Juhás, P., Boin, E., Taylor, R., Florence, A. J., & Shankland, K. (2010). Characterisation of amorphous and nanocrystalline molecular materials by total scattering. *CrystEngComm*, *12*(5), 1366–1368. https://doi.org/10.1039/b915453a

Billinge, S. J. L., & Kanatzidis, M. G. (2004). Beyond crystallography: The study of disorder, nanocrystallinity and crystallographically challenged materials with pair distribution functions. *Chemical Communications*, *4*(7), 749–760. https://doi.org/10.1039/b309577k

Biswas, R. K., Khan, P., Mukherjee, S., Mukhopadhyay, A. K., Ghosh, J., & Muraleedharan, K. (2018). Study of short range structure of amorphous Silica from PDF using Ag radiation in laboratory XRD system, RAMAN and NEXAFS. *Journal of Non-Crystalline Solids*, *488*, 1–9. https://doi.org/10.1016/j.jnoncrysol.2018.02.037

Bøtker, J. P., Karmwar, P., Strachan, C. J., Cornett, C., Tian, F., Zujovic, Z., … Rades, T. (2011). Assessment of crystalline disorder in cryo-milled samples of indomethacin using atomic pair-wise distribution functions. *International Journal of Pharmaceutics*, *417*(1–2), 112–119. https://doi.org/10.1016/j.ijpharm.2010.12.018

Boultif, A., & Louër, D. (2004). Powder pattern indexing with the dichotomy method. *Journal of Applied Crystallography*, *37*(5), 724–731. https://doi.org/10.1107/S0021889804014876

Bragg, W. H., & Bragg, W. L. (1913). The Reflection of X-rays by Crystals. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *88*(605), 428–438. https://doi.org/10.1098/rspa.1913.0040

Braun, E., Gilmer, J., Mayes, H. B., Mobley, D. L., Monroe, J. I., Prasad, S., & Zuckerman, D. M. (2019). Best Practices for Foundations in Molecular Simulations. *Living Journal of Computational Molecular Science*, *1*(1), 5957. https://doi.org/10.33011/livecoms.1.1.5957

Braun, E., Moosavi, S. M., & Smit, B. (2018). Anomalous Effects of Velocity Rescaling Algorithms: The Flying Ice Cube Effect Revisited. *Journal of Chemical Theory and Computation*, *14*(10), 5262–5272. https://doi.org/10.1021/acs.jctc.8b00446

Brittain, H. G. (2016). *Polymorphism in Pharmaceutical Solids: Second edition*. *Polymorphism in Pharmaceutical Solids: Second Edition*. Informa Healthcare. https://doi.org/10.3109/9781420073225

Bruno, I. J., Cole, J. C., Edgington, P. R., Kessler, M., Macrae, C. F., McCabe, P., … Taylor, R. (2002a). New software for searching the Cambridge Structural Database and visualizing crystal structures. *Acta Crystallographica Section B: Structural Science*, *58*(3 PART 1), 389–397. https://doi.org/10.1107/S0108768102003324

Bruno, I. J., Cole, J. C., Edgington, P. R., Kessler, M., Macrae, C. F., McCabe, P., … Taylor, R. (2002b). New software for searching the Cambridge Structural Database and visualizing crystal structures. *Acta Crystallographica Section B: Structural Science*, *58*(3 PART 1), 389–397. https://doi.org/10.1107/S0108768102003324

Bruno, I. J., Cole, J. C., Kessler, M., Luo, J., Momerwell, W. D. S., Purkis, L. H., … Orpen, A. G. (2004). Retrieval of crystallographically-derived molecular geometry information. *Journal of Chemical Information and Computer Sciences*, *44*(6), 2133–2144. https://doi.org/10.1021/ci049780b

Bruno, I. J., Cole, J. C., Lommerse, J. P. M., Rowland, R. S., Taylor, R., & Verdonk, M. L. (1997). IsoStar: A library of information about nonbonded interactions. *Journal of Computer-Aided Molecular Design*, *11*(6), 525–537. https://doi.org/10.1023/A:1007934413448

Burley, J. C., Duer, M. J., Stein, R. S., & Vrcelj, R. M. (2007). Enforcing Ostwald's rule of stages: Isolation of paracetamol forms III and II. *European Journal of Pharmaceutical Sciences*, *31*(5), 271–276. https://doi.org/10.1016/j.ejps.2007.04.002

Cervellino, A., & Frison, R. (2020). Texture corrections for total scattering functions. *Acta Crystallographica Section A Foundations and Advances*, *76*(3), 302–317. https://doi.org/10.1107/s2053273320002521

Chapman, K. W., Lapidus, S. H., & Chupas, P. J. (2015). Applications of principal component analysis to pair distribution function data. *Journal of Applied Crystallography*, *48*(6), 1619–1626. https://doi.org/10.1107/S1600576715016532

Chen, S., Sheikh, A. Y., & Ho, R. (2014). Evaluation of Effects of Pharmaceutical Processing on Structural Disorders of Active Pharmaceutical Ingredient Crystals Using Nanoindentation and High-Resolution Total Scattering Pair Distribution Function Analysis. *Journal of Pharmaceutical Sciences*, *103*(12), 3879–3890. https://doi.org/10.1002/jps.24178

Chen, Y., Wang, S., Wang, S., Liu, C., Su, C., Hageman, M., … Qian, F. (2016). Initial Drug Dissolution from Amorphous Solid Dispersions Controlled by Polymer Dissolution and Drug-Polymer Interaction. *Pharmaceutical Research*, *33*(10), 2445–2458. https://doi.org/10.1007/s11095-016-1969-2

Chieng, N., Trnka, H., Boetker, J., Pikal, M., Rantanen, J., & Grohganz, H. (2013). Detecting phase separation of freeze-dried binary amorphous systems using pair-wise distribution function and multivariate data analysis. *International Journal of Pharmaceutics*, *454*(1), 167–173. https://doi.org/10.1016/J.IJPHARM.2013.07.030

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

Clegg, W., Blake, A. J., Cole, J. M., Evans, J. S. O., Main, P., Parsons, S., & Watkin, D. J. (2009). *Crystal Structure Analysis: Principles and Practice*. *Crystal Structure Analysis: Principles and Practice* (Vol. 9780199219469). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199219469.001.0001

Cline, J. P., Mendenhall, M. H., Black, D., Windover, D., & Henins, A. (2015). The Optics and alignment of the divergent beam laboratory x-ray powder diffractometer and its calibration using nist standard reference materials. *Journal of Research of the National Institute of Standards and Technology*, *120*, 173–222. https://doi.org/10.6028/jres.120.013

Coelho, A. A. (2003). Indexing of powder diffraction patterns by iterative use of singular value decomposition. *Journal of Applied Crystallography*, *36*(1), 86–95. https://doi.org/10.1107/S0021889802019878

Coelho, A. A., Chater, P. A., & Kern, A. (2015). Fast synthesis and refinement of the atomic pair distribution function. *Journal of Applied Crystallography*, *48*(3), 869–875. https://doi.org/10.1107/S1600576715007487

Coelho, Alan A. (2018). TOPAS and TOPAS-Academic: An optimization program integrating computer algebra and crystallographic objects written in C++: An. *Journal of Applied Crystallography*, *51*(1), 210–218. https://doi.org/10.1107/S1600576718000183

Couillaud, B. M., Espeau, P., Mignet, N., & Corvis, Y. (2019). State of the Art of Pharmaceutical Solid Forms: from Crystal Property Issues to Nanocrystals Formulation. *ChemMedChem*, *14*(1), 8–23. https://doi.org/10.1002/cmdc.201800612

Craig, D. Q. M., Royall, P. G., Kett, V. L., & Hopton, M. L. (1999). The relevance of the amorphous state to pharmaceutical dosage forms: Glassy drugs and freeze dried systems. In *International Journal of Pharmaceutics* (Vol. 179, pp. 179–207). Elsevier. https://doi.org/10.1016/S0378-5173(98)00338-X

David, W. I. F., Shankland, K., McCusker, L. B., & Baerlocher, C. (2010a). *Structure Determination from Powder Diffraction Data*. (W.I.F. David, K. Shankland, L. B. McCusker, & C. Bärlocher, Eds.), *Structure Determination from Powder Diffraction Data* (Vol. 9780199205). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199205530.001.0001

David, William I.F., Shankland, K., McCusker, L. B., & Baerlocher, C. (2010b). *Structure Determination from Powder Diffraction Data*. *Structure Determination from Powder Diffraction Data* (Vol. 9780199205). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199205530.001.0001

David, William I.F., Shankland, K., Van De Streek, J., Pidcock, E., Motherwell, W. D. S., & Cole, J. C. (2006). DASH: A program for crystal structure determination from powder diffraction data. *Journal of Applied Crystallography*, *39*(6), 910–915. https://doi.org/10.1107/S0021889806042117

De Oliveira, D. N., De Menezes, M., & Catharino, R. R. (2015). Thermal degradation of sucralose: A combination of analytical methods to determine stability and chlorinated byproducts. *Scientific Reports*, *5*(1), 1–5. https://doi.org/10.1038/srep09598

Di Martino, P., Conflant, P., Drache, M., Huvenne, J. P., & Guyot-Hermann, A. M. (1997). Preparation and physical characterization of forms II and III of paracetamol. *Journal of Thermal Analysis*, *48*(3), 447–458. https://doi.org/10.1007/bf01979491

Dinnebier, R. E. (2013). Rigid bodies in powder diffraction. A practical guide. *Powder Diffraction*. https://doi.org/10.1017/S0885715600010356

Dinnebier, R. E., & Billinge, S. J. L. (2008). Powder diffraction: Theory and practice. *Theory and Practice*, 2. https://doi.org/10.1039/9781847558237

Dinnebier, R. E., Leineweber, A., & Evans, J. S. O. (2019). *Rietveld Refinement*. *Rietveld Refinement*. De Gruyter. https://doi.org/10.1515/9783110461381

Dollase, W. A. (1986). CORRECTION OF INTENSITIES OF PREFERRED ORIENTATION IN POWDER DIFFRACTOMETRY: APPLICATION OF THE MARCH MODEL. *Journal of Applied Crystallography*, *19*(pt 4), 267–272. https://doi.org/10.1107/S0021889886089458

Dykhne, T., Taylor, R., Florence, A., & Billinge, S. J. L. (2011). Data Requirements for the Reliable Use of Atomic Pair Distribution Functions in Amorphous Pharmaceutical Fingerprinting. *Pharmaceutical Research*, *28*(5), 1041–1048. https://doi.org/10.1007/s11095-010-0350-0

Ericsson, H., Nelander, K., Lagerstrom-Fermer, M., Balendran, C., Bhat, M., Chialda, L., … Skrtic, S. (2018). Initial Clinical Experience with AZD5718, a Novel Once Daily Oral 5-Lipoxygenase Activating Protein Inhibitor. *Clinical and Translational Science*, *11*(3), 330–338. https://doi.org/10.1111/cts.12546

Farrow, C. L., Juhas, P., Liu, J. W., Bryndin, D., Božin, E. S., Bloch, J., … Billinge, S. J. L. (2007). PDFfit2 and PDFgui: computer programs for studying nanostructure in crystals. *Journal of Physics: Condensed Matter*, *19*(33), 335219. https://doi.org/10.1088/0953-8984/19/33/335219

Fini, A., Cavallari, C., & Ospitali, F. (2008). Raman and thermal analysis of indomethacin/PVP solid dispersion enteric microparticles. *European Journal of Pharmaceutics and Biopharmaceutics*, *70*(1), 409–420. https://doi.org/10.1016/j.ejpb.2008.03.016

Finney, J. L., Hallbrucker, A., Kohl, I., Soper, A. K., Soper, A. K., Bowron, D. T., & Bowron, D. T. (2002). Structures of high and low density amorphous ice by neutron diffraction. *Physical Review Letters*, *88*(22), 225503/1-225503/4. https://doi.org/10.1103/PhysRevLett.88.225503

Florence, A. J., Bedford, C. T., Fabbiani, F. P. A., Shankland, K., Gelbrich, T., Hursthouse, M. B., … Fernandes, P. (2008). Two-dimensional similarity between forms I and II of cytenamide, a carbamazepine analogue. *CrystEngComm*, *10*(7), 811–813. https://doi.org/10.1039/b719717a

Florence, Alastair J., Baumgartner, B., Weston, C., Shankland, N., Kennedy, A. R., Shankland, K., & David, W. I. F. (2003). Indexing Powder Patterns in Physical Form Screening: Instrumentation and Data Quality. *Journal of Pharmaceutical Sciences*, *92*(9), 1930–1938. https://doi.org/10.1002/JPS.10459

Florence, Alastair J., Shankland, N., Shankland, K., David, W. I. F., Pidcock, E., Xu, X., … Frampton, C. S. (2005). Solving molecular crystal structures from laboratory X-ray powder diffraction data with DASH: The state of the art and challenges. *Journal of Applied Crystallography*, *38*(2), 249–259. https://doi.org/10.1107/S0021889804032662

Fluitt, A. M., & De Pablo, J. J. (2015). An Analysis of Biomolecular Force Fields for Simulations of Polyglutamine in Solution. *Biophysical Journal*, *109*(5), 1009–1018. https://doi.org/10.1016/j.bpj.2015.07.018

Gaisford, S., Buanz, A. B. M., & Jethwa, N. (2010). Characterisation of paracetamol form III with rapid-heating DSC. *Journal of Pharmaceutical and Biomedical Analysis*, *53*(3), 366–370. https://doi.org/10.1016/j.jpba.2010.04.017

Geddes, H. S., Blade, H., McCabe, J. F., Hughes, L. P., & Goodwin, A. L. (2019). Structural characterisation of amorphous solid dispersions: Via metropolis matrix factorisation of pair distribution function data. *Chemical Communications*, *55*(89), 13346–13349. https://doi.org/10.1039/c9cc06753a

Groom, C. R., Bruno, I. J., Lightfoot, M. P., & Ward, S. C. (2016). The Cambridge structural database. *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials*, *72*(2), 171–179. https://doi.org/10.1107/S2052520616003954

Groom, C. R., Bruno, I. J., Lightfoot, M. P., Ward, S. C., & IUCr. (2016). The Cambridge Structural Database. *Acta Crystallographica Section B Structural Science, Crystal Engineering and Materials*, *72*(2), 171–179. https://doi.org/10.1107/S2052520616003954

Gupta, J., Nunes, C., & Jonnalagadda, S. (2013). A molecular dynamics approach for predicting the glass transition temperature and plasticization effect in amorphous pharmaceuticals. *Molecular Pharmaceutics*, *10*(11), 4136–4145. https://doi.org/10.1021/mp400118v

Gupta, J., Nunes, C., Vyas, S., & Jonnalagadda, S. (2011). Prediction of solubility parameters and miscibility of pharmaceutical compounds by molecular dynamics simulations. *Journal of Physical Chemistry B*, *115*(9), 2014–2023. https://doi.org/10.1021/jp108540n

Halgren, T. A. (1992). Representation of van der Waals (vdW) Interactions in Molecular Mechanics Force Fields: Potential Form, Combination Rules, and vdW Parameters. *Journal of the American Chemical Society*, *114*(20), 7827–7843. https://doi.org/10.1021/ja00046a032

Hancock, B. C., & Parks, M. (2000). What is the true solubility advantage for amorphous pharmaceuticals? *Pharmaceutical Research*, *17*(4), 397–404. https://doi.org/10.1023/A:1007516718048

Hancock, B. C., Shalaev, E. Y., & Shamblin, S. L. (2002). Polyamorphism: a pharmaceutical science perspective. *Journal of Pharmacy and Pharmacology*, *54*(8), 1151–1152.

https://doi.org/10.1211/002235702320266343

Hancock, B. C., Shamblin, S. L., & Zografi, G. (1995). Molecular Mobility of Amorphous Pharmaceutical Solids Below Their Glass Transition Temperatures. *Pharmaceutical Research*, *12*(6), 799–806. https://doi.org/10.1023/A:1016292416526

Hauptman, H. A. (1997). The Phase Problem of X-ray Crystallography: Overview. In *Electron Crystallography* (pp. 131–138). Springer Netherlands. https://doi.org/10.1007/978-94-015-8971-0_9

Hédoux, A. (2016, May). Recent developments in the Raman and infrared investigations of amorphous pharmaceuticals and protein formulations: A review. *Advanced Drug Delivery Reviews*. https://doi.org/10.1016/j.addr.2015.11.021

Hess, B., Kutzner, C., Van Der Spoel, D., & Lindahl, E. (2008). GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation*, *4*(3), 435–447. https://doi.org/10.1021/ct700301q

Hilden, L. R., & Morris, K. R. (2004). Physics of amorphous solids. *Journal of Pharmaceutical Sciences*, *93*(1), 3–12. https://doi.org/10.1002/jps.10489

Hilfiker, R., & Raumer, M. von (Eds.). (2018). *Polymorphism in the Pharmaceutical Industry*. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA. https://doi.org/10.1002/9783527697847

Hodge, R. M., Bastow, T. J., Edward, G. H., Simon, G. P., & Hill, A. J. (1996). Free volume and the mechanism of plasticization in water-swollen poly(vinyl alcohol). *Macromolecules*, *29*(25), 8137–8143. https://doi.org/10.1021/ma951073j

Hou, D., Zhao, C., Paterson, A. R., Li, S., & Jones, J. L. (2018). Local structures of perovskite dielectrics and ferroelectrics via pair distribution function analyses. *Journal of the European Ceramic Society*, *38*(4), 971–987. https://doi.org/10.1016/j.jeurceramsoc.2017.12.003

Ilja Siepmann, J., & Frenkel, D. (1992). Conflgurational bias monte carlo: A new sampling scheme for flexible chains. *Molecular Physics*, *75*(1), 59–70. https://doi.org/10.1080/00268979200100061

Iuraş, A., Scurr, D. J., Boissier, C., Nicholas, M. L., Roberts, C. J., & Alexander, M. R. (2016). Imaging of Crystalline and Amorphous Surface Regions Using Time-of-Flight Secondary-Ion Mass Spectrometry (ToF-SIMS): Application to Pharmaceutical Materials. *Analytical Chemistry*, *88*(7), 3481–3487. https://doi.org/10.1021/acs.analchem.5b02621

Jacques, S. D. M., Di Michiel, M., Kimber, S. A. J., Yang, X., Cernik, R. J., Beale, A. M., & Billinge, S. J. L. (2013). Pair distribution function computed tomography. *Nature Communications*, *4*. https://doi.org/10.1038/ncomms3536

Jiang, B., Grande, T., & Selbach, S. M. (2017). Local Structure of Disordered Bi0.5K0.5TiO3 Investigated by Pair Distribution Function Analysis and First-Principles Calculations. *Chemistry of Materials*, *29*(10), 4244–4252. https://doi.org/10.1021/acs.chemmater.7b00276

Johari, G. P., Kim, S., & Shanker, R. M. (2005). Dielectric Studies of Molecular Motions in Amorphous Solid and Ultraviscous Acetaminophen. *Journal of Pharmaceutical*

Sciences, *94*(10), 2207–2223. https://doi.org/10.1002/jps.20455

Jorgensen, W. L., Maxwell, D. S., & Tirado-Rives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, *118*(45), 11225–11236. https://doi.org/10.1021/ja9621760

Juhás, P., Davis, T., Farrow, C. L., & Billinge, S. J. L. (2013). PDFgetX3: A rapid and highly automatable program for processing powder diffraction data into total scattering pair distribution functions. *Journal of Applied Crystallography*, *46*(2), 560–566. https://doi.org/10.1107/S0021889813005190

Juhás, P., Granlund, L., Gujarathi, S. R., Duxbury, P. M., Billinge, S. J. L., M., M., … A., L. B. (2010). Crystal structure solution from experimentally determined atomic pair distribution functions. *Journal of Applied Crystallography*, *43*(3), 623–629. https://doi.org/10.1107/S002188981000988X

Jurkiewicz, K., Kamiński, M., Glajcar, W., Woźnica, N., Julienne, F., Bartczak, P., … Burian, A. (2018). Paracrystalline structure of gold, silver, palladium and platinum nanoparticles. *Journal of Applied Crystallography*, *51*(2), 411–419. https://doi.org/10.1107/S1600576718001723

Kamberaj, H. (2020). *Molecular Dynamics Simulations in Statistical Physics: Theory and Applications*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-35702-3

Kamińska, E., Minecka, A., Tarnacka, M., Hachuła, B., Kamiński, K., & Paluch, M. (2020). Influence of Annealing in the Close Vicinity of Tg on the Reorganization within Dimers and Its Impact on the Crystallization Kinetics of Gemfibrozil. *Molecular Pharmaceutics*, *17*(3), 990–1000. https://doi.org/10.1021/acs.molpharmaceut.9b01244

Karmwar, P., Graeser, K., Gordon, K. C., Strachan, C. J., & Rades, T. (2011). Investigation of properties and recrystallisation behaviour of amorphous indomethacin samples prepared by different methods. *International Journal of Pharmaceutics*, *417*(1–2), 94–100. https://doi.org/10.1016/j.ijpharm.2010.12.019

Karmwar, P., Graeser, K., Gordon, K. C., Strachan, C. J., & Rades, T. (2012). Effect of different preparation methods on the dissolution behaviour of amorphous indomethacin. In *European Journal of Pharmaceutics and Biopharmaceutics* (Vol. 80, pp. 459–464). Elsevier. https://doi.org/10.1016/j.ejpb.2011.10.006

Keen, D. A. (2001). A comparison of various commonly used correlation functions for describing total scattering. *Journal of Applied Crystallography*, *34*(2), 172–177. https://doi.org/10.1107/S0021889800019993

Khachan, J. (2018). *Thermal Properties of Matter*. *Thermal Properties of Matter*. IOP Publishing. https://doi.org/10.1088/978-1-6817-4585-5

Kong, Y., & Hay, Y. N. (2002). The measurement of cristalinity of polymers by DSC. *Polymer*, *43*(1), 3873–3878. https://doi.org/https://doi.org/10.1016/S0032-3861(02)00235-5

Kratochvíl, B. (2011). Solid Forms of Pharmaceutical Molecules (pp. 129–140). https://doi.org/10.1007/978-90-481-2882-2_8

Kumara, L. S. R., Sakata, O., Kohara, S., Yang, A., Song, C., Kusada, K., … Kitagawa, H. (2016).

Structural studies of metal nanoparticles using high-energy x-ray diffraction. In *AIP Conference Proceedings* (Vol. 1741, p. 050016). American Institute of Physics Inc. https://doi.org/10.1063/1.4952936

L'vov, B. V. (2007). *Thermal Decomposition of Solids and Melts* (Vol. 7). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-1-4020-5672-7

Laggner, P., & Paudel, A. (2018). Density fluctuations in amorphous pharmaceutical solids. Can SAXS help to predict stability? *Colloids and Surfaces B: Biointerfaces*, *168*, 76–82. https://doi.org/10.1016/j.colsurfb.2018.05.003

Laitinen, R., Loƀmann, K., Strachan, C. J., Grohganz, H., & Rades, T. (2013, August). Emerging trends in the stabilization of amorphous drugs. *International Journal of Pharmaceutics*. https://doi.org/10.1016/j.ijpharm.2012.04.066

Lamparter, P., & Kniep, R. (1997). Structure of amorphous Al2O3. *Physica B: Condensed Matter*, *234–236*, 405–406. https://doi.org/10.1016/S0921-4526(96)01044-7

Latry, C., & Rouge, B. (2000). Optimized sampling for CCD instruments: The supermode scheme. In *International Geoscience and Remote Sensing Symposium (IGARSS)* (Vol. 5, pp. 2322–2324). IEEE. https://doi.org/10.1109/igarss.2000.858396

Le Bail, A., Duroy, H., & Fourquet, J. L. (1988). Ab-initio structure determination of LiSbWO6 by X-ray powder diffraction. *Materials Research Bulletin*, *23*(3), 447–452. https://doi.org/10.1016/0025-5408(88)90019-0

Lerbret, A., Mason, P. E., Venable, R. M., Cesàro, A., Saboungi, M. L., Pastor, R. W., & Brady, J. W. (2009). Molecular dynamics studies of the conformation of sorbitol. *Carbohydrate Research*, *344*(16), 2229–2235. https://doi.org/10.1016/j.carres.2009.08.003

Liengme, B. V. (2014). *Modelling Physics with Microsoft Excel*. Morgan & Claypool Publishers. https://doi.org/10.1088/978-1-627-05419-5

Lim, W., Feng, Y. P., & Liu, X. Y. (2005). Molecular dynamics simulation of paracetamol molecules ordering around glycogen. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *71*(5), 051604. https://doi.org/10.1103/PhysRevE.71.051604

Liu, J., He, X., & Zhang, J. Z. H. (2017). Structure of liquid water – a dynamical mixture of tetrahedral and 'ring-and-chain' like structures. *Physical Chemistry Chemical Physics*, *19*(19), 11931–11936. https://doi.org/10.1039/C7CP00667E

Ma, X., & Yu, J. (2004). Formamide as the plasticizer for thermoplastic starch. *Journal of Applied Polymer Science*, *93*(4), 1769–1773. https://doi.org/10.1002/app.20628

Mallamace, F. (2009, September 8). The liquid water polymorphism. *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences. https://doi.org/10.1073/pnas.0908198106

Maple, J. R., Dinur, U., & Hagler, A. T. (1988). Derivation of force fields for molecular mechanics and dynamics from ab initio energy surfaces. *Proceedings of the National Academy of Sciences*, *85*(15), 5350–5354. https://doi.org/10.1073/pnas.85.15.5350

Mariedahl, D., Perakis, F., Späh, A., Pathak, H., Kim, K. H., Camisasca, G., … Amann-Winkel, K. (2018). X-ray Scattering and O-O Pair-Distribution Functions of Amorphous Ices.

Journal of Physical Chemistry B, *122*(30), 7616–7624. https://doi.org/10.1021/acs.jpcb.8b04823

Masadeh, A. S. (2016). Total scattering atomic pair distribution function: new methodology for nanostructure determination. *Journal of Experimental Nanoscience*, *11*(12), 951–974. https://doi.org/10.1080/17458080.2016.1184769

Mauritz, K. A., Storey, R. F., & George, S. E. (1990). A General Free Volume Based Theory for the Diffusion of Large Molecules in Amorphous Polymers above tg. 1. Application to Di-n-alkyl Phthalates in PVC. *Macromolecules*, *23*(2), 441–450. https://doi.org/10.1021/ma00204a016

Mavračić, J., Mocanu, F. C., Deringer, V. L., Csányi, G., & Elliott, S. R. (2018). Similarity between Amorphous and Crystalline Phases: The Case of TiO2. *Journal of Physical Chemistry Letters*, *9*(11), 2985–2990. https://doi.org/10.1021/acs.jpclett.8b01067

McGreevy, R. L., & Pusztai, L. (1988). Reverse Monte Carlo Simulation: A New Technique for the Determination of Disordered Structures. *Molecular Simulation*, *1*, 359–367. https://doi.org/10.1080/08927028808080958

Meng, F., Trivino, A., Prasad, D., & Chauhan, H. (2015). Investigation and correlation of drug polymer miscibility and molecular interactions by various approaches for the preparation of amorphous solid dispersions. *European Journal of Pharmaceutical Sciences*, *71*, 12–24. https://doi.org/10.1016/J.EJPS.2015.02.003

Miriyala, N., Ouyang, D., Perrie, Y., Lowry, D., & Kirby, D. J. (2017). Activated carbon as a carrier for amorphous drug delivery: Effect of drug characteristics and carrier wettability. *European Journal of Pharmaceutics and Biopharmaceutics*, *115*, 197–205. https://doi.org/10.1016/j.ejpb.2017.03.002

Moinuddin, S. M., Ruan, S., Huang, Y., Gao, Q., Shi, Q., Cai, B., & Cai, T. (2017). Facile formation of co-amorphous atenolol and hydrochlorothiazide mixtures via cryogenic-milling: Enhanced physical stability, dissolution and pharmacokinetic profile. *International Journal of Pharmaceutics*, *532*(1), 393–400. https://doi.org/10.1016/j.ijpharm.2017.09.020

Moultos, O. A., Tsimpanogiannis, I. N., Panagiotopoulos, A. Z., Trusler, J. P. M., & Economou, I. G. (2016). Atomistic Molecular Dynamics Simulations of Carbon Dioxide Diffusivity in n -Hexane, n -Decane, n -Hexadecane, Cyclohexane, and Squalane. *Journal of Physical Chemistry B*, *120*(50), 12890–12900. https://doi.org/10.1021/acs.jpcb.6b04651

Mullen, K., & Levin, I. (2011). Mitigation of errors in pair distribution function analysis of nanoparticles. *Journal of Applied Crystallography*, *44*(4), 788–797. https://doi.org/10.1107/S0021889811014749

Müllertz, A., Perrie, Y., & Rades, T. (Eds.). (2016). *Analytical Techniques in the Pharmaceutical Sciences*. New York, NY: Springer New York. https://doi.org/10.1007/978-1-4939-4029-5

Munjal, B., & Suryanarayanan, R. (2021, March 1). Applications of synchrotron powder X-ray diffractometry in drug substance and drug product characterization. *TrAC - Trends in Analytical Chemistry*. Elsevier B.V. https://doi.org/10.1016/j.trac.2021.116181

Murdande, S. B., Pikal, M. J., Shanker, R. M., & Bogner, R. H. (2011). Solubility Advantage of Amorphous Pharmaceuticals, Part 3: Is Maximum Solubility Advantage Experimentally Attainable and Sustainable? *Journal of Pharmaceutical Sciences*, *100*(10), 4349–4356. https://doi.org/10.1002/jps.22643

Myers, J. L., Well, A. D., & Lorch, R. F. (2013). *Research design and statistical analysis, third edition*. *Research Design and Statistical Analysis, Third Edition* (Vol. 9780203726631). Taylor and Francis. https://doi.org/10.4324/9780203726631

Nanubolu, J. B., & Burley, J. C. (2012). Investigating the recrystallization behavior of amorphous paracetamol by variable temperature Raman studies and surface raman mapping. *Molecular Pharmaceutics*, *9*(6), 1544–1558. https://doi.org/10.1021/mp300035g

Narten, A. H., Venkatesh, C. G., & Rice, S. A. (1976). Diffraction pattern and structure of amorphous solid water at 10 and 77 °K. *The Journal of Chemical Physics*, *64*(3), 1106–1121. https://doi.org/10.1063/1.432298

Naumov, D. Y., Vasilchenko, M. A., & Howard, J. A. K. (1998). The monoclinic form of acetaminophen at 150 K. *Acta Crystallographica Section C: Crystal Structure Communications*, *54*(5), 653–655. https://doi.org/10.1107/S0108270197018386

Neilson, J. R., & McQueen, T. M. (2015). Representational analysis of extended disorder in atomistic ensembles derived from total scattering data. *Journal of Applied Crystallography*, *48*(5), 1560–1572. https://doi.org/10.1107/S1600576715016404

Neumann, M. A. (2003). X-Cell: A novel indexing algorithm for routine tasks and difficult cases. *Journal of Applied Crystallography*, *36*(2), 356–365. https://doi.org/10.1107/S0021889802023348

Newman, A., Hastedt, J. E., & Yazdanian, M. (2017). New directions in pharmaceutical amorphous materials and amorphous solid dispersions, a tribute to Professor George Zografi – Proceedings of the June 2016 Land O'Lakes Conference. *AAPS Open*, *3*(1), 1–14. https://doi.org/10.1186/s41120-017-0017-6

Nguyen Thi, Y., Rademann, K., & Emmerling, F. (2015). Direct evidence of polyamorphism in paracetamol. *CrystEngComm*, *17*(47), 9029–9036. https://doi.org/10.1039/C5CE01583A

Nichols, G., & Frampton, C. S. (1998a). Physicochemical characterization of the orthorhombic polymorph of paracetamol crystallized from solution. *Journal of Pharmaceutical Sciences*, *87*(6), 684–693. https://doi.org/10.1021/js970483d

Nichols, G., & Frampton, C. S. (1998b). Physicochemical Characterization of the Orthorhombic Polymorph of Paracetamol Crystallized from Solution. *Journal of Pharmaceutical Sciences*, *87*(6), 684–693. https://doi.org/10.1021/JS970483D

Norberg, S. et al. (2019). *No Title*.

Nosé, S. (1984). A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics*, *52*(2), 255–268. https://doi.org/10.1080/00268978400101201

Ostwald, W. (1897). Studien über die Bildung und Umwandlung fester Körper. *Zeitschrift Für Physikalische Chemie*, *22U*(1). https://doi.org/10.1515/zpch-1897-2233

Palatinus, L. (2013). The charge-flipping algorithm in crystallography. *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials*, *69*(1), 1–16. https://doi.org/10.1107/S2052519212051366

Palomäki, E. A. K., Lipiäinen, T., Strachan, C. J., & Yliruusi, J. K. (2020). Effect of trehalose and melibiose on crystallization of amorphous paracetamol. *International Journal of Pharmaceutics*, *590*, 119878. https://doi.org/10.1016/j.ijpharm.2020.119878

Papon, P., Leblond, J., & Meijer, P. H. E. (2006). *The physics of phase transitions: Concepts and applications*. *The Physics of Phase Transitions: Concepts and Applications*. Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-33390-8

Parrinello, M., & Rahman, A. (1980). Crystal structure and pair potentials: A molecular-dynamics study. *Physical Review Letters*, *45*(14), 1196–1199. https://doi.org/10.1103/PhysRevLett.45.1196

Patel, R., Baria, A., & Patel, N. (2008). An overview of size reduction technologies in the field of pharmaceutical manufacturing. *Asian Journal of Pharmaceutics*, *2*(4), 216. https://doi.org/10.4103/0973-8398.45033

Patterson, A. L. (1934). A fourier series method for the determination of the components of interatomic distances in crystals. *Physical Review*, *46*(5), 372–376. https://doi.org/10.1103/PhysRev.46.372

Patterson, J. E., James, M. B., Forster, A. H., Lancaster, R. W., Butler, J. M., & Rades, T. (2005). The influence of thermal and mechanical preparative techniques on the amorphous state of four poorly soluble compounds. *Journal of Pharmaceutical Sciences*, *94*(9), 1998–2012. https://doi.org/10.1002/jps.20424

Pawley, G. S. (1981). Unit-cell refinement from powder diffraction scans. *Journal of Applied Crystallography*, *14*(6), 357–361. https://doi.org/10.1107/s0021889881009618

Pecharsky, V. K., & Zavalij, P. Y. (2005). *Fundamentals of powder diffraction and structural characterization of materials*. *Fundamentals of Powder Diffraction and Structural Characterization of Materials*. Springer US. https://doi.org/10.1007/978-0-387-09579-0

Perlovich, G. L., Volkova, T. V., & Bauer-Brandl, A. (2007). Polymorphism of paracetamol. *Journal of Thermal Analysis and Calorimetry*, *89*(3), 767–774. https://doi.org/10.1007/s10973-006-7922-6

Perrin, M.-A., Neumann, M. A., Elmaleh, H., & Zaske, L. (2009). Crystal structure determination of the elusive paracetamol Form III. *Chemical Communications*, *0*(22), 3181. https://doi.org/10.1039/b822882e

Peterson, P. F., Božin, E. S., Proffen, T., & Billinge, S. J. L. (2003). Improved measures of quality for the atomic pair distribution function. *Journal of Applied Crystallography*, *36*(1), 53–64. https://doi.org/10.1107/S0021889802018708

Petkov, V., Ren, Y., & Suchomel, M. (2012). Molecular arrangement in water: random but not quite. *Journal of Physics: Condensed Matter*, *24*(15), 155102. https://doi.org/10.1088/0953-8984/24/15/155102

Priemel, P. A., Grohganz, H., & Rades, T. (2016, May 1). Unintended and in situ amorphisation of pharmaceuticals. *Advanced Drug Delivery Reviews*. Elsevier B.V.

https://doi.org/10.1016/j.addr.2015.12.014

Priemel, Petra A., Laitinen, R., Barthold, S., Grohganz, H., Lehto, V. P., Rades, T., & Strachan, C. J. (2013). Inhibition of surface crystallisation of amorphous indomethacin particles in physical drug-polymer mixtures. *International Journal of Pharmaceutics*, *456*(2), 301–306. https://doi.org/10.1016/j.ijpharm.2013.08.046

Prill, D., Juhás, P., Billinge, S. J. L., & Schmidt, M. U. (2016). Towards solution and refinement of organic crystal structures by fitting to the atomic pair distribution function. *Acta Crystallographica Section A Foundations and Advances*, *72*(1), 62–72. https://doi.org/10.1107/S2053273315022457

Prill, D., Juhás, P., Schmidt, M. U., & Billinge, S. J. L. (2015). Modelling pair distribution functions (PDFs) of organic compounds: Describing both intra- and intermolecular correlation functions in calculated PDFs. *Journal of Applied Crystallography*, *48*(1), 171–178. https://doi.org/10.1107/S1600576714026454

Proffen, T., & Billinge, S. J. L. (1999). PDFFIT , a program for full profile structural refinement of the atomic pair distribution function. *Journal of Applied Crystallography*, *32*(3), 572–575. https://doi.org/10.1107/S0021889899003532

Pyramides, G., Robinson, J. W., & William Zito, S. (1995). The combined use of DSC and TGA for the thermal analysis of atenolol tablets. *Journal of Pharmaceutical and Biomedical Analysis*, *13*(2), 103–110. https://doi.org/10.1016/0731-7085(94)00112-F

Qian, F., Huang, J., & Hussain, M. A. (2010). Drug–Polymer Solubility and Miscibility: Stability Consideration and Practical Challenges in Amorphous Solid Dispersion Development. *Journal of Pharmaceutical Sciences*, *99*(7), 2941–2947. https://doi.org/10.1002/JPS.22074

Qiu, X., Božin, E. S., Juhas, P., Proffen, T., & Billinge, S. J. L. (2004). Reciprocal-space instrumental effects on the real-space neutron atomic pair distribution function. *Journal of Applied Crystallography*, *37*(1), 110–116. https://doi.org/10.1107/S0021889803026670

Rapaport, D. C. (2004). *The Art of Molecular Dynamics Simulation*. *The Art of Molecular Dynamics Simulation*. Cambridge University Press. https://doi.org/10.1017/cbo9780511816581

Reiss, C. A., Van Mechelen, J. B., Goubitz, K., & Peschar, R. (2018a). Reassessment of paracetamol orthorhombic Form III and determination of a novel low-temperature monoclinic Form III-m from powder diffraction data. *Acta Crystallographica Section C: Structural Chemistry*, *74*(3), 392–399. https://doi.org/10.1107/S2053229618002619

Reiss, C. A., Van Mechelen, J. B., Goubitz, K., & Peschar, R. (2018b). Reassessment of paracetamol orthorhombic Form III and determination of a novel low-temperature monoclinic Form III-m from powder diffraction data. *Acta Crystallographica Section C: Structural Chemistry*, *74*(3), 392–399. https://doi.org/10.1107/S2053229618002619

Rietveld, H. M. (1969a). A profile refinement method for nuclear and magnetic structures. *Journal of Applied Crystallography*, *2*(2), 65–71. https://doi.org/10.1107/S0021889869006558

Rietveld, H. M. (1969b). A profile refinement method for nuclear and magnetic structures.

Journal of Applied Crystallography, *2*(2), 65–71. https://doi.org/10.1107/s0021889869006558

Robertson, M. J., Qian, Y., Robinson, M. C., Tirado-Rives, J., & Jorgensen, W. L. (2019). Development and Testing of the OPLS-AA/M Force Field for RNA. *Journal of Chemical Theory and Computation*, *15*(4), 2734–2742. https://doi.org/10.1021/acs.jctc.9b00054

Rumondor, A. C. F., Ivanisevic, I., Bates, S., Alonzo, D. E., & Taylor, L. S. (2009). Evaluation of drug-polymer miscibility in amorphous solid dispersion systems. *Pharmaceutical Research*, *26*(11), 2523–2534. https://doi.org/10.1007/s11095-009-9970-7

Sagarik, K. P., & Ahlrichs, R. (1987). A test particle model potential for formamide and molecular dynamics simulations of the liquid. *The Journal of Chemical Physics*, *86*(9), 5117–5126. https://doi.org/10.1063/1.452630

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, *36*(8), 1627–1639. https://doi.org/10.1021/ac60214a047

Shalaev, E., Soper, A., Zeitler, J. A., Ohtake, S., Roberts, C. J., Pikal, M. J., … Boldyreva, E. (2019, January 1). Freezing of Aqueous Solutions and Chemical Stability of Amorphous Pharmaceuticals: Water Clusters Hypothesis. *Journal of Pharmaceutical Sciences*. Elsevier B.V. https://doi.org/10.1016/j.xphs.2018.07.018

Shankland, K., David, W. I. F., & Sivia, D. S. (1997). Routine ab initio structure determination of chlorothiazide by X-ray powder diffraction using optimised data collection and analysis strategies. *Journal of Materials Chemistry*, *7*(3), 569–572. https://doi.org/10.1039/a606998c

Sharma, S. (2019). *Molecular dynamics simulation of nanocomposites using BIOVIA materials studio, lammps and gromacs*. *Molecular Dynamics Simulation of Nanocomposites using BIOVIA Materials Studio, Lammps and Gromacs*. Elsevier. https://doi.org/10.1016/C2017-0-04396-3

Shtukenberg, A. G., Tan, M., Vogt-Maranto, L., Chan, E. J., Xu, W., Yang, J., … Kahr, B. (2019). Melt Crystallization for Paracetamol Polymorphism. *Crystal Growth & Design*, *19*(7), 4070–4080. https://doi.org/10.1021/acs.cgd.9b00473

Sibik, J., Löbmann, K., Rades, T., & Zeitler, J. A. (2015). Predicting Crystallization of Amorphous Drugs with Terahertz Spectroscopy. *Molecular Pharmaceutics*, *12*(8), 3062–3068. https://doi.org/10.1021/acs.molpharmaceut.5b00330

Sibik, J., Sargent, M. J., Franklin, M., & Zeitler, J. A. (2014). Crystallization and phase changes in paracetamol from the amorphous solid to the liquid phase. *Molecular Pharmaceutics*, *11*(4), 1326–1334. https://doi.org/10.1021/mp400768m

Siddique, H. et al. (2019). No Title.

Skinner, L. B., Huang, C., Schlesinger, D., Pettersson, L. G. M., Nilsson, A., & Benmore, C. J. (2013). Benchmark oxygen-oxygen pair-distribution function of ambient water from x-ray diffraction measurements with a wide *Q* -range. *The Journal of Chemical Physics*, *138*(7), 074506. https://doi.org/10.1063/1.4790861

Smith, S. J., Bishop, M. M., Montgomery, J. M., Hamilton, T. P., & Vohra, Y. K. (2014). Polymorphism in Paracetamol: Evidence of Additional Forms IV and V at High

Pressure. *The Journal of Physical Chemistry A*, *118*(31), 6068–6077. https://doi.org/10.1021/jp411810y

Soper, A.K. (1996). Empirical potential Monte Carlo simulation of fluid structure. *Chemical Physics*, *202*(2–3), 295–306. https://doi.org/10.1016/0301-0104(95)00357-6

Soper, A K. (2011). *GudrunN and GudrunX: programs for correcting raw neutron and X-ray diffraction data to differential scattering cross section*.

Soper, Alan K. (2011). Empirical Potential Structure Refinement - EPSRshell: a user's guide. Version 18 - May 2011. *Science and Technology Facilities Council*, RAL-TR-2011-012.

Soper, Alan K., & Barney, E. R. (2011). Extracting the pair distribution function from white-beam X-ray total scattering data. *Journal of Applied Crystallography*, *44*(4), 714–726. https://doi.org/10.1107/s0021889811021455

Soper, Alan K., & Barney, E. R. (2012). On the use of modification functions when Fourier transforming total scattering data. *Journal of Applied Crystallography*, *45*(6), 1314–1317. https://doi.org/10.1107/S002188981203960X

Soper, Alan K. (2011). *GudrunN and GudrunX : programs for correcting raw neutron and X-ray diffraction data to differential scattering cross section*. *Rutherford Appleton Laboratory Technical Report*. Didcot: Science & Technology Facilities Council.

Stachurski, Z. H. (2011). On structure and Properties of amorphous materials. *Materials*, *4*(9), 1564–1598. https://doi.org/10.3390/ma4091564

Starr, F. W., Bellissent-Funel, M.-C., & Stanley, H. E. (1998). Structural evidence for the continuity of liquid and glassy water.

Strachan, C. J., Rades, T., & Gordon, K. C. (2007). A theoretical and spectroscopic study of γ-crystalline and amorphous indometacin. *Journal of Pharmacy and Pharmacology*, *59*(2), 261–269. https://doi.org/10.1211/jpp.59.2.0012

Sun, H. (1998). Compass: An ab initio force-field optimized for condensed-phase applications - Overview with details on alkane and benzene compounds. *Journal of Physical Chemistry B*, *102*(38), 7338–7364. https://doi.org/10.1021/jp980939v

Sun, Huai, Jin, Z., Yang, C., Akkermans, R. L. C., Robertson, S. H., Spenley, N. A., … Todd, S. M. (2016a). COMPASS II: extended coverage for polymer and drug-like molecule databases. *Journal of Molecular Modeling*, *22*(2), 1–10. https://doi.org/10.1007/s00894-016-2909-0

Sun, Huai, Jin, Z., Yang, C., Akkermans, R. L. C., Robertson, S. H., Spenley, N. A., … Todd, S. M. (2016b). COMPASS II: extended coverage for polymer and drug-like molecule databases. *Journal of Molecular Modeling*, *22*(2), 47. https://doi.org/10.1007/s00894-016-2909-0

Sun, Huai, Mumby, S. J., Maple, J. R., & Hagler, A. T. (1994). An ab Initio CFF93 All-Atom Force Field for Polycarbonates. *Journal of the American Chemical Society*, *116*(7), 2978–2987. https://doi.org/10.1021/ja00086a030

Sun, Y., Zhu, L., Wu, T., Cai, T., Gunn, E. M., & Yu, L. (2012, September). Stability of amorphous pharmaceutical solids: Crystal growth mechanisms and effect of polymer additives. *AAPS Journal*. Springer. https://doi.org/10.1208/s12248-012-9345-6

Szczepańska, A. (2011). Research Design and Statistical Analysis, Third Edition by Jerome L. Myers, Arnold D. Well, Robert F. Lorch, Jr. *International Statistical Review*, *79*(3), 491–492. https://doi.org/10.1111/j.1751-5823.2011.00159_12.x

Takeshi, E., & Billinge, S. J. L. (2012a). The Method of Total Scattering and Atomic Pair Distribution Function Analysis. In *Pergamon Materials Series* (Vol. 16, pp. 55–111). https://doi.org/10.1016/B978-0-08-097133-9.00003-4

Takeshi, E., & Billinge, S. J. L. (2012b). Underneath the Bragg Peaks - Structural Analysis of Complex Materials. In *Pergamon Materials Series* (Vol. 16, pp. 441–453). https://doi.org/10.1016/B978-0-08-097133-9.00013-7

Tanaka, M., Girard, G., Davis, R., Peuto, A., & Bignell, N. (2001). Recommended table for the density of water between 0 °c and 40 °c based on recent experimental reports. *Metrologia*, *38*(4), 301–309. https://doi.org/10.1088/0026-1394/38/4/3

Taylor, L. S., & Zografi, G. (1997). Spectroscopic Characterization of Interactions Between PVP and Indomethacin in Amorphous Molecular Dispersions. *Pharmaceutical Research*, *14*(12), 1691–1698. https://doi.org/10.1023/A:1012167410376

Terban, M. W., Cheung, E. Y., Krolikowski, P., & Billinge, S. J. L. (2016). Recrystallization, Phase Composition, and Local Structure of Amorphous Lactose from the Total Scattering Pair Distribution Function. *Crystal Growth and Design*, *16*(1), 210–220. https://doi.org/10.1021/acs.cgd.5b01100

Trasi, N. S., Purohit, H. S., & Taylor, L. S. (2017). Evaluation of the Crystallization Tendency of Commercially Available Amorphous Tacrolimus Formulations Exposed to Different Stress Conditions. *Pharmaceutical Research*, *34*(10), 2142–2155. https://doi.org/10.1007/s11095-017-2221-4

Trasi, N. S., & Taylor, L. S. (2012). Effect of polymers on nucleation and crystal growth of amorphous acetaminophen. *CrystEngComm*, *14*(16), 5188–5197. https://doi.org/10.1039/c2ce25374g

Treacy, M. M. J., & Borisenko, K. B. (2012). The local structure of amorphous silicon. *Science*, *335*(6071), 950–953. https://doi.org/10.1126/science.1214780

Tucker, M. G., Keen, D. A., Dove, M. T., Goodwin, A. L., & Hui, Q. (2007). RMCProfile: Reverse Monte Carlo for polycrystalline materials. *Journal of Physics Condensed Matter*, *19*(33), 335218. https://doi.org/10.1088/0953-8984/19/33/335218

Tulk, C. A., Benmore, C. J., Urquidi, J., Klug, D. D., Neuefeind, J., Tomberli, B., & Egelstaff, P. A. (2002). Structural studies of several distinct metastable forms of amorphous ice. *Science*, *297*(5585), 1320–1323. https://doi.org/10.1126/science.1074178

Verlet, L. (1967). Computer “experiments” on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Physical Review*, *159*(1), 98–103. https://doi.org/10.1103/PhysRev.159.98

Vranić, E. (2004). Amorphous pharmaceutical solids. *Bosnian Journal of Basic Medical Sciences / Udruzenje Basicnih Mediciniskih Znanosti = Association of Basic Medical Sciences*. Association of Basic Medical Sciences of Federation of Bosnia and Herzegovina. https://doi.org/10.17305/bjbms.2004.3383

Walton, F., Bolling, J., Farrell, A., MacEwen, J., Syme, C. D., Jiménez, M. G., … Wynne, K.

(2020). Polyamorphism Mirrors Polymorphism in the Liquid-Liquid Transition of a Molecular Liquid. *Journal of the American Chemical Society*, *142*(16), 7591–7597. https://doi.org/10.1021/jacs.0c01712

Wang, B., & Pikal, M. J. (2010). The impact of thermal treatment on the stability of freeze dried amorphous pharmaceuticals: I. Dimer formation in sodium ethacrynate. *Journal of Pharmaceutical Sciences*, *99*(2), 663–682. https://doi.org/10.1002/jps.21959

Wang, X., Ramírez-Hinestrosa, S., Dobnikar, J., & Frenkel, D. (2020). The Lennard-Jones potential: When (not) to use it. *Physical Chemistry Chemical Physics*, *22*(19), 10624–10633. https://doi.org/10.1039/c9cp05445f

Waseda, Y., Matsubara, E., & Shinoda, K. (2011). *X-Ray Diffraction Crystallography*. *X-Ray Diffraction Crystallography*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-16635-8

Werner, P. E., Eriksson, L., & Westdahl, M. (1985). TREOR, a semi-exhaustive trial-and-error powder indexing program for all symmetries. *Journal of Applied Crystallography*, *18*(5), 367–370. https://doi.org/10.1107/s0021889885010512

White, C. E., Provis, J. L., Proffen, T., Riley, D. P., & Van Deventer, J. S. J. (2010). Combining density functional theory (DFT) and pair distribution function (PDF) analysis to solve the structure of metastable materials: The case of metakaolin. *Physical Chemistry Chemical Physics*, *12*(13), 3239–3245. https://doi.org/10.1039/b922993k

Widmann, A. H., Laso, M., & Suter, U. W. (1995). Optimized atomic Lennard-Jones 6-12 parameters for simulating pVT properties of a realistic polymethylene melt. *The Journal of Chemical Physics*, *102*(14), 5761–5769. https://doi.org/10.1063/1.469307

Wilson, C. C. (2000). *Single Crystal Neutron Diffraction From Molecular Materials* (Vol. 2). WORLD SCIENTIFIC. https://doi.org/10.1142/4029

Wolberg, G. (1994). *Digital Image Warping*. *IEEE Computer Society Press*.

Woolfson, M. M. (1971). Direct methods in crystallography. *Reports on Progress in Physics*, *34*(2), 369–434. https://doi.org/10.1088/0034-4885/34/2/301

Wyttenbach, N., & Kuentz, M. (2017). Glass-forming ability of compounds in marketed amorphous drug products. *European Journal of Pharmaceutics and Biopharmaceutics*, *112*, 204–208. https://doi.org/10.1016/j.ejpb.2016.11.031

Xiang, T. X., & Anderson, B. D. (2013). Molecular dynamics simulation of amorphous indomethacin. *Molecular Pharmaceutics*, *10*(1), 102–114. https://doi.org/10.1021/mp3000698

Xiang, T. X., & Anderson, B. D. (2014). Molecular dynamics simulation of amorphous hydroxypropyl-methylcellulose acetate succinate (HPMCAS): Polymer model development, water distribution, and plasticization. *Molecular Pharmaceutics*, *11*(7), 2400–2411. https://doi.org/10.1021/mp500135f

Yang, X., Juhas, P., Farrow, C. L., & Billinge, S. J. L. (2014). xPDFsuite: an end-to-end software solution for high throughput pair distribution function transformation, visualization and analysis.

Yong, X., & Zhang, L. T. (2013). Thermostats and thermostat strategies for molecular

dynamics simulations of nanofluidics. *Journal of Chemical Physics*, *138*(8), 084503. https://doi.org/10.1063/1.4792202

Yoshioka, M., Hancock, B. C., & Zografi, G. (1994). Crystallization of indomethacin from the amorphous state below and above its glass transition temperature. *Journal of Pharmaceutical Sciences*, *83*(12), 1700–1705. https://doi.org/10.1002/jps.2600831211

Young, C. A., & Goodwin, A. L. (2011). Applications of pair distribution function methods to contemporary problems in materials chemistry. *Journal of Materials Chemistry*, *21*(18), 6464–6476. https://doi.org/10.1039/c0jm04415f

Yu, L. (2016, May 1). Surface mobility of molecular glasses and its importance in physical stability. *Advanced Drug Delivery Reviews*. Elsevier B.V. https://doi.org/10.1016/j.addr.2016.01.005

Zhang, G. G. Z., & Zhou, D. (2009). Crystalline and Amorphous Solids. In *Developing Solid Oral Dosage Forms* (pp. 25–60). Elsevier Inc. https://doi.org/10.1016/B978-0-444-53242-8.00002-3

Zhao, M., Barker, S. A., Belton, P. S., McGregor, C., & Craig, D. Q. M. (2012). Development of fully amorphous dispersions of a low Tg drug via co-spray drying with hydrophilic polymers. *European Journal of Pharmaceutics and Biopharmaceutics*, *82*(3), 572–579. https://doi.org/10.1016/j.ejpb.2012.07.012

Zhu, M., Wang, J.-Q., Perepezko, J. H., & Yu, L. (2015). Possible existence of two amorphous phases of d -mannitol related by a first-order transition. *The Journal of Chemical Physics*, *142*(24), 244504. https://doi.org/10.1063/1.4922543

Zhu, M., & Yu, L. (2017). Polyamorphism of D-mannitol. *The Journal of Chemical Physics*, *146*(24), 244503. https://doi.org/10.1063/1.4989961

Zimmermann, B., & Baranović, G. (2011). Thermal analysis of paracetamol polymorphs by FT-IR spectroscopies. *Journal of Pharmaceutical and Biomedical Analysis*, *54*(2), 295–302. https://doi.org/10.1016/j.jpba.2010.08.023

Zimmermann, B., & Kohler, A. (2013). Optimizing savitzky-golay parameters for improving spectral resolution and quantification in infrared spectroscopy. *Applied Spectroscopy*, *67*(8), 892–902. https://doi.org/10.1366/12-06723

Zografi, G., & Crowley, K. J. (2002). Cryogenic grinding of indomethacin polymorphs and solvates: Assessment of amorphous phase formation and amorphous phase physical stability. *Journal of Pharmaceutical Sciences*, *91*(2), 492–507. https://doi.org/10.1002/jps.10028

Zografi, G., & Newman, A. (2017). Interrelationships Between Structure and the Properties of Amorphous Solids of Pharmaceutical Interest. *Journal of Pharmaceutical Sciences*, *106*(1), 5–27. https://doi.org/10.1016/j.xphs.2016.05.001

Appendices

Appendix A4 (Chapter 4)

A4.1 Experimental methods

A4.1.1. Differential scanning calorimetry and thermogravimetric analysis

All differential scanning calorimetry (DSC) and thermogravimetric analysis (TGA) measurements were collected using a Netzsch DSC214 Polyma and Netzsch Jupiter F1 449 instrument, respectively. Buoyancy correction was carried out with an empty pan prior to sample analysis. Temperature and sensitivity calibrations were carried out over the temperature range –93 – 605 °C with indium, tin, bismuth, zinc, and caesium chloride standard materials for DSC. All data analysis was carried out with the Netzsch Proteus Analysis software v7.0.1 for DSC and v6.1.0 for TGA. Unless otherwise stated, all samples were prepared and measured in triplicate (i.e. 3 lots of the same sample) for DSC and TGA measurements.

Based on results from a heating rate study, samples were heated from 25 – 225 °C for TGA and –20 – 225 °C for DSC at a heating rate of 5 °C/min under flowing helium purge gas at a flow rate of 40 ml/min. To ensure accurate determination of the sample T_g, two heating cycles were carried out for the DSC to obtain the T_g from the second heat absent from the influence of moisture. The sample weight loss over the temperature range 40 °C – 130 °C was determined. The T_g (onset, midpoint, endset, ΔC_p and ΔH_{relax}) was determined for DSC, in addition to crystallisation and melting peak temperatures (T_c and T_m) and enthalpies (ΔH_c and ΔH_m).

Standard samples were allowed to equilibrate for 30 min in a petri dish on a laboratory bench at room temperature (21 °C) and laboratory RH to minimise variation in residual moisture content. Standard samples were weighed into aluminium pans, sealed with pierced lids, after which DSC and TGA analysis was started immediately. Samples for

dry conditions were placed in a vacuum desiccator over P_2O_5 with pan lids removed to dry for 48 hours at 60 °C. Samples for moist conditions were placed in a humidifying glove bag (Aldrich® AtmosBag) with pan lids removed to humidify for 24 hours at 65 – 71% RH. Samples were hermetically sealed with lids immediately upon removal from a desiccator/glove bag, then weighed, and lids were only pierced immediately before insertion into the DSC and TGA instruments for analysis.

A4.1.2. Thermo gravimetric analysis

TGA data were collected using a Netzsch Jupiter F1 449. 3-5 mg of sample were weighed into aluminium pans and heated between 25 °C and 225 °C at 5 °C per minute, under a flowing helium purge. Buoyancy correction was carried out on an empty pan prior to analysis. Data analysis was performed using the Netzsch Proteus Analysis Software V6.1.0. Mass loss over the temperature range 35 - 100 °C was determined.

A4.1.3. Fourier transform infra-red spectroscopy

All Fourier transform infra-red (FTIR) data were collected on a Bruker Tensor II ATR FTIR spectrometer in a spectral range of 400 – 4000 cm^{-1}. The resolution was 4 cm^{-1}, and the final spectra were a mean of 16 scans. Data processing was performed using the Bruker OPUS version 7.5 (Bruker Optik GmbH, UK) software. All analysed samples were transported to the benchtop instrument in sealed glass vials and scanned within 1 min.

A4.2. Characterisation of day 0 reference materials and ASDs

All reference and day 0 powder materials were characterised using various solid-state analytical techniques to capture solid-state information about freshly prepared samples prior any changes upon storage or additional processing.

A4.2.1. FTIR characterisation

The region of interest for the IND-PVP system is between 1525 – 1800 cm^{-1} where the IND benzoyl C=O stretch can shift depending on whether IND is crystalline or amorphous, and also where hydrogen bonding between IND and PVP occurs in the ASD (Taylor & Zografi, 1997) (Figure S.1 and Figure S.2).

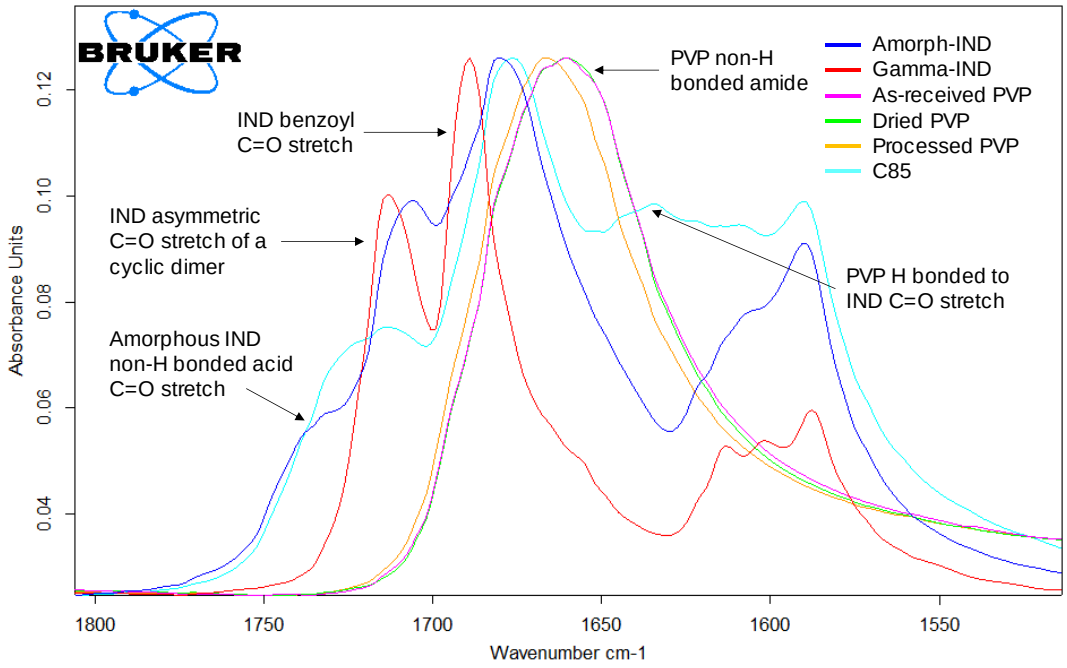

Figure A4.1: Overlaid FTIR spectra of day 0 powders of reference materials with labelled key bands.

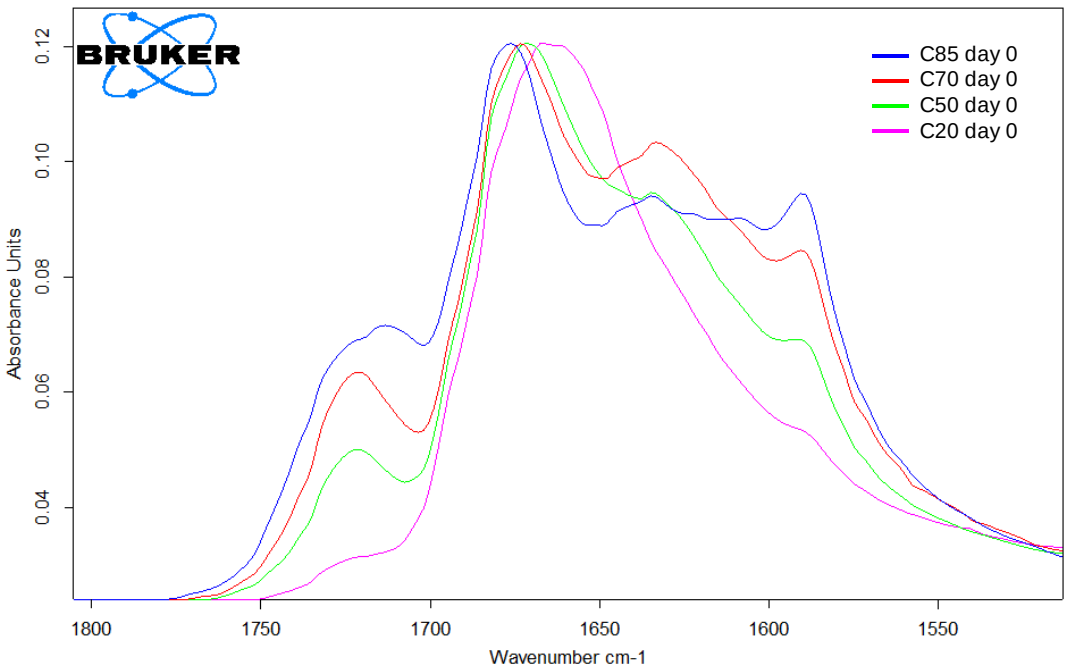

Figure A4.2: Overlaid FTIR spectra of day 0 powders of IND-PVP ASDs. Note that C20, C50, and C70 were prepared by HME, while C85 was prepared by melt cooling.

Table A4.1: Tabulated FTIR data of reference materials and ASD containing key bands in the C=O region.

Material	IR band	Experimental (cm^{-1})	Taylor and Zografi, 1997 (cm^{-1})	Observations
IND-PVP ASD	Amorphous IND non-H bonded acid C=O stretch	1726 (shoulder)	1726 (shoulder)	Peak intensity increases with increasing IND wt% in the ASD, and goes from a shoulder at low concentrations of IND, where the absorbance of the

			PVP carbonyl dominates the spectrum, into a single broad peak at high IND content.
Amorphous IND asymmetric acid C=O stretch of a cyclic dimer	1713 (medium)		
Amorphous IND benzoyl C=O stretch	1676 (strong)		As Amorph-IND content decreases in the ASD, a corresponding shift from left to right of this peak is observed.
PVP H bonded C=O stretch	1635 (weak)	1636 (band intensity varies depending on IND-PVP composition)	Peak is present in IND-PVP ASDs containing 5 – 50 wt% PVP. As the concentration of PVP increases further this

				peak merges with the non-hydrogen bonded carbonyl peak at 1666 cm^{-1} and is apparent only as a shoulder at (60 – 70 wt% PVP). At 80 wt% PVP (C20) it is barely visible.
Gamma-IND	Crystalline γ-IND asymmetric acid C=O stretch of a cyclic dimer	1713 (strong)	1717 (strong)	
	Crystalline γ-IND benzoyl C=O stretch	1689 (strong)	1692 (strong)	
Dried PVP	PVP non-H bonded amide C=O stretch	1660 (strong)†	1679 (strong)‡	

As-received PVP	PVP non-H bonded amide C=O stretch	1661 (strong)		
Processed PVP	PVP non-H bonded amide C=O stretch	1666 (strong)		
Amorph-IND	Amorphous IND non-H bonded acid C=O stretch	1732 (shoulder)	1735 (shoulder)	
	Amorphous IND asymmetric acid C=O stretch of a cyclic dimer	1706 (strong)	1710 (strong)	
	Amorphous IND benzoyl C=O stretch	1679 (strong)	1684 (strong)	
IND-PVP physical mixture	Crystalline γ-IND asymmetric	1713 (strong)		This peak is indicative of the presence of

	acid C=O stretch of a cyclic dimer		crystallinity. The intensity of this peak is observed to increase with increasing crystalline content in the ASD
	Crystalline γ-IND benzoyl C=O stretch	1689 (strong)	The emergence of this peak, initially as a shoulder in the ASD is indicative of the presence of crystallinity. The progressive shift of this peak to the left is observed to occur with increasing crystalline content in the ASD

[†]PVP-K25 [‡]PVP-K90

A4.2.1. Laboratory molybdenum XRPD characterisation

Figure S.3 shows the X-ray diffraction patterns of day 0 reference powder materials measured with a lab Molybdenum source. As anticipated, intense Bragg peaks identified

the crystalline Gamma-IND, while the amorphous materials showed a characteristic broad halo. It is clear from the patterns of the amorphous samples that they do not contain detectable traces of crystallinity, and as such, they can be regarded as X-ray amorphous. The patterns observed are very similar for as-received PVP, Dried PVP, and Processed PVP, giving a distinct double halo between 4 – 10 °2θ.

Figure S.4 shows the stacked plot of MoXRPD patterns of day 0 ASDs. A double halo characteristic of PVP was exhibited by the amorphous EXTD-C20 powder between 4 – 10 °2θ, indicating the high PVP content. However, the halo at ~4 °2θ decreased in intensity with increasing IND content.

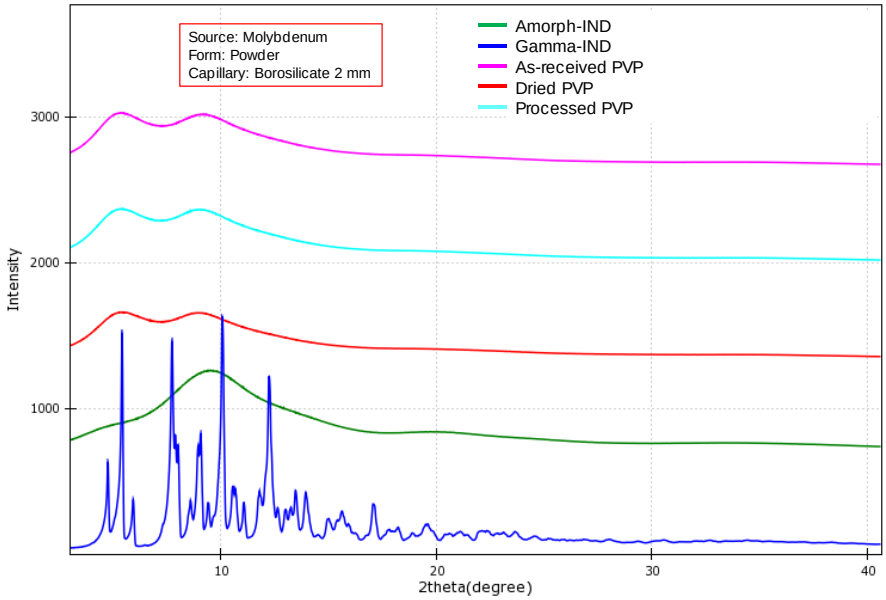

Figure A4.3: Stacked plot of MoXRPD patterns of day 0 reference materials.

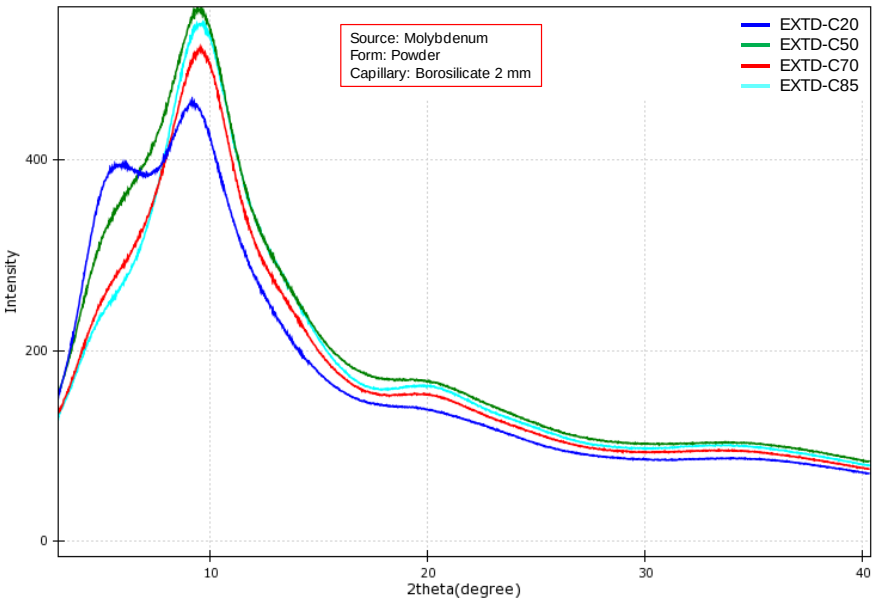

Figure A4.4: Stacked plot of MoXRPD scans of day 0 ASDs.

A4.3. Reproducibility of ASD

A4.3.1. FTIR analysis

Figure S.5 shows the overlaid IR absorption spectra of all three day 0 Melt cooled C85 powder samples (ASD-R1, ASD-R2, and ASD-R3). The close match of the spectra from the powder samples is indicative of a high similarity in the compositions, however the spectra do not completely overlap, and this is likely due to measurement artefacts arising from sample effects (e.g. particle size, orientation etc.). To eliminate such effects and amplify any subtle differences between samples, a first derivative of all three absorption spectra was taken and compared. As shown in Figure S.5, all three first derivative spectra overlap indicates a high reproducibility between samples.

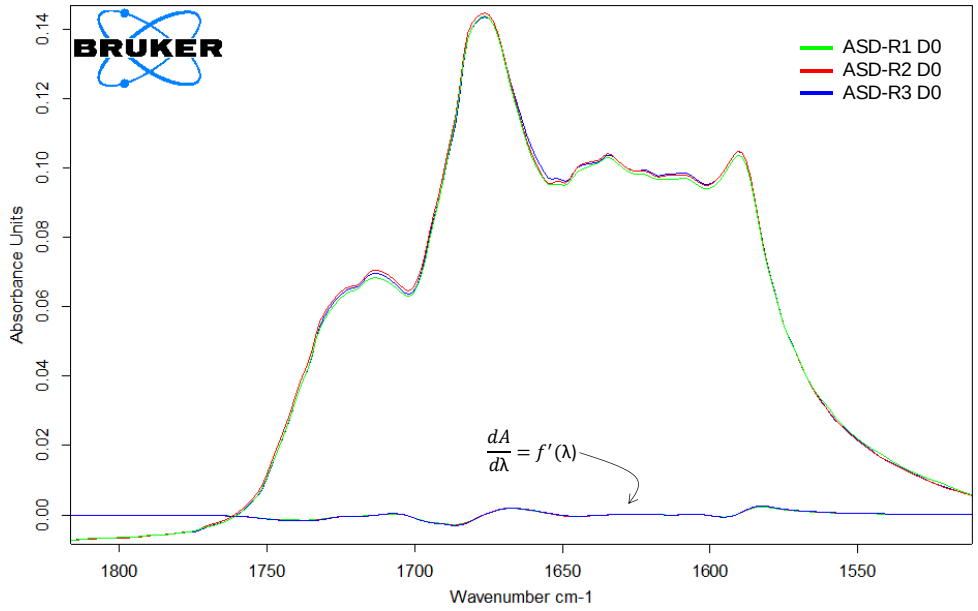

Figure A4.5: FTIR analysis showing absorbance and derivative spectra of day 0 Melt cooled C85 powders.

A.4.3.2. PSD analysis

Figure S.6 shows the Morphologi measurements obtained for all three Melt cooled C85 powder samples after gently triturating the clear, dark yellow glass for 4 min. The particle size distribution (PSD) of ASD-R1 was broadest, followed by ASD-R2, with both having similar $d_{4,3}$ values of 162 µm and 164 µm respectively. The volume mean diameter, $d_{4,3}$ is chosen as a definition of the mean size since it is very sensitive to the presence of large particles in the distribution. ASD-R3 had larger particles with $d_{4,3}$ of 175 µm, highlighting the variability in PSD that is introduced by a manual trituration operation. Small variations like these in PSD are unlikely to cause significant differences in sample behaviour, however very broad or bi-modal distributions may lead to notably different sample behaviours during the dynamic heating step of DSC measurements.

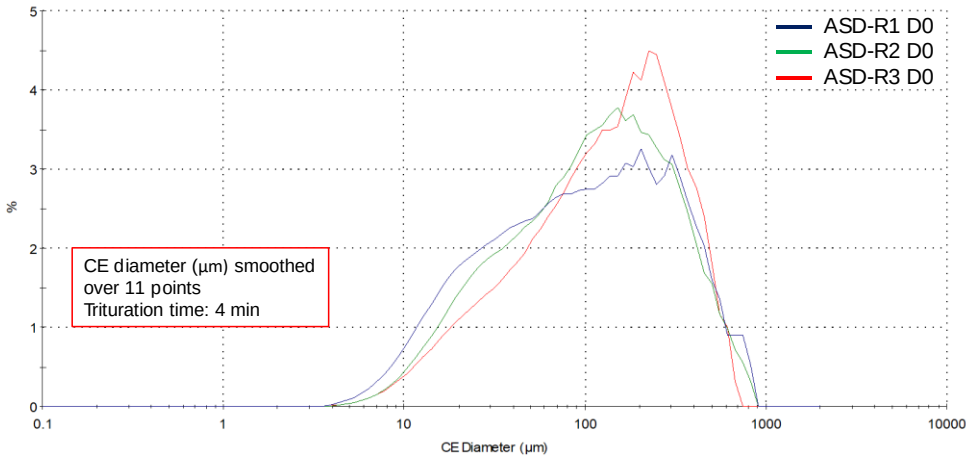

Figure A4.6: Morphologi PSD measurements of day 0 Melt cooled C85 powders.

A4.3.3. DSC analysis

In Figure S.7 all three samples showed very similar DSC traces, with an average mid-point T_g of 48 $\pm$ 1.0 °C and a T_m of 154 $\pm$ 0.2 °C confirming good reproducibility of sample compositions.

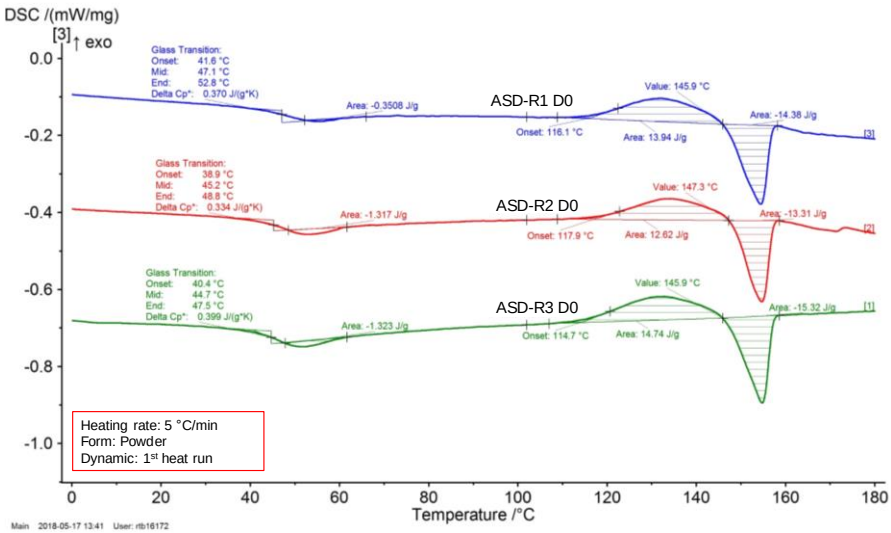

Figure A4.7: Stack plot of DSC measurements from day 0 Melt cooled C85 powders.

A4.4. Detecting levels of crystallinity

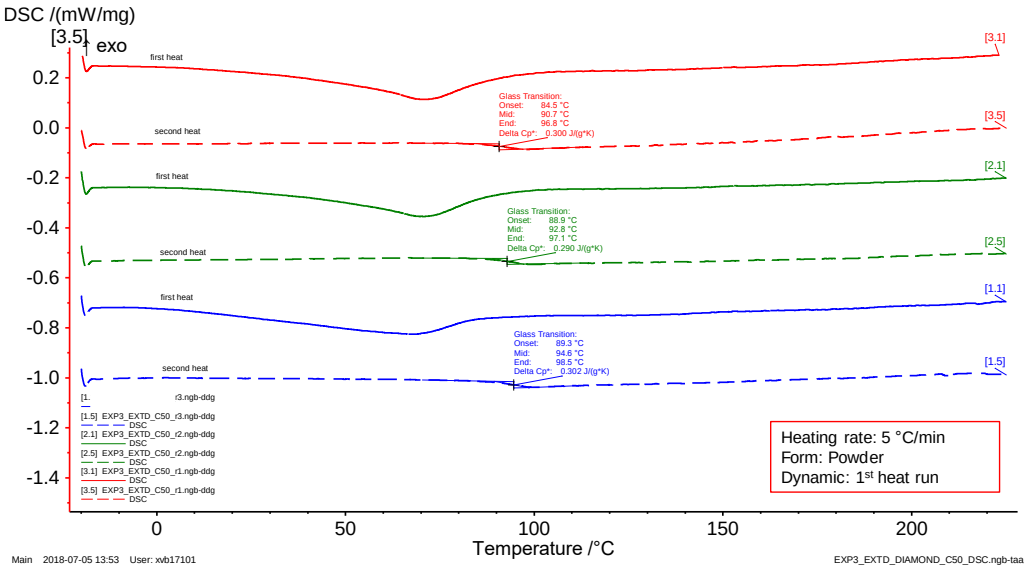

Figure A4.8: DSC trace of EXTD-C50 powder D125 showing a single T_g with no signs of phase separation or recrystallisation after extended storage.

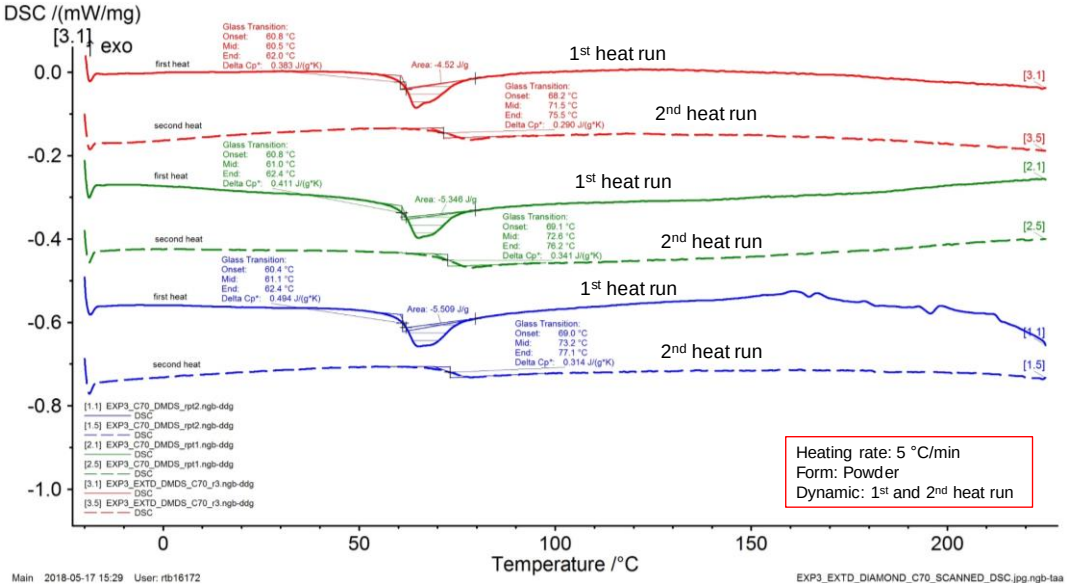

Figure A4.9: DSC trace of EXTD-C70 powder D127 showing the early occurrence of two T_g values.

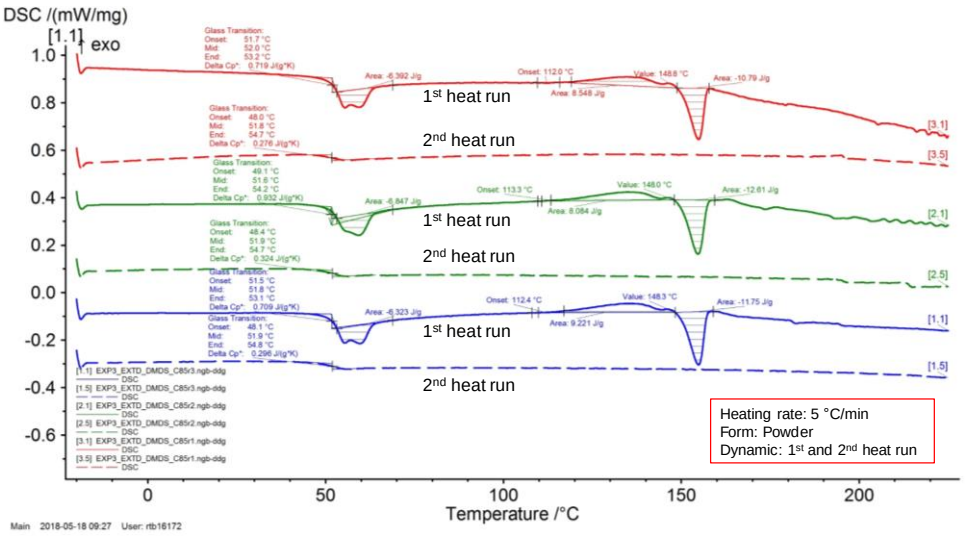

Figure A4.10: DSC trace of EXTD-C85 powder D128 showing the occurrence of two T_gs.

A4.5. Calculation of percentage crystallinity

The amount of crystalline IND, X_c, already present in the EXTD-C85 powder D128 sample was determined using the equation below (Kong & Hay, 2002) to make allowance for amorphous IND crystallising on heating.

$$X_c = (\Delta H_m - \Delta H_c)/\Delta H_m^0 \qquad \textbf{(A4.1)}$$

where ΔH_m is the enthalpy of fusion of the sample at the melting temperature, T_m, ΔH_c is the enthalpy of crystallisation of the sample, and ΔH_m^0 is the heat of fusion of the completely crystalline IND at the equilibrium melting temperature, T_m^0. ΔH_m^0 was 118.6 J/g as determined for Gamma-IND.

A4.6. Investigating structural changes

A4.6.1. Preparation of C20 samples

Freshly extruded ASD using HME technique was ground into a fine powder using a mortar and pestle in a nitrogen filled glove bag with humidity <4%. Ground ASD was isolated in a vacuum oven for drying under P2O5 in petri-dish for 24 hours at room temperature, afterwards it was separated into two batches Batch A and Batch B in a dry glove bag. Batch A was isolated in a sealed vial pending loading into a 2 mm capillary within the glove bag for EMPYREAN X-ray measurement. DSC, IR, and TGA measurements were also carried out. Batch B was isolated in a humidifying glove bag using a petri dish/ weighing boat for 48 hours to humidify sufficiently. Humidification was provided by a NaCl supersaturated solution to give a maximum of 67 - 71% RH in the glove bag. The wet ASD (Batch B) was then sampled for DSC, TGA and PDF measurements in a 2mm capillary. Finally, the wet ASD (Batch B) was re-dried in a vacuum oven at room temperature for 48 hours with P2O5 to return it as close as possible to the original dry state.

A4.6.2. Preparation of Processed PVP Samples

5g of PVP-K25 was weighed out into the stainless steel beaker and was immersed into hot oil at 175 °C. Processed PVP was thus prepared using Method C by heating the PVP powder for 25 min while stirring intermittently with an overhead impeller. The processed PVP bulk was isolated into a glove bag and the steel containment beaker was covered with parafilm to prevent the uptake of moisture while allowing the powder to cool down to room temperature. The processed PVP was then spread across a petri dish and dried in a vacuum oven for 19 hours at ~80 °C. After drying, a sample was isolated into TGA pans alongside the bulk material, and further dried in a desiccator at room temperature for 21 hours with P2O5. After drying, the DWD-PVP-DRIED was sampled for PDF, TGA, and IR. The remaining

bulk was placed in a humidifying glove bag with 65 - 71% RH to humidify for over 22 hours. After humidification, DWD-PVP-WET was sampled and analysed by PDF, IR, TGA, and then re-dried in a desiccator with P2O5 at 60 °C for ~22 hours. The final DWD-PVP-REDRIED was then analysed by the above techniques and PDF to assess influence of moisture on PDF trace.

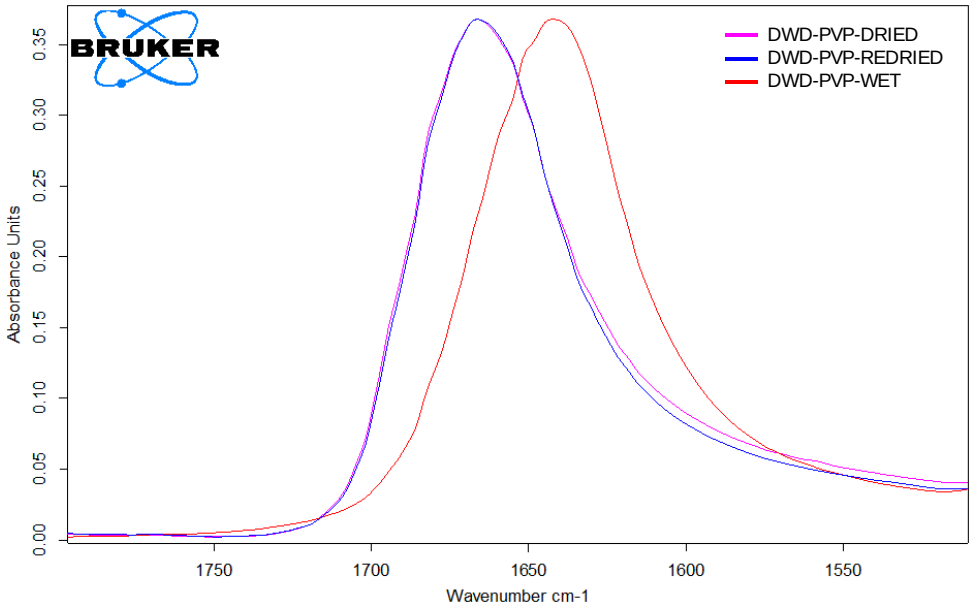

Figure A4.11: FTIR absorption spectra of Processed PVP for dried, wet, and re-dried powders.

Table A4.2: Weight loss in Processed PVP and Extrudate C20 samples from TGA analysis

Sample	Experiment name	Mean weight loss (%)
Processed PVP	DWD-PVP-DRIED	6.7 ± 0.85
	DWD-PVP-WET	15.2 ± 1.56

	DWD-PVP-REDRIED	2.5 ± 0.59
Extrudate C20	DWD-C20-DRIED	2.9 ± 0.88
	DWD-C20-WET	11.2 ± 0.91
	DWD-C20-REDRIED	3.18 ± 0.14

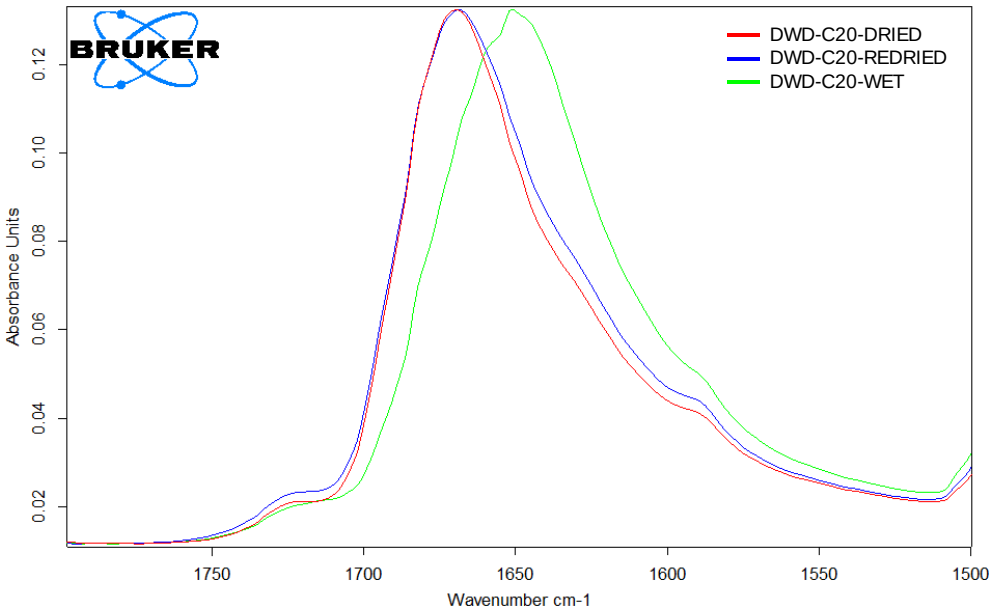

Figure A4.12: FTIR absorption spectra of Extrudate C20 for dried, wet, and re-dried powders.

A4.7. Ripples from Fourier transformation: Relationship with Q_{max} and $F(Q)$

Table A4.3: Comparison of calculated ripple wavelength in the $G(r)$ with observed ripple wavelength in flat line $G(r)$ at different Q_{max} values.

Q_{max} (Å^{-1})	Measured flat line ripple	Calculated ripple wavelength,

	wavelength (Å)	$2\pi/Q_{max}$ (Å)
1	6.325	6.283
2	3.168	3.142
4	1.569	1.571
6	1.05	1.047
8	0.787	0.785
10	0.627	0.628
12	0.527	0.524
14	0.45	0.449
16	0.395	0.393
18	0.349	0.349
20	0.313	0.314
22	0.288	0.286
24	0.261	0.262
30	0.21	0.210
35	0.178	0.179

A4.8. Error minimisation algorithm

```
close all

clear all

clc

%import files

Fqraw = xlsread('FQ.xlsx','Amorph IND','C2:C858');
```

```
Q = xlsread('FQ.xlsx','Amorph IND','B2:B858');


%Minimise F(Q) and find corresponding Qmax

Qrange = find(Q>=15); %sets lower limit of Qmax range as 15

Qval = Q(Qrange);

[Fqmin,I] = min(abs(Fqraw(Qrange))); %Finds minimum value of F(Q)

Qmax = Qval(I); %Finds corresponding Q value for minimum F(Q)

disp ('Qmax for F(Q) =')

disp (Qmax)


%plot F(Q)

figure(3)

x=0;

ax1=subplot(2,1,1);

plot(ax1,Q,Fqraw,'r-');

hold on

plot(ax1,Qmax,(min(Fqraw):0.001:max(Fqraw)),'g.-','Linewidth',0.5)

hold on

plot(ax1,Q, x*ones(size(Q)),'k-', 'Linewidth', 0.5)
```

```
title(ax1,'Reduced Structure Function F(Q)')

ylabel(ax1,'F(Q)','fontsize', 12)

xlabel(ax1,'Q (A^-^1)','fontsize', 12)

legend (ax1,'F(Q) raw');
```