# Hydration, polymorphism and disorder in organic solids, including materials of pharmaceutical relevance

by

Talbir Kaur Austin

A thesis submitted to Cardiff University

for the requirement of the degree of

Doctor of Philosophy

School of Chemistry

Cardiff University

July 2006





ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

## DECLARATION

This work has not been previously accepted in substance for any degree and is not being concurrently submitted in candidature for any degree

Signed .....(candidate)

Date .....29/08/06.....

## STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated.

Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed .....(candidate)

Date .....29/08/06.....

## STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations

Signed .....(candidate)

Date .....29/08/06.....

# Summary

The studies described in this thesis are concerned with understanding and rationalising the inter-relationships between structural and other physical properties of organic solids; encompassing hydrates, polymorphs and homologous series. This is achieved through assessment of the crystal packing arrangements and the nature of the intermolecular interactions, which may have a profound impact on the stability and physical properties of organic materials. Such an understanding is critically important when selecting or designing materials that demonstrate specific defined properties.

The first part of the thesis describes the relationships between hydrate and anhydrate phases, exemplified by tetraphenyl phosphonium bromide as a model system. The relative inter-relationships between the two main phases of this material are rationalised by characterisation of both structural and dynamic properties.

The second aspect is focused on the influence of short-range interactions on the long-range periodicity and subsequent properties of materials using 4-hydroxy benzoic acid esters as a homologous series to explore odd-even alternations, seen for other series containing long chain functionalities. In this study, the nature of short-range interactions is found to correlate with the observed odd-even alternation in material properties, with deviations from these observations shown to correspond to the presence of significant disorder in the structures.

The third aspect concerns the properties of a polymorphic pharmaceutical material, AZD7140. The crystal structures for both polymorphs were solved and assessment of hydrogen bonding motifs and thermodynamic evaluation using phase diagrams allowed the selection of a robust polymorph as a suitable development material. The final aspect of the thesis concerns strategies that may be employed to solve structures of disordered systems directly from XRPD data using direct space methodology.

I dedicate this thesis to my late Father

Sardar Raghbir Singh Bhogal

A truly great man who inspired me to pursue a career in science

## Acknowledgements

My sincere gratitude goes to Professor K. D. M. Harris for his unstinted supervision, guidance and encouragement throughout this work. Thanks are also given to Dr. F. Leveiller and Dr. G. Steele for guidance and assistance as Industrial supervisors, and to AstraZeneca for sponsorship. I would also especially like to express my gratitude to Professor R. Ceolin from Paris-5 University for direction and supervision in constructing the p/T diagrams and for determining the liquid density measurements of AZD7140.

I would like to thank members of the Structural Organic Chemistry research group, especially Dr. S. Kitchin to whom I'm extremely grateful for help in SSNMR spectroscopy measurements, and Dr. E. Cheung and Dr. S. Habershon for helpful discussion and guidance in the use of the GA program and Rietveld refinement. Thanks also to Dr. B. Kariuki for supplying me with structural information of the reported structure of DBEB. Thanks are also due to AZ colleagues including Mr. G. Allsop and Dr. M. Stocks for assistance in synthesising perdeuterated TPPB and paraben esters respectively. Mrs H. Pancholi is thanked for assisting in elemental analysis. I would also like to express thanks to Dr. T. Page and Dr. J. Murray for their words of encouragement and support in allowing me time off to conduct my research work. Members of the P&B group at AZ are thanked for providing the lighter moments!

Finally I would like to express special thanks to my Parents, Parents-in-law and family for their unfaltering love and support. To my Mother and (late) Father and especially my brothers Satnam and Onkar who provided a continuous source of enthusiasm, support and blessing. Finally, I would like to express my sincerest gratitude to my husband Rupert Austin without whom all of this would have been impossible, for his unflagging love and inspiration that have kept me going throughout this work.

# Table of Contents

# List of Abbreviations

| | |
|---|---|
| ADP | Atomic Displacement Parameter(s) |
| ATR | Attenuated Total Reflectance |
| CCD | Charge Coupled Device |
| CP | Cross Polarisation |
| CSA | Chemical Shift Anisotropy |
| CSD | Cambridge Structural Database |
| ct | Contact Time |
| DBEB | 1,3-DiBromo-5-Ethynyl Benzene |
| DSC | Differential Scanning Calorimetry |
| DVS | Dynamic Vapour Sorption |
| EAGER | Evolutionary Algorithm Generalised for Energy and R-factor |
| EFG | Electric Field Gradient |
| E/T | Energy-Temperature |
| FID | Frequency Induction Decay |
| FOM | Figure Of Merit |
| FTIR | Fourier Transform Infra Red |
| GA | Genetic algorithm |
| ITC | Iso-Thermal micro-Calorimetry |
| MAS | Magic Angle Spinning |
| NMR | Nuclear Magnetic Resonance |
| NQC | Non-Quaternary Carbon(s) |
| NQS | Non-Quaternary Suppression |
| PAS | Principal Axis System |
| p/T | Pressure-Temperature |

| | |
|---|---|
| QSAR | Quantitative Structure Activity Relationship(s) |
| rf | Radio Frequency |
| RH | Relative Humidity |
| rmsd | Root Mean Square Difference |
| SSNMR | Solid State Nuclear Magnetic Resonance |
| SXRD | Single crystal X-Ray Diffraction |
| TGA | Thermo-Gravimetric Analysis |
| TOSS | Total Sideband Suppression |
| $TPP^+$ | Tetra Phenyl Phosphonium cation |
| TPPB | Tetra Phenyl Phosphonium Bromide |
| TPPB-d20 | Perdeuterated Tetra Phenyl Phosphonium Bromide |
| UV | Ultraviolet |
| VT | Variable Temperature |
| XRD | X-Ray Diffraction |
| XRPD/PXRD | X-Ray Powder Diffraction/ Powder X-Ray Diffraction |

# 1. Introduction

## 1.1 Introduction

Molecular crystals are a class of solids that are composed of discrete molecules arranged in a structural framework. The structures of molecular crystals are influenced by both intramolecular and intermolecular interactions. Intramolecular forces determine molecular shape, which in turn contributes to the way the molecules pack in the crystal (Wright 1995). Intermolecular forces are relatively weak and their effect is largely short-range. As a consequence of this short-range effect, diversity in the arrangement of molecules within the molecular crystals is brought about, which also gives rise to differences in properties and performance of the molecular crystals. Furthermore, a variation in spatial arrangement can give rise to the enhanced possibility of structural dynamics within molecular crystals, leading to a variation in the performance and behaviour of the resultant material. Thus an understanding of molecular crystals, and in particular the intermolecular interactions driving the molecular packing within the structure, allows an understanding of the molecular properties.

A significant interest in molecular crystals originates from the ability to use molecular level 'crystal engineering' strategies to rationally design crystal packing in order to control specific physical properties (Ward *et al.* 1997). The crystal engineering approach utilises additives and other molecules to direct the self-assembly of the parent molecules to give a desired solid-state motif. Thus, control or understanding of arrangements in molecular crystals leads to control or understanding of various physical properties. Molecular crystals cover a diverse range of materials used in dyes and speciality chemicals, conductors, non-linear optical materials, agrochemicals and also pharmaceuticals.

Solid phases are defined in thermodynamic terms as a state of matter that is uniform throughout in chemical composition and also in physical state (Wunderlich 1999). A change in the physical or spatial arrangement of the molecules or inclusion of other molecule types (to give an heterogeneous material) results in the formation of different phases – termed polymorphs and hydrates/co-crystals respectively. This change or variation in spatial arrangement, consequently giving rise to differences in molecular interactions, can give rise to dramatically different bulk properties. For instance, a variation can occur in optical properties

of organic dyes, conductivity of one-dimensional conductors, magnetism of organic ferromagnets and frequency doubling characteristics of non-linear optical materials (Ward *et al.* 1997). For Pharmaceutical and Agrochemical materials, differences in solubility, stability and mechanical properties can arise. As a large number of Pharmaceutical materials are delivered as solid dosage forms, the physical and chemical properties are important to understand.

Studies that aim to establish an insight into the aspects of a structure that influence particular behaviours such as physicochemical properties, biological activity and so forth are called quantitative structure X relationships (where X can represent property, biological activity, toxicity etc.) (Grover *et al.* 2000). These relationships, in turn, can be used to design new chemical entities with more advantageous properties.

Quantitative structure activity relationships (QSAR) have been successfully used in the design of molecules with improved biological profiles (Hansch and Leo 1995). From a pharmaceutical perspective QSAR relies on both empirical data such as molecular and crystal structure and more theoretical (quantum mechanical) descriptors. Considerable work has been reported on the use of both empirical and theoretically derived descriptors for correlations of simple organic compounds. In one study, a series of topological descriptors were used to correlate molecular structure to the infinite dilution activity coefficients ($\gamma^{\infty}$) in aqueous solutions. This coefficient is used to provide a measure of the thermodynamic behaviour of dilute solutions (as a direct measure of the chemical and physical forces that exist between solute and solvent molecules) for phase equilibrium behaviour (Mitchell and Jurs 1998).

The use of a series of homologous organic compounds to probe structure property relations includes the effect of specific functional group variation with properties under investigation. Variation of substituents in homologous structures gives rise to changes in, amongst other parameters, topology and significant electronic effects. Specifically, when the properties relate to the solid-state the existence of polymorphism, whilst being seen as a possible source for complication in deriving models, can provide a useful insight to discrete crystal structure property relationships (Bernstein 1993) - since, by definition, different polymorphs will have the same chemical composition, but different crystal structures and hence different properties.

## 1.2 Polymorphism

Structural diversity is present in almost every facet of nature, and crystal polymorphism is one manifestation of this diversity (Bernstein 2002). Polymorphism, in a chemical sense, is a solid-state phenomenon where the crystal structures of a chemical entity are different but correspond to identical liquid and vapour states (McCrone 1965).

A variation in crystal structure is brought about by differences in molecular packing and intermolecular interactions within the three-dimensional framework of the crystalline state. The way the molecules pack is defined in part by the molecular structure itself and also the possibility of forming stable intermolecular interactions such as hydrogen bonds, giving rise to structures with differences in density. Consequently, polymorphs will have different lattice energies, which in turn govern the physical properties and behaviours of the material.

An understanding and control of this phenomenon is of paramount importance in the fields of crystal engineering or material selection, crucial to the pharmaceutical, chemical, food and agrochemical industries. Polymorphism is a common phenomenon in small organic molecules and the occurrence of polymorphs has been documented extensively (Borka and Haleblain 1990, Byrn *et al.* 1999, Bernstein 2002). Specifically, in the area of pharmaceutical material selection, polymorphs are selected on the basis of physical and chemical stability, behaviour to processing and formulation, and biopharmaceutical properties. Knowledge of the relative behaviour of the polymorphs with respect to the properties outlined above allows a rationalised selection. For instance, differences in solubility and dissolution rate between polymorphs can have a pronounced impact on the oral bioavailability (i.e. dissolution and absorption from the gastro-intestinal tract) of pharmaceuticals as exemplified by investigations of formulations of tolbutamide (Kimura *et al.* 1999). Other differences in properties also include thermodynamic and kinetic variations between polymorphs. Such differences encompass distinctions in reactivity involving both physical changes (e.g. involving interconversion of a metastable to a stable form) and chemical changes. Physical changes can occur in either the solid-state or via a solution-mediated process, but are driven in accordance to Ostwald's law of stages (Ostwald 1897), which states that a highly metastable form should transform to the most stable form via a series of thermodynamically driven phase transitions. Differences in chemical reactivity, such as those exemplified by the three

physical forms of trans-2-ethoxy cinnamic acid (Cohen and Green 1973), for which the α and β forms dimerise under U.V. irradiation whereas, the γ form gives no reaction, illustrate the importance of selecting a stable and robust polymorph. Furthermore, diversity in bulk properties can also be extended to polychromism, as exemplified by the three main polymorphs of 5-XII (an aromatic carbonitrile) each of which exhibits a different colour; red, yellow and orange (Yu *et al.* 2000). The polymorphism in this case, and hence the different colouration, is due to a variation in the molecular conformation.

McCrone stated, in 1965, that the number of polymorphs identified is directly proportional to the time and effort spent looking for them. Many approaches can be taken to induce polymorphic changes in order to explore its occurrence. These include solution-mediated transitions such as recrystallisation and solution maturation studies (Cardew and Davey 1985), thermally induced (Giron 1995) and mechanical/pressure induced changes such as those exhibited by chlorpropamide (Wildfong *et al.* 2005) and *p*-dichlorobenzene (Sankaran *et al.* 1986). Other solvent free methods of isolating polymorphs involve quenching from the molten liquid or gaseous state (sublimation experiments), as used to isolate polymorphs of venlafaxine hydrochloride (Roy *et al.* 2005). The occurrence of polymorphism can also be explored using computational methodology (Beyer *et al.* 2001, Neumann 2005). The basis of these approaches involves *in-silico* generation of all plausible crystal structures, which are subsequently ranked in order of calculated lattice energies. While the applicability has been demonstrated for small rigid structures, there are many limitations in the wider use of this approach – in particular for structures with significant conformational flexibility. Furthermore, the veracity of such approaches depends on the quality of the force fields used to model thermodynamic and kinetic properties satisfactorily (Gavezzotti 2002), which renders the current approaches applicable only to a small subset of organic structures.

Once polymorphs have been identified, an understanding of the relative stabilities of different polymorphic forms is important, particularly in relation to primary and secondary processing of the material. Ostwald's law of stages (1897) can be exploited in the use of solution maturation studies, in which the most thermodynamically stable polymorph will crystallise at the expense of dissolution of the metastable form. Hence, solution maturation studies as a function of temperature can be performed to assess relative thermodynamic stability. Furthermore, microcalorimetry can be used to provide a measure of the lattice

energy, as the enthalpy of dissolution is directly proportional to the lattice energy. While these techniques offer an insight into the relative thermodynamic stability under isothermal conditions, it is important to identify polymorph stability relationships as a function of both temperature and pressure. One approach to establish such relationships is in the construction and use of energy/temperature (E/T) and pressure/temperature (p/T) diagrams extracted from Gibbs' fundamental equation (Eq 2.24), as discussed in Section 2.3. Defining polymorphs as enantiotropic (for which the relative stability of polymorphs changes as a function of temperature) or monotropic (for which one polymorph is stable over the entire temperature range up to the melt) allows the relative stability as a function of temperature and pressure to be derived. Various empirical rules have also been applied in the assignment of relative stability, including density, heat of fusion and IR rules (Burger and Ramberger 1979). Figure 1.1 illustrates E/T diagrams for monotropic and enantiotropic dimorphic systems. In the case of enantiotropy, a transition temperature exists below the melting temperatures of the polymorphs evaluated. This transition temperature represents a point at which the difference in free energy between the two forms is equal to zero. It also defines the temperature at which the stability hierarchy changes.

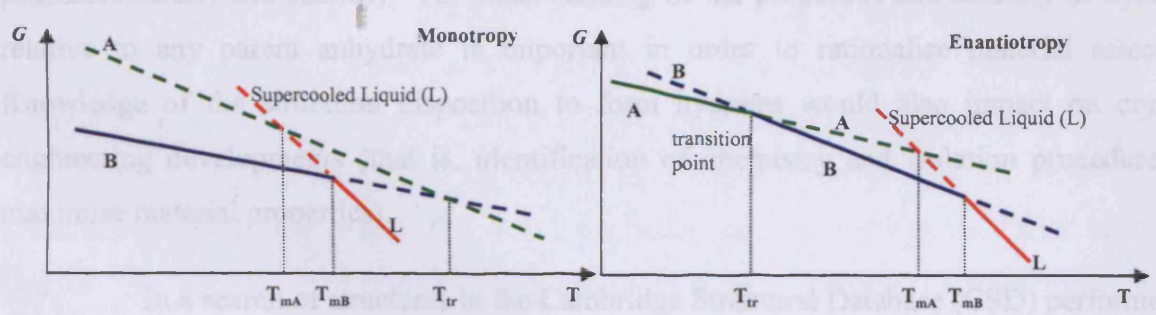



Figure 1.1: E/T diagrams showing monotropy and enantiotropy

In most industries in which polymorphism plays an important role in materials and their properties, there are several business drivers for polymorph characterisation and selection. Firstly, there is a need to understand the external effect on structural behaviour, enabling the selection of a robust and stable material that will not interconvert to a less desirable polymorph upon storage or processing. Secondly, it is important to have as much of the polymorph "hypersurface" mapped to ensure that all plausible low energy structures, which could represent developable forms, are isolated and characterised. Information on structural relationships and the ease of interconversion (exploring both kinetics and mechanisms) allows the selection of the most optimum or developable form. Leading on from this is another business driver, which relates to intellectual property. Upon identifying

all possible developable polymorphs, it is important to have patent coverage in order to protect intellectual property. In the area of Pharmaceuticals there have been several important polymorph litigation cases (Bernstein 2002).

## 1.3 Hydrates and understanding processes of hydration-dehydration

Hydrates represent structures that incorporate stoichiometric or non-stoichiometric amounts of water of crystallisation within the crystal structure. Generally speaking, four main roles are fulfilled by solvents in crystal structures (Van der Sluis and Kroon 1989). These are: (i) participation as acceptors and/or donors in hydrogen bonding schemes, (ii) filling void spaces, (iii) completing coordination around metal ions, and (iv) bridging polar and non-polar regions in the crystal. The small polar nature of the water molecule dictates that for hydrates, the first three of these four roles are often fulfilled.

Inclusion of water of crystallisation can alter the free energy of a crystal structure and consequently, as with polymorphism, can have a profound impact on physicochemical properties such as solubility, dissolution (and hence bioavailability in the case of pharmaceuticals) and stability. An understanding of the properties and stability of hydrates relative to any parent anhydrate is important in order to rationalise material selection. Knowledge of the structural disposition to form hydrates would also impact on crystal-engineering developments (that is, identification of chemistry and isolation procedures to maximise material properties).

In a search of structures in the Cambridge Structural Database (CSD) performed in 1999 (Senthil Kumar *et al.*), only 25% of all reported crystals of small organic molecules were solvates. This observation was attributed in part to the increase in free energy of the structure if solvents were included (as a result of loss of entropy by including solvent molecules), particularly if scope for favourable intermolecular interactions exits in the parent structure. The formation and occurrence of hydrates, as with all other processes, results from a fine thermodynamic balance, i.e. compensation between enthalpy and entropy of the system. Generally, the formation of hydrates is governed by a net increase in favourable intermolecular interactions and the requirement of water is to satisfy specific roles in order to stabilise the crystal structure. Furthermore, the presence of water may serve to increase packing efficiency within the three-dimensional framework, thus maintaining a stable low

energy structure in accordance with the edict on the stability of the structure being related to density (Kitaigorodskii 1961).

In instances for which water plays a crucial role in maintaining the crystal structure via the formation of a hydrogen-bonding network, dehydration can often lead to complete structural collapse giving rise to an amorphous anhydrate, as observed with eprosartan mesylate dihydrate (Sheng *et al.* 1999). In this particular, case the water of crystallisation forms an hydrogen-bonding framework directly to the parent drug and the salt counter-ion. Dehydration results in an amorphous material, which becomes annealed upon heating giving rise to a crystalline anhydrate. Such hydrates are considered to be very stable and represent developable materials.

Hydrates in which water acts as a "space filler" occupying voids (often referred to as channel hydrates) can dehydrate to give isomorphous anhydrates or undergo a change of structure to give a more densely packed arrangement. Examples of these types of hydrates include cephalexin (Kennedy *et al.* 2003), erythromycin A and cefaclor (Stephenson 1997) all of which give rise to isomorphous anhydrates as determined by XRPD analysis. Generally, these types of hydrates are non-stoichiometric and the number of equivalent water molecules in the structure is directly related to the water activity in the surrounding environment. The geometry and size of the solvent channels in these structures can vary significantly from long, wide and rigid structures that are maintained by a robust hydrogen-bonded framework, to small interweaving arrangements for which the water may interact with the "host" structure. Dehydration from the long rigid channels results in minimal structural disruption and hence the resultant hydrate is structurally identical to the parent. Instances in which dehydration occurs from inter-weaving channels can give rise to disruption of the structure, creating a high-energy arrangement that may undergo relaxation in the form of anisotropic contraction of the lattice, evident by a shift in the XRPD lines from larger to smaller d-spacings, as observed in cephalexin (Stephenson *et al.* 2000). In both cases however, the parent anhydrate is regarded as an hygroscopic material. Typically, this category of hydrates is regarded as less stable and less desirable as a developable material.

## 1.4 Disorder in molecular solids

A distinguishing structural feature of crystalline solids is that they have a regular, long-range three-dimensional periodic ordering of atoms or molecules. Crystals are loosely

defined by static lattice positions, thus implying regular and defined rigid sites occupied by part of, or the entire molecule. Disordered molecular crystals are such that there is some degree of interruption in periodicity, to the extent that the arrangement of molecules in different unit cells may differ in some aspect from each other. Such a disruption of perfect periodic order can be separated into several types: orientational disorder (in which molecules or parts of molecules can assume several distinguishable orientations), translational disorder (in which there is periodic discontinuity in the translational distribution of molecules in the structure), and positional disorder (in which there are an excess of crystallographic sites in the structure capable of accommodating molecules or parts of the molecule). Furthermore, disorder in a structure can arise from either structural mobility/dynamics (in which the molecular arrangement in a given unit cell changes as a function of time), or static disorder. In the case of static disorder, molecules or parts of the molecules within a structure remain in fixed positions (over an appropriate timescale), while dynamic disorder refers to movement of all or part of the molecule between permitted positions and/or orientations within the structure. Such motions include rigid body torsions, librations, translations and rotations (Wright 1995, Harris and Aliev 1995).

Disorder, and in particular dynamic disorder, can contribute to the explanation of the nature of bulk properties and reactivity of materials. For instance, the flexibility and performance of a bulk polymer is highly dependant upon the flexibility and degrees of freedom of the constituent molecules in the structure, i.e. rigidity at the molecular level results in rigidity of bulk properties (Duer 2002). Furthermore, physical and chemical transformations can be related to the existence of disorder. For example, most solid-to-solid phase transitions arise from either the onset (as in the case of glass transitions) or quenching of molecular mobility. The chemical stability of a material can be related quite closely to the magnitude of molecular mobility. Mobility (or molecular loosening) increases the probability of collisions occurring between reactive functional groups, which increases the likelihood of chemical instability. Assessment of the configurational entropy barrier (i.e. the entropy difference between amorphous and crystalline states), for example, has shown that for amorphous materials with high entropy barriers, there is low molecular mobility and in turn these amorphous material are far more chemically stable (Zhou *et al.* 2002).

An understanding of the specifics of any disordered system, be it dynamic or static, is important in order to comprehend the nature of any bulk property or the propensity for physical and chemical transformations. The crystal structure and any disorder are

mutually dependent. By studying the geometry of the structure and rates of any motion, insights into the relationships between structure and properties can be obtained. Several methods can be employed to evaluate the mode and mechanism of disorder, including isothermal and thermal assessments using spectroscopy and diffraction techniques. By using an holistic approach, a complete assessment of the mode, mechanism and timescale of any disorder can be obtained.

Solid-state NMR spectroscopy is an excellent technique for investigating structural dynamics, primarily due to the fact that all nuclear spin interactions are anisotropic, i.e. they depend upon molecular orientations with respect to the applied magnetic field. Molecular motions are incoherent processes (Duer 2002) and as such can be described by autocorrelation functions. These are used to describe the speed of the molecular motions via measurements of correlation times ($\tau_C$ which are inversely related to the rate of the mobility). In the case of $^2H$ NNMR spectroscopy, for example, very slow motions (of the order of $\tau_C > 10^{-3}$ s) can be studied using two-dimensional exchange methods. Motions in the intermediate regime (where the correlation rates are of the order of the nuclear spin interaction isotropy) can be studied using powder lineshape methods, which are sensitive to motions with $\tau_C^{-1}$ of the order of the width of the powder pattern. Finally, fast motions ($\tau_C$ in the range $< 10^{-8}$ s) can be studied using spin relaxation studies. The specifics of these investigations are discussed in more detail in Chapter 2.

Diffraction studies can also be used to evaluate disorder, but are not as robust in differentiating between static and dynamic disorder. The structure obtained from diffraction studies represents a space and time averaged representation of the actual structure, with the refined atomic coordinates reflecting the mean atomic positions. Information can also be obtained on the probability density functions of the atomic displacements from the mean atomic positions. These are represented by atomic displacement parameters (ADP). ADP values are expressed as mean square displacements and include contributions from thermally induced and positional disorder and may be associated with occupation factors (a measure of the mole fraction of the different atomic arrangements in the disorder). Occupancy (or concentration) in most cases is highly dependant upon temperature. The use of diffraction data recorded at a single temperature can provide a measure of disorder, but cannot allow the distinction between mean square amplitude of motion and unresolved crystallographic disorder (Bürgi 2002). Single crystal diffraction experiments as a function of temperature can

be used to probe low frequency, large amplitude vibrations. Variable temperature single crystal diffraction studies and the impact on ADP measurements can be used to calculate vibration frequencies and the coordination of atomic motions (Trueblood and Dunitz 1983). Exploring dynamics using powder diffraction data however, is less straightforward and requires appropriate input models based on adjunct data such as spectroscopy. The input disorder structure can be used to generate a trial structure, which is then refined with appropriate ADP values that describe the extent of the disorder. As with single temperature single crystal studies, this methodology does not allow differentiation between static and dynamic disorder. Thus, the use of multiple techniques to understand and rationalise structure and disorder is clearly advantageous.

## 1.5 Objectives of this research

The primary objective of this research work is to understand the influence of molecular structure on crystal structure, and the occurrence of disorder. The ability to probe and understand these properties allows appropriate materials selection, a criterion critical in the Pharmaceutical industry. Polymorphic and hydrated materials (which represent pharmaceutically relevant molecular solids) are used to explore disorder and dynamics, and there relationship to crystal structure. Such investigations have the potential to provide a better understanding of the relative stabilities of these materials. Investigations carried out include spectroscopic and diffraction studies, and the use of fundamental thermodynamics.

## 1.6 References


- Bernstein J., (1993), *J. Phys. D: Appl. Phys.*, **26**, B66
- Bernstein J., (2002), *Polymorphism in Molecular Crystals*, IUCr monographs on crystallography 14, Clarendon Press Oxford
- Beyer T., Lewis T., Price S. L., (2001), *CrystEng. Comm.*, **44**, 1
- Borka L., Haleblain J. K., (1990), *Acta. Pharm. Jugosl.*, **40**, 71
- Burger A., Ramberger R., (1979), *Mikrochim. Acta*, **11**, 259 and 273
- Bürgi H-B., (2002), *Faraday Discussions*, **122**, 41
- Byrn S. R., Pfeiffer R. R., Stowell J. G., (1999), *Solid-state Chemistry of Drugs*, SSCI inc
- Cardew P. T., Davey, R. J., (1985), *Proc. R. Soc.*, **A398**, 415
- Cohen M. D., Green B. S., (1973), *Chem. Brit.*, **9**, 490
- Duer M. J., (2002), *Solid-state NMR Spectroscopy Principles and Applications*, Blackwell Science
- Gavezzotti A., (2002), *CrystEngComm.*, **4**, 343
- Giron D., (1995), *Thermochim. Acta*, **248**, 1
- Grover M., Singh B., Bakshi M., Singh S., (2000), *Research Focus PSTT*, **3**, 50
- Hansch C., Leo A., (1995), *Exploring QSAR*, American Chemical Society
- Harris K. D. M., Aliev A. E., (1995), *Chem. Brit.*, 132
- Kennedy A. R., Okoth M. O., Sheen D. B., Sherwood, J. N., Teat S. J., Vroelj R. M., (2003), *Acta Cryst.* **C59**, 650
- Kimura K., Hirayama F., Uekama K., (1999), *J. Pharm. Sci.*, **88**, 385

- Kitaigorodskii A. I., (1961), *Organic Chemical Crystallography*, Consultants Bureau New York
- McCrone W. C., (1965), *Physics & Chemistry of the Organic Solid State*, Volume 2 (Ed D. Fox, M. M. Labes and A. Weissberger), Wiley Interscience New York
- Mitchell B. E., Jurs P. C., (1998), *J. Chem. Inf. Comput. Sci.*, **38**, 200
- Neumann M., (2005), *J. Phys. Chem. B*, **109**, 15531
- Ostwald W. F., (1897), *Z. Phys. Chem.*, **22**, 289
- Roy S., Aitipamula S., Nangia A., (2005), *Cryst. Growth Des.*, **5**, 2269
- Sankaran H., Sharma S. M., Sikka S. K., Chidambaram R., (1986), *Pramana - J. Phys.*, **27**, 835
- Senthil Kumar V. S., Kuduva S. S., Desiraju G. S., (1999), *J. Chem. Soc. Perkin Trans 2*, 1069
- Shen J., Venkatesh G. M., Duddu S. P., Grant D. J. W., (1999), *J. Pharm. Sci.*, **88**, 1021
- Stephenson G. A., Groleau E. G., Kleemann R. L., Xu W., Rigsbee D. R., (1997), *J. Pharm. Sci.*, **87**, 536
- Stephenson G. A., Diseroad B. A., (2000), *Int. J. Pharm.*, **198**, 167
- Trueblood K. N., Dunitz J. D., (1983), *Acta Cryst.*, **B39**, 120
- Van der Sluis P., Kroon J., (1989), *J. Cryst. Growth*, **97**, 645
- Ward M. D., Bonafede S. J., Hillier A. C., (1997), *Modular Chemistry* (Ed. J. Michl), Kluwer Academic Publishers
- Wildfong P. L. D., Morley N. A., Moore M. D., Morris K. R., (2005), *J. Pharm. Biomed. Anal.*, **39**, 1
- Wright J.D., (1995), *Molecular Crystals*, Cambridge University Press
- Wunderlich B., (1999), *Thermochim. Acta*, **340 – 341**, 37

- Yu L., Stephenson G. A., Mitchell C. A., Bunnell C. A., Snorek S. V., Bowyer J. J., Borchardt T. B., Stowell J. G., Byrn S. R., (2000), *J. Am. Chem. Soc.*, **122**, 585

- Zhou D., Zhang G. G. Z., Law D., Grant D. J. W., Schmitt E. A., (2002), *J. Pharm. Sci.*, **91**, 1863

# 2. Introduction to Experimental Techniques

## 2.1 Molecular Spectroscopy

### 2.1.1 Solid State NMR Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy was invented in the 1940s and, within ten years of the first detection of signals, became established as a crucial technique for chemical structure analysis (Abraham *et al.* 1988). Notwithstanding the fact that NMR spectroscopy is most familiar as a high-resolution spectroscopic technique for the study of liquids and solutions, primarily to determine molecular structure, solid-state NMR has seen a steady increase in use. The initial lack of utility of solid-state NMR was attributed to an inherent deficiency in spectral resolution of a solid sample placed in a conventional high-resolution spectrometer suitable for liquids. The onset of ancillary techniques and sub-routines enabled higher resolution to be achieved, resulting in a proportional increase in the use and application of solid-state NMR spectroscopy (Duer 2002). However, solid-state NMR (SSNMR) spectra are composed of broader lines; $^{13}C$ NMR line broadening is still typically of the order of 30 - 60 Hz for a crystalline organic solid, as compared to typical solution $^{13}C$ NMR line-widths of only a few Hz. Nevertheless, it remains a very powerful technique for investigating the solid state and the resulting spectra are, in principle, far more information-rich than those of solutions.

In addition to gleaning information on molecular structure, SSNMR is recognised as an important adjunct to powder diffraction studies in the elucidation of crystal structure from powder data. For instance, this technique can be used to provide insights into the number of molecules present in the asymmetric unit and also information on site symmetry within the crystal structure. Furthermore, SSNMR can be used to assess the incidence and mechanism of dynamic disorder within the ordered solid-state - a property that is not amenable to exploration using diffraction studies (data from which represent a time averaged representation). Dynamic processes can be studied by SSNMR using, amongst other methods, the measurement of relaxation times as a function of temperature.

In the case of SSNMR it is necessary to consider certain important interactions encountered by a nuclear spin within a solid. These include dipole-dipole and quadrupolar interactions together with chemical shift anisotropy. These interactions, whilst also being

present in the solution state, are either averaged to zero (for the direct dipole-dipole and quadrupolar interactions) or averaged to their isotropic value (for the chemical shift and indirect spin-spin interaction). The reduction or complete elimination of these anisotropic interactions in the solution state is due to rapid isotropic molecular motion, which is restricted or absent in the solid state.

#### 2.1.1.1 General background and theory

##### *2.1.1.1.1 Nuclear spin interactions: Applied magnetic fields*

All nuclei possess a nuclear magnetic moment, defined as $\mu$, which can be described in terms of the nuclear spin (I) and the gyromagnetic ratio ($\gamma$) and shown in Eq 2.1.

$$\mu = \gamma I \qquad \text{Eq 2.1}$$

In the presence of an applied static magnetic field, all nuclear magnetic moments align with the applied field. This process is known as the Zeeman interaction. In a uniform static applied field (denoted $B_0$) the magnetic moment precesses about this fixed field at a constant frequency defined as the Lamor frequency denoted by $\omega_0$, as described by Eq 2.2.

$$\omega_0 = - \gamma B_0 \qquad \text{Eq 2.2}$$

The Zeeman interaction yields an overall net magnetisation (M). Furthermore, the application of an additional external field in the form of a radiofrequency (rf) wave with its oscillating magnetic field also interacts with the nuclei. The frequency of this rf field (denoted as $B_1$) is defined as $\omega_{rf}$. On applying an rf pulse, the vector of the precession of the net magnetisation is perturbed in accordance with the angle of pulse applied. Measurement of the perturbed system, and its return to equilibrium, gives rise to the NMR spectrum.

##### *2.1.1.1.2 Internal sources of magnetic fields*

The spin Hamiltonian for a solid, H, describing the interactions of the spins with the magnetic field $B_0$ and their surroundings is shown in Eq 2.3.

$$H = H_z + H_{rf} + H_{cs} + H_D + H_Q + H_{other} \qquad \text{Eq 2.3}$$

where $H_z$ is the Zeeman interaction, $H_{rf}$ is the radiofrequency Hamiltonian, $H_{cs}$ is the chemical shift, $H_D$ relates to the dipolar interactions and $H_Q$ is the quadrupolar interaction, which is

zero for nuclei with I = 1/2 but may dominate the Hamiltonian for nuclei with I > 1/2. The external applied static field ($B_0$) is much larger than any local internal magnetic fields, but intermolecular interactions such as dipolar and quadrupolar interactions and chemical shielding nevertheless, result in significant line broadening. For instance, when the local fields have components that are parallel or antiparallel to $B_0$, or have components that precess in a plane perpendicular to $B_0$ at or near the resonance frequency of the spin system, then an effect (as described above) is exerted on the spin state.

*2.1.1.1.2.1 The chemical shift and chemical shift anisotropy (CSA)*

A nucleus in a static magnetic field experiences a secondary field brought about by the electrons circulating around the nucleus. This field has a cumulative effect on the overall field observed at the nucleus and consequently changes the resonance frequency of that nucleus. This phenomenon is known as the chemical shift and is specific to the particular environment of the nucleus within a given molecule.

Furthermore, the electron distribution around a nucleus is not spherically symmetric and the chemical shift is therefore highly dependant upon the orientation of the molecule relative to $B_0$. The chemical shift (CS) is described by a tensor denoted as $\sigma$ (a 3 x 3 matrix) that shows how the size of the shielding varies with molecular orientation. The CS tensor can be represented by an ellipsoid centred on the nucleus in question. The axes of this ellipsoid represent a principal axis system (PAS) describing the orientation relative to $B_0$ or the laboratory frame, as illustrated in Figure 2.1 (Duer 2002). A change in the molecular orientation relative to the laboratory frame results in a change in the orientation of the ellipsoid. Thus, in a powder (as opposed to a single crystal), all molecular orientations are present, as are all values for $\theta$ (the Euler angle which relates the laboratory frame to the principal axis frame). The resulting powder pattern is composed of a series of overlapping lines (from these orientations) that cover a range of frequencies to give a continuous and characteristic lineshape. The resultant intensity at any given frequency is proportional to the number of orientations that have that particular chemical shift. As the magnitude of the CSA is proportional to $B_0$, an increase in the magnetic field merely serves to increase the magnitude of the CSA, giving rise to broader lineshapes.

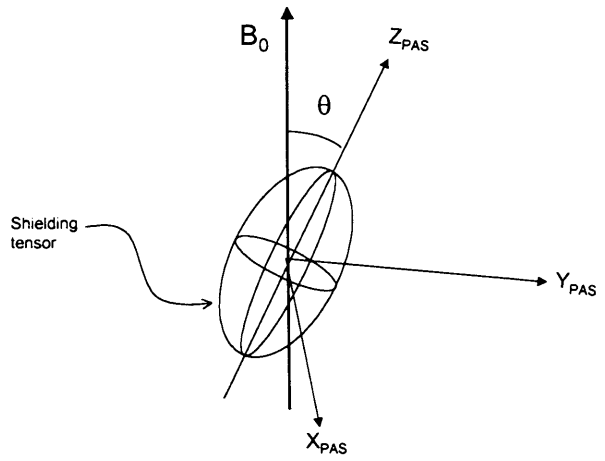



Figure 2.1: Definition of the principal axis system of the shielding tensor ($\sigma$) relative to the laboratory frame ($B_0$)

The principal values of the shielding tensor are described as $\sigma_{XX}^{PAS}$, $\sigma_{YY}^{PAS}$ and $\sigma_{ZZ}^{PAS}$. These values describe the main components of the lineshape, that is the isotropic value $\sigma_{iso}$, the anisotropy value $\Delta$ and the asymmetry function $\eta$ as described in Eq 2.4 to 2.6 respectively.

$$\sigma_{iso} = \frac{1}{3}(\sigma_{XX}^{PAS} + \sigma_{YY}^{PAS} + \sigma_{ZZ}^{PAS}) \qquad \text{Eq 2.4}$$

$$\Delta = \sigma_{ZZ}^{PAS} - \sigma_{iso} \qquad \text{Eq 2.5}$$

$$\eta = \frac{(\sigma_{XX}^{PAS} - \sigma_{YY}^{PAS})}{\sigma_{ZZ}^{PAS}} \qquad \text{Eq 2.6}$$

If the nucleus is coincident with a crystallographic site of symmetry then the shielding tensor will reflect this. In this case, the $Z_{PAS}$ axis will coincide with the axis of symmetry and the principal values will be as described in Eq 2. 7.

$$\sigma_{XX}^{PAS} = \sigma_{YY}^{PAS} \neq \sigma_{ZZ}^{PAS} \qquad \text{Eq 2.7}$$

The powder lineshape is very characteristic and reflects the site symmetry at the nucleus in question; therefore, this can be exploited by extracting symmetry information at the site of the nucleus. The lineshape for a powder sample has discontinuities that relate to the principal values of the shielding tensor.

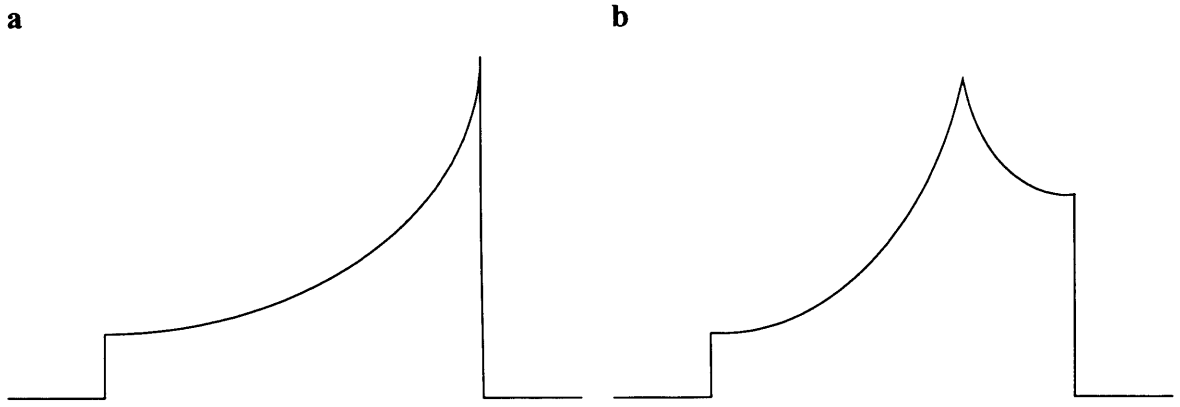



Figure 2.2: Chemical shift anisotropy powder patterns of samples with (a) axial symmetry and (b) non-axial symmetry

In order to obtain orientation information and details of site symmetry, static CSA powder lineshape experiments are conducted and the experimental data compared to simulated lineshapes calculated using the program SOLIDS (Eichele and Wasylishen 1994). As discussed in the preceding section, the principal components of the PAS of the chemical shielding tensor are used to calculate the powder lineshape using the expression shown in Eq 2. 8. For powder samples all possible orientations of the polar angles $\theta$ and $\phi$ (the polar angle between $X^{PAS}$ and $Y^{PAS}$ and $B_0$) contribute to the observed pattern.

$$\nu_{CS} = (\frac{\gamma_I B_0}{2\pi})[1 - (\sigma_{xx} \sin^2\theta\cos^2\varphi + \sigma_{YY} \sin^2\theta\sin^2\varphi + \sigma_{ZZ} \cos^2\theta)] \quad \text{Eq 2.8}$$

*2.1.1.1.2.2 The direct dipolar interaction*

As described earlier, nuclear spins with $I > 0$ possess a magnetic moment, and the interaction of one nuclear spin with another is described as the dipolar interaction or coupling. The type, distance and orientation of one nuclear spin and another govern the magnitude and nature of this interaction. The general Hamiltonian $H_D$ representing the interaction between two magnetic moments $\mu_I$ and $\mu_S$ is shown in Eq 2. 9.

$$H_D^{IS} = -\gamma_I \gamma_S \hbar \sum_{i,j} r_{i,j}^{-3} I_z^i S_z^j (3\cos^2\theta_{i,j} - 1) \quad \text{Eq 2.9}$$

where $\theta_{ij}$ is the angle between the static magnetic field and the internuclear vector $r_{i,j}$ as shown in Figure 2.3 (Duer 2002).

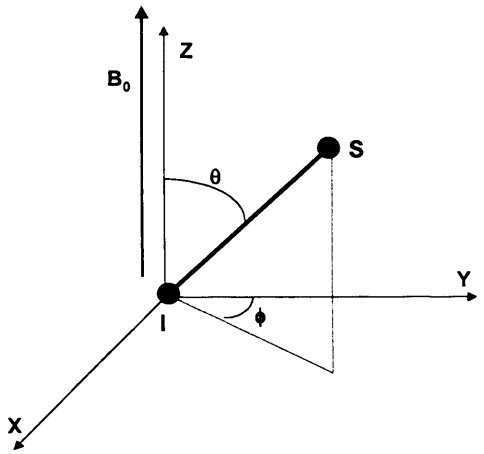



Figure 2.3: Definition of orientation of the inter-nuclear vector relative to the applied magnetic field showing the polar angles θ and ϕ

In a static sample such as a solid, the magnetic moments interact directly, producing a relatively large dipolar interaction. The widths of the lines of the resultant spectra are of the order of the magnitude of this interaction, which could potentially render spectra featureless. In more mobile samples, such as liquids and gases, these interactions are averaged to zero by the presence of rapid molecular tumbling. As a direct consequence, the resultant line widths are significantly narrower. Dipolar interactions in general can be separated into two types: homonuclear and heteronuclear couplings. As alluded to in Eq 2.9, the magnitude of the interaction (and hence the linewidth) is directly related to the product of the nuclear magnetic moments in question, and thus a $^{13}C$-$^{13}C$ interaction is considerably smaller than a $^{13}C$-$^{1}H$ interaction, which in turn is smaller than an $^{1}H$-$^{1}H$ interaction (for the same internuclear distance). Furthermore, homonuclear coupling in $^{13}C$ spectra is not as prevalent, primarily due to the relatively low natural abundance of $^{13}C$.

For many nuclei, such as $^{1}H$ and $^{19}F$, the dipolar interaction dominates, rendering the interpretation of spectra and relaxation behaviour difficult, since dipolar coupling is a multi-spin interaction with both inter and intramolecular contributions. The result is often a broad, almost featureless spectrum from which very little information can be extracted. The effects of dipolar coupling are largely removed by employing techniques such as high-power decoupling or magic angle spinning, both of which will be discussed in section 2.1.1.2.

*2.1.1.1.2.3 The quadrupolar interaction*

All nuclei with spin I > 1/2 possess a quadrupole moment that is caused by a non-spherical charge distribution of the nucleus. The quadrupole interaction is a coupling between

the electric quadrupole moment of the nucleus and the electric field gradient (EFG) present at the nucleus. The EFG is a tensor, present in all solids at all nuclei, and the magnitude and nature of the interaction is governed in part by any site symmetry present at the nucleus. As with dipolar coupling, the quadrupolar interaction results in varying degrees of line broadening depending upon the magnitude of this interaction. Further discussion of this interaction and its uses and applications is included in section 2.1.1.3.

#### 2.1.1.2 High resolution techniques for spin I = 1/2 nuclei

Small organic compounds contain inequivalent nuclear sites and would give rise to severe peak overlap if the NMR spectrum were recorded under conditions used for conventional solution state NMR, due to the severe line broadening brought about by the anisotropic interactions discussed earlier. The overall effect on the spectra is to give rise to a diffuse profile devoid of meaningful information. In order to enhance resolution to give almost liquid like spectra, these anisotropic interactions can be removed or minimised by a series of experimental techniques discussed in this section.

Low abundance spins such as $^{13}C$ and $^{15}N$ when observed unedited exhibit low signal-to-noise ratios and, generally speaking, long relaxation times. The lengthy relaxation times are due to the absence of strong homonuclear dipolar interactions that serve to enhance the relaxation behaviour. Additionally, in the absence of molecular mobility, the above effects result in long acquisition times and poor signal intensity. To overcome this issue, a technique involving polarisation transfer from the dilute to a more highly polarised and abundant spin such as $^{1}H$ is employed. The specifics of this technique are discussed in the next section.

##### *2.1.1.2.1 Cross polarisation*

The process of polarisation transfer between abundant and dilute spins is termed cross polarisation. Pines and co-workers (1973) demonstrated that by simultaneously irradiating a sample with two radio-frequency fields, for a period referred to as the contact time, at the Lamor frequencies of spin I and spin S a flow of magnetisation from the abundant I spins to the dilute S spins can occur. This cross polarisation sequence is summarised in Figure 2.4.

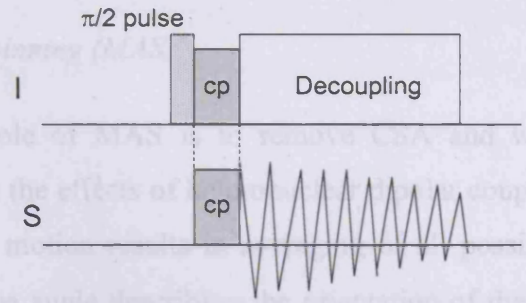



Figure 2.4: The cross polarisation pulse sequence where cp denotes the contact pulse

Polarisation transfer is permitted if the Hartmann-Hahn condition (as described by Eq 2.10) is satisfied.

$$\gamma_I B_{1I} = \gamma_S B_{1S} \qquad \text{Eq 2.10}$$

where $\gamma_I$ and $\gamma_S$ are the gyromagnetic ratio for spins I and S, and $B_{1I}$ and $B_{1S}$ are the rf fields applied for each type of spin. The process involves a $\pi/2$ pulse being applied to the abundant I spins to create transverse magnetisation. A second pulse, phase shifted by $\pm \pi/2$, is then applied to the I spins for the duration of the contact time to spin-lock the I spins. Simultaneously, a spin-lock field is applied to the S spins and polarisation or magnetisation transfer can then take place during the contact time if $B_{1I}$ and $B_{1S}$ satisfy the Hartmann-Hahn condition. As the magnetisation associated with spin I is larger than that associated with spin S, then cross relaxation from the I spins to the S spins will result in an increase in the S magnetisation. The S signal can be enhanced by a factor equal to $\gamma_I/\gamma_S$ (a factor of 4 for I = $^1$H and S = $^{13}$C) compared with simple direct excitation for S. Furthermore, because the S signal is generated by cross-polarisation from the I spins, the rate at which the experiment may be repeated is determined by the spin-lattice relaxation rate of the I spins rather than that of the S spins.

As almost every small organic molecule contains protons (which are highly polarised) the magnetisation transfer is mediated efficiently between the $^1$H spins and the low abundant X spins (e.g. $^{13}$C). The rate of this magnetisation depends upon the strength of the $^1$H and X dipolar coupling.

*2.1.1.2.2 Magic angle spinning (MAS)*

The primary role of MAS is to remove CSA and weak homonuclear dipolar coupling, and to minimise the effects of heteronuclear dipolar coupling. In the solution state, rapid isotropic molecular motion results in averaging of all possible molecular orientations (and hence values of θ, the angle describing the orientation of the CS tensor with respect to $B_0$). This results in the angular dependence of transition frequencies ($3\cos^2\theta - 1$) averaging to zero. As discussed previously, in the solid state molecular reorientation is either non-existent or, if it does occur, it is usually anisotropic resulting in substantial signal broadening. In 1959, Lowe and Andrew independently proposed that by spinning the solid sample about an axis inclined at an angle of 54°44' with respect to $B_0$ then the term ($3\cos^2\theta - 1$) averages to zero in accordance with the angular dependence of the anisotropic parameters as shown in Eq 2.11

$$\langle 3\cos^2\theta - 1\rangle = \frac{1}{2}(3\cos^2\theta_R - 1)(3\cos^2\beta - 1) \qquad \text{Eq 2.11}$$

where the angles θ, $\theta_R$ and β (as shown in Figure 2.5) represent the angle between the z-axis of the interaction tensor and $B_0$, the angle defining the spinning axis relative to $B_0$, and the angle between the z-axis of the interaction tensor and the spinning axis, respectively (Duer 2002).

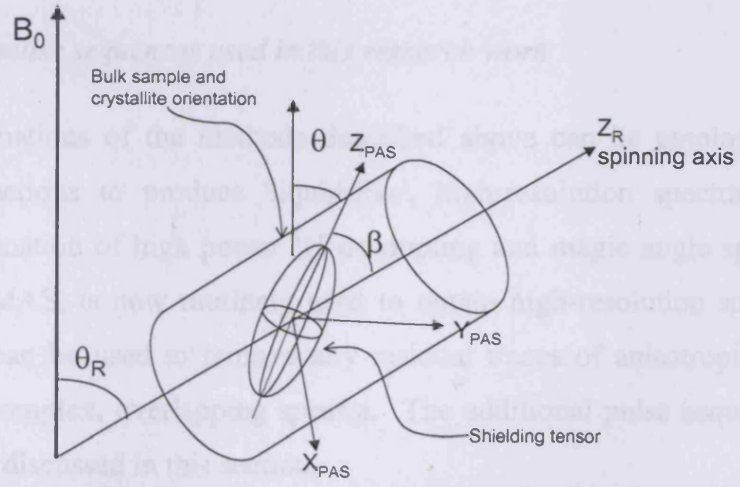



Figure 2.5: The process of MAS

The CSA component is removed completely by spinning the sample in a rotor at the magic angle ($\theta_R$ = 54°44') and at a frequency that is greater than the magnitude of the anisotropy, whilst the dipolar broadening is reduced depending upon the magnitude of the

dipolar coupling. If the CSA is greater than the achievable spinning frequency, a series of satellite peaks are observed, known as spinning side bands, which require additional pulse sequences to remove them completely (e.g. the TOSS sequence which is discussed later).

#### 2.1.1.2.3 *High power decoupling*

Analysis of low abundant spins such as $^{13}C$ with closely associated $^{1}H$ (or other abundant) spins provides a weak broadened signal. As already discussed in Section 2.1.1.2.2, MAS fails to completely eliminate the effects of heteronuclear dipolar coupling in such cases. Consequently, a pulse sequence is introduced in which a continuous high power field at the resonance frequency of the abundant spin is applied whilst the FID for the less abundant spin is acquired. Figure 2.6 represents graphically the sequence involved in this technique. For most small organic compounds high-power proton decoupling is used, where the decoupling field is typically of the order of around 40 – 50 kHz and is used to decouple the $^{1}H$ from the $^{13}C$.

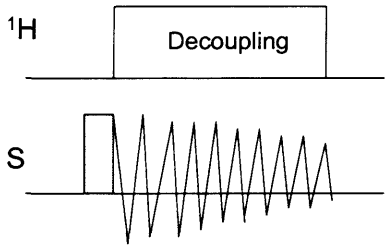



Figure 2.6: High power proton decoupling pulse sequence

#### 2.1.1.2.4 *Other pulse sequences used in this research work*

Combinations of the methods described above can be employed to remove all anisotropic interactions to produce 'liquid-like', high-resolution spectra. With $^{13}C$, for example, a combination of high power $^{1}H$ decoupling and magic angle spinning with cross polarisation, CP/MAS, is now routinely used to obtain high-resolution spectra. Additional pulse sequences can be used to remove any residual traces of anisotropic interactions and further simplify complex, overlapping spectra. The additional pulse sequences used in this research work are discussed in this section.

*2.1.1.2.4.1 Total side band suppression (TOSS) sequence*

As a result of the anisotropy of the molecular orientation to $B_0$, the spinning speeds (MAS) required to reduce the broadened peaks to a single isotropic peak must be at a frequency higher than the interaction itself. For some interactions, such as CSA, spinning at a slower rate gives rise to additional satellite peaks known as spinning side bands. These side bands radiate out from either side of the main isotropic peak and are separated by a frequency equal to the spinning speed. The two main ways to eliminate these satellite peaks are firstly to spin the sample faster (which may be restricted by the magnitude of the CSA and the instrumentation available) or secondly to use an additional pulse sequence known as the Total Side band Suppression (TOSS) sequence first introduced by Shaefer, Sefick and Stejskal in 1982. The sequence (represented graphically in Figure 2.7) consists of a 90° pulse followed by a series of four 180° pluses at specific points prior to FID acquisition. In some instances, when there is a significant number of spinning side bands, TOSS may be unable to remove them all efficiently leaving some residual bands with phase distortions.

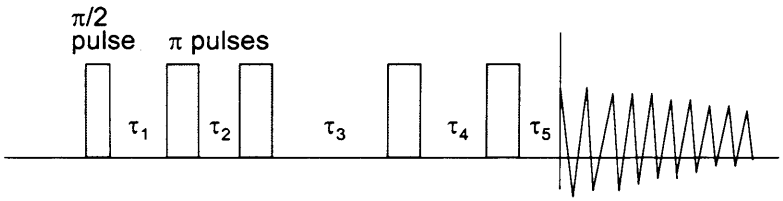



Figure 2.7: TOSS pulse sequence used to remove spinning side bands, τ denotes a rotor period

*2.1.1.2.5 Non-quaternary suppression (NQS)*

The NQS, or dipolar dephasing, pulse sequence is used to simplify solid-state NMR spectra of complex structures, providing a distinction between S spins that experience strong heteronuclear dipolar interactions from those that do not experience these interactions. Using $^{13}C$ NMR as an example, a clear distinction between quaternary and non-quaternary carbons can be made, by allowing transverse magnetisation on carbons directly attached to protons to decay rapidly by exploiting the strong dipolar coupling. The technique is shown schematically in Figure 2.8, where ct denotes contact time. The resultant spectra consist largely of lines corresponding to the quaternary carbons. In some instances, signals from non-quaternary carbons may be detected when using this technique, due to the occurrence of molecular motion, which serves to weaken the dipolar coupling for these non-quaternary

carbons. A common example includes methyl carbons, for which dipolar coupling is averaged by rapid methyl group reorientation.

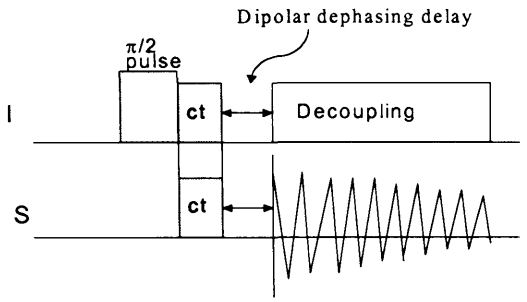



Figure 2.8: NQS pulse sequence

### 2.1.1.3 Techniques for spin I = 1 nuclei

All nuclei with spin I > 1/2 possess a nuclear quadrupole moment (as introduced in section 2.1.1.1.2.3) by virtue of a non-spherical nuclear charge distribution. Deuterium ($^2$H) possesses a spin I = 1 and thus exhibits a quadrupolar interaction, giving rise to two important effects. Firstly, the spectra may be split into several lines by this interaction and, secondly, modulation of the quadrupole interaction by molecular motion provides a very efficient relaxation mechanism. The following sections summarise the theory and applications of $^2$H NMR spectroscopy.

#### *2.1.1.3.1 $^2$H quadrupolar interaction*

The quadrupolar interaction dominates other anisotropic interactions; consequently, the total Hamiltonian may be approximated by Eq 2.12, since the terms relating to dipolar interactions, chemical shift anisotropy and spin-spin coupling may essentially be neglected

$$H = H_Z + H_Q \qquad \text{Eq 2.12}$$

where $H_Z$ refers to the Zeeman Hamiltonian and $H_Q$ refers to the quadrupolar Hamiltonian.

The basic form of the quadrupolar Hamiltonian is described in Eq 2.13

$$H_Q = \frac{eQ}{6I(2I-1)\hbar} I.V.I \qquad \text{Eq 2.13}$$

where I is the nuclear spin vector, V is the electric field gradient (EFG) tensor at the nucleus and Q is the nuclear quadrupole moment. Nuclei with spin I = 1 have three quantised (Zeeman) energy states in the presence of an external static magnetic field. The interaction of the quadrupole moment with the electric field gradient at the nucleus results in a perturbation of the Zeeman levels as shown in Figure 2.9 (Jelinski 1985).

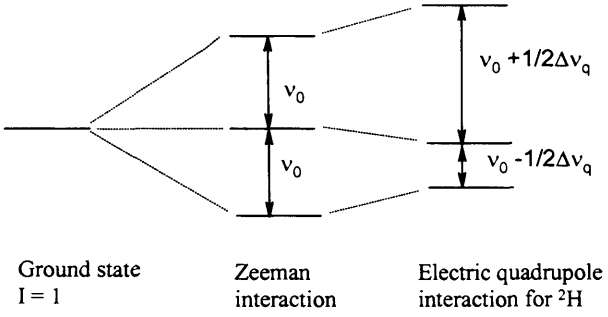



Figure 2.9: Schematic of the quantised energy states

The frequency difference, $\Delta\nu$, between the two transitions between adjacent spin states is then given by the expression shown in Eq 2.14

$$\Delta\nu = \frac{3}{4}\left(\frac{V_{ZZ}e^2Q}{\hbar}\right)(3\cos^2\theta - 1 - \eta\sin^2\theta\cos 2\psi) \qquad \text{Eq 2.14}$$

where $\chi = e^2V_{ZZ}Q/h$ is the quadrupole coupling constant and $\eta$ is the asymmetry parameter of the electric field gradient tensor V. In the principal axis system (PAS) the electric field gradient (EFG) tensor is diagonal with components $|V_{ZZ}| \geq |V_{YY}| \geq |V_{XX}|$. In the case of a C-D or O-D bond, the principal component $V_{ZZ}$ is approximately collinear with the bond direction. The asymmetry parameter, which represents a measure of variation from axial symmetry, is represented in Eq 2.15 (Aliev *et al.* 1998).

$$\eta = \frac{(|V_{YY}| - |V_{XX}|)}{|V_{ZZ}|} \qquad \text{Eq 2.15}$$

The EFG tensor is traceless, and as such $0 \leq \eta \leq 1$. The angle $\theta$ defines the orientation of $V_{ZZ}$ and $\psi$ is the angle defining the orientation of $V_{YY}$ and $V_{XX}$ with respect to the static magnetic field $B_0$, as shown in Figure 2.10.

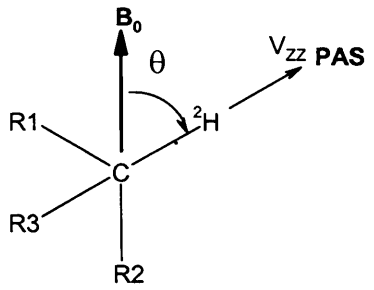



Figure 2.10: Schematic of the quadrupole interaction

where $B_0$ represents the external magnetic field, PAS is the principal axis system and, $R_1$, $R_2$ and $R_3$ represent functional groups.

The $^2H$ NMR spectrum of a single crystal containing one deuteron site in the unit cell is composed of a sharp symmetrical doublet, equally spaced about zero frequency; the splitting within the doublet depends (via $\theta$ and $\psi$) upon the orientation of the crystal with respect to the applied magnetic field $B_0$. A typical spectrum for a single crystal is shown in Figure 2.11 (Facey *et al.* 1996).

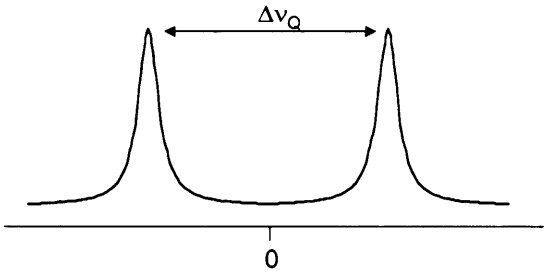



Figure 2.11: Spectrum for a single crystal with one orientation of the X-D bond in the unit cell

The spectrum of a polycrystalline powder sample, for which all orientations of crystals are represented, is therefore a sum of doublets arising from all orientations ($\theta$, $\psi$). In general, the spectra are symmetric about the centre of mass with three pairs of singularities separated by the frequencies $\Delta_1$, $\Delta_2$ and $\Delta_3$, where $\Delta_3 > \Delta_2 > \Delta_1$ and represented by Eq 2.16 – 2.18.

$$\Delta_1 = \frac{3}{4}\chi(1-\eta) \qquad \text{Eq 2.16}$$

$$\Delta_2 = \frac{3}{4}\chi(1+\eta) \qquad \text{Eq 2.17}$$

$$\Delta_3 = \frac{3}{2}\chi \qquad \text{Eq 2.18}$$

The powder pattern, also known as the Pake pattern, for a static sample with axial symmetry is shown in Figure 2.12.

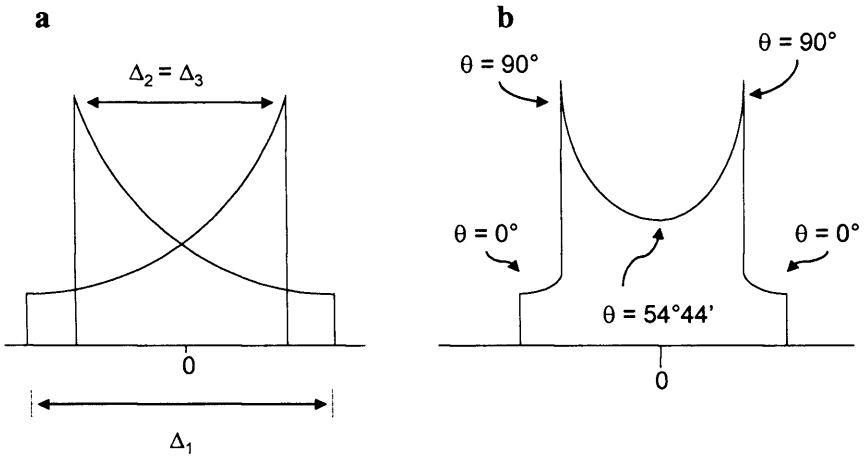



Figure 2.12: Powder or Pake pattern for a static sample where $\eta = 0$, showing (a) the lineshapes arising from each of the transitions, and (b) the composite spectrum detailing the contributions from different crystal orientations

The quadrupolar interaction is extremely sensitive to molecular motions, which give rise to an averaged EFG tensor. This in turn modifies the shape of the powder pattern in accordance with the rate and mechanism of the motion observed at the site of the deuteron. As the motion causes averaging of the lineshape in a specific way, information on mechanism and rate can be extracted from assessment of the lineshape. Solid-state $^2H$ NMR spectra are sensitive to local rotational motions with frequencies comparable to or greater than *ca.* $10^3$ $s^{-1}$. If the rates of reorientation are between *ca.* $10^3$ $s^{-1}$ and $10^8$ $s^{-1}$, known as the intermediate motion regime, the $^2H$ NMR lineshape is extremely sensitive to the rate and geometry of the molecular motion. If the rates of reorientation are greater than *ca.* $10^8$ $s^{-1}$, known as the fast motion regime, the lineshape is independent of the rate of reorientation but information can still be obtained on the geometry of the motion. Furthermore, the modulation of the quadrupole interaction by rotational molecular motion also dominates relaxation, and hence the dynamic range over which molecular motion may be studied can be extended up to around $10^{12}$ $s^{-1}$, or greater in some cases, by measurement of spin-lattice relaxation times.

### 2.1.1.3.2 *The Quadrupole-Echo technique*

$^2$H NMR spectra are composed of extremely broad lines, of the order of up to 350 kHz. Such broad lines are associated with a rapidly decaying FID (typically of the order of tens of microseconds). In conventional pulse sequences there is a requirement to have a delay, known as the dead time, between the application of an rf pulse and the data acquisition. In the case of rapidly decaying FID, this delay results in loss of a significant proportion of the signal. To overcome this issue, a specific quadrupole echo pulse sequence was introduced by Davis and co-workers (1976). This pulse sequence (shown schematically in Figure 2.13) involves the application of a $\pi/2$ pulse followed by a time delay $\tau$ and a second phase-shifted $\pi/2$ pulse, which causes the spins to be refocused at time $\tau' \approx \tau$. The length of $\tau$' is adjusted such that data acquisition begins at the echo maximum.

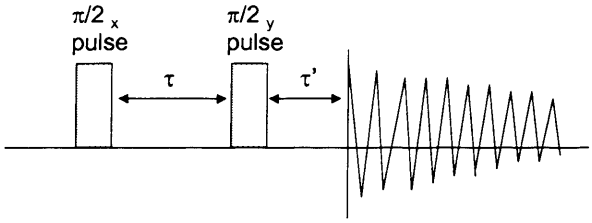



Figure 2.13: Quadrupole echo pulse sequence

### 2.1.1.3.3 *Relaxation time measurements using the saturation recovery technique*

The spin-lattice relaxation time, $T_1$, can be measured using two experiments - saturation recovery or inversion recovery, in which spectra are recorded as a function of relaxation delay with quadrupole echo detection. In this research work, the saturation recovery technique was used, and is described in this section.

This technique, as devised by Look and co-workers (1969), involves the saturation of the spin system by a series of pulses. Once saturated, the quadrupole echo is acquired over a range of relaxation delay ($\tau_r$) periods. The sequence shown schematically in Figure 2.14 was used.

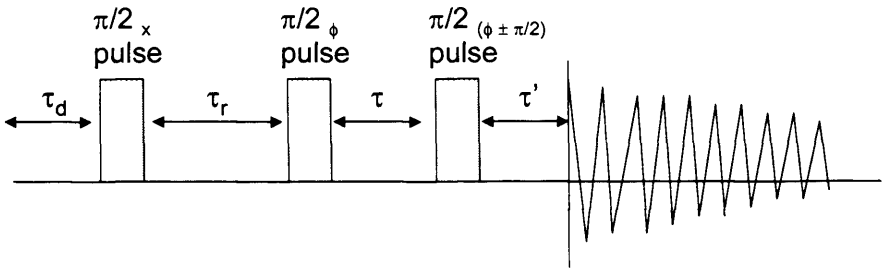



Figure 2.14: Saturation recovery pulse sequence

The variable relaxation delay $\tau_r$ is increased geometrically according to $\tau_r = t_0[(10)^{1/10}]^{(N-1)}$, where N is the delay number and $t_0$ is the value of the first delay (i.e. $\tau_r = t_0$ when N = 1). A recovery curve is obtained by plotting the signal intensity against the relaxation delay. Relaxation times are determined by fitting the recovery curves to multiple exponential recovery functions.

#### *2.1.1.3.4 Chemical exchange and $^2H$ line shape analysis*

The $^2H$ quadrupolar interaction is sufficiently sensitive to molecular motion that information on the rate and mechanism of motion at the deuteron site can be extracted from the analysis of the $^2H$ NMR powder pattern. Motions with correlation times in the range $10^{-3}$ to $10^{-8}$s (the intermediate motion regime) give line-shapes that depend critically on the rate and mechanism of the motion. For molecular motions on timescales less than *ca.* $10^{-8}$s (the fast-motion regime), an average spectrum is observed and details of the exact rate of the motion cannot be established, although line-shape analysis can still provide detailed information on the geometry and mechanism of the motion. Finally, if molecular motions are slow relative to the $^2H$ NMR time-scale with the motional timescale more than $10^{-3}$ s (slow-motion regime) then a static $^2H$ NMR spectrum is obtained, irrespective of the rate and geometry of the motion.

In the present work, the mechanism and rates of motion were established by comparing spectra obtained from simulating different motional models with the experimental data. Simulations were performed using the MXQET program (Greenfield *et al.* 1987). The motional models are defined as Markov type instantaneous jumps between N sites. The sites are in turn defined by rotations of the structure/functional group through specified angles (termed Euler angles) that describe the orientation (relative to a space fixed reference frame) of the principal-axis system of the electric field gradient tensor ($V^{PAS}$) at the deuteron site. In

this instance, the z-axis of the $V^{PAS}$ is assumed to lie along the direction of the X-D bond. The components of $V^{PAS}$ are set such that $|V_{ZZ}| \geq |V_{YY}| \geq |V_{XX}|$. Furthermore, the values of the asymmetry parameter $\eta$ and the quadrupole-coupling constant $\chi$ are also varied to provide an appropriate model in the simulations. The motional model contributes to the calculation of the time evolution of the total transverse magnetisation of the deuteron; this leads to the calculation of the simulated FID.

### 2.1.2 Fourier transform infra-red (FTIR) and Raman spectroscopy

Vibrational spectroscopy techniques such as FTIR and Raman spectroscopy provide information on quantised vibrational energy levels in a molecule. Subjecting a sample to monochromatic incident radiation (of frequency $\nu_0$) results in a perturbation of the system, for example by absorption (in the case of FTIR spectroscopy) or scattering (in the case of Raman spectroscopy) of the incident radiation as summarised in Figure 2.15.

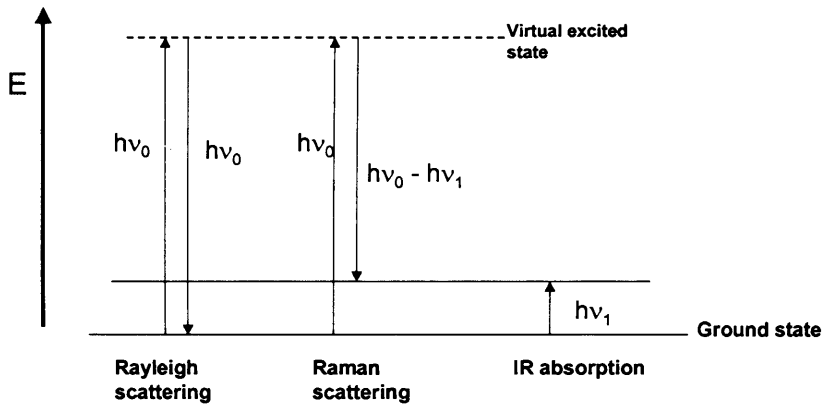



Figure 2.15: Schematic diagram of the spectroscopic transitions underlying IR and Raman spectroscopy

The virtual state represented in Figure 2.15 is a transient distortion of the electron distribution by the electric field of the incident radiation at an energy state lower than the first electronic excited state (McCreery 2000). Rayleigh scattering involves elastic scattering of the incident light; i.e. the light scattered from a sample has the same frequency as the incident radiation, while Raman scattering is inelastic.

If a coil or a spring, as in the classical simple harmonic motion, represents a bond in a molecule, then displacement of the bond (coil) from an equilibrium state requires a force with a proportional force constant (Hendra *et al.* 1991). Generally, a stronger bond has a larger force constant, and a weaker bond has a proportionally smaller force constant. Overall

a spectrum therefore consists of a set of bands of force constants corresponding to the displacements of bonds within a molecule. The spectral frequencies can be calculated from the expression given in Eq 2.19.

$$\Delta E = h\upsilon_{vib} = \frac{h}{2\pi}\sqrt{\frac{k}{\mu}} \qquad \text{Eq 2.19}$$

Therefore, the spectrum provides bond-specific characteristic frequencies giving rise to a molecular fingerprint. The spectra are expressed in terms of reciprocal wavelength ($\lambda$), that is wavenumber ($\upsilon$). IR and Raman spectroscopy differ in the way the excitation radiation interacts with the bonds in a molecule, giving rise to different excitation mechanisms, and hence to different selection rules. Thus, the two techniques provide complementary information on the molecular vibrations.

In the case of IR spectroscopy, the absorption spectrum arises from the coupling of the incident radiation with a specific type of motion associated with the bonds in a molecule. For a vibration to be IR active (i.e. for a strong absorption band to be observed), the vibration must produce an oscillating dipole moment, which interacts with the oscillating electric field of the incident radiation, provided the frequency of the incident radiation equals that of the vibration. Homo-atomic bonds (such as C-C bonds) possess a poor dipole moment and hence either give rise to a weak absorption band or are IR inactive. This technique therefore provides information on mainly hetero-atomic interactions and is useful in extracting details such as hydrogen bonding arrangements. IR spectroscopy using standard FTIR spectrometers allows information to be extracted from 4000 – 500 $cm^{-1}$.

Raman, however, relies on molecular polarisability; a bond must be anisotropically polarisable in order for the vibration of the bond to be Raman active. The distortion of a bond in an electric field gives rise to an induced dipole, the magnitude of which is determined by the extent of polarisability. As this polar state is more energetic than the relaxed state, spontaneous relaxation is accompanied by a release of energy. It is this relaxation or emission of radiation that is termed scattering. A high degree of polarisability (such as that found in most homo-atomic bonds) gives rise to Raman active bands. Furthermore, the Raman spectrometer design allows information to be obtained in the frequency region 4000 – 100 $cm^{-1}$, thus allowing lattice vibrational modes (i.e. translation and rotation motions of the entire molecule within the crystalline lattice) to be evaluated (Bugay 2001).

## 2.2 X-ray powder diffraction (XRPD)

### 2.2.1 Theory and background to diffraction from crystalline samples

X-rays are electromagnetic radiation with wavelengths of the order of approximately 100 pm. Generally they are produced from rapid bombardment of metals (such as copper) with high-energy electrons. This bombardment results in emission of excess energy in the form of x-rays. Diffraction of x-radiation by crystals was first discovered in 1912 by Von Laue and others (e.g. Bragg and Bragg), and subsequently by Debye and Scherrer in 1916 and Hull in 1917 (Azaroff and Buerger 1958). The principle behind the method relies on the phenomenon of interference of light in the form of waves scattered from different points in a sample. The incident radiation (in the form of a beam of monochromatic x-rays) penetrates a sample and interacts with the electron cloud associated with each atom. This interaction leads to a scattering process, and the interference of x-rays scattered from parts of the sample gives rise to the phenomenon of diffraction.

A crystalline sample can be represented as a three-dimensional periodic array of atoms or molecules. The periodic repeat unit is called the "unit cell". Within the unit cell, atoms or molecules may be related to each other by a set of symmetry operations called the space group. Diffraction from a crystalline sample is based on the ability of crystals to diffract radiation for which the wavelength is similar to the periodic distances, analogous to the diffraction of visible light by an optical grating. Hence, a crystal behaves as a three-dimensional grating capable of diffracting appropriate radiation, such as x-rays, neutron beams or electron beams.

A particular crystal structure generates a unique and characteristic diffraction pattern. Bragg defined a crystal as being composed of reflecting planes, such that the lattice is composed of layers of these planes. This simple description of the geometry of crystals and the geometry of the diffraction process gave rise to Bragg's Law. The basis of the geometry of diffraction (as represented in Figure 2.16) is that each plane reflects or diffracts part of the beam, with the angle of reflection equal to the angle of incidence. As there are many parallel planes involved in the scattering process, reflections from successive planes interfere with each other. Constructive interference (as shown in Figure 2.16) is the term used when the difference in path length between rays reflected from successive planes is an integral number of wavelengths, as described by Eq 2.20 (used to calculate the positions of the diffraction

intensity maxima). The intersections of a plane with three crystallographic axes may be expressed as multiples of the unit cell with dimensions a, b and c. The reciprocals of these intercepts represent the Miller indices h, k and l, which are used to label each set of lattice planes and may also be used to label the reflections. (The reciprocal lattice has unit cell dimensions a*, b* and c*).

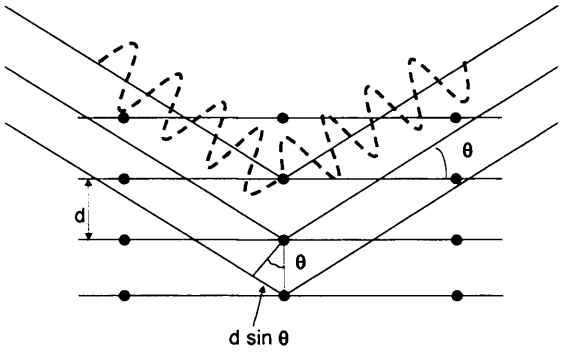



Figure 2.16: Graphical representation of Bragg's Law describing the geometry of diffraction from crystal planes

$$n\lambda = 2d \sin\theta \qquad \text{Eq 2.20}$$

In a crystalline sample, scattering occurs from atoms or molecules, and since each atom type has a specific scattering factor, the scattering amplitude is established by summing contributions from scattering from electrons in all atoms in the molecules in the unit cell. This summation also takes into account phase differences that arise from the fact that different atoms are located in different positions in the unit cell and is expressed as the structure factor $F_{hkl}$ described in Eq 2.21

$$F_{hkl} = \sum_{j}^{N} f_j \exp(2\pi i[hx_j + ky_j + lz_j]) \exp(-2\pi^2 U_j [\frac{\sin^2\theta}{\lambda^2}]) \qquad \text{Eq 2.21}$$

where N is the number of atoms in the unit cell, $f_j$ is the scattering amplitude for atom j, $x_j$, $y_j$ and $z_j$ are the fractional coordinates of atom j within the unit cell, and $U_j$ is the displacement parameter of atom j (which represents its average distribution around the point ($x_j$, $y_j$, $z_j$). The displacement parameter can be considered either as isotropic or anisotropic.

The intensity of diffracted radiation by a set of planes with Miller indices (hkl) is proportional to the square of the amplitude of its structure factor $F_{hkl}$. The structure factor characterises completely the amplitude and phase of each diffracted beam, and is required to calculate electron density maps from which the positions of the atoms can be determined. For

this reason, the intensities measured in a diffraction experiment are converted to a set of structure factors as described in Eq 2.22.

$$|F_{hkl}|^2 = \frac{K.I_{hkl}}{L.p.A} \qquad \text{Eq 2.22}$$

where K is a scale factor, L is the Lorentz factor (specific to the geometry of the technique used) and p and A are correction terms for polarisation and absorption respectively. The parameters K, L and p are instrument specific whilst A depends on the shape of the crystal and the absorption coefficient (which varies with the nature of the atoms present and the wavelength of the incident radiation).

### 2.2.2 Comparison of single crystal and powder x-ray diffraction

Single crystal diffraction and molecular spectroscopy, especially NMR spectroscopy, are important techniques to gain a full insight into the structure and dynamics of a system (structure analysis by diffraction techniques only provides a space and time averaged representation). However, it is often very difficult to grow single crystals of suitable quality, particularly for long chain flexible molecules. Additionally, it can be impossible to obtain and analyse crystals of metastable phases at ambient temperature. Consequently, solving the crystal structure from powder diffraction data acquired under non-ambient conditions is desirable giving insight into the changes associated with phase transitions. A further advantage of obtaining the crystal structure directly from the polycrystalline sample is that it is representative of the powder rather than of one specific single crystal selected for analysis.

The protocol used to establish the crystal structure (either directly from single crystals or from powder data methods) involves data collection from a good quality sample, followed by unit cell determination and space group assignment, and then structure solution, and finally structure refinement. In the case of single crystal x-ray diffraction (SXRD), the structural information is distributed in three-dimensional space, whereas for powder diffraction (PXRD), the three-dimensional diffraction data are compressed into one-dimensional space (i.e. the powder pattern). As a consequence, the powder diffraction pattern often has severe peak overlap, leading to difficulties in reliable peak intensity ($I_{hkl}$) extraction. Furthermore, peak overlap may also give rise to ambiguities in indexing the powder pattern (in order to define the unit cell parameters), and space group assignment. Notwithstanding these limitations, PXRD allows assessment of structures for which growth of a suitable single

crystal is not feasible and is particularly important in determining structures of non-ambient phases. Figure 2.17 illustrates a schematic representation of the two (SXRD and PXRD) routes to crystal structure determination.

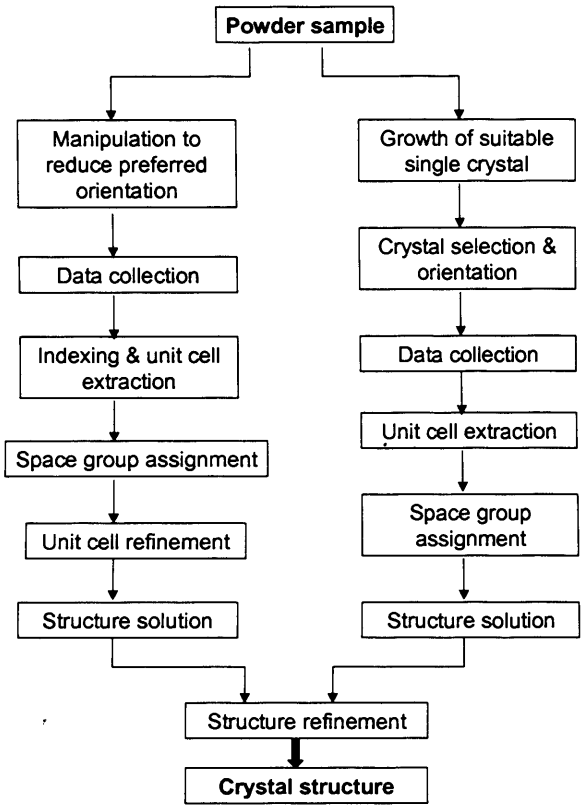



Figure 2.17: Schematic representation of the procedure for structure determination

### 2.2.2.1 Instrumentation

Single crystal samples are mounted on a goniometer and aligned in order to centralise the crystal within the beam. Most diffractometers allow rotation of the sample around three axes with the detector (typically a scintillation counter or area detector, such as a CCD) rotated in the plane of the crystal – these instruments are referred to as four-circle diffractometers. In this instrument, the detector measures the intensity of each individual reflection. The orientation of the crystal is determined by indexing initial reflections, and then the instrumental setting required to measure any particular reflection (hkl) may be computed automatically.

Powder diffractometers incorporate a detector to measure the scattered intensity as a function of diffraction angle (2θ), resulting in a diffraction pattern composed of a series of peaks at specific values of 2θ. Data can be collected in either transmission or reflection mode

geometries. In the case of transmission geometry, the sample is presented as either thin films or packed into capillary tubes. Whilst this method offers advantages in the reduction of preferred orientation, brought about by some powdered samples, material that is prone to absorb x-radiation require the use of an absorption correction. In reflection mode, the powder sample is presented as a flat specimen that is spun in its plane to minimise effects of preferred orientation. Furthermore, the reflection geometry is more amenable to non-ambient (variable temperature and humidity) assessment, and for the purposes of this research work the reflection geometry was used for all powder data collection. For most instruments with a Bragg-Brentano geometry (as shown in Figure 2.18) the source is kept static whilst the sample and detector rotate in the arc of a circle centred on the sample.

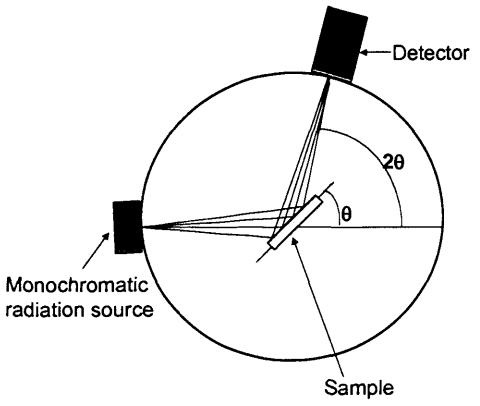



Figure 2.18: Bragg-Brentanto geometry of a powder diffractometer

### 2.2.3 Structure determination from PXRD

When single crystals of suitable size and quality cannot be prepared for conventional SXRD assessment, structure determination from PXRD data provides a viable route to the crystal structure. There are several limitations to approaching structure solution with powder methods, as discussed in section 2.2.2.

#### 2.2.3.1 Data acquisition to minimise preferred orientation effects

The data should be collected from a powder sample that exhibits random orientation of the crystallites. In such instances the powder pattern generated and more specifically the peak intensities are truly representative of the intensity-weighted reciprocal lattice. Polycrystalline samples that exhibit plate or acicular morphologies have a tendency to align with the plane of the sample holder. This gives rise to a disproportionate increase in the relative intensity of the reflections from these orientations, an occurrence known as preferred

orientation effects. In order to ensure that the collected data is devoid of preferred orientation effects, it is important to take appropriate measures, including procedures such as particle size reduction (grinding, milling, micronisation), spray drying (Smith *et al.* 1979), or co-mixing with an inert or amorphous material (Mason 1955). The extent of preferred orientation in a sample can be tested using a simple diffraction procedure prior to high-resolution data collection (Cheung *et al.* 2003).

#### 2.2.3.2 Indexing the powder pattern and space group assignment

The peak positions of a powder pattern are used to extract information on the unit cell parameters, known as indexing of the powder pattern. Accurate determination of the peak positions in the powder diffraction pattern is required, followed by the use of auto-indexing programs such as ITO (Visser 1969), TREOR (Werner *et al.* 1985), DICVOL (Boultif and Louer 1991) or X-cell (Neumann 2003).

The next procedure is to carry out pattern decomposition or profile fitting, which also serves to assign the space group. This procedure is performed using a least squares refinement method. There are two approaches that have been developed for this purpose, the Pawley (1981) and Le Bail (1988) fitting methods. The least squares refinement procedure involves refinement of the unit cell parameters, the zero point error, peak shapes and peak widths, the background function, and the use of arbitrary intensities for individual reflections. This procedure is of critical importance in defining reliable parameters for use in the subsequent structure solution process. In the first instance, the space group is assigned manually on the basis of systematic absences. The final discrimination and assignment involves least squares refinement and assessment of the goodness of fit of the extracted powder pattern, and the newly calculated cell is then assessed against the experimental data. The problem of peak overlap can limit unambiguous assignment. Consequently, several space groups may have to be considered in the structure solution calculations.

#### 2.2.3.3 Structure Solution

The crystal structure is defined by a series of structure factors, $F_{hkl}$, of each atom or molecule type in the unit cell. $F_{hkl}$ can be split into two components; the structure factor amplitude $|F_{hkl}|$, and the phase component $\alpha_{hkl}$. The amplitude values can be obtained (although not necessarily reliably) from the experimental powder diffraction data, but the

phase data cannot be obtained. This is known as the phase problem and as such, direct solutions of the crystal structure are not possible.

Several methodologies exist that allow structure solution from powder data to be performed, encompassing traditional approaches and direct-space approaches. The traditional approaches such as Patterson or direct methods, involve extraction of the intensities of individual reflections $I_{hkl}$ directly from the powder diffraction pattern. These methods then use algorithms in order to calculate and refine an electron density map using either Fourier summation or probability distribution based logic (Christensen *et al.* 1985, McCusker 1991, Cheetham *et al.* 1991, Harris *et al.* 1996).

Direct space techniques involve the generation of trial structures in direct space and using the calculated powder pattern of each trial structure, a goodness of fit against the experimental powder pattern is assessed (Harris and Cheung 2004). This protocol bypasses the need to use the diffraction-integrated intensities, and hence the phase problem, and operates on the basis that the diffraction pattern (the $|F_{hkl}|$ data) of any trial structure can always be determined automatically. Algorithms implemented within the direct-space approach include; simulated annealing (including also energy information, not just diffraction data) (Newsam *et al.* 1992), Monte Carlo (Harris *et al.* 1994, Harris et *al.* 1996), combined Monte-Carlo/simulated annealing (Engel *et al.* 1999, David *et al.* 1998) and a genetic algorithm method (Harris *et al.* 1998, Kariuki *et al.* 1999). The genetic algorithm approach was used to solve structures in this research work and as such will be discussed in more detail.

#### 2.2.3.3.1 *The genetic algorithm approach*

The Genetic Algorithm (GA) is a stochastic global optimisation technique based on the principles of evolution, and is utilised in many applications of global optimisation (Goldberg 1989, Devillers 1996). The mechanics of a simple GA follow the theory of natural selection, which inevitably leads to the survival and progression of the fittest members of a population. In following the principles of natural selection and evolution the GA manipulates a population of potential trial structures, in order to continually improve the quality of the structures present. As in nature, the manipulations involve series of combinations, mutations and crossovers that eventually lead to new sub-populations. All potential solutions that result from the optimisation are then ranked with a goodness of fit value, which depends on the quality of the fit to the experimental PXRD pattern.

Each trial structure generated in direct space is defined by a set of variables representing the asymmetric unit (i.e. position, orientation and geometry of all structural fragments). The structural fragment has a molecular geometry that is based on reliably known information on bond lengths and angles. A typical definition of the variables for a fragment containing 'n' torsion angles ($\tau$) (represented by 6+n components) is shown in Figure 2.19.

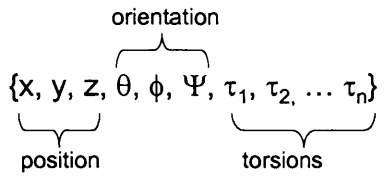



Figure 2.19: Definition of the variables to describe the structural fragment

The genetic code of the structure is defined as the string of variables (Harris *et al.* 2001). Optimisation of an initial population of trial structures involves a series of mating and mutation (of the variables) operations to give rise to the next generation from which the powder data is simulated. An assessment of the goodness of fit is made using the value of $R_{wp}$ – the weighted profile factor (defined in section 2.2.3.4) where a low value indicates a better fit. The evolutionary process (represented schematically in Figure 2.20) involves the mating of pairs of structures ("parents") selected from the population, with the probability of selecting a given structure proportional to its fitness (Harris *et al.* 2002). Combining genetic information from the two parents then generates new structures or "offspring". In the mutation procedure, a few random structures are introduced within each generation. In practice, a few structures are selected from the current population and random changes are made to parts of their genetic information (note that the original structure from which the mutant is derived is still retained within the population). The introduction of mutations is important to ensure that diversity is maintained within the population throughout the evolutionary process. The process of natural selection ensures that only the best structures survive on moving from one generation to the next generation. As a consequence, the overall quality of the population continually improves. After the population has evolved for a sufficiently large number of generations, the fittest member in the population (i.e. the structure with lowest $R_{wp}$ representing a global minimum on the $R_{wp}$ hypersurface) should, in principle, be close to the correct crystal structure.

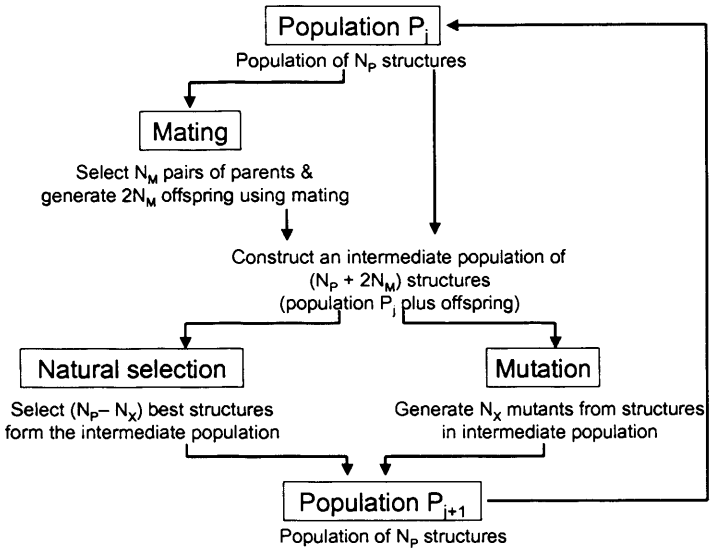



Figure 2.20: Schematic representation of the evolutionary process

### 2.2.3.4 Rietveld Refinement

Once a solution of the crystal structure has been obtained, the parameters describing this approximate structural model are refined against the experimental powder diffraction pattern using the Rietveld refinement method. In this method (Rietveld 1969; Young *et al.* 1987, Young 1993) each point on the profile is taken as an intensity measurement. The powder diffraction profile for the structural model is calculated on the basis of the following information (and also incorporates a description of the background intensity and corrections for preferred orientation). From the structure, the unit cell parameters are used to determine the peak positions, the atomic positions and atomic displacement parameters determine the peak intensities, and analytical functions describe the peak shapes, the peak widths and the background intensity distribution. The preferred orientation correction is only instigated, if at all, when a sufficiently good structural model has been established.

Rietveld refinement involves least squares fitting to minimise the difference between the observed and calculated powder diffraction profiles. The intensity $y_{io}$ at the $i^{th}$ step of the experimental profile is compared with the corresponding intensity $y_{ic}$ calculated for the model. Hence, a comparison of the whole powder diffraction profile, point by point, is performed. The progress of the refinement is followed by the profile R-factor, the most widely used of which is the weighted profile R-factor $R_{wp}$ as described in Eq 2.23.

$$R_{wp} = 100x \frac{\sqrt{[\Sigma_i w_i (y_{io} - y_{ic})^2]}}{\sqrt{\Sigma w_i y_{io}^2}} \quad \text{Eq 2.23}$$

The value of $R_{wp}$ decreases as the level of agreement between the experimental and calculated powder diffraction patterns improves (i.e. as the structural model gets closer to the correct structure). The structural model may be subjected to constraints and/or restraints in order to allow stable refinement. Restraints are imposed as soft conditions and must be satisfied approximately, whereas constraints are imposed as absolute conditions and must be satisfied exactly.

## 2.3 Phase diagrams for assigning stability hierarchy of polymorphs

Assigning relative stability hierarchy of polymorphs, especially over temperature and pressure space, is important in industries where polymorphism plays an important role in product integrity. The stability hierarchy, definition of enantiotropism or monotropism, is related to the differences in free energy ($\Delta G$) between pairs of polymorphs. Assessing the variation of free energy over temperature and pressure space provides increased confidence that a robust polymorph has been selected, which is stable to both primary and secondary processing. One way of achieving this is to represent the stability profile of all isolated polymorphs as a function of pressure and temperature in the form of energy-temperature (E/T) and pressure-temperature (p/T) diagrams, which are topological two-dimensional representations of polymorph thermodynamic space utilising the Gibbs fundamental equation shown in Eq 2.24.

$$dG = V\,dP - S\,dT + \sum_B \mu_B dn_B \quad \text{Eq 2.24}$$

For polymorphs, the last component of Eq 2.24, relating to changes in chemical composition can be neglected and hence the topological representations consider the first two terms.

### 2.3.1 The E/T diagram

An E/T diagram is a topological representation of enthalpy and free energy of polymorphs as a function of temperature, extracted from DSC data and extrapolated to 0 K. The approach assumes that any contribution of pressure to phase transitions is negligible and

solely related to the temperature, enthalpy and differences in heat capacity of melting of polymorphs (Yu 1995).

### 2.3.2 The p/T diagram

A more rigorous assessment of free energy differences between polymorphs incorporates an assessment not only over temperature but also pressure (Ceolin *et al.* 1992, Espeau *et al.* 2005). A p/T approach is based on the fact that each polymorph is capable of co-existing in the three states of matter; solids, liquid and vapour. As such, the p/T diagram is composed of triple points representing the equilibrium points of the three states of matter, and equilibrium curves which represent the equilibrium boundary between two phases. The diagram is constructed from parameters obtained from melting thermodynamics, temperature related volume variation in the solid and liquid states, and information on sublimation characteristics.

The number of triple points of a one-component system capable of existing under more than one solid phase has been demonstrated by Riecke (1890) in accordance with the expression shown in Eq 2.25

$$N = \frac{n\,(n-1)\,(n-2)}{(n-1)!} \qquad \text{Eq 2.25}$$

where N represents the number of triple points to be found in the diagram, and n represents the total number of phases (liquid, vapour and all solid phases) under consideration.

The slopes of the phase equilibrium curves are obtained from the Claypeyron equation, shown in Eq 2.26

$$\frac{dp}{dT} = \frac{\Delta H}{T\Delta V} \qquad \text{Eq 2.26}$$

where $\Delta H$ represents the change in enthalpy for the phase transition (for instance melting or sublimation), and $\Delta V$ represents the change in volume for the phase transition.

## 2.4 Microcalorimetry and other thermal methods

### 2.4.1 Thermal methods

Differential scanning calorimetry (DSC) methods utilise the fact that a thermally induced change in physical state of a material is accompanied by significant heat flow, which when integrated as a function of time gives an enthalpic term for the transition under investigation. The basis of the approach involves measurement of a temperature differential between a sample cell relative to a reference cell, and the heat flow required to maintain both the reference and sample at the same temperature is measured. DSC is a sensitive technique that allows information on polymorphic transitions, solvent loss and presence of chemical and physical impurities (mixed polymorphs) to be extracted. The resulting thermogram shows a measure of heat flow as a function of temperature, and an example thermogram is illustrated in Figure 2.21.

Thermogravimetric analysis (TGA) is used to assess mass loss as a function of temperature. This technique (involving the use of a precision balance coupled to a highly regulated furnace) is useful in qualitatively identifying the occurence of dehydration, desolvation and degradation. Furthermore, the mass loss can be used to establish hydrate or solvate stoichiometry. TGA when used either in thermal or isothermal mode can be important in the assessment of the kinetics of various events such as dehydration. A typical TGA thermogram is shown in Figure 2.21. When DSC and TGA are used in conjunction with each other, the nature of any thermally induced transitions from both qualitative and potentially quantitative perspectives can be explored.

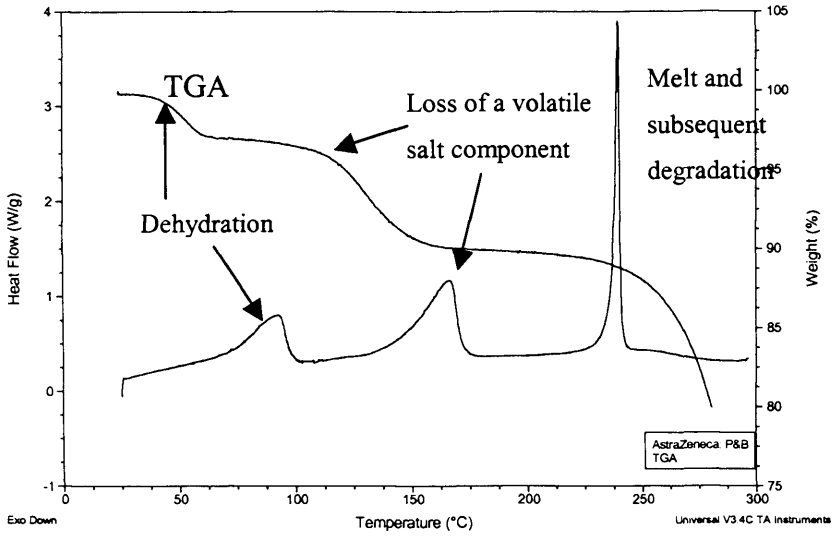



Figure 2.21: DSC and TGA thermogram showing complimentarity of the two techniques in identifying the nature of various thermal events

## 2.4.2 Isothermal methods

### 2.4.2.1 Isothermal reaction kinetics

The kinetics of a reaction such as a dehydration event can be investigated using isothermal TGA. The data (plotting fraction reacted against time) obtained from each measurement can be modelled to a number of solid-state reaction kinetic models (Byrn *et al.* 1999) to investigate the model that best describes the reaction.

### 2.4.2.2 Solution microcalorimetry

Solution calorimetry provides a direct measure of the thermodynamics of dissolution. On mixing a material with an appropriate solvent, heat flow is measured as a function of time and integrated to give the molar enthalpy of solution ($\Delta_{sol}H$). The cumulative enthalpy of solution encompasses a measure of the energy associated with wetting, breaking of intermolecular forces within the crystal and solvation. Classification of thermodynamic stability by solution calorimetry relies upon the energetics of wetting and solvation across polymorphs to be constant, thereby providing a measure of the internal energy of a crystal. Whilst the component associated with heat of solvation may be correctly regarded as constant between polymorphs, the heat of wetting may vary as a function of crystal habit and changes in surface characteristics, and thus could give rise to some variability.

## 2.5 Non-ambient analysis

Non-ambient analysis involves the use of standard analytical instrumentation with a non-ambient sample chamber or the use of a sample holder designed to maintain a sealed environment, permitting analysis as a function of temperature and humidity and in-situ assessment of metastable non-ambient phases. Phase transitions that may be followed include; crystalline solid-to-solid transitions, amorphous to crystalline modifications, and hydration (solvation)-dehydration (desolvation) assessments. Such experiments could provide information on the nature, kinetics and mechanisms of temperature and humidity induced phase changes. Furthermore, this also provides a method for crystal structures of non-ambient phases to be determined from PXRD data.

Combining a temperature chamber with XRPD (Karjalainen *et al.* 2005) and spectroscopic techniques such as SSNMR or vibrational spectroscopy (Zhu *et al.* 1997) allows the molecular and structural consequences of temperature-induced transitions to be evaluated.

Vapour perfusion experiments allow the influence of a variation in water vapour pressure on material property and behaviour to be explored, achieved by combining with an environmental/relative humidity (RH) chamber. In most cases, variable humidity assessments are performed under isothermal conditions. Experiments such as gravimetric vapour sorption, in which mass uptake as a function of %RH is investigated, provides information on the mechanism and stoichiometry of uptake. Vapour perfusion microcalorimetry (Ramos *et al.* 2005, Puipeddi *et al.* 1996) provides a thermodynamic assessment of the nature of the phase transitions. The use of variable humidity in conjunction with molecular spectroscopy and powder diffraction (Li *et al.* 2000, Stephenson *et al.* 2000) allows further insight into the molecular and structural consequences of any vapour induced transitions.

## 2.6 References


- Abraham R. J., Fisher J., Loftus P., (1988) *Introduction to NMR Spectroscopy,* John Wiley & Sons Limited
- Aliev A. E., MacLean E. J., Harris K. D. M., Kariuki B. M., Glidewell C., (1998), *J. Phys. Chem. B*, **102**, 2165
- Andrew E.R., Bradbury A., Eades R.G., (1959), *Nature,* **183**, 1802
- Azaroff L. V., Buerger M. J., (1958), *The Powder Method*, Mcgraw-Hill)
- Boultif A., Louër D., (1991), *J. Appl. Crystallogr.*, **24**, 987
- Bugay D. E., (2001), *Adv. Drug Deliv. Rev.*, **48**, 43
- Byrn S. R., Pfeiffer R. R., Stowell J. G., (1999), *Solid-state Chemistry of Drugs*, SSCI, Inc
- Ceolin R., Toscani S., Agafonov V., Dugue J., (1992), *J. Solid State Chem.*, **98**, 366
- Cheetham A. K., Wilkinson A. P., (1991), *J. Phys. Chem. Solids*, **52**, 1199
- Cheung E. Y., Harris K. D. M., Foxman B. M., (2003), *Cryst. Growth Des.*, **3**, 705
- Christensen A. N., Lehmann M. S., Nielsen M., (1985), *Aust. J. Phys.*, **38**, 497
- David W. I. F., Shankland K., Shankland N., (1998), *Chem. Comm.*, 931
- Davis J. H., Jeffrey K. R., Bloom M., Valic M. I., Higgs T. P., (1976), *Chem. Phys. Lett.*, **42**, 390
- Devillers J. (Ed), (1996), *Genetic Algorithm in Molecular Modelling*, Academic Press
- Duer M. J., (2002), *Solid State NMR Spectroscopy: Principles and Applications,* Blackwell Science
- Eichele K., Wasylishen R. E., (1994), *J. Magn. Reson. A*, **106**, 46

- Engel G. E., Wilke S., König O., Harris K. D. M., Leusen F. J. J., J. (1999), *J. Appl. Cryst.*, **32**,1169
- Espeau P., Ceolin R., Tamarit J-L., Perrin M-A., Leveiller F., (2005), *J. Pharm. Sci.*, **94**, 524
- Facey G. A., Connolly T. J., Bensimon C., Durst T., (1996), *Can. J. Chem.*, **74**, 1844
- Goldeberg D. E., (1989), *Genetic Algorithms in Search, Optimisation and Machine Learning*, Addison-Wesley, Reading (MA)
- Greenfield M. S., Ronemus A. D., Vold R. L., Vold R. R., Ellis P. D., Raidy T. R., (1987), *J. Magn. Reson.*, **72**, 89
- Hammond C., (1997), *The Basics of Crystallography and Diffraction*, Oxford University Press
- Harris K. D. M., Tremayne M., Lightfoot P., Bruce P. G., (1994), *J. Am. Chem. Soc.*, **116**, 3543
- Harris K. D. M., Tremayne M., (1996), *Chem. Mater.*, **8**, 2554
- Harris K. D. M., Johnston R. L., Kariuki B. M., (1998), *Acta Cryst.*, **A54**, 632
- Harris K. D. M., Tremayne M., Kariuki B. M., (2001), *Angew. Chem. Int. Ed.*, **40**, 1626
- Harris K. D. M., Johnston R. L., Cheung E. Y., Turner G. W., Habershon S., Albesa-Jove D., Tedesco E., Kariuki B. M., (2002), *CrystEngComm*, **4**, 356
- Harris K. D. M., Cheung E. Y., (2004), *Chem. Soc. Rev.*, **33**, 526
- Hendra P., Jones C., Warnes G., (1991), *Fourier Transform Raman Spectroscopy*, Ellis Horwood
- Jelinski L. W., (1985), *Ann. Rev. Mater. Sci.*, **15**, 35
- Kariuki B. M., Calcagno P., Harris K. D. M., Philp D., Johnston R. L., (1999), *Angew. Chem. Int. Ed.*, **38**, 831

- Karjalainen M., Airaksinen S., Rantanen J., Aaltonen J., Yliruusi J., (2005), *J. Pharm. Biomed. Anal.*, **39**, 27
- Le Bail A., Duroy H., Fourquet J. L., (1988), *Mater. Res. Bull.*, **23**, 447
- Li Y., Zhang G. G. Z., Grant D. J. W., Suryanarayanan R., (2000), *Pharm. Dev. Technol.*, **5**, 257
- Look D.C., Locke D.R., (1969), *Rev. Sci. Instrum.*, **41**, 250
- Lowe I. J., (1959), *Phys. Rev. Lett.*, **2**, 285
- Manson J.E., (1955), *J. Appl. Phys.*, **26**, 1254
- McCreery R. L., (2000), *Raman Spectroscopy for Chemical Analysis*, Wiley-interscience
- McCusker L. B., (1991), *Acta Cryst.*, **A47**, 297
- Neumann M., (2003), *J. Appl. Cryst.*, **36**, 356
- Newsam J .M., Deem M. W., Freeman C.M., (1992), *Accuracy in Powder Diffraction II: NIST Special Publication,* **846,** 80
- Pawley G. S., (1981), *J. Appl. Crystallogr.*, **14**, 357
- Pines A., Gibby M.G., Waugh J.S., (1973), *J. Chem. Phys.*, **59**, 569
- Puipeddi M., Sokoloski T. D., Duddu S. P., Carstensen J. T., (1996), *J. Pharm. Sci.*, **85**, 381
- Ramos R., Gaisford S., Buckton G., (2005), *Int. J. Pharm.*, 13
- Riecke E., (1890), *Z. Phys. Chem.*, **6**, 411
- Rietveld H. M., (1969), *J. Appl. Crystallogr.*, **59**, 569
- Schaefer J., Sefick M.D., Stejskal E.O., McKay R.A., (1982), *J. Magn. Reson.*, **49**, 341
- Smith S. T., Snyder R. L., Brownell W. E., (1979), *Adv. X-ray Anal.*, **22**, 77

- Stephenson G. A., Diseroad B. A., (2000), *Int. J. Pharm.*, **198**, 167
- Visser J.W., (1969), *J.Appl. Crystallogr.*, **2**, 89
- Werner P.E., Eriksson L., Westdhal M., (1985), *J. Appl. Crystallogr.*, **18**, 367
- Young R.A., Wiles D.H., (1987), *J. Appl. Crystallogr.*, **20**, 411
- Young R.A. (Ed.), (1993), *The Rietveld Refinement Method*, IUCr. Oxford University Press, New York
- Yu L., (1995), *J. Pharm. Sci.*, **84**, 966
- Zhu H., Padden B. E., Munson E. J., Grant D. J. W., (1997), *J. Pharm. Sci.*, **86**, 418

# 3. Tetraphenyl phosphonium bromide: understanding hydration and dehydration through structure and dynamics

## 3.1 Introduction

Tetraphenyl phosphonium bromide (TPPB) is used commercially as a phase transfer catalyst (Onkubo *et al.* 1977). In the present study it represents a simple tetrahedral ionic compound that is present as both crystalline anhydrate and hydrated phases. This molecule was known to crystallise as a polymorphic anhydrate (Müller 1980, Alcock *et al.* 1985, Schweizer *et al.* 1989), a monohydrate and a dihydrate (Vincent *et al.* 1988) phase. Although the crystal structures of two anhydrous forms have been reported, they have so far remained elusive under ambient conditions. The dihydrate packs in an orthorhombic Pnma structure, with one phenyl ring perpendicular to and another ring in the mirror plane. The $Br^-$ anion is at an inversion centre coordinated by the oxygen atoms from the water of crystallisation. The $Br^-$ and water complexes form an almost infinite planar chain in between alternating chains of the $TPP^+$ cations. The propensity of TPPB to form anhydrate and dihydrate phases is exploited here in investigations of dehydration-rehydration mechanisms. An investigation into the driving force and molecular consequence of dehydration and rehydration was initiated to include a study of the dynamics of the molecule in the crystal structure as a function of relative humidity and temperature.

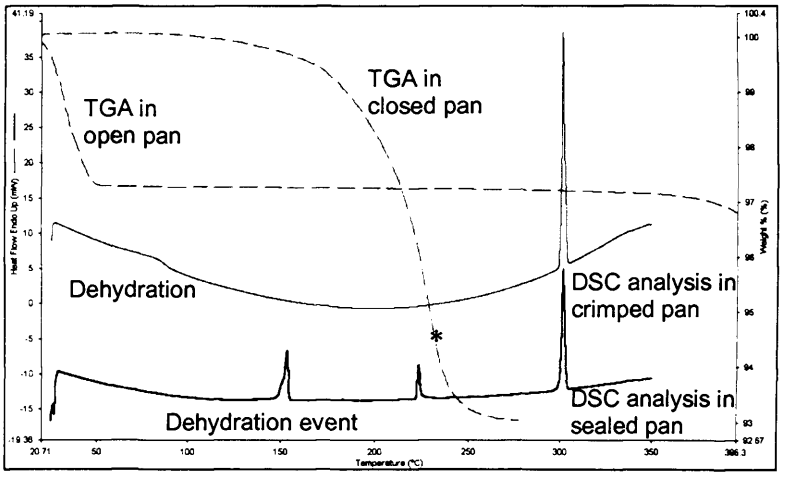
It has been suggested that the propensity to form hydrates stems from the need to redress the balance between hydrogen bond donors and acceptors (Desiraju 1991). Whilst this notion holds true for structures in which the water molecules form the basis of the hydrogen-bonding network, it may not be applicable in situations in which water is present as a void filler. Therefore an understanding of the molecular predisposition to form such hydrates and their relative stabilities is sought. An insight into the fundamental properties that govern dehydration and rehydration of hydrates is advantageous to the development of such materials.

The present study utilises TPPB as a model compound to investigate issues governing the formation and stability of hydrates in which the water acts as a void filler.

TPPB is stable as the dihydrate under ambient conditions, readily converting to a metastable anhydrate phase in a high temperature or low relative humidity environment. It has been stated that structures containing large empty channels are generally not stable (Gorbitz and Hersleth 2000), and either form hydrates/solvates or undergo structural collapse/phase transitions. However, this study serves to investigate the role of water in affecting the general structure, in the absence of significant favourable interactions and given the large expected cost in entropy.

## 3.2 Characterisation of pure anhydrate and dihydrate phases

A sample of the pure dihydrate was prepared by grinding TPPB and storing the ground material under an atmosphere of 100% RH. The pure dihydrate phase was then characterised using a range of techniques. Attempts to prepare and isolate the anhydrate phase, including recrystallisation from dried solvents and quench cooling of the dehydrated material, under ambient conditions were unsuccessful. Figure 3.1 illustrates the thermal profile of TPPB dihydrate. Differential scanning calorimetry (DSC) analysis in open versus sealed pans revealed a profound change in dehydration temperature as a direct effect of localised water vapour pressure. For instance, the dehydration temperature changed from 353 K in an open pan to 423 K in a sealed pan. This was also observed by thermogravimetric analysis (TGA). The mass loss observed by TGA was shown to give a TPPB hydrate stoichiometry of 2:1 water:TPPB confirming that the material is indeed a dihydrate.



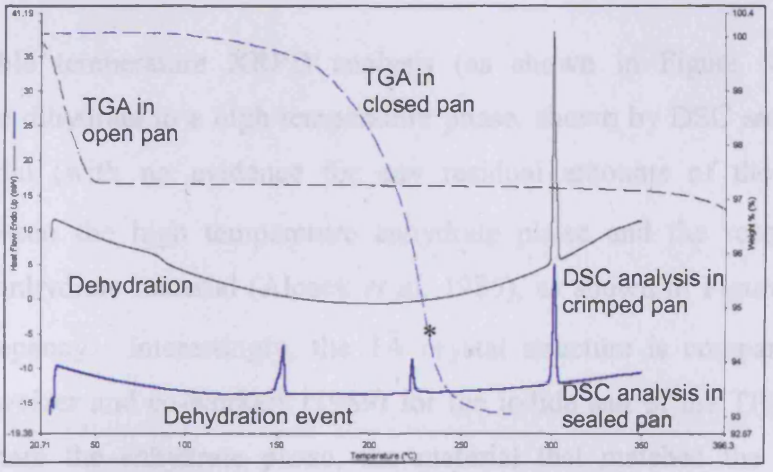

* denotes artefact due to sealed pan lid exploding due to increased pressure

Figure 3.1: Thermal profile of TPPB showing the variation in dehydration temperature as a function of localised water vapour pressure

TPPB is stable as the dihydrate under ambient conditions, readily converting to a metastable anhydrate phase in a high temperature or low relative humidity environment. It has been stated that structures containing large empty channels are generally not stable (Gorbitz and Hersleth 2000), and either form hydrates/solvates or undergo structural collapse/phase transitions. However, this study serves to investigate the role of water in affecting the general structure, in the absence of significant favourable interactions and given the large expected cost in entropy.

## 3.2 Characterisation of pure anhydrate and dihydrate phases

A sample of the pure dihydrate was prepared by grinding TPPB and storing the ground material under an atmosphere of 100% RH. The pure dihydrate phase was then characterised using a range of techniques. Attempts to prepare and isolate the anhydrate phase, including recrystallisation from dried solvents and quench cooling of the dehydrated material, under ambient conditions were unsuccessful. Figure 3.1 illustrates the thermal profile of TPPB dihydrate. Differential scanning calorimetry (DSC) analysis in open versus sealed pans revealed a profound change in dehydration temperature as a direct effect of localised water vapour pressure. For instance, the dehydration temperature changed from 353 K in an open pan to 423 K in a sealed pan. This was also observed by thermogravimetric analysis (TGA). The mass loss observed by TGA was shown to give a TPPB hydrate stoichiometry of 2:1 water:TPPB confirming that the material is indeed a dihydrate.




* denotes artefact due to sealed pan lid exploding due to increased pressure

Figure 3.1: Thermal profile of TPPB showing the variation in dehydration temperature as a function of localised water vapour pressure

The XRPD pattern of the dihydrate matches the XRPD pattern simulated from the reported crystal structure of the dihydrate, as shown in Figure 3.2, from which it can be concluded that the "dihydrate" prepared in the present work is a single phase with the same structure as that reported previously (Vincent *et al.* 1988).

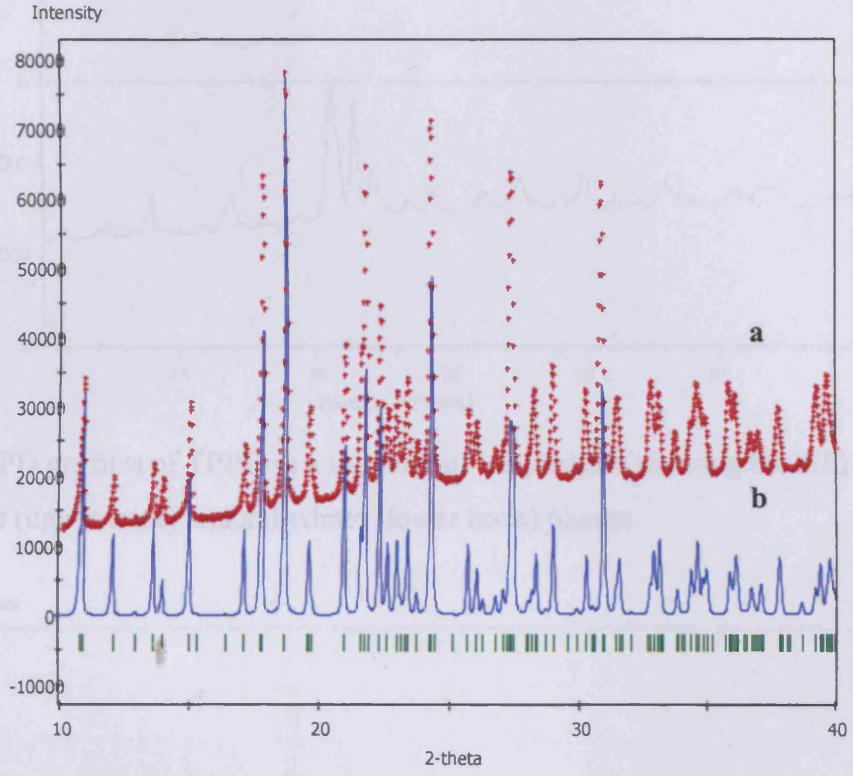



Figure 3.2: Overlay of XRPD diffractograms of (a) the experimental powder pattern and (b) the simulated powder pattern for the reported crystal structure

Variable temperature XRPD analysis (as shown in Figure 3.3) verified the conversion of the dihydrate to a high temperature phase, shown by DSC and TGA to be the anhydrate material (with no evidence for any residual amounts of the dihydrate). A comparison between the high temperature anhydrate phase and the reported $I\bar{4}$ crystal structure of the anhydrate material (Alcock *et al.* 1985), as shown in Figure 3.4, revealed a significant discrepancy. Interestingly, the $I\bar{4}$ crystal structure is comparable to the one reported by Schweizer and co-workers (1989) for the iodide salt of the $TPP^+$ cation. In all attempts to prepare the anhydrate phase, the material that matched the reported crystal structure was never obtained. Other structures of the anhydrate reported in the literature (Müller 1980) were comparable to the monohydrate structure also reported.

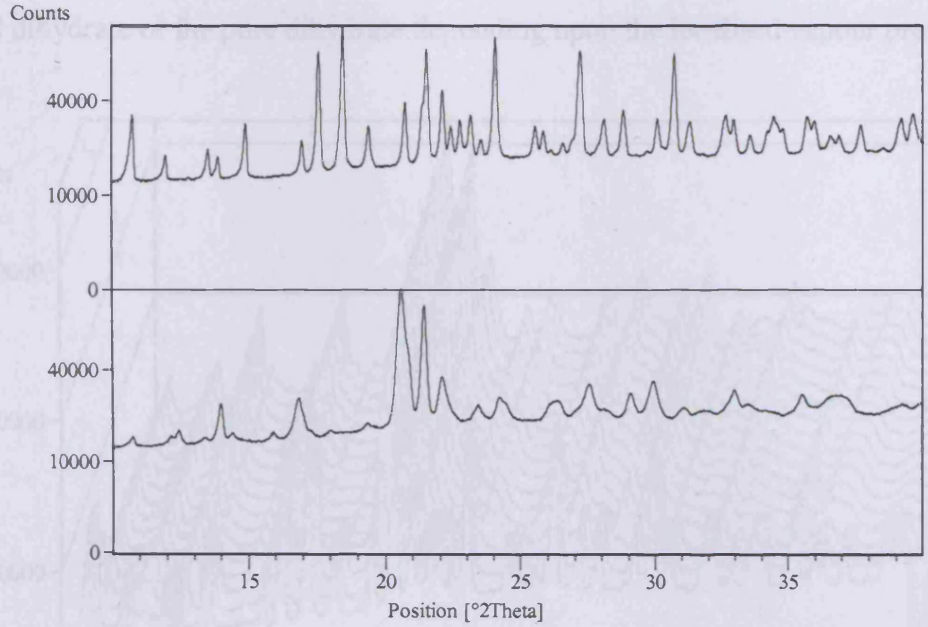



Figure 3.3: XRPD profiles of TPPB as a function of temperature showing the diffractograms of the dihydrate (upper trace) and anhydrate (lower trace) phases

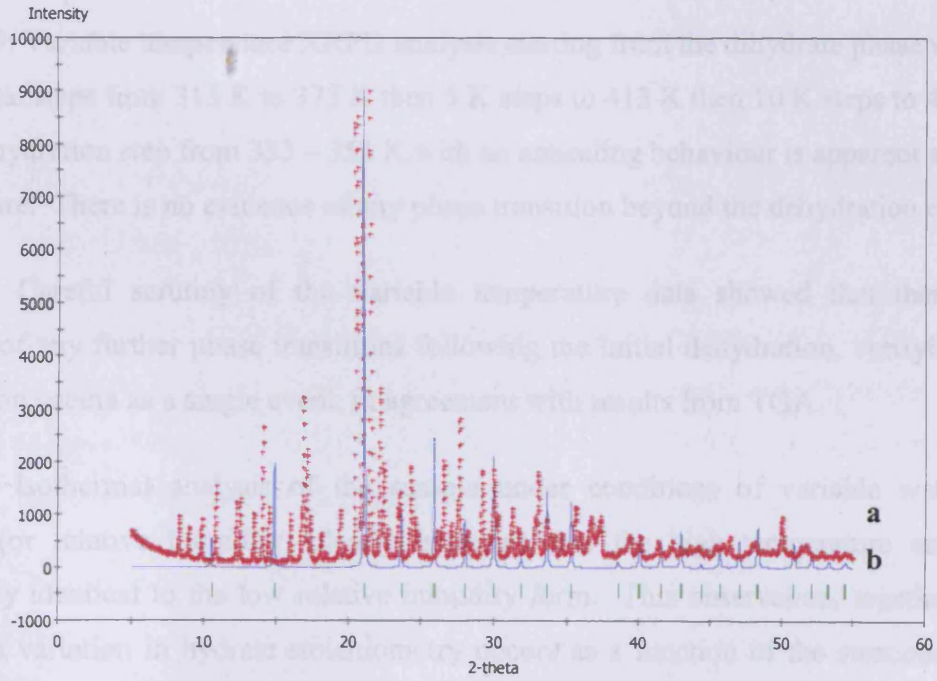



Figure 3.4: Overlay of XRPD diffractograms of (a) the experimental powder pattern and (b) the simulated powder pattern from the reported $I\bar{4}$ crystal structure

The anhydrate phase could only be observed as an in-situ high temperature or low relative humidity (RH) phase as exemplified by the in-situ non-ambient XRPD profiles shown in Figures 3.5 and 3.6. In all cases the anhydrate phase was shown to be a relatively unstable

entity under ambient temperature and relative humidity, converting to either mixtures of the mono and dihydrate or the pure dihydrate depending upon the localised vapour pressure.

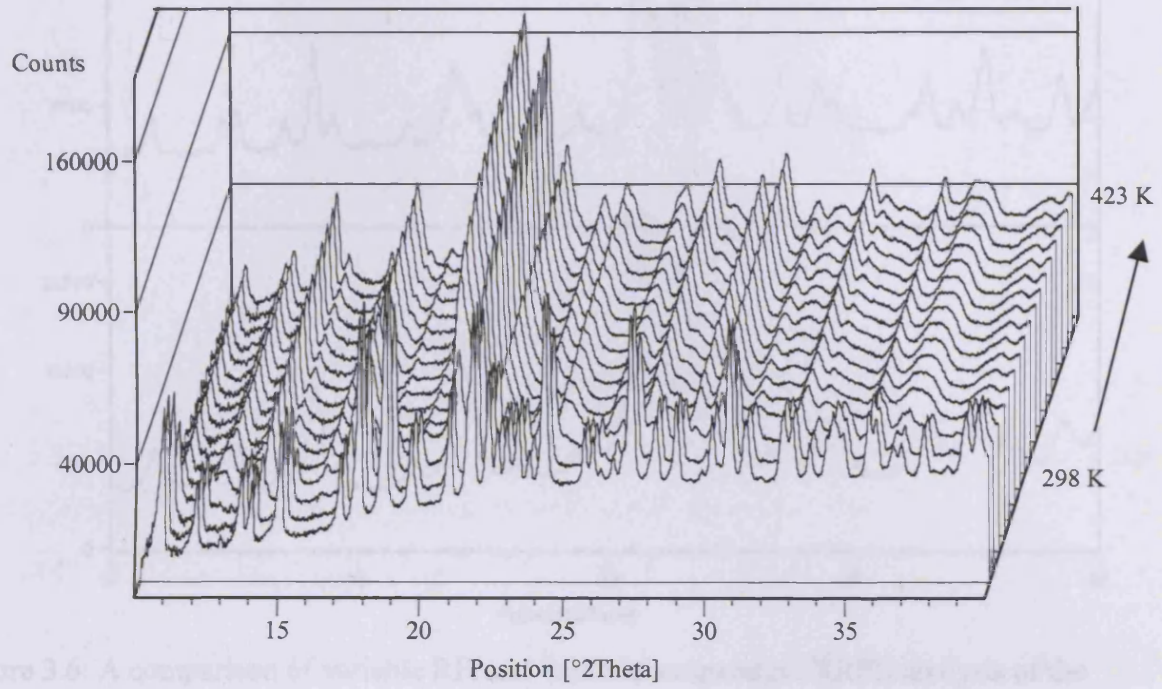



Figure 3.5: Variable temperature XRPD analysis starting from the dihydrate phase with 10 K incremental steps from 313 K to 373 K then 5 K steps to 413 K then 10 K steps to 423 K. A single dehydration step from 333 – 353 K with an annealing behaviour is apparent above this temperature. There is no evidence of any phase transition beyond the dehydration event

Careful scrutiny of the variable temperature data showed that there was no evidence of any further phase transitions following the initial dehydration, verifying that the dehydration occurs as a single event, in agreement with results from TGA.

Isothermal analysis of the sample under conditions of variable water vapour pressure (or relative humidity) clearly indicates that the high temperature anhydrate is structurally identical to the low relative humidity form. This observation, together with the fact that a variation in hydrate stoichiometry occurs as a function of the surrounding water activity, is consistent with TPPB forming a channel hydrate.

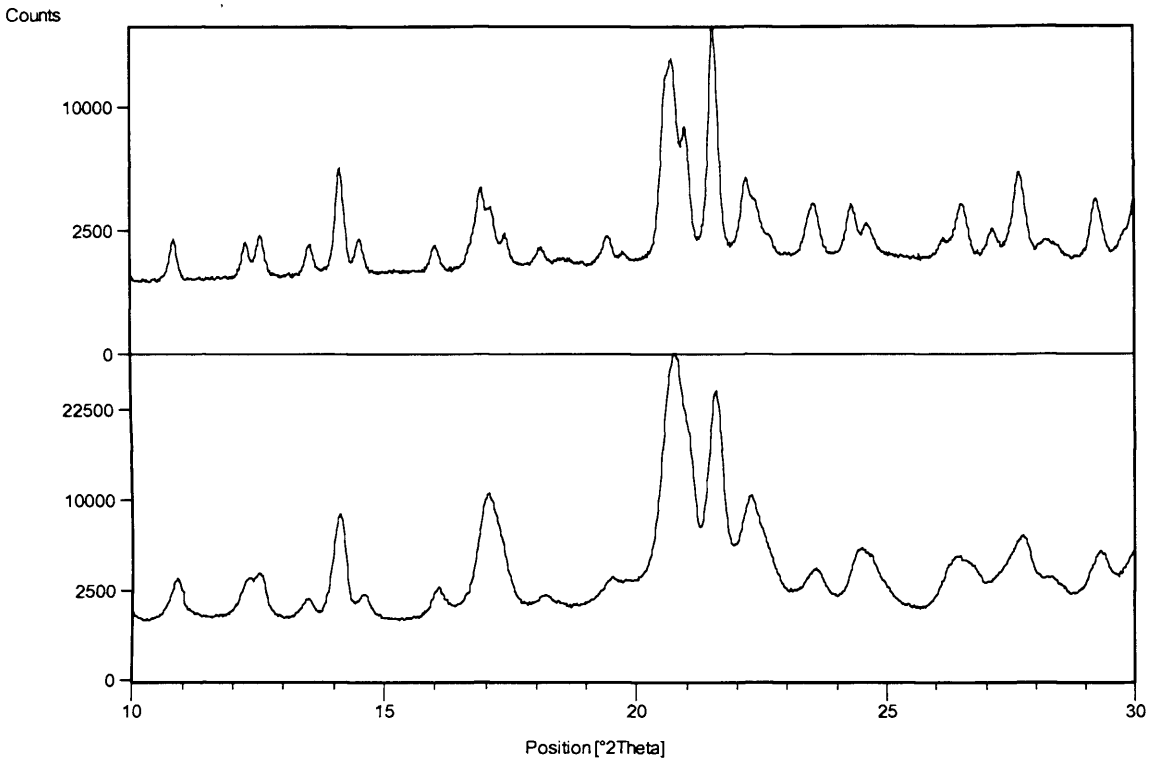



Figure 3.6: A comparison of variable RH and variable temperature XRPD analysis of the dihydrate, showing the formation of the same anhydrate phase under conditions of high temperature (upper trace) and low RH (lower trace)

## 3.3 Thermodynamic relationships between the dihydrate and anhydrate phases

The relative inter-relationships between the two phases, and hence the relative stability, was explored further. Analysis of the influence of water vapour pressure (or relative humidity) on mass change of TPPB was investigated by gravimetric vapour sorption at 298 K. Figure 3.7 illustrates the sorption/desorption isotherms for TPPB.

The formation of the dihydrate was shown to be highly dependent upon the localised water vapour pressure, as exemplified by the variable humidity XRPD studies, and proceeded via a hemi-hydrate and monohydrate. The mechanism of dehydration, however, was shown to advance via a single step, directly to the anhydrate under conditions of low relative humidity (< 10% w/w). This observation corroborated the thermal analysis results, in that dehydration was shown to proceed directly from the dihydrate to the anhydrous form.

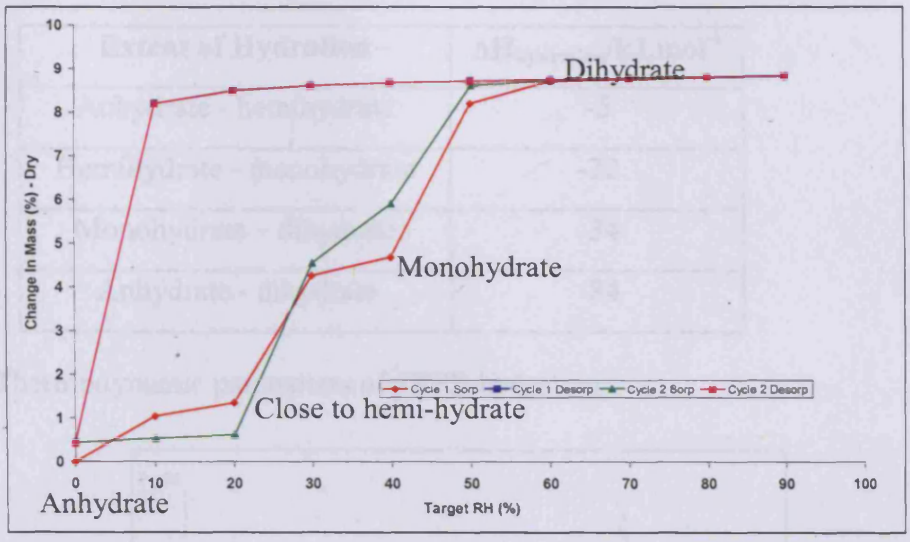



Figure 3.7: Water Sorption-Desorption Isotherm for TPPB at 298 K showing the % w/w water uptake giving rise to the different hydrate stoichiometry. Two sorption-desorption cycles were performed showing consistency in uptake and loss

The thermodynamics of the hydration/dehydration were explored further using linear perfusion isothermal microcalorimetry (ITC) experiments. The relative humidity in the sample chamber (at 298 K) was varied linearly, and any heat flow was measured as a function of time. Using a very slow perfusion rate, hydration was shown to proceed via a hemihydrate. An increase in the perfusion rate resulted in the formation of a monohydrate prior to the formation of a dihydrate, as exemplified in Figure 3.8. A further increase in rate showed that TPPB proceeds directly to the dihydrate from the anhydrous form in one single step. The thermodynamic parameters of hydrate formation as a function of the different mechanisms are summarised in Table 3.1. Figure 3.9 summarises the thermodynamic inter-relationship of all forms of TPPB thus far identified. The slight discrepancy in the enthalpy of hydration ($\Delta H_{hydration}$) on going directly from the anhydrate to the dihydrate or via the intermediate hydrate phases is attributed to variations within particle size and wettability between samples. The values nonetheless are of roughly the same order of magnitude.

| Extent of Hydration | $\Delta H_{hydration}$/kJ.mol$^{-1}$ |
|---|---|
| Anhydrate - hemihydrate | -5 |
| Hemihydrate - monohydrate | -22 |
| Monohydrate – dihydrate | -34 |
| Anhydrate - dihydrate | -84 |

Table 3.1: Thermodynamic parameters of TPPB hydration

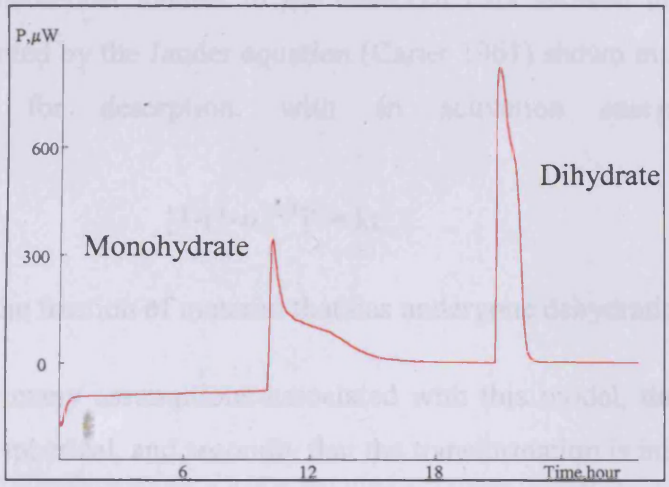



Figure 3.8: ITC of the formation of hydrates of TPPB using an intermediate perfusion rate. The graph shows the heat output associated with the formation of the hydrates

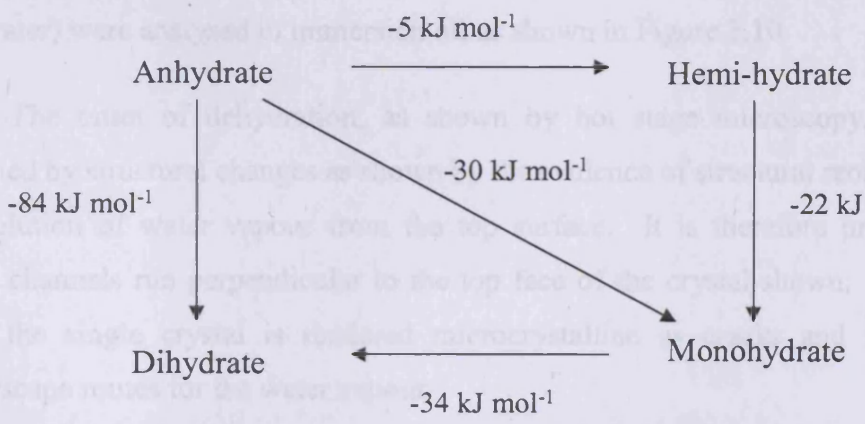



Figure 3.9: Thermodynamic relationship of all the hydrated forms of TPPB at 298 K from isothermal perfusion microcalorimetry assessments

From the DSC experiments, the $\Delta H_{dehydration}$ at 65°C was shown to be 35.5 kJ mol$^{-1}$. Using the Kirchoff equation shown in Eq 3.1, and assuming that $\Delta C_p$ is roughly equal to

1 J $mol^{-1}$ $K^{-1}$, then from the DSC experiment the enthalpy of dehydration at 25°C is ~ 44 kJ $mol^{-1}$, which is comparable (or of the same order of magnitude when taking into consideration enthalpy contributions from adsorption, wetting, etc., in the ITC data) to the isothermal microcalorimetry data.

$$\Delta H_{T2} - \Delta H_{T1} = \Delta C_p (T_2 - T_1) \qquad \text{Eq 3.1}$$

The kinetics of dehydration were investigated using isothermal TGA. A least squares fit of the theoretical models to the observed data showed that three-dimensional diffusion, as represented by the Jander equation (Carter 1961) shown in Eq 3.2, was the most likely mechanism for desorption, with an activation energy of dehydration $E_A = 11$ kJ $mol^{-1}$.

$$[1-(1-\alpha)^{1/3}]^2 = kt \qquad \text{Eq 3.2}$$

Where $\alpha$ represents the fraction of material that has undergone dehydration.

There are many assumptions associated with this model, the first being that the particles/crystals are spherical, and secondly that the transformation is initiated from the entire surface. From the crystal structure of the dihydrate (Vincent *et al.* 1988) it is evident that the dihydrate forms planar infinite chains along one axis. An investigation into the anisotropy of dehydration was initiated using hot stage microscopy. Single crystals of TPPB (grown from purified water) were analysed in immersion oil as shown in Figure 3.10.

The onset of dehydration, as shown by hot stage microscopy, appears to be accompanied by structural changes as shown by the evidence of structural reorganisation prior to the evolution of water vapour from the top surface. It is therefore proposed that the interstitial channels run perpendicular to the top face of the crystal shown. As dehydration proceeds, the single crystal is rendered microcrystalline as cracks and fissures appear, allowing escape routes for the water vapour.

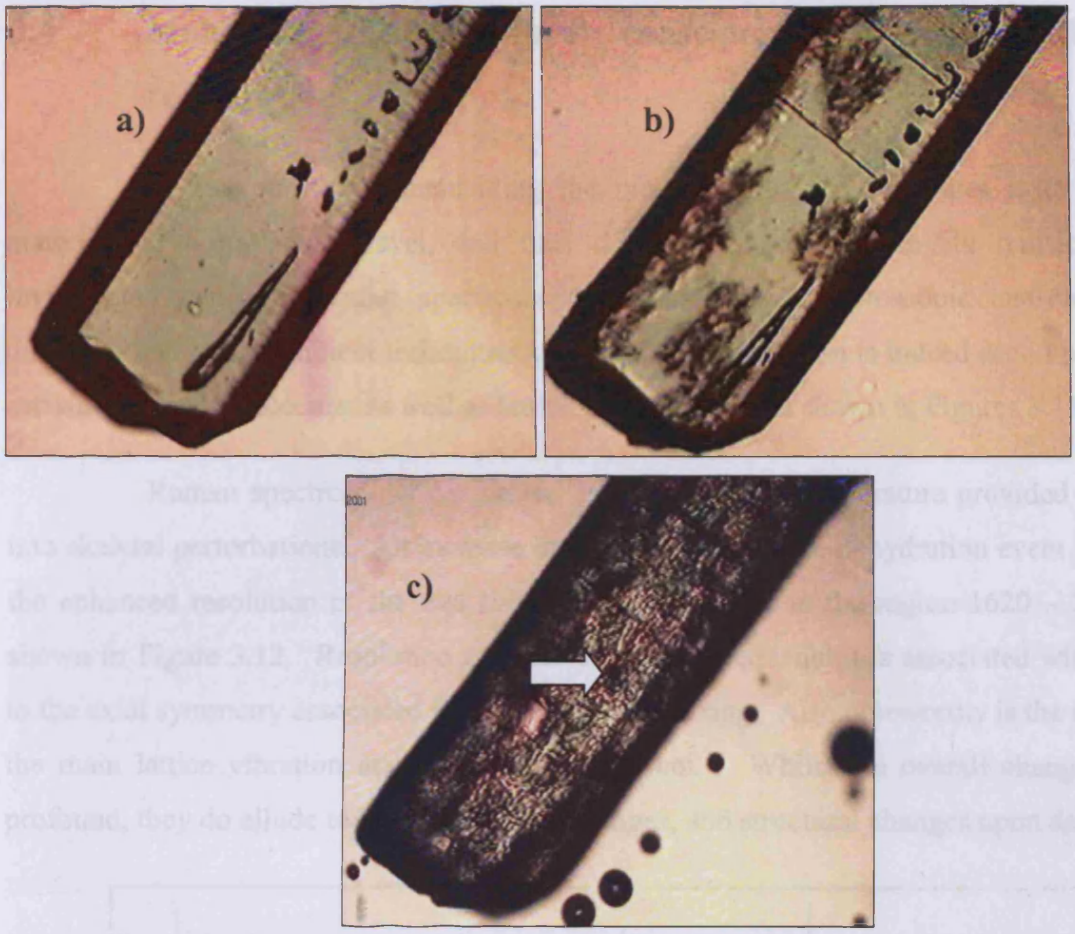



Figure 3.10: Hot stage micrographs of a single crystal of TPPB in immersion oil at (a) ambient temperature, (b) onset of dehydration and (c) dehydration showing accompanying structural collapse. The arrow indicates evolution of water vapour on the top surface of the crystals

The driving force for dehydration under normal conditions may involve an increase in entropy (with regard to the solvent molecules) accompanied by structural rearrangement. However, the present study suggests the dihydrate is the most stable form under ambient conditions. This implies that anhydrous TPPB acts as a "molecular sponge" primarily due to the interstitial voids created by the parallel stacks of the phosphonium cations. It had been inferred that a low energy structure is one that has maximum packing density thus minimising void space (Kitiagorodskii 1973). In the case of the anhydrate there are no strong hydrogen bonds to maintain a framework for a less densely packed lattice.

## 3.4 Molecular and structural consequences of dehydration and rehydration

A first step to understanding the propensity to form hydrates is to study the materials at a molecular level, and thus the pure phases and in-situ transitions were investigated using molecular spectroscopy. Vibrational spectroscopic analyses on the dihydrate using non-ambient techniques verified that dehydration is indeed accompanied by a certain degree of molecular as well as lattice perturbations, as shown in Figures 3.11 and 3.12.

Raman spectroscopy performed as a function of temperature provided an insight into skeletal perturbations. An increase in temperature past the dehydration event resulted in the enhanced resolution of the two ring quadrant stretches in the region 1620 – 1565 $cm^{-1}$, shown in Figure 3.12. Resolution of these stretching frequencies is associated with changes in the axial symmetry associated in the vicinity of the ring. Also noteworthy is the splitting of the main lattice vibration at approximately 200 $cm^{-1}$. Whilst the overall changes are not profound, they do allude to local symmetry changes, and structural changes upon dehydration.

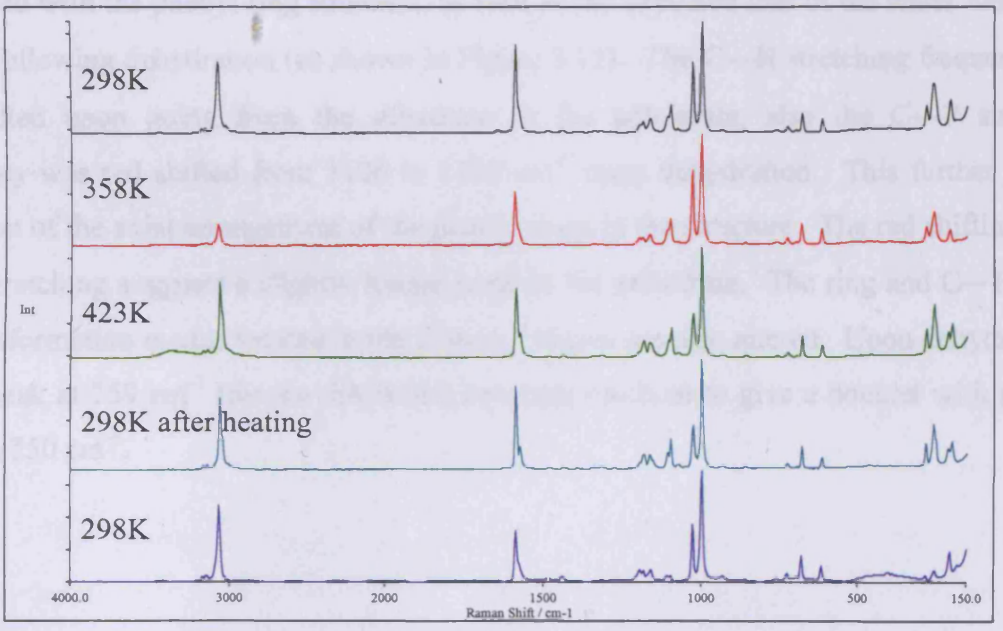



Figure 3.11: Variable Temperature Raman Spectra of TPPB dihydrate

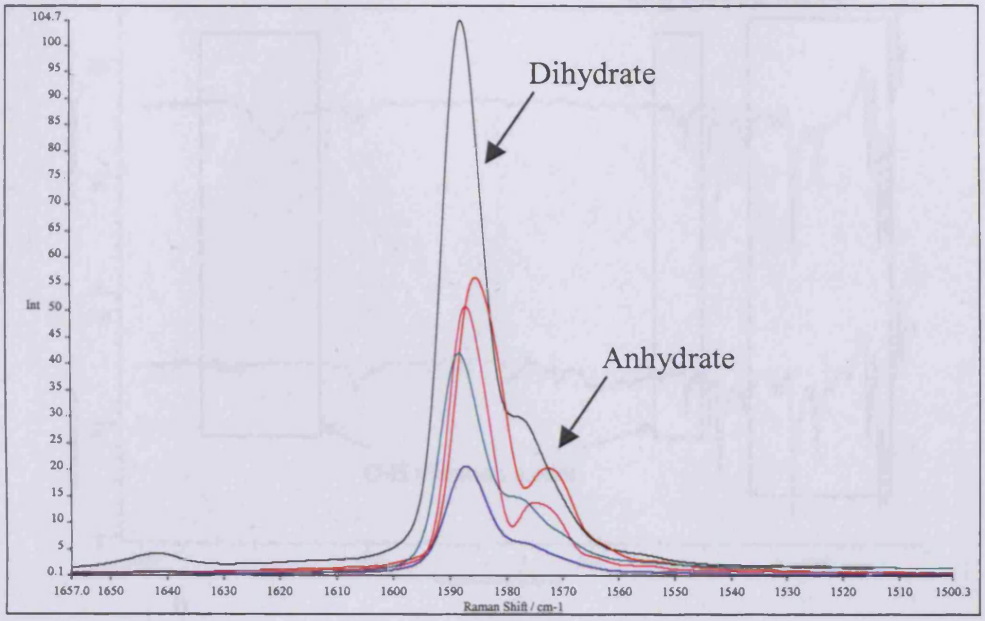



Figure 3.12: Variable Temperature Raman Spectra of Phenyl Ring Stretch Region

Furthermore, assessment of the dihydrate, and the in-situ formation of the anhydrate, by variable temperature FTIR spectroscopy showed small but significant changes associated with the phenyl ring stretches, as well as the expected loss of the water vibrational modes following dehydration (as shown in Figure 3.13). The C—H stretching frequency was red shifted upon going from the dihydrate to the anhydrate, also the C—P stretching frequency was red shifted from 1106 to 1100 $cm^{-1}$ upon dehydration. This further verifies alteration of the axial arrangement of the phenyl rings in the structure. The red shifting of the C—P stretching suggests a slightly longer bond in the anhydrate. The ring and C—H out of plane deformation modes located in the 750 $cm^{-1}$ region are also altered. Upon dehydration, a single peak at 759 $cm^{-1}$ (for the dihydrate) becomes resolved to give a doublet with peaks at 762 and 750 $cm^{-1}$.

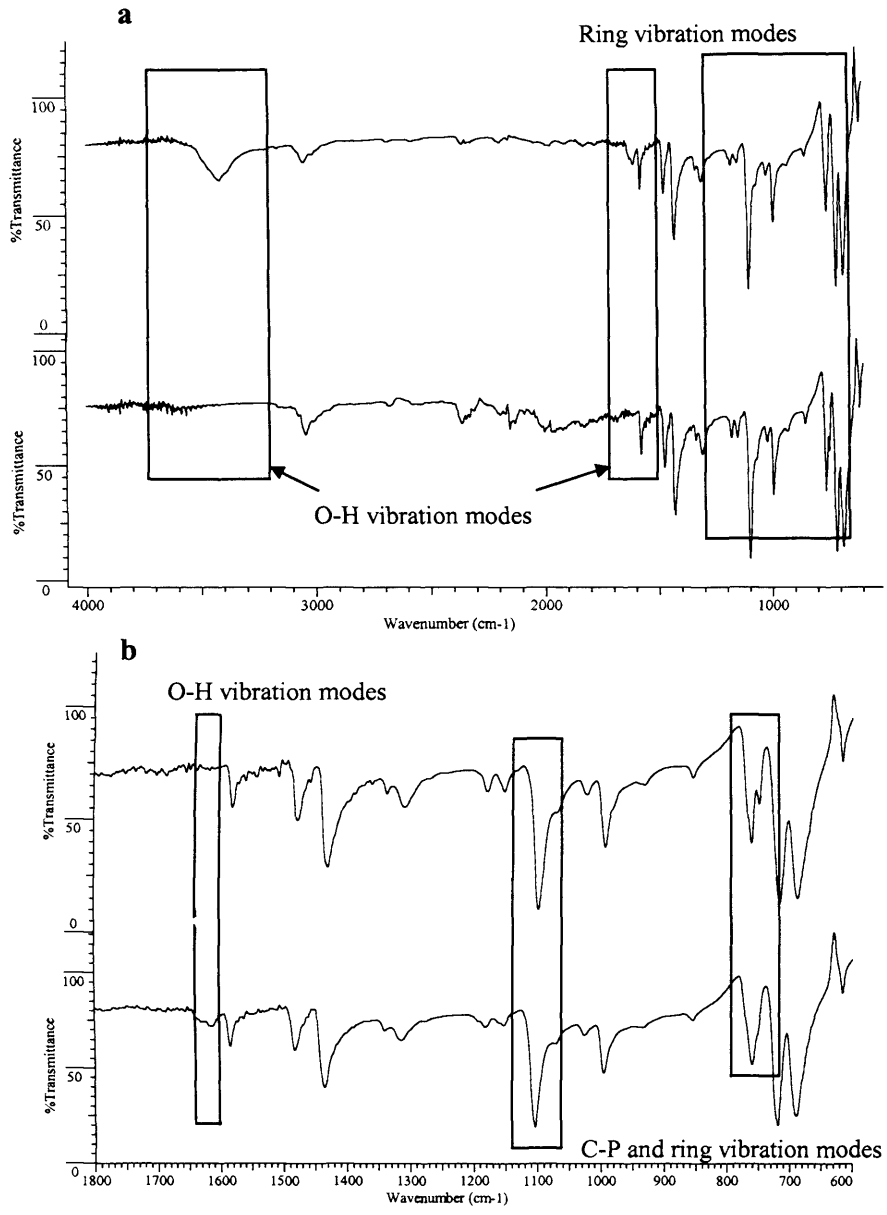



Figure 3.13: FTIR spectra of the anhydrate and dihydrate, showing (a) the full spectral region highlighting the areas of significant differences between the two phases and (b) the expanded finger print region

The dehydration/rehydration transitions were shown to be completely reversible under favourable humidity environments. Based on this evidence, coupled with vibrational spectroscopy data and the visual assessment of dehydration by optical microscopy, a proposed mechanism involves initiation of structural rearrangement prior to the loss of all the water of crystallisation. The apparent annealing event observed in the VT-XRPD experiments (as shown in Figure 3.5) supports this proposed mechanism of dehydration.

Dehydration as a function of temperature and relative humidity was also explored using non-ambient XRPD as shown in Figures 3.14 and 3.15.

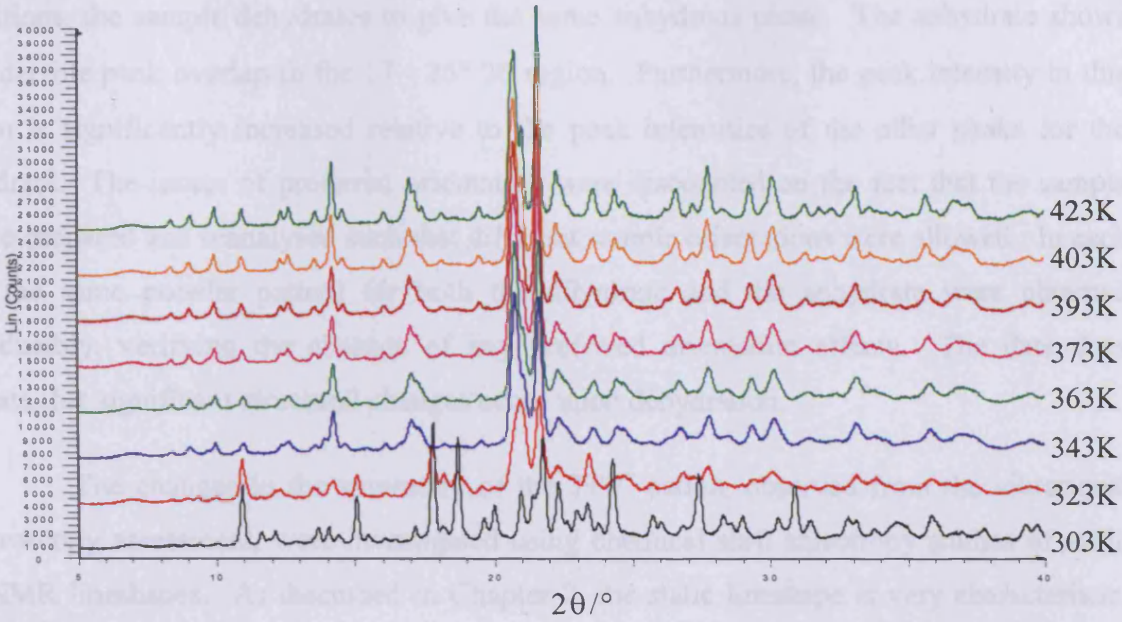



Figure 3.14: Variable Temperature XRPD of TPPB dihydrate at Ambient RH (heating from 303 to 423K)

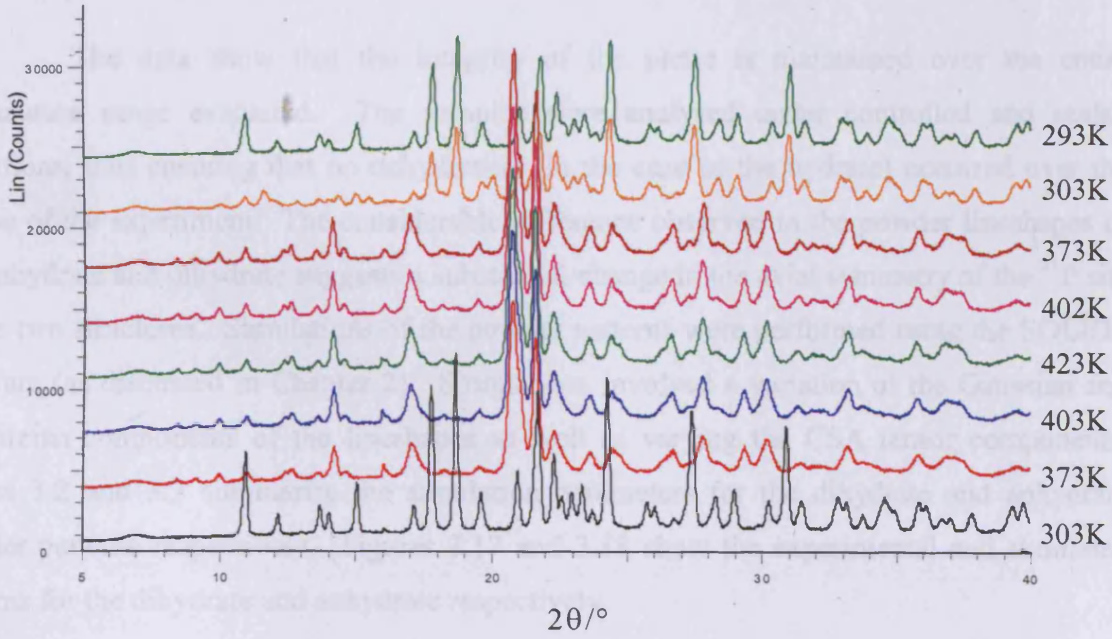



Figure 3.15: Variable Temperature XRPD of TPPB dihydrate at approximately 90% RH (heating from 303 to 423 K and then cooling from 423 to 293 K)

Under ambient humidity conditions dehydration is initiated at 323 K, whilst under an atmosphere of approximately 90% RH, dehydration started at 373 K. However, what is clearly evident is the fact that thermally induced dehydration occurs irrespective of the relative humidity. It is also interesting to note that, under both high and low humidity

conditions, the sample dehydrates to give the same anhydrous phase. The anhydrate shows considerable peak overlap in the 17 - 25° 2θ region. Furthermore, the peak intensity in this region is significantly increased relative to the peak intensities of the other peaks for the anhydrate. The issues of preferred orientation were discounted on the fact that the sample was re-prepared and reanalysed such that different sample orientations were allowed. In each case the same powder pattern for both the dihydrate and the anhydrate were obtained reproducibly, verifying the absence of any preferred orientation effects. The data does indicate that significant structural changes occur upon dehydration.

The changes in the symmetry of the $TPP^+$ cation, observed from the vibrational spectroscopy assessment, were investigated using chemical shift anisotropy studies of static $^{31}P$ NMR lineshapes. As discussed in Chapter 2, the static lineshape is very characteristic, and reflects the site symmetry at the nucleus under evaluation. Data were collected on sealed samples of the pure phases as a function of temperature; Figure 3.16 shows the experimental powder lineshape data for both the anhydrate and dihydrate phases.

The data show that the integrity of the phase is maintained over the entire temperature range evaluated. The samples were analysed under controlled and sealed conditions, thus ensuring that no dehydration (in the case of the hydrate) occurred over the course of the experiment. The considerable difference observed in the powder lineshapes of the anhydrate and dihydrate suggests a substantial change in the axial symmetry of the $^{31}P$ site in the two structures. Simulations of the powder patterns were performed using the SOLIDS software (as discussed in Chapter 2). Simulations involved a variation of the Gaussian and Lorentzian components of the lineshapes as well as varying the CSA tensor components. Tables 3.2 and 3.3 summarise the simulation parameters for the dihydrate and anhydrate powder patterns respectively. Figures 3.17 and 3.18 show the experimental and simulated patterns for the dihydrate and anhydrate respectively.

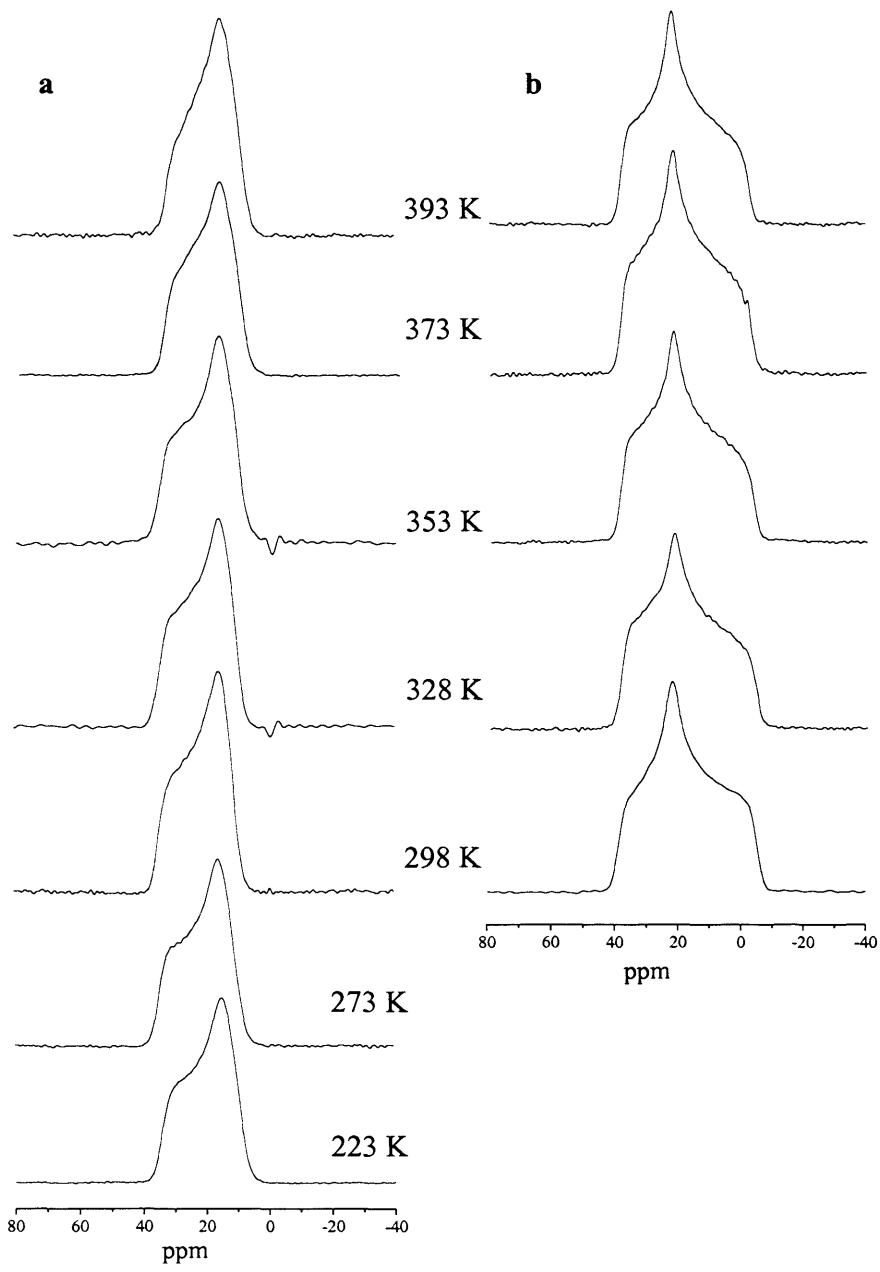



Figure 3.16: Static powder $^{31}P$ NMR lineshapes for (a) anhydrate and (b) dihydrate as a function of temperature

| T/K and model | G/L % | GB/Hz | LB/Hz | $\sigma_{XX}$ | $\sigma_{YY}$ | $\sigma_{ZZ}$ | σiso/ppm | Δ/ppm | η/ppm |
|---|---|---|---|---|---|---|---|---|---|
| 298 a | 90 | 400 | 200 | 22 | 0.5 | 40 | 20.83 | 19.17 | 0.54 |
| 298 b | 90 | 400 | 200 | 22 | 0.3 | 41 | 21.10 | 19.90 | 0.53 |
| 298 c | 100 | 500 | - | 22 | 0.5 | 40 | 20.83 | 19.17 | 0.54 |
| | | | | | | | | | |
| 328 a | 50 | 600 | 400 | 22 | 0.8 | 42 | 21.60 | 20.40 | 0.50 |
| 328 b | 50 | 800 | 300 | 23 | 0.5 | 40 | 21.17 | 18.83 | 0.56 |
| 328 c | 50 | 800 | 300 | 22 | 0.5 | 40 | 20.83 | 19.17 | 0.54 |
| | | | | | | | | | |
| 353 a | 50 | 600 | 400 | 22 | 0.5 | 40 | 20.83 | 19.17 | 0.54 |
| 353 b | 50 | 600 | 400 | 22.5 | 0.3 | 40.5 | 21.10 | 19.40 | 0.55 |
| 353 c | 50 | 600 | 400 | 22 | 0.3 | 40 | 20.77 | 19.23 | 0.54 |
| | | | | | | | | | |
| 373 a | 50 | 600 | 400 | 22 | 0.3 | 40 | 20.77 | 19.23 | 0.54 |
| 373 b | 50 | 600 | 300 | 21 | 0.3 | 39 | 20.10 | 18.90 | 0.53 |
| 373 c | 50 | 600 | 300 | 22 | 0.3 | 39 | 20.43 | 18.57 | 0.56 |
| | | | | | | | | | |
| 393 a | 50 | 600 | 400 | 22 | 0.5 | 40 | 20.83 | 19.17 | 0.54 |
| 393 b | 50 | 600 | 400 | 22.5 | 0.3 | 40.5 | 21.10 | 19.40 | 0.55 |
| 393 c | 50 | 600 | 300 | 22.5 | 0.3 | 40.5 | 21.10 | 19.40 | 0.55 |

Table 3.2: Summary of simulation parameters used for simulating TPPB dihydrate powder patterns using three different models

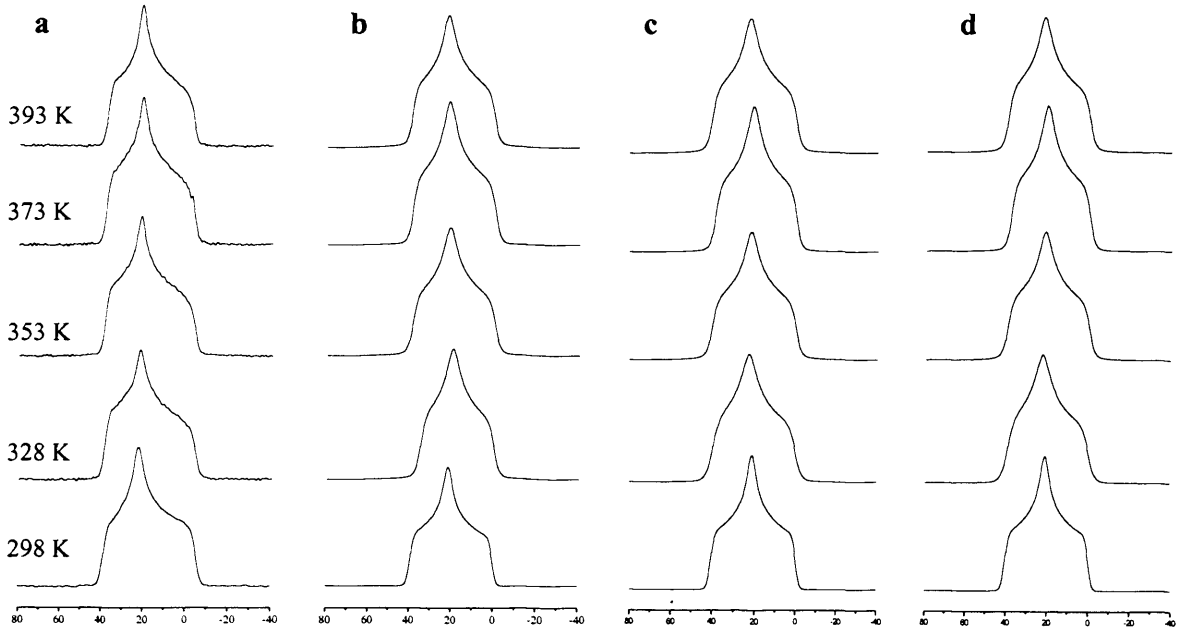



Figure 3.17: Experimental (a) and simulated powder patterns for TPPB dihydrate as a function of temperature. Three different models were considered (as summarised in the text) (b) model a, (c) model b and (d) refers to model c

| T/K and model | G/L % | GB/Hz | LB/Hz | $\sigma_{XX}$ | $\sigma_{YY}$ | $\sigma_{ZZ}$ | σiso/ppm | Δ/ppm | η/ppm |
|---|---|---|---|---|---|---|---|---|---|
| 223 a | 90 | 900 | 250 | 12.7 | 12.7 | 34.0 | 19.8 | 14.2 | 0 |
| 223 b | 90 | 600 | 100 | 15.5 | 8.5 | 35.0 | 19.7 | 15.3 | 0.20 |
| | | | | | | | | | |
| 273 a | 90 | 900 | 250 | 12.2 | 12.2 | 35.0 | 19.8 | 15.2 | 0 |
| 273 b | 100 | 700 | - | 15.5 | 10 | 34.0 | 19.8 | 14.17 | 0.16 |
| | | | | | | | | | |
| 298 a | 90 | 900 | 250 | 12.4 | 12.4 | 35.0 | 19.9 | 15.1 | 0 |
| 298 b | 100 | 900 | - | 15.5 | 10 | 34.0 | 19.8 | 14.17 | 0.16 |
| | | | | | | | | | |
| 328 a | 90 | 900 | 250 | 13.0 | 13.0 | 34.0 | 20.0 | 14.0 | 0 |
| 328 b | 90 | 1000 | 200 | 15.5 | 12 | 34.0 | 20.5 | 13.50 | 0.10 |
| | | | | | | | | | |
| 353 a | 90 | 900 | 250 | 13.1 | 13.1 | 34.0 | 20.1 | 13.9 | 0 |
| 353 b | 50 | 800 | 500 | 15.5 | 10 | 36.0 | 20.5 | 15.50 | 0.15 |
| | | | | | | | | | |
| 373 a | 90 | 900 | 250 | 13.2 | 13.2 | 34.0 | 20.1 | 13.9 | 0 |
| 373 b | 90 | 1000 | 500 | 16 | 12 | 34.0 | 20.7 | 13.33 | 0.12 |
| | | | | | | | | | |
| 393 a | 90 | 900 | 250 | 12.7 | 12.7 | 34.0 | 19.8 | 14.2 | 0 |
| 393 b | 50 | 900 | 500 | 15.3 | 9 | 32.5 | 18.9 | 13.57 | 0.19 |

Table 3.3: Summary of simulation parameters used for simulating TPPB anhydrate powder patterns using two different models

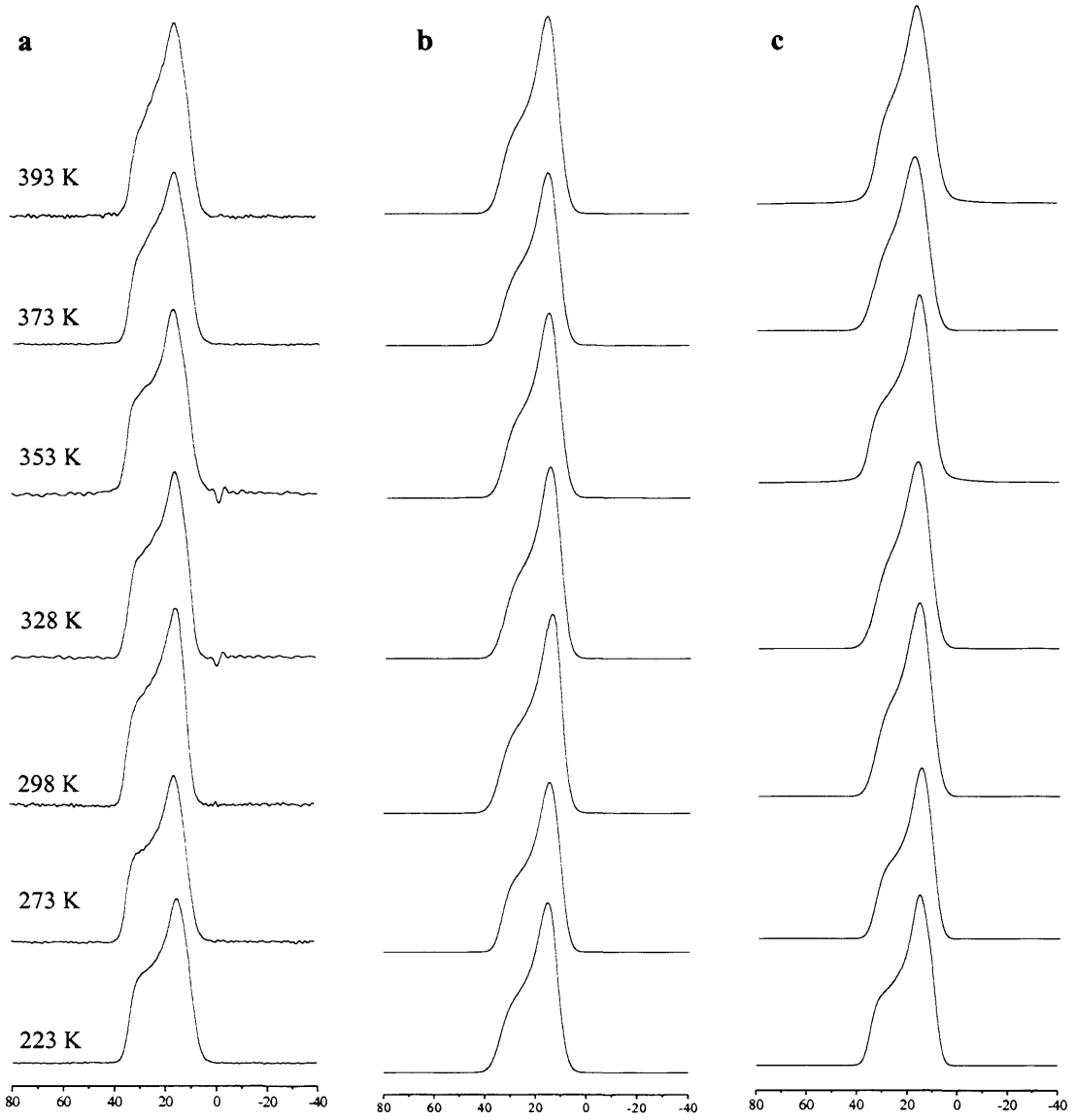



Figure 3.18: Experimental (a) and simulated static powder $^{31}P$ NMR lineshapes for TPPB anhydrate as a function of temperature. Two different models were considered (as summarised in the text) (b) model a (where $\eta = 0$) and (c) model b (where $\eta > 0$)

Previous variable temperature NMR studies on TPPB in open samples (Bach 1999) suggest that an increase in temperature (past the dehydration event) caused the $^{31}P$ chemical shift anisotropy to change from non-axially symmetric (or axially asymmetric) to axially symmetric. In the present study, all models considered for the dihydrate patterns indicate that the dihydrate exhibits axial asymmetry with an asymmetry parameter ($\eta$) of 0.5 to 0.56. Of the three simulation models considered for the dihydrate, model 'b' appears to offer the best fit of the experimental data. However, irrespective of model type, the conclusion regarding the structure is that the powder lineshape is not axially symmetric (i.e. $\eta \neq 0$). In the case of

the anhydrate, two models were considered; one with $\eta = 0$ (Figure 3.18b), and the other with $\eta$ in the range ~ 0.1 to 0.2 (Figure 3.18c). In general, a better fit of the experimental and simulated data is obtained with model b (Figure 3.18c), suggesting that whilst the TPP cation does not exhibit complete axial symmetry, the extent of asymmetry is reduced upon dehydration. A further observation from both the anhydrate simulations is that a significant Gaussian component contributes to the line broadening. This form of line broadening is attributed to the presence of homonuclear ($^{31}P$-$^{31}P$) dipolar coupling (Eichele and Wasylishen 1994) which may interfere with accurate lineshape simulations. To understand this further, experiments at very high magnetic fields or the use of homonuclear decoupling techniques may be required.

Assessment of the dihydrate (analysed as an open sample) by $^{13}C$ CP/MAS spectroscopy, as shown in Figure 3.19, reveals a broad and complex spectrum. The spectrum shows at least four different types of carbon, each split into at least five broad peaks. The broad and diffuse profile suggests possible disorder.

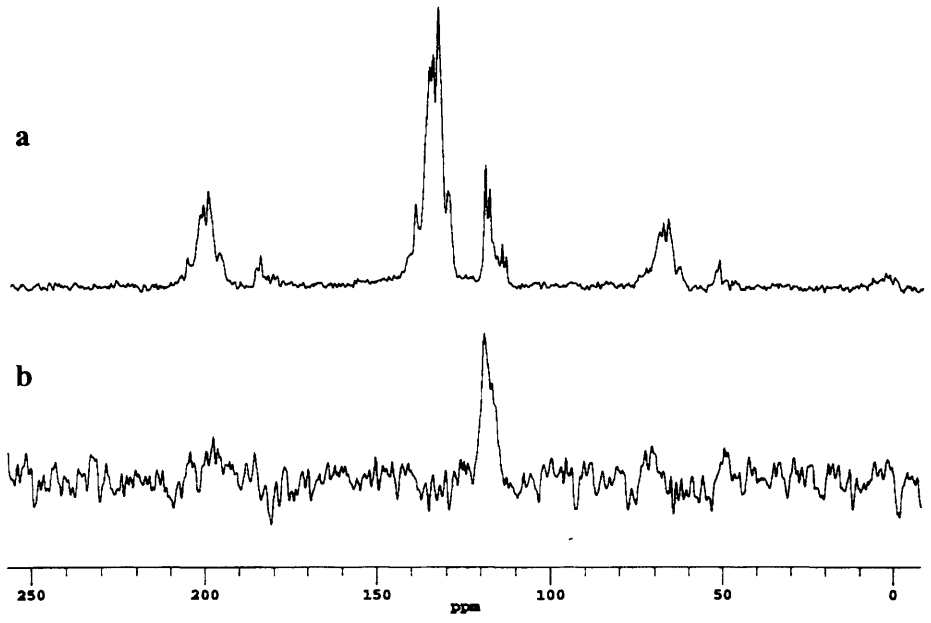



Figure 3.19: $^{13}$CP/MAS spectrum of TPPB dihydrate, (a) full spectrum and (b) NQS spectrum

Disorder and, in particular, molecular motion in TPPB dihydrate and anhydrous TPPB was investigated using $^{2}H$ NMR spectroscopy. The rates and mechanisms of any molecular reorientation were studied both by measuring spin-lattice relaxation times and by assessment of lineshapes as a function of temperature. The samples used in these experiments were pure phases of the dihydrate and anhydrous TPPB, deuterated at the ortho, meta and para

positions of the aromatic rings (referred to as TPPB-$d_{20}$ dihydrate and anhydrous TPPB-$d_{20}$ respectively). Relaxation times were measured using the saturation recovery technique (Chapter 2) at temperatures between 298 K and 368 K. The recovery curves were fitted to single and bi-exponential recovery functions to obtain relaxation times. It was found that fitting the data to a single exponential function gave very unsatisfactory results. A bi-exponential function, however, was found to give a very good fit of the data, indicating that there are at least two components that relax independently. The relaxation times for the two components (labelled fast and slow) are shown in Table 3.4, and the Arrhenius plot (of $\ln(T_1/s)$ versus K/T) is shown Figure 3.20.

| T/K | $T_1^{slow}$/s | $T_1^{fast}$/s |
|---|---|---|
| **298** | 4.60 | 2.18 |
| **308** | 3.47 | 1.03 |
| **318** | 2.77 | 0.35 |
| **338** | 2.65 | 0.23 |
| **348** | 2.41 | 0.14 |
| **358** | 2.20 | 0.09 |
| **368** | 1.60 | 0.06 |

Table 3.4: Relaxation parameters of TPPB dihydrate

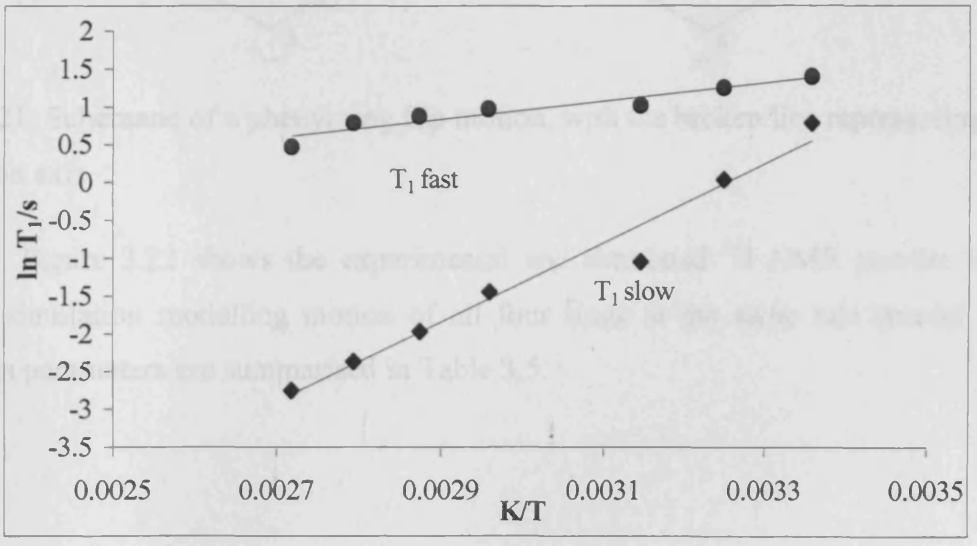



Figure 3.20: Arrhenius type plot for TPPB dihydrate, showing the changes of the fast and slow components of the spin-lattice relaxation times as a function of temperature

The linear fits of the two components of the $^2H$ spin lattice relaxation of TPPD-$d_{20}$ dihydrate, $T_1^{fast}$ and $T_1^{slow}$, give activation energies of $E_A^{fast}$ = 43.6 kJ $mol^{-1}$ and $E_A^{slow}$ = 11.1 kJ $mol^{-1}$. Anhydrous TPPB-$d_{20}$ is found to have far longer spin lattice relaxation times, for the fast component, at all temperatures studied. For example, at T = 25°C, a $T_1$ is *ca.* 40 s, compared with to *ca.* 2 s for TPPD-$d_{20}$ dihydrate.

The $^2H$ NMR powder lineshape of TPPB-$d_{20}$ dihydrate at 298 K is consistent with 180° phenyl ring flips about the phenyl C1-C4 axes, (refer to Figure 3.21) and therefore the spectra were simulated initially with a dynamic model that involves ring flips for all four phenyl rings. The lineshape simulations represent a summation of lineshapes for different components: one lineshape represents the ortho and meta deuterons (*i.e.* ortho and meta deuterons were considered to be equivalent with the same quadrupole coupling constant $\chi$, asymmetry parameter $\eta$, exchange rate k, and angle between the C—D bonds and the C1 to C4 axes of the phenyl rings), whereas a second effectively static lineshape was simulated for the para deuterons (with fast small amplitude librational motion accounted for by variation of the static quadrupole coupling constant and asymmetry parameter). The two simulated lineshapes were combined with a 4:1 scaling to account for the different numbers of ortho/meta and para deuterons. This model is referred to as model 1.

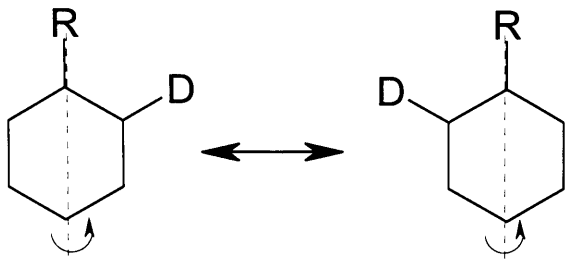



Figure 3.21: Schematic of a phenyl ring flip motion, with the broken line representing the C1-C4 rotation axis

Figure 3.22 shows the experimental and simulated $^2H$ NMR powder lineshape, with the simulation modelling motion of all four rings at the same rate (model 1). The simulation parameters are summarised in Table 3.5.

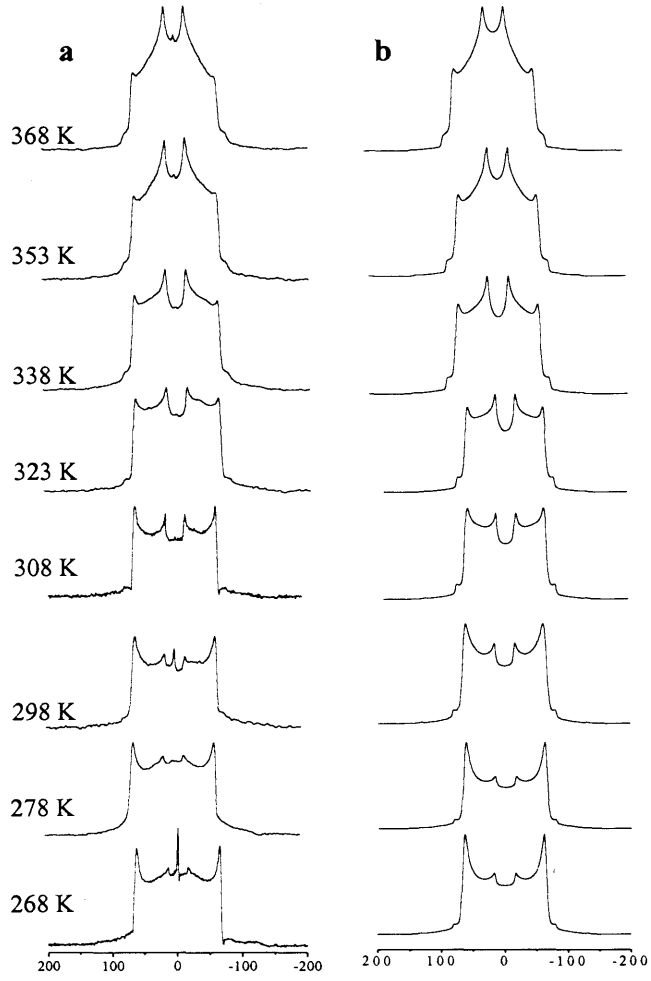



Figure 3.22: $^{2}H$ NMR spectra of TPPB-$d_{20}$ dihydrate, (a) experimental, (b) simulated using model 1 as described in the main text

| T/K | χ/kHz | | η | | $k/s^{-1}$ |
|---|---|---|---|---|---|
| | **ortho/meta** | **para** | **ortho/meta** | **para** | |
| 268 | 169.0 | 175.0 | 0.03 | 0.03 | $1 \times 10^5$ |
| 278 | 169.0 | 175.0 | 0.03 | 0.03 | $1.1 \times 10^5$ |
| 298 | 170.5 | 175.0 | 0.04 | 0.05 | $3.5 \times 10^5$ |
| 308 | 170.0 | 175.0 | 0.04 | 0.04 | $9 \times 10^5$ |
| 323 | 168.3 | 174.0 | 0.03 | 0.03 | $1.8 \times 10^6$ |
| 338 | 168.3 | 173.0 | 0.03 | 0.03 | $3.0 \times 10^6$ |
| 353 | 168.3 | 171.0 | 0.03 | 0.03 | $9 \times 10^6$ |
| 368 | 168.3 | 171.0 | 0.03 | 0.03 | $5 \times 10^7$ |

Table 3.5: Parameters for simulations of the $^2H$ NMR spectra of TPPB-$d_{20}$ dihydrate using model 1 described in the text.

Comparison of the simulated and experimental lineshapes indicates that model 1 is a good representation of the dynamics at temperatures greater than 308 K. At T = 278 K, for example, the intensity in the centre of the spectrum (between the outer horns) is clearly too high. These observations suggest that although some aspects of the model proposed for the motion of the aromatic rings are broadly correct, the exact details of the model may require further elaboration. The possibility that a phase change occurs within the temperature range of these experiments, and thus the mechanism of the molecular motion changes, was also considered. However, full characterisation (using XRPD and DSC) under varying conditions of relative humidity verified that no phase transitions occur in this temperature range. It should be noted that the analysis of the spin lattice relaxation data involves two components and is thus consistent with model 1 in which one component is associated with the ortho and meta deuterons ($T_1^{fast}$), whereas the other component is associated with the small amplitude librations of the para deuterons ($T_1^{slow}$).

Other motional models were also considered in which the phenyl rings all undergo 180° flips, but with a different exchange rate for each ring. The crystal structure of the dihydrate at ambient temperature is orthorhombic, spacegroup Pnma with 2 1/2 phenyl rings in the asymmetric unit (Vincent *et al.* 1988). This fact suggests the possibility that the three crystallographically independent phenyl rings, in the ratio 2:1:1, could undergo different rates of motion. Simulations describing three independent motions involving a combination of four

calculated lineshapes were therefore considered. Three of the lineshapes, representing the ortho and meta deuterons were calculated using the 180° ring flip model as previously described for model 1, but in this case the rates of the ring flips for the three lineshapes were independent. The fourth lineshape was a static simulation representing the para deuterons. The four calculated lineshapes were combined, with appropriate scaling to account for the different numbers of deuterons in each case, to obtain final simulations for comparison with the experimental spectra. Figure 3.23 shows the simulations in which three of the phenyl rings undergo ring flips in the fast motion regime ($k > 10^8$ $s^{-1}$), whereas the fourth ring is in the slow motion regime ($k < 10^3$ $s^{-1}$), designated model 2. These simulations reproduce the intensity in the centre of the experimental spectra for the patterns at 268 and 278 K (Figure 3.23a and b). However, this model insufficiently describes the motion of the sample at higher temperature (as in Figure 3.23c), indicating that this motional model is unsatisfactory in describing motion of the dihydrate as a function of temperature, and therefore does not reflect a true representation of the motion.

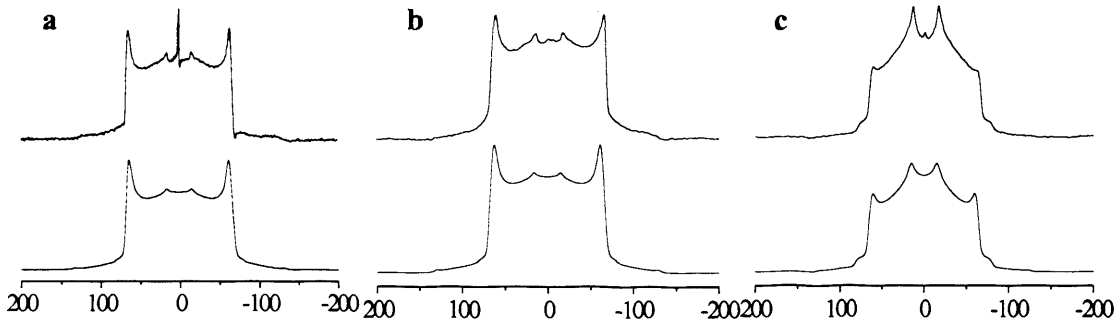



Figure 3.23: Simulated lineshape profiles of TPPB dihydrate with internal hygrostat (~100% RH) using model 2 at (a) 268 K, (b) 278 K and (c) 368 K

In conclusion, from the data generated from the analysis of the spin lattice relaxation data, involving two components rather than up to four components (as would be expected for model 2), and the evidence that no phase transitions in sealed systems over this temperature range occurs, model 1 best describes the motion of the dihydrate phase.

Simulations for the $^2H$ NMR powder lineshapes of the anhydrate were also performed and are shown in Figure 3.25. Tables 3.6 to 3.8 summarise the simulation parameters for each of the models used to perform the simulations. Model 'a' involves a combination of two calculated lineshapes: one representing the ortho and meta deuterons on one phenyl ring undergoing the ring flip motion in the intermediate regime, and the second

lineshape describing static ortho and meta deuterons on the remaining three phenyl rings and static para deuterons on all rings. As before, the two calculated lineshapes were combined, with appropriate scaling to account for the different numbers of deuterons in each case, to obtain the final simulations for comparison with the experimental spectra. Model 'b' was similar to model 'a' in that it describes one ring undergoing ring flip motion, while the remaining three are static, however, the rate of motion is reduced. Model 'c' involves the combination of three lineshapes: one lineshape describes the 180° ring flip motion of one phenyl ring at a given rate, whereas the other three phenyl rings undergo the ring flip motion at an independent rate, the third lineshape describes the static para deuterons. The primary rationale for assessing these motional models was motivated by the possible presence of axial symmetry associated with the $TPP^+$ cation (as discussed previously). Figure 3.24 illustrates a three-fold axis of symmetry, which would give rise to axial symmetry for the $TPP^+$ cation. For this symmetry, it is reasonable that one phenyl ring (in Figure 3.24) may undergo motion at a different rate to the remaining three rings. In order for the symmetry to remain, it is plausible for the three rings related by the 3-fold axis, to be either static or to undergo motion at the same rate as each other.

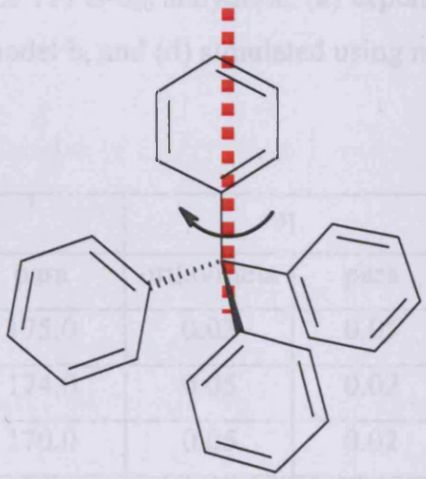


Figure 3.24: $TPP^+$ cation showing a three-fold axis of symmetry (denoted by the red broken line)

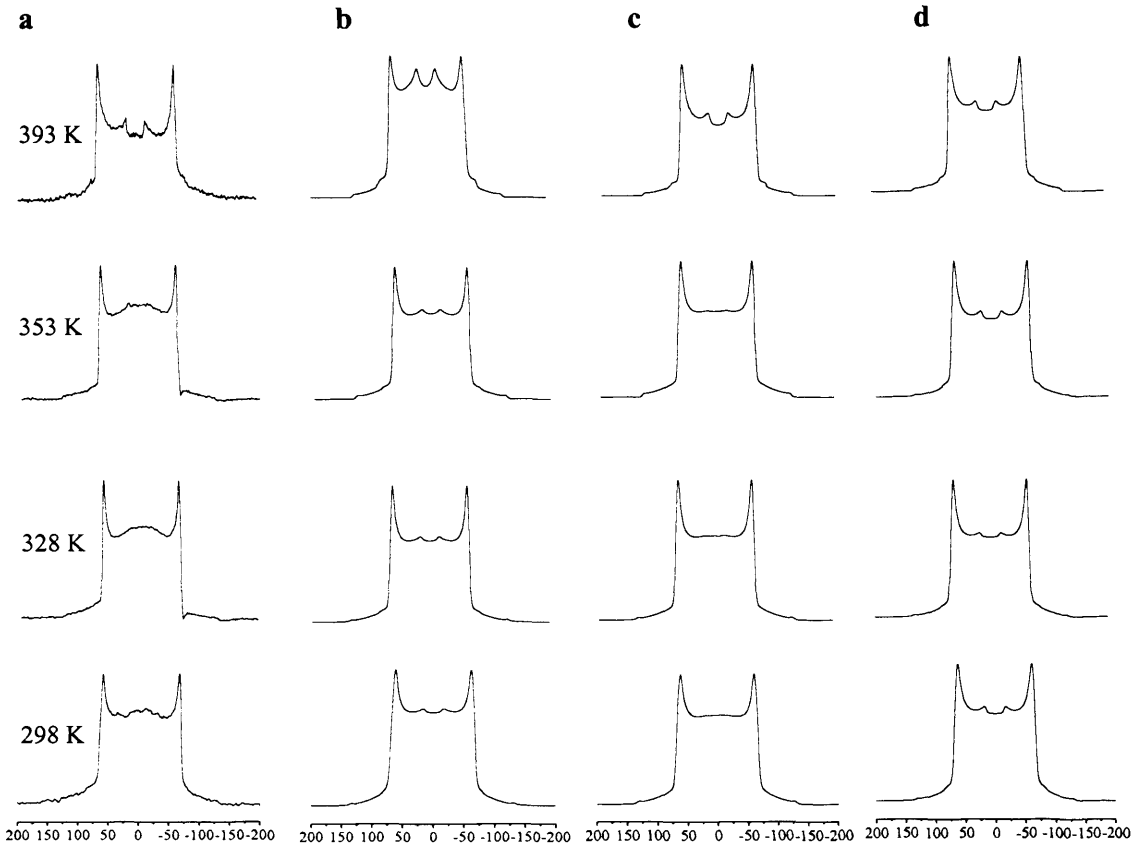



Figure 3.25: $^2H$ NMR spectra of TPPB-$d_{20}$ anhydrate, (a) experimental, (b) simulated using model a, (c) simulated using model b, and (d) simulated using model c, as described in the text

| T/K | χ/kHz | | η | | $k/s^{-1}$ |
|---|---|---|---|---|---|
| | ortho/meta | para | ortho/meta | para | |
| 298 | 169.0 | 175.0 | 0.03 | 0.05 | $1 \times 10^6$ |
| 328 | 174.0 | 174.0 | 0.05 | 0.02 | $3 \times 10^6$ |
| 353 | 173.0 | 170.0 | 0.05 | 0.02 | $5.5 \times 10^6$ |
| 393 | 173.0 | 169.0 | 0.05 | 0.03 | $5.5 \times 10^6$ |

Table 3.6: Parameters for the simulations of the $^2H$ NMR spectra of TPPB-$d_{20}$ anhydrate using model a (as described in the text)

| T/K | χ/kHz | | η | | $k/s^{-1}$ |
|---|---|---|---|---|---|
| | ortho/meta | para | ortho/meta | para | |
| 298 | 172 | 173 | 0.03 | 0.05 | $2 \times 10^4$ |
| 328 | 170 | 174 | 0.05 | 0.02 | $4 \times 10^4$ |
| 353 | 173 | 170 | 0.05 | 0.02 | $8 \times 10^4$ |
| 393 | 168 | 169 | 0.04 | 0.03 | $10 \times 10^4$ |

Table 3.7: Parameters for the simulations of the $^2H$ NMR spectra of TPPB-$d_{20}$ anhydrate using model b (as described in the main text)

| T/K | Ortho/meta fast | | | Ortho/meta medium | | | Para | |
|---|---|---|---|---|---|---|---|---|
| | χ/kHz | η | $k/s^{-1}$ | χ/kHz | η | $k/s^{-1}$ | χ/kHz | η |
| 298 | 172 | 0.03 | $3 \times 10^4$ | 172 | 0.03 | $2 \times 10^4$ | 173 | 0.05 |
| 328 | 175 | 0.02 | $5 \times 10^4$ | 174 | 0.02 | $2 \times 10^4$ | 174 | 0.02 |
| 353 | 173 | 0.04 | $1 \times 10^5$ | 173 | 0.03 | $2 \times 10^4$ | 174 | 0.03 |
| 393 | 169 | 0.04 | $2 \times 10^5$ | 169 | 0.03 | $2 \times 10^4$ | 169 | 0.03 |

Table 3.8: Parameters for simulations of the $^2H$ NMR spectra of TPPB-$d_{20}$ anhydrate using model c (as described in the main text)

Model 'a' does not sufficiently describe the motion observed at 393 K, although the simulations at lower temperatures do show a good fit to the observed data. Rates of motion are known to increase monotonically with temperature, except when a discontinuity occurs due to the occurrence of a phase transition. As it is known that no phase transitions occur over this temperature range (from non-ambient spectroscopic and diffraction studies), other models for simulating the 393 K powder lineshape (describing the same motion regime but different rates) were considered. The results are summarised in Table 3.9 and Figure 3.26. From this data the model termed b2 was deemed to give the best simulation of the 393 K lineshape and was thus used to simulate the lineshapes at the lower temperatures. This regime formed the basis of model 'b' illustrated in Figure 3.25 and Table 3.7.

| Model | χ/kHz | | η | | $k/s^{-1}$ |
|---|---|---|---|---|---|
| | ortho/meta | para | ortho/meta | para | |
| b1 | 173 | 169 | 0.05 | 0.03 | $5.5 \times 10^6$ |
| b2 | 168 | 169 | 0.04 | 0.03 | $1 \times 10^5$ |
| b3 | 168 | 170 | 0.04 | 0.04 | $2 \times 10^5$ |
| b4 | 168 | 170 | 0.03 | 0.04 | $3 \times 10^5$ |
| b5 | 168 | 169 | 0.03 | 0.04 | $2.5 \times 10^5$ |
| b6 | 168 | 169 | 0.05 | 0.03 | $6 \times 10^6$ |

Table 3.9: Parameters for simulations of the $^2H$ NMR spectra of TPPB-$d_{20}$ anhydrate at 393 K using different motional models (b1 to b6)

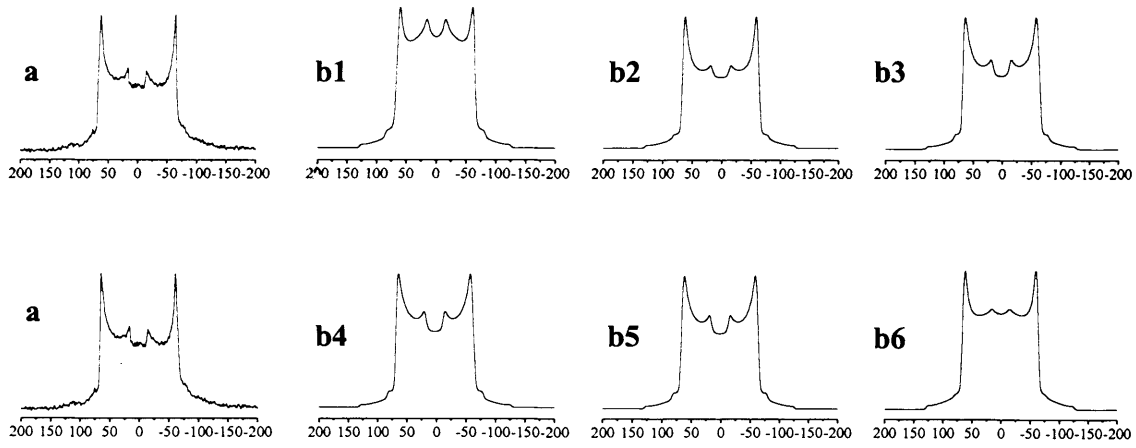



Figure 3.26: Simulated lineshape profiles of TPPB anhydrate at 393K using different simulation models labelled as b1 to b6 (refer to text for details of the models used), the experimental data is shown in (a)

The results indicate that both model 'b' and 'c' (as shown in Figure 3.25) give satisfactory simulations of the powder lineshapes as a function of temperature. It can therefore be concluded, on the basis of spectroscopic evidence, that the $TPP^+$ cation in the anhydrate phase exhibits either axial symmetry or an arrangement that is close to axial symmetry. The different temperature dependencies of the powder lineshapes of the dihydrate and the anhydrate are therefore attributed to a significant difference in motion regime of the phenyl rings in the two phases.

### 3.4.1 Structure solution from powder XRD of the anhydrate phase of TPPB and comparison with the reported dihydrate structure

Growth of single crystals of the anhydrate phase was not feasible, as the sample under ambient conditions was always present as the dihydrate (or a mixed phase composed of the monohydrate and dihydrate). The anhydrate phase cannot be isolated and is only formed under conditions of high temperature or low relative humidity. As such, powder XRPD data was used to determine the structure of this high temperature phase. Data acquired at 423 K were used to index the structure giving a triclinic cell with parameters; a = 9.49 Å, b = 10.29 Å, c = 11.43 Å, $\alpha$ = 75.1°, $\beta$ = 72.4°, $\gamma$ = 79.2°, and cell volume V = 1021 Å$^3$. Numerous high temperature data sets were collected (from two instruments) and each powder pattern was indexed using TREOR, ITO, DICVOL and XCELL (as described in Chapter 2). In 23 of 32 cases, the same the cell as shown above was obtained.

Space group assignment was achieved during Pawley refinement. Due to density considerations, refinement was performed using the $P\bar{1}$ spacegroup. A best-fit using Pawley refinement was achieved with the $P\bar{1}$ spacegroup, giving $R_{wp}$ = 3.40% (with the Pawley fit shown in Figure 3.27) using the Berar-Baldinozzi asymmetry correction algorithm. However, the Rietveld asymmetry correction algorithm was used, as this was the asymmetry parameter used in the GA software EAGER. Refinement in the $P\bar{1}$ space group gives $R_{wp}$ = 10.37%. The Pawley refined plot is shown in Figure 3.28.

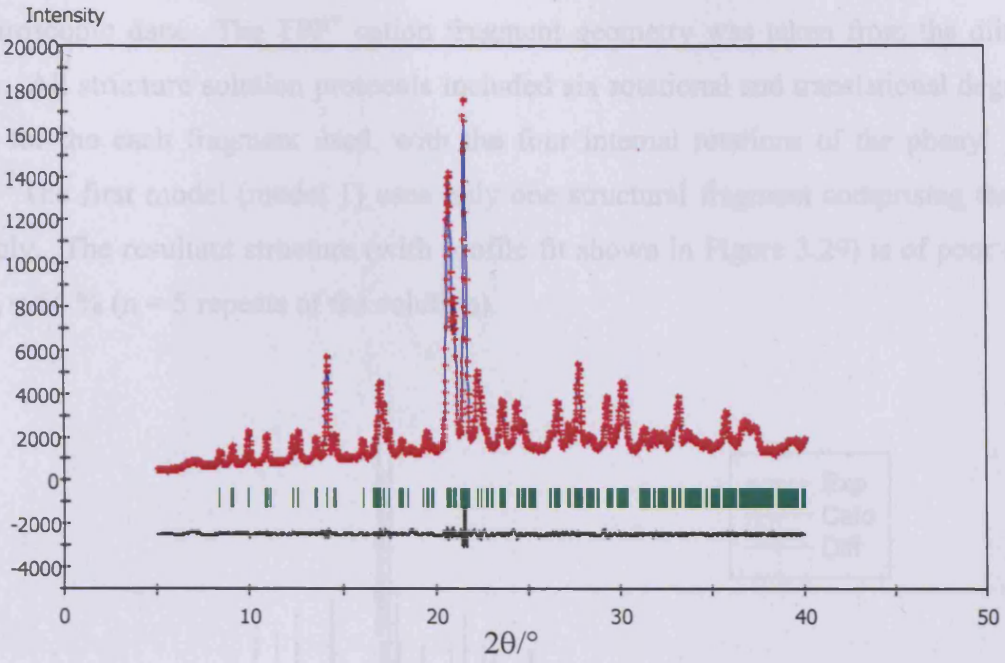



Figure 3.27: Pawley refinement for TPPB anhydrate following indexing by TREOR of data collected on a Bruker AXS D5000 (showing experimental data in red, calculated data in blue and the difference plot in the lower trace in black)

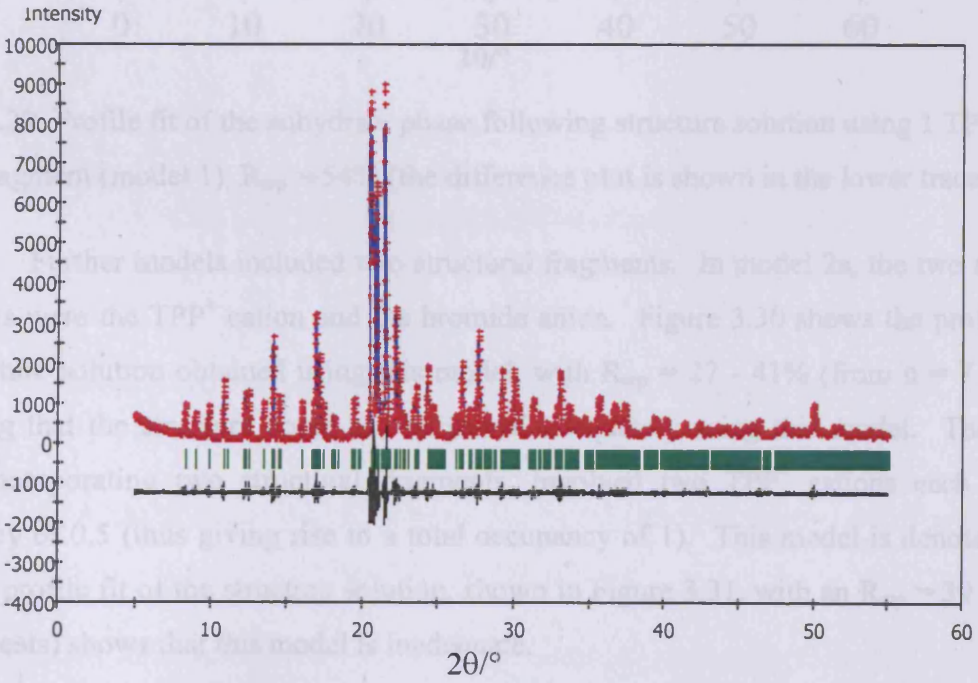



Figure 3.28: Pawley refinement for TPPB anhydrate following indexing by TREOR of data collected on a Panalytical Xpert XRPD (showing experimental data in red, calculated data in blue and the difference plot in the lower trace in black)

Structure solution was performed using different input models in accordance with the spectroscopic data. The $TPP^+$ cation fragment geometry was taken from the dihydrate structure. All structure solution protocols included six rotational and translational degrees of freedom for the each fragment used, with the four internal rotations of the phenyl groups allowed. The first model (model 1) uses only one structural fragment comprising the $TPP^+$ cation only. The resultant structure (with profile fit shown in Figure 3.29) is of poor quality with $R_{wp}$ = 54 % (n = 5 repeats of the solution).

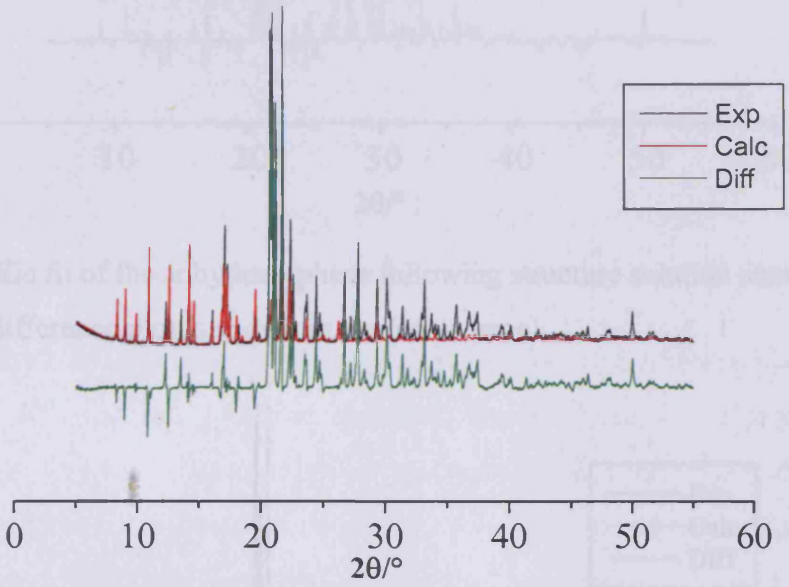



Figure 3.29: Profile fit of the anhydrate phase following structure solution using 1 $TPP^+$ cation fragment (model 1), $R_{wp}$ = 54% (the difference plot is shown in the lower trace)

Further models included two structural fragments. In model 2a, the two structural fragments were the $TPP^+$ cation and the bromide anion. Figure 3.30 shows the profile fit of the structure solution obtained using this model, with $R_{wp}$ = 27 - 41% (from n = 7 repeats), indicating that the structure could not be solved adequately using this model. The second model incorporating two structural fragments, involved two $TPP^+$ cations each with an occupancy of 0.5 (thus giving rise to a total occupancy of 1). This model is denoted model 2b. The profile fit of the structure solution, shown in Figure 3.31, with an $R_{wp}$ = 39 % (from n = 3 repeats) shows that this model is inadequate.

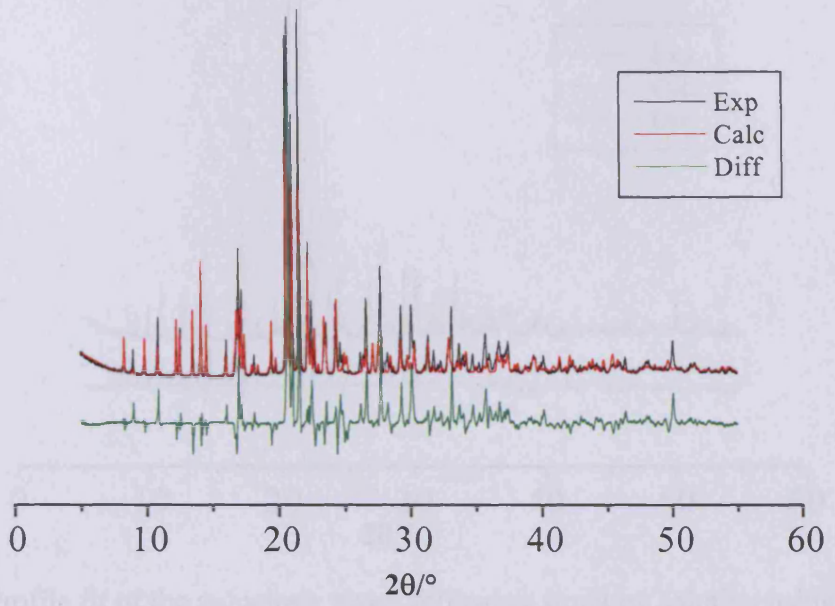



Figure 3.30: Profile fit of the anhydrate phase following structure solution using model 2a, $R_{wp}$ = 27% (the difference plot is shown in the lower trace)

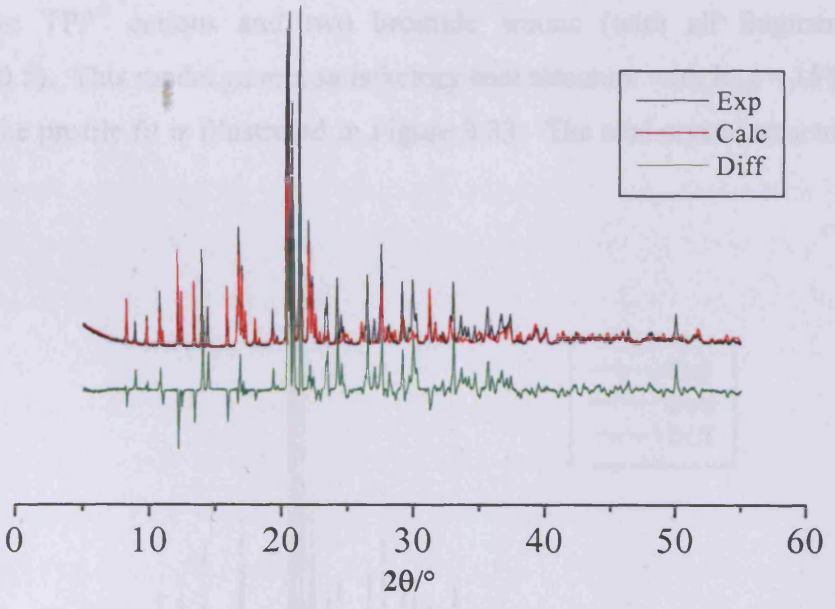



Figure 3. 31: Profile fit of the anhydrate phase following structure solution using model 2b, $R_{wp}$ = 39% (the difference plot is shown in the lower trace)

Using this systematic approach, to establish the appropriate input model for structure solution, three structural fragments were used in model 3, involving two $TPP^+$ cations, each with occupancy 0.5, and one bromide anion with unit occupancy. The resultant structure (with profile fit shown in Figure 3.32) is of better quality with $R_{wp}$ = 23 % (from n = 5 repeats of the solution).

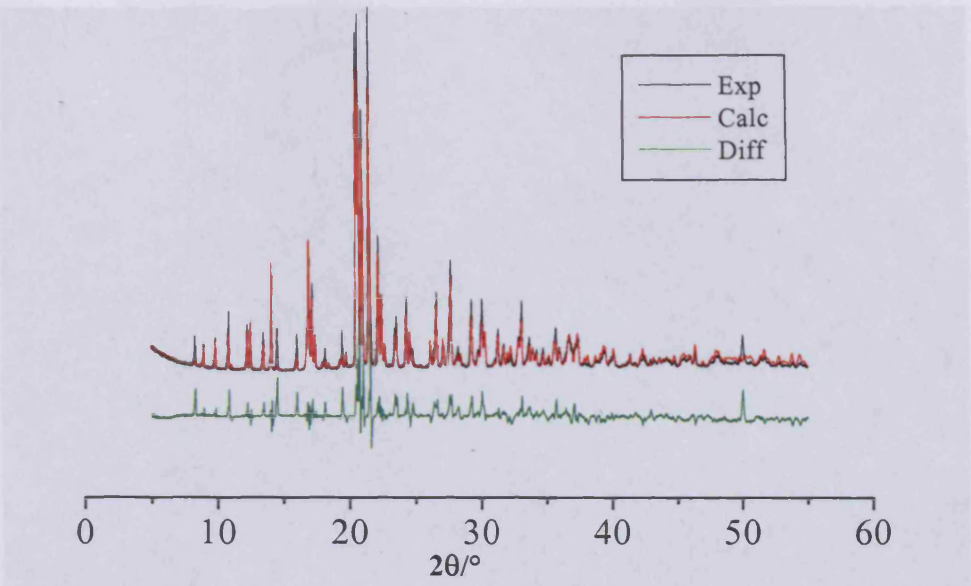



Figure 3.32: Profile fit of the anhydrate phase following structure solution using model 3, $R_{wp}$ = 23% (the difference plot is shown in the lower trace)

The final model that was considered, model 4, incorporated four structural fragments; two $TPP^+$ cations and two bromide anions (with all fragments having an occupancy of 0.5). This model gave a satisfactory trial structure with $R_{wp}$ = 15% (from n = 10 repeats), and the profile fit is illustrated in Figure 3.33. The trial crystal structure is shown in Figure 3.34.

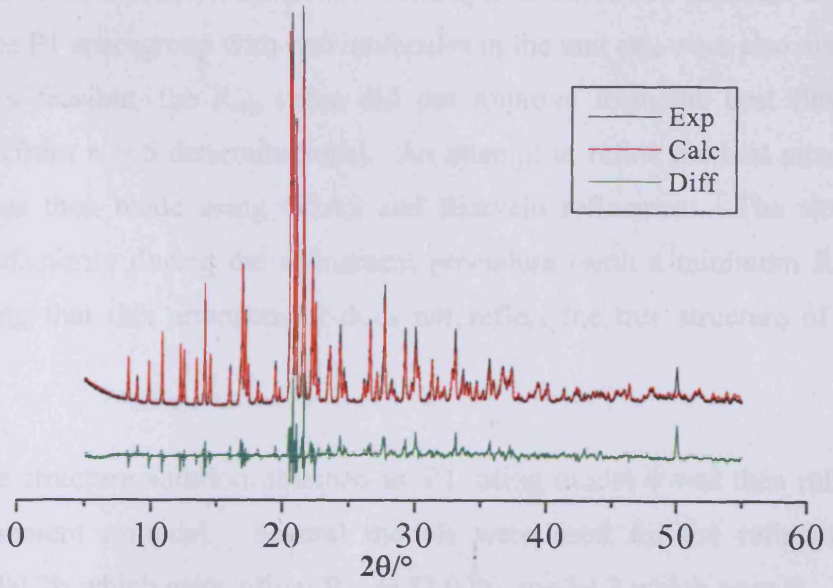



Figure 3.33: Profile fit of the anhydrate phase following structure solution using model 4, $R_{wp}$ = 15% (the difference plot is shown in the lower trace)

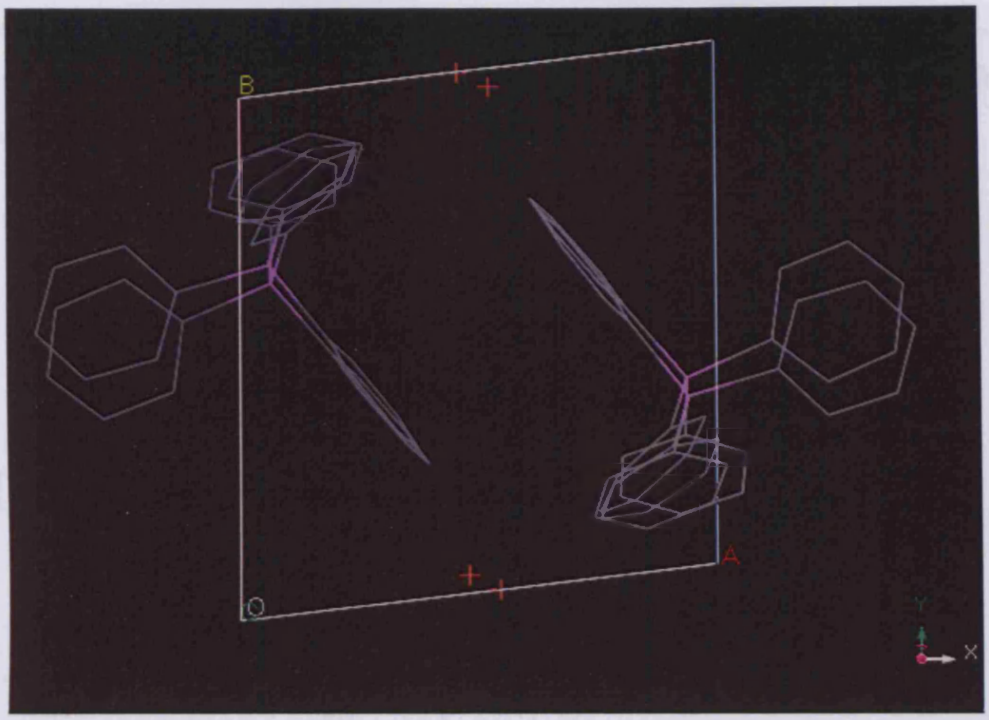



Figure 3.34: Trial structure of TPPB anhydrate solved using 4 structural fragments described by model 4

The structure indicates an apparent disorder for both the cation and the anion. The systematic approach of using different input models to obtain a suitable trial structure shows the requirement of two molecules of TPPB in the asymmetric unit (each with 0.5 occupancy) in order to arrive at a sensible structure. Attempts to solve the structure of the anhydrate phase using the P1 spacegroup with two molecules in the unit cell were also made. Whilst the structure looks feasible, the $R_{wp}$ value did not improve from the best (lowest) value of $R_{wp}$ = 21.6% (from n = 5 determinations). An attempt to refine the best structure from this assessment was then made using GSAS and Rietveld refinement. The structure did not converge satisfactorily during the refinement procedure (with a minimum $R_{wp}$ of 21.98%) thus concluding that this arrangement does not reflect the true structure of the anhydrate material.

The structure solution obtained in $P\bar{1}$ using model 4 was then refined using the Rietveld refinement protocol. Several models were used for the refinement procedure including model 2b which gave a final $R_{wp}$ = 52.02%, model 3 which gave $R_{wp}$ = 30.18% and finally, the correct structure which gave $R_{wp}$ = 13.48%. Figure 3.35 illustrates the profile fit obtained from the best refinement protocol, which uses 4 structural fragments during the refinement (based on model 4). Preferred orientation did not serve to influence the nature of

the refinement and was therefore excluded from the final refined structure reported. Figure 3.36 shows the final refined structure of TPPB anhydrate, and Table 3.10 gives the fractional coordinates for this refined structure.

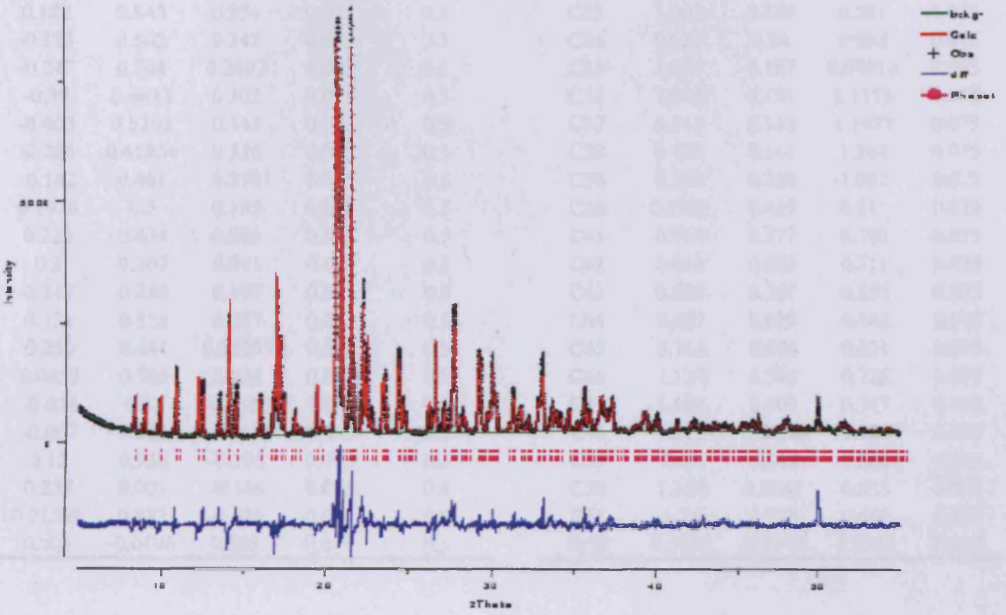


Figure 3.35: Profile fit with $R_{wp}$ = 13.48% ($R_p$ = 9.81%) following Rietveld refinement of TPPB anhydrate (showing the experimental data in black, the calculated data in red and the difference plot in blue in the lower trace)

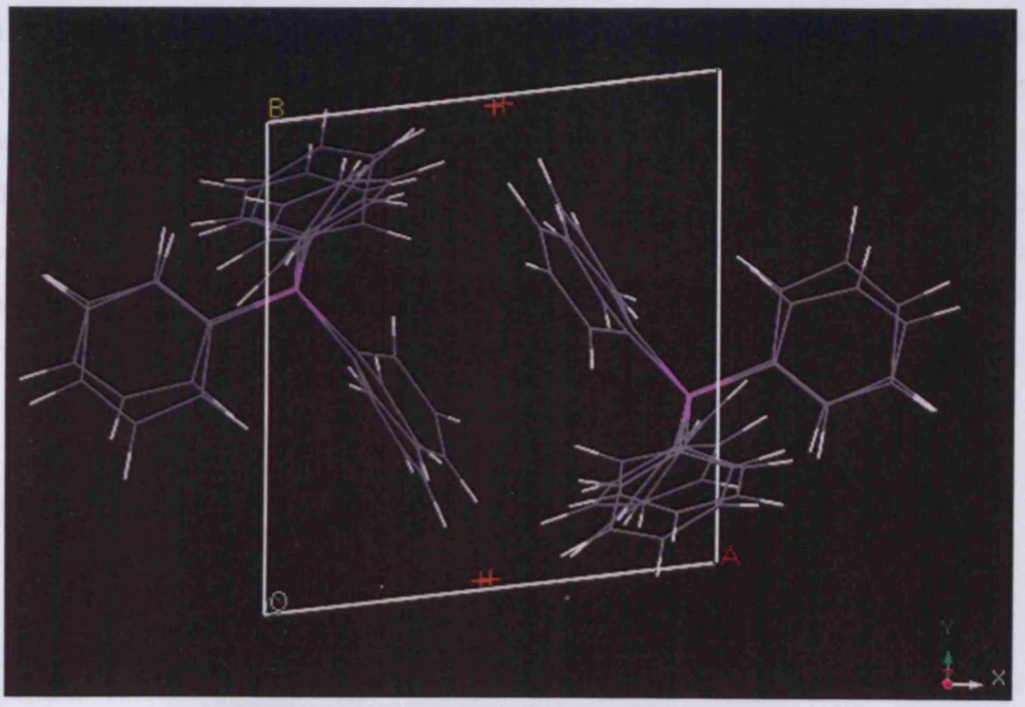



Figure 3.36: Rietveld refined structure of TPPB anhydrate

| Atom | x | y | z | Uiso | Occupancy |
|---|---|---|---|---|---|
| P1 | 0.0656 | 0.653 | 0.188 | 0.075 | 0.5 |
| C2 | 0.088 | 0.746 | 0.295 | 0.075 | 0.5 |
| C3 | 0.025 | 0.706 | 0.422 | 0.075 | 0.5 |
| C4 | 0.074 | 0.754 | 0.505 | 0.075 | 0.5 |
| C5 | 0.161 | 0.857 | 0.467 | 0.075 | 0.5 |
| C6 | 0.23 | 0.89 | 0.34 | 0.075 | 0.5 |
| C7 | 0.182 | 0.845 | 0.254 | 0.075 | 0.5 |
| C8 | -0.123 | 0.602 | 0.242 | 0.075 | 0.5 |
| C9 | -0.247 | 0.704 | 0.2692 | 0.075 | 0.5 |
| C10 | -0.39 | 0.6633 | 0.302 | 0.075 | 0.5 |
| C11 | -0.403 | 0.5202 | 0.342 | 0.075 | 0.5 |
| C12 | -0.286 | 0.41854 | 0.328 | 0.075 | 0.5 |
| C13 | -0.142 | 0.461 | 0.278 | 0.075 | 0.5 |
| C14 | 0.1936 | 0.5 | 0.183 | 0.075 | 0.5 |
| C15 | 0.222 | 0.434 | 0.085 | 0.075 | 0.5 |
| C16 | 0.3 | 0.307 | 0.091 | 0.075 | 0.5 |
| C17 | 0.347 | 0.246 | 0.197 | 0.075 | 0.5 |
| C18 | 0.334 | 0.314 | 0.287 | 0.075 | 0.5 |
| C19 | 0.256 | 0.441 | 0.28297 | 0.075 | 0.5 |
| C20 | 0.0853 | 0.765 | 0.034 | 0.075 | 0.5 |
| C21 | -0.034 | 0.81 | -0.0165 | 0.075 | 0.5 |
| C22 | -0.007 | 0.882 | -0.1412 | 0.075 | 0.5 |
| C23 | 0.12 | 0.938 | -0.198 | 0.075 | 0.5 |
| C24 | 0.237 | 0.907 | -0.146 | 0.075 | 0.5 |
| C25 | 0.21549 | 0.822 | -0.026 | 0.075 | 0.5 |
| Br26 | 0.503 | -0.0196 | 0.503 | 0.075 | 0.5 |

| Atom | x | y | z | Uiso | Occupancy |
|---|---|---|---|---|---|
| P27 | 0.938 | 0.343 | 0.7942 | 0.075 | 0.5 |
| C28 | 0.903 | 0.244 | 0.697 | 0.075 | 0.5 |
| C29 | 0.77 | 0.188 | 0.732 | 0.075 | 0.5 |
| C30 | 0.7262 | 0.146 | 0.642 | 0.075 | 0.5 |
| C31 | 0.8245 | 0.137 | 0.533 | 0.075 | 0.5 |
| C32 | 0.9653 | 0.172 | 0.502 | 0.075 | 0.5 |
| C33 | 1.002 | 0.239 | 0.581 | 0.075 | 0.5 |
| C34 | 0.927 | 0.24 | 0.952 | 0.075 | 0.5 |
| C35 | 1.057 | 0.187 | 0.98916 | 0.075 | 0.5 |
| C36 | 1.045 | 0.131 | 1.1175 | 0.075 | 0.5 |
| C37 | 0.911 | 0.112 | 1.1977 | 0.075 | 0.5 |
| C38 | 0.783 | 0.161 | 1.164 | 0.075 | 0.5 |
| C39 | 0.787 | 0.228 | 1.037 | 0.075 | 0.5 |
| C40 | 0.7983 | 0.489 | 0.811 | 0.075 | 0.5 |
| C41 | 0.764 | 0.577 | 0.706 | 0.075 | 0.5 |
| C42 | 0.668 | 0.693 | 0.719 | 0.075 | 0.5 |
| C43 | 0.605 | 0.721 | 0.836 | 0.075 | 0.5 |
| C44 | 0.627 | 0.629 | 0.942 | 0.075 | 0.5 |
| C45 | 0.716 | 0.506 | 0.931 | 0.075 | 0.5 |
| C46 | 1.127 | 0.391 | 0.725 | 0.075 | 0.5 |
| C47 | 1.164 | 0.509 | 0.747 | 0.075 | 0.5 |
| C48 | 1.311 | 0.5255 | 0.729 | 0.075 | 0.5 |
| C49 | 1.42 | 0.442 | 0.666 | 0.075 | 0.5 |
| C50 | 1.389 | 0.3247 | 0.633 | 0.075 | 0.5 |
| C51 | 1.24 | 0.296 | 0.666 | 0.075 | 0.5 |
| Br52 | 0.5235 | -0.0179 | 1.0113 | 0.075 | 0.5 |

Table 3.10: Final refined atomic coordinates and isotropic displacement factors for TPPB anhydrate (for non-hydrogen atoms only)

The refined structure shows clear disorder in the phenyl rings, with the phenyl rings tilted out of plane from each other by up to 1 Å. The P···P distance between the $TPP^+$ cations in the disordered structure (shown in Figure 3.36) is 0.21 Å. The root mean squared difference (rmsd) between corresponding atoms in the phenyl rings are summarised in table 3.11. The two molecules in the asymmetric unit are shown in Figure 3.37, the overlay of the two crystallographically independent $TPP^+$ cations in the unit cell illustrates a slight variation in the planar arrangements of the rings in the tetrahedral arrangement.

| Phenyl ring | Rmsd/Å |
|---|---|
| 1 | 0.134 |
| 2 | 0.197 |
| 3 | 0.183 |
| 4 | 0.145 |

Table 3.11: The rmsd values for the phenyl rings describing the disorder in TPPB anhydrate

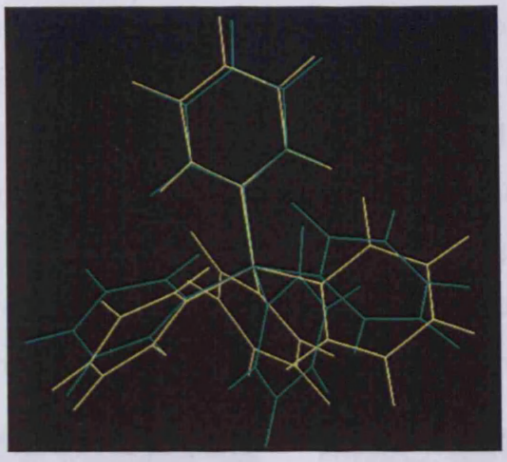


Figure 3.37: Summary of the asymmetric unit containing two molecules of TPPB showing an overlay of the two crystallographically independent $TPP^+$ cations

Figure 3.38 illustrates the perspective along the three crystallographic axes. Tetraphenyl phosphonium salts are known to pack such that the $TPP^+$ cations form parallel stacks with the anions occupying the interstitial spaces between the stacks (Bogaard and Rae 1982). The view along crystallographic b and c axes clearly show parallel stacks giving rise to a zigzag arrangement allowing edge-to-face and offset face-to-face phenyl interactions.

It is apparent that there is a requirement for two molecules (of both the $TPP^+$ cation and bromide anion) with half occupancy to be present within the asymmetric unit in the average structure in order to sufficiently describe the disorder within the anhydrate structure. As crystallography provides a space and time averaged representation of the structure, it is difficult to discern the exact mode of disorder exhibited by this structure solely on the basis of the XRPD data. NMR spectroscopy indicates the possibility for the cation to exhibit near axial symmetry, and that there is some dynamic disorder present, associated with the phenyl rings involved in ring flip motion. The phenyl disorder, though described as complete ring flips, could in fact be better represented by librational oscillations coupled with 180° jumps. Moreover, the 180° ring flip motion, as deduced from $^2H$ NMR, interconvert structures that are indistinguishable from each other so should not be described by the XRPD structure. The disorder as observed in this structure must be in addition to the previously described 180° ring flip motion. The relatively small distances between like-phenyl rings (as exemplified by the rmsd metric) would suggest dynamic disorder is exhibited. Furthermore, it has been reported that the tetramethyl phosphonium bromide salt undergoes librational oscillation of the entire cation as a whole (Ang and Dunell 1976). The disorder shown in the crystal structure may also be due to small amplitude librations of the whole cation. $^2H$ NMR lineshape simulations,

however, show a good fit of the experimental data to a bimodal motional model (i.e. one ring is involved in one motional regime whilst the remaining three are involved in another). Also encompassed in the model are simulations of small amplitude librations, by virtue of the temperature variation in the both the quadrupole coupling constant, $\chi$, and asymmetry parameter, $\eta$. Whilst the disorder associated with the $TPP^+$ cation can be attributed to molecular mobility, the bromide anion is disordered over two sites with a Br···Br distance of 5.6 Å. This suggests that there is static disorder primarily in the bromide positions.

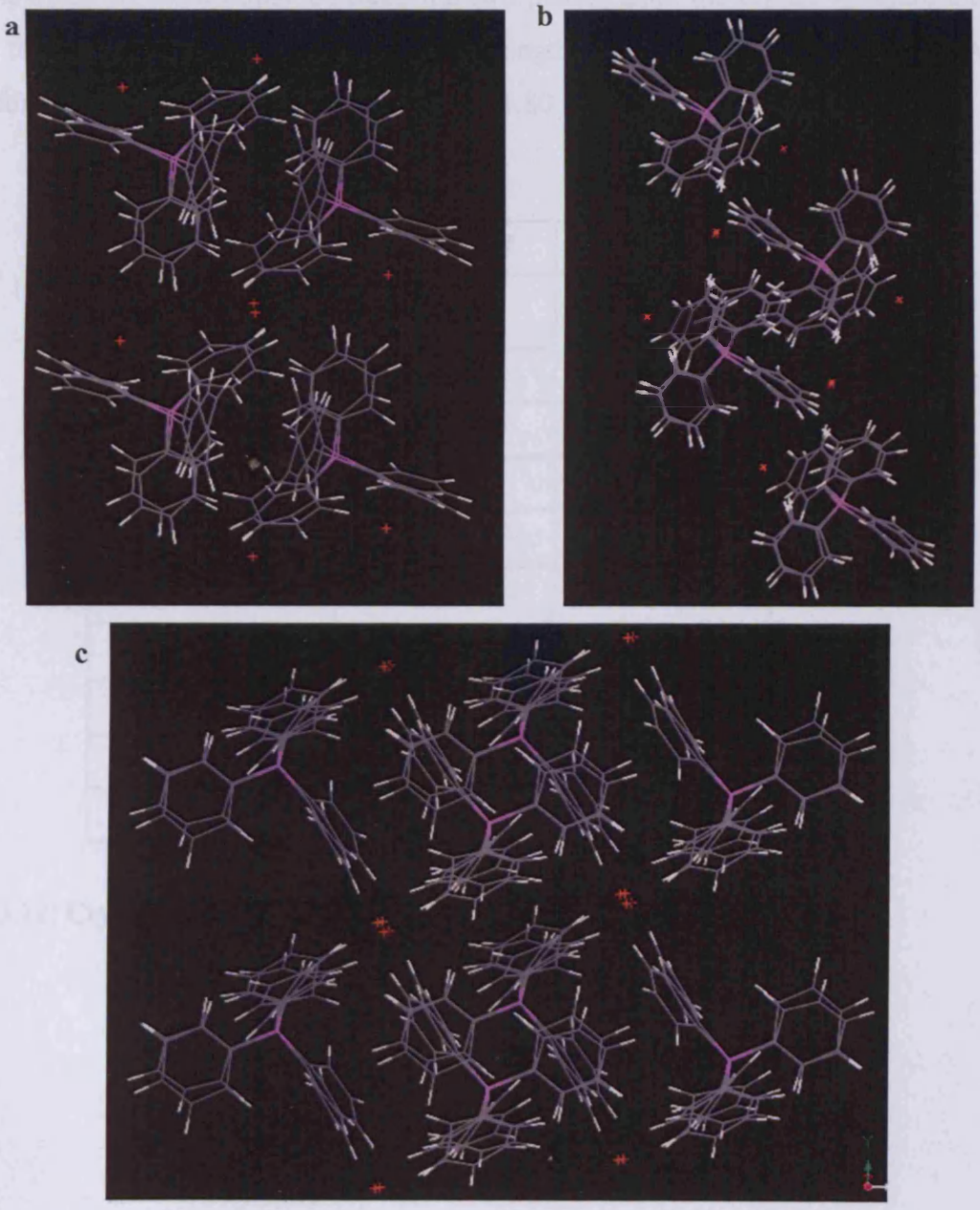



Figure 3.38: Refined structure of TPPB anhydrate, showing the perspective along (a) a-axis and (b) b-axis (c) c-axis

In addition, the structures of the anhydrate and the dihydrate were compared. Table 3.12 summarises the structural parameters of both pure phases, and Figure 3.39 shows the conformational properties of the $TPP^+$ cation in both the anhydrate and dihydrate structures. The overall tetrahedral arrangement remains essentially comparable (with angles in the range 107 to 112°) but the ring torsion angles in the two structures differ. FTIR spectroscopy, and specifically the C—P stretching bond were shown to be red shifted in the anhydrate phase relative to the dihydrate, indicative of longer C—P bonds in the anhydrate. A comparison of these bonds between the two phases from the crystal structure did indeed reveal that for the anhydrate the C—P bond lengths are in the range 1.81 to 1.84 Å, whilst those for the dihydrate are in the range 1.79 to 1.80 Å.

| Parameter | Dihydrate | Anhydrate |
|---|---|---|
| **Spacegroup** | Pnma | $P\bar{1}$ |
| **a / Å** | 16.255 | 9.496 |
| **b / Å** | 10.810 | 10.292 |
| **c / Å** | 12.667 | 11.440 |
| **α / °** | 90 | 75.138 |
| **β / °** | 90 | 72.360 |
| **γ / °** | 90 | 79.184 |
| **Volume / $Å^3$** | 2226 | 1023 |
| **Z** | 4 | 2 |
| **Density / g $cm^{-3}$** | 1.358 | 1.427 |

Table 3.12: Crystal structure data of TPPB anhydrate and dihydrate phases

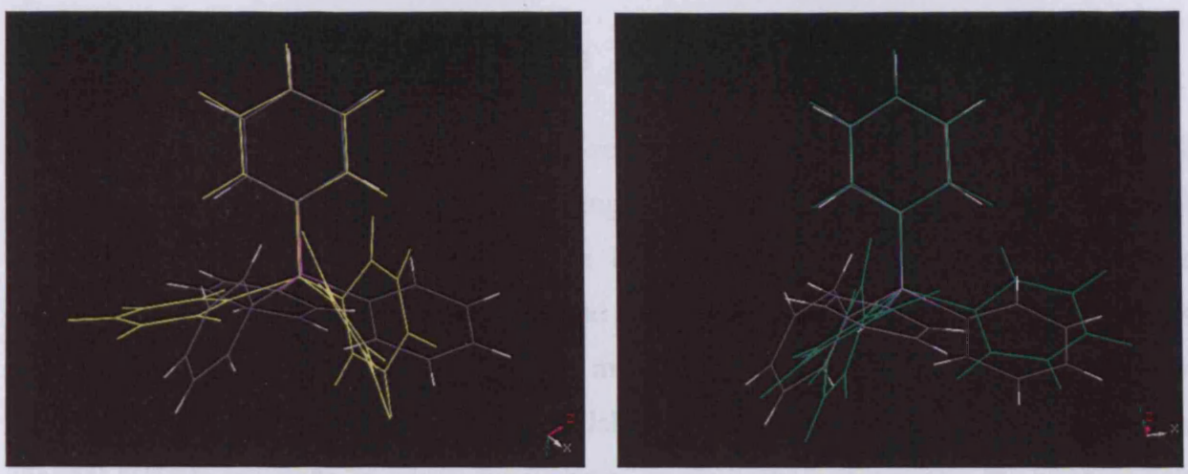


Figure 3.39: Comparison of conformational properties of the $TPP^+$ cation in the anhydrate (in yellow and green) and dihydrate structures

The specific volume occupied within the dihydrate structure is 557 $Å^3$, whilst that for the anhydrate is 511 $Å^3$. The anhydrate structure is more densely packed, reflected in metrics such as the phosphorus to bromine distance. In the anhydrate, the shortest P···Br distance is 5.58 Å, whilst for the dihydrate it is 6.23 Å. The $TPP^+$ cations in the anhydrate structure pack in alternating chains, which are also seen in the dihydrate structure, with the disordered $Br^-$ forming infinite anion chains in between the parallel cation stacks. The $Br^-$ anion in the dihydrate structure is on an inversion centre. In the anhydrate phase, one of the disordered anions is also on an inversion centre whilst the other is not. The phenyl rings in the anhydrate structure are not related by any symmetry operations unlike those in the dihydrate structure, which possesses 2.5 phenyl rings in the asymmetric unit. Whilst the two structures appear to pack in a similar systematic way, forming infinite parallel stacks of both the cations and anions, the local symmetry differences are apparent and give rise to the significant variation in structures and properties of the two pure phases.

## 3.5 Discussion and Conclusions

The structure, dynamics and inter-relationships of the pure anhydrate and dihydrate phases of TPPB have been investigated using a range of experimental techniques. Whilst other hydrate phases are known and have been reported, they often represent transient intermediate phases that are difficult to isolate and study in a controllable way. The dihydrate phase of TPPB has been shown to be the more stable phase under conditions of ambient temperature and relative humidity (standard laboratory conditions). The dihydrate exists as a channel hydrate structure, in which the water of crystallisation coordinates the bromide anion (forming planar infinite chains) and occupies the void space between parallel stacks of the $TPP^+$ cations. The anhydrate structure reported here was only attainable under non-ambient conditions.

The hydrate stoichiometry is directly related to the water activity of the surrounding environment, with a maximum water stoichiometric ratio corresponding to the dihydrate. Dehydration was shown to occur under conditions of very low water activity (low RH). Assessment of humidity and temperature induced dehydration showed that a single step is involved in this process, resulting in an anisotropic contraction of the structure. Whilst the cations remain as parallel stacks, the structure was shown to contract to form a more densely packed arrangement upon dehydration. The activation energy associated with the dehydration event is relatively small, and thus reflects the relative ease with which dehydration occurs.

The structural changes that were shown to accompany dehydration largely involve a variation in local symmetry as well as in structural dynamics, established from $^2H$ NMR spectroscopy. The exact mechanism of structural dynamics exhibited by the anhydrate phase has been shown to be difficult to establish in detail, although it is clear that the phenyl rings undergo 180° flips as in the dihydrate, but at a slower rate than the dihydrate. Thus, the dynamics are predominantly associated with the motion of the phenyl rings, which are significantly more restricted in the anhydrate than the dihydrate. The overall anisotropic contraction of the structure that accompanies dehydration is clearly a contributory factor to the restriction of the dynamics in the anhydrate phase.

With regard to thermodynamics, for a reaction or process to be favourable there needs to be some degree of enthalpy-entropy compensation (Dunitz 1995). That is, when a process involves a large expense in enthalpic terms, the process needs to be supplemented by

an entropic gain; conversely a loss in entropy needs to be accompanied by a balancing thermodynamic component. In the case of hydration, transferring water from a bulk state to a crystallisation site in a solid hydrate incurs an entropic penalty of approximately 40 J $mol^{-1}$ $K^{-1}$ per molecule (Dunitz 1994). The resultant dihydrate however, does not contain significant hydrogen bonds that would be of sufficient magnitude to offset the entropic cost of forming the hydrated phase. Hence, the fact that there is enhanced scope for structural dynamics might provide an opportunity to compensate and balance the thermodynamic cost incurred.

Furthermore, consideration of volume differences and the relative hygroscopicity of the anhydrate phase should be made. For instance, the specific volume of the dihydrate and the anhydrate phase are 556 and 511 $Å^3$ respectively. An estimate of the specific or Van der Waals volume of the water molecules is approximately 30 $Å^3$. Therefore, it could be estimated that the volume once two molecules of water of crystallisation (from the dihydrate) have been removed should actually be closer to 496 $Å^3$ rather than the observed 511 $Å^3$. This indicates that whilst the anhydrate structure is more densely packed, it does not necessarily represent the most efficient packing arrangement. Consequently, the additional void space, and the scope for enhanced dynamics, renders the anhydrate structure hygroscopic and the dihydrate the most stable structural arrangement under standard ambient conditions.

## 3.6 Experimental

### 3.6.1 Thermal analysis

Differential scanning calorimetry (DSC) experiments were conducted using a Perkin-Elmer DSC-7. Samples were analysed under a nitrogen atmosphere in 50 μl aluminium pans either left open, vented or sealed. Mass loss as a function of temperature was investigated using a Perkin-Elmer TGA-7 under a nitrogen atmosphere. Samples were analysed in open and semi-sealed pans to investigate the effects of localised vapour pressure. TGA was used in thermal mode to quantify mass loss as a function of dehydration with respect to temperature. Isothermal experiments were also performed in open pans to establish the kinetics of dehydration over the temperature range 25 - 70°C.

Hot-stage microscopy (HSM) experiments were performed on a Zeiss Axioskope microscope fitted with a Linkham hot stage unit and controlled by Linksys software. Single crystals of TPPB (grown from purified water using vapour sublimation) were immersed in silicone oil and heated from ambient temperature to 150°C at a rate of 10 °C $min^{-1}$.

### 3.6.2 XRPD Analysis

XRPD experiments were performed using a Bruker AXS D5000 diffractometer fitted with a SiLi scintillation detector and a TTK camera attachment to allow variable temperature and humidity experiments to be performed. A Cu Kα monochromatic radiation source was used (λ = 1.54056 Å), with target tube conditions of 40 kV and 25 mA. Variable slits were used and data were collected over the range 2 – 70° 2θ with a step size of 0.02°. High-resolution XRPD data were collected on a Philips Panalytical Xpert1 diffractometer in reflection mode (using Bragg-Brentano geometry). Typically, data were collected over the range 5 - 50° 2θ using a step size of 0.02 to 0.03° 2 θ and a counting time of 10 - 100 seconds per step.

### 3.6.3 Spectroscopy

#### 3.6.3.1 Vibrational spectroscopy

FTIR spectroscopy was performed on pure samples using a Perkin-Elmer Spectrum-GX spectrometer fitted with a single bounce SpecAc variable temperature ATR

attachment. Samples were analysed neat and spectra were acquired over the range 4000 – 600 $cm^{-1}$ using a spectral resolution of 4 $cm^{-1}$ with 64 accumulations. Raman spectroscopy was performed using a Perkin-Elmer NIR FT-Raman spectrometer fitted with a NdYAG (1064 nm) laser. Data were collected over the range 4000 – 100 $cm^{-1}$ using 4 $cm^{-1}$ resolution and 64 accumulations. The samples were spun in a direction perpendicular to that of the incident radiation to minimise preferred orientation and laser heating effects. Data were collected using isothermal and non-isothermal conditions using a Ventacon non-ambient sample unit.

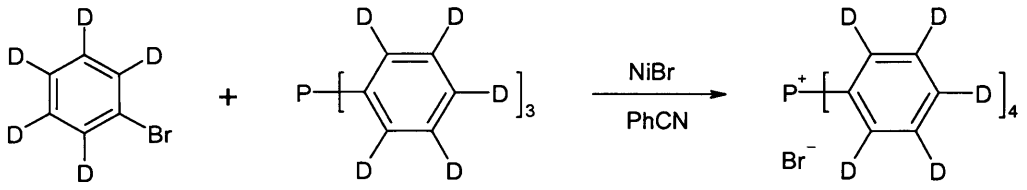
### 3.6.3.2 $^{13}C$ CP/MAS experiments

High-resolution solid-state $^{13}C$ NMR spectra were recorded at 75.5 MHz using a Chemagnetics Infinity 300 spectrometer and a Chemagnetics triple resonance probe. A standard CP and $^{1}H$ decoupling sequence was used (as described in Chapter 2) and samples were packed in 4 mm zirconia rotors and spun at typical speeds of 5 kHz.

### 3.6.3.3 Solid state $^{2}H$ NMR spectroscopy

#### 3.6.3.3.1 *Synthesis of perdeuterated tetraphenyl phosphonium bromide*

A previously reported synthetic scheme for preparing TPPB was modified to prepare the perdeuterated analogue (Affandi *et al.* 1987).




Perdeuterated bromobenzene and triphenyl phosphine were heated to reflux (under nitrogen), with nickel bromide and dried benzonitrile, for a total of 24 hours. The resultant dark green residue was allowed to cool to ambient temperature under a nitrogen atmosphere. The solid material was extracted by filtration under vacuum, and is designated sample A. The filtrate was distilled under vacuum to remove residual solvent; the residue from this is designated sample B. To each residue the following procedure was performed to extract the nickel. Water and dichloromethane were added and the mixture transferred to a separating funnel. The lower organic layers were extracted and the aqueous layer was washed with

further aliquots of dichloromethane. The organic washes were combined and the solvent removed under vacuum. To each of the solid samples, hot water was added and two repeated recrystallisations were performed (one from cyclohexane and one from hot water) to give two batches of material in moderate to poor yield. The two batches were analysed by reverse phase HPLC and the main peak identified by mass spectrometry detection. Sample A was shown to have a chemical purity of 95% and sample B of 88%. Sample A was progressed for all subsequent studies.

| | | |
|---|---|---|
| MS m/z: | 359 ($M^+$) | |
| Melting point $T_M$: | 301.2°C | |
| Analysis: | Theoretical | C (57.3 %), D (8.0 %), Br (15.9 %) |
| | Found | C (57.9 %), D (8.1 %), Br (16.0 %) |

*3.6.3.3.2 Spectroscopic assessment*

Solid-state $^2H$ NMR experiments were carried out at 46.080 MHz, using a Chemagnetics Infinity 300 spectrometer, and a Chemagnetics goniometer probe fitted with a 6 mm internal diameter solenoid coil. The samples were sealed in glass tubes as either the dihydrate with an internal hygrostat (of cotton wool soaked in water separated from the sample by glass wool, maintaining an environment of around 100% RH), or as the dried anhydrate phase. Spectra were recorded using the standard quadrupole echo pulse sequence described in Chapter 2. Typically around 500 to 1000 transients were recorded with recycle delays of around five times $T_1$. Simulated lineshapes were calculated using the program MXQET, also as described Chapter 2.

$^2H$ NMR spin lattice relaxation times were measured using a saturation-recovery pulse sequence with quadrupole echo detection (as described in Chapter 2): $[-\tau_d - (\pi/2)_x]_\nu - \tau_r - (\pi/2)_\phi - \tau - (\pi/2)_{(\phi \pm \pi/2)} - \tau' - \text{Acq}$. The number of $\pi/2$ pulses used for saturation was typically n = 30, with $\tau_d = 250$ $\mu$s, $\tau$ = 30 $\mu$s, and $\tau' < \tau$. The variable relaxation delay $\tau_r$ was increased geometrically according to $\tau_r = t_0[(10)^{1/10}]^{(N-1)}$, where N is the delay number and $t_0$ is the value of the first delay (i.e. $\tau_r = t_0$ when N = 1). Relaxation times were determined from the recovery curves by fits to multiple exponential recovery functions using the spectrometer software.

### 3.6.4 Solid state $^{31}$P NMR spectroscopy

Solid-state $^{31}$P NMR experiments were conducted on a Chemagnetics Infinity 300 spectrometer at 121.5 MHz. A Chemagnetics goniometer probe fitted with a 6 mm internal diameter solenoid coil was used for all measurements. The samples were sealed in glass tubes as either the dihydrate with an internal hygrostat (of cotton wool soaked in water separated from the sample by glass wool, maintaining an environment of around 100% RH), or as the dried anhydrate phase. Spectra were recorded from static samples over a temperature range of 273 to 393 K using the proton-decoupled sequence described in Chapter 2. $^{31}$P chemical shifts were referenced with respect to an external standard of 85% $H_3PO_4$. Simulated lineshapes were calculated using the using the SOLIDS program, also as described in Chapter 2.

### 3.6.5 Structure solution from X-ray powder diffraction data

Materials Studio Reflex module (version 1.2) from Accelerys was used for indexing and Pawley refinement. ITO (Visser 1969), TREOR (Werner *et al.* 1985), DICVOL (Boultif and Louer 1991) or X-cell (Neumann 2003) were used simultaneously to index the powder pattern of the anhydrate phase. The best-fit cell was obtained from the TREOR analysis. The Pawley fit procedure was used to refine the powder pattern with respect to the profile and background parameters (as described in Chapter 2). The measure of similarity used between the simulated and experimental profiles was $R_{wp}$. Structure solution was performed using the GA program EAGER, using the multi-population parallel genetic algorithm technique. Typically, the structure solution calculations involved 60 to 100 starting populations with 50 to 100 generations. The mating rate (i.e. the number of mating events per generation) was set at 15, and the mutation rate was typically set at 8 per generation.

Refinement was performed using the GSAS software as described in Chapter 2. Following an initial Le Bail refinement, the positions of the atoms were refined using the Rietveld protocol with standard geometrical restraints applied to bond lengths, bond angles and planar arrangements. For all non-hydrogen atoms, the isotropic displacement factors were constrained to a single value. Hydrogen atoms were incorporated into the structure in the final stages of refinement but were not refined.

### 3.6.6 Variable humidity analysis

A surface Measurement System DVS-1 was used to determine moisture uptake as a function of relative humidity. The samples, analysed relative to an empty pan, were exposed to two cycles of 0 – 90% RH and 90 – 0% RH regimes in 10% RH incremental steps. At each increment the mass was allowed to stabilise prior to the change of humidity. Isothermal Microcalorimetry (ITC) experiments were performed using a Thermometric TAM-1 fitted with Thermometric perfusion calorimeters. The samples were placed in an ampoule and equilibrated at 0% RH prior to analysis. Linear perfusion experiments were performed using incremental increase in the flow of nitrogen pre-saturated with moisture to give the desired RH value. All experiments were performed at 298 K.

Variable humidity XRPD analysis was performed using a Bruker AXS D5000 diffractometer (as described in Section 3.2.2) fitted with an in-line humidity generator.

## 3.7 References


- Affandi S., Green R. L., Hsieh B. T., Holt M. S., Nelson J. H., Alyea E. C., (1987), *Synth. React. Inorg. Met.-org. Chem.*, **17**, 307
- Alcock N. W., Pennington M., Willey G. R., (1985), *Acta Cryst.*, **C41**, 1549
- Ang T. T., Dunell B. A., (1976), *Can. J. Chem.*, **54**, 1985
- Bach M., (1999), PhD. Thesis, The University of Birmingham
- Bogaard M. P., Rae A. D., (1982), *Cryst. Struc. Comm.*, **11**, 175
- Boultif A., Louer D., (1991), *J. Appl. Crystallogr.*, **24**, 987
- Carter R. E., (1961), *J. Chem. Phys.*, **34**, 2010
- Desiraju G. R., (1991), *J. Chem. Soc., Chem. Comm*, 426
- Dunitz J. D., (1994), *Science*, **264**, 670
- Dunitz J. D., (1995), *Chem. & Biol.*, **2**, 709
- Eichele K., Wasylishen R. E., (1994), *J. Phys. Chem.*, **98**, 3108
- Gorbitz C. H., Hersleth H-P., *Acta Cryst*, (2000), **B56**, 526
- Kitiagorodski A. I., (1973), *Molecular Crystals and Molecules*, 29, Academic Press
- Müller U., *Acta Cryst.*, (1980), **B36**, 107
- Neumann M., (2003), *J. Appl. Crystallogr.*, **36**, 356
- Onkubo K. *et al.*, (1977), *Bull. Jpn. Pet. Inst.*, **19**, 73
- Schweizer E. E., Baldacchini C. J., Rheingold A. L., (1989), *Acta Cryst.*, **C45**, 1236
- Vincent B. R., Knop O., Linden A., Cameron T. S., Robertson K. N., (1988), *Can. J. Chem.*, **66**, 3060

- Visser J.W., (1969), *J.Appl. Crystallogr.*, **2**, 89
- Werner P.E., Eriksson L., Westdhal M., (1985), *J. Appl. Crystallogr.*, **18**, 367

# 4 Structural aspects of a homologous series of paraben esters

## 4.1 Introduction

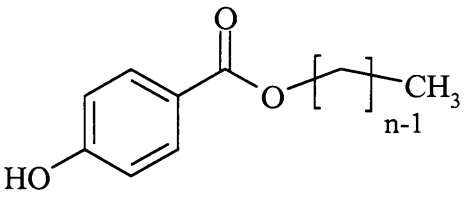
### 4.1.1 Use of a homologous series to explore structure property relationship

Homologous series of organic compounds have been used to probe structure-activity or structure-property relationships. Such studies involve investigations on the effects of varying specific functional group(s) on the activity or property under investigation. Variation of substituents in homologous structures gives rise to changes in, amongst other parameters, topology and electronic effects. Specifically, when the properties relate to the solid state, the existence of polymorphism, whilst being seen as a possible source for complication in deriving models, can provide a useful insight into discrete crystal structure property relationships. Considerable work has been performed in the investigation of physical properties of homologous series of long chain molecules (especially n-alkanes) in order to obtain predictive models. Structure-property relationships fall into two main categories, crystal structure relationships and molecular structure relationships. In the former case, the properties intrinsically depend upon both the nature of the molecules and the way they interact with each other in the solid state; investigations on polymorphic systems can be particularly informative in this regard.

It is well documented that n-alkanes exhibit an odd-even alternation in many of their physical properties such as melting behaviour (Small 1986), which is attributed to differences in the crystal structures of the odd and even homologues; the chains pack in lamellae arrangements in orthorhombic and triclinic cells respectively (Dorset *et al.* 1991). Furthermore, as the chain length increases, the chain ends exert a less significant effect on the overall packing of the structure (McGann *et al.* 1999). The structures of the shorter chain homologues generally have a decreased packing density attributed to the energetic effects associated with the poorer packing of the chain ends. Upon increasing the chain length, the chain ends represent a much smaller fraction of the molecule, with the main chain largely dominating the packing. Other structures containing long chains such as p-amino benzoic acid esters are also reported to have a chain length dependence of various physical properties (Yalkowsky *et al.* 1972). In this case, rather than an odd-even alternation, there is an

inflection in the linear relationship of physical properties, such as melting, with chain length. This inflection has been attributed to the variation in structure as a function of chain length, governed by the effects of the aromatic polar head group versus the aliphatic chain. The polar head groups were referred to as the loading group and dominate the crystal structure until the polymethylene unit reaches a critical size, after which it can exert greater influence on the nature of the structure. An odd-even alternation has also been reported for many other substituted n-alkanes such as n-alkan-1-ols, α,ω-diamine alkanes, n-alkane-α,ω-dithiols, n-alkane-α,ω-carboxylic acids, and dialkyl oxybenzoquinones (Dall'Acqua *et al.* 2002). Additionally, co-crystals of pyrazine with n-carboxylic acids (from methyl to nonyl homologues) also show an odd-even alternation upon increasing the chain length, which correlate with the crystal densities (Bond 2003).

### 4.1.2 Paraben esters




Where n = 1,2,3, ... ,15

Paraben esters (4-hydroxy benzoic acid esters) are commercially used as excipients in pharmaceutical, food and cosmetic products, specifically as preservatives. This group of compounds represents a simple homologous series of aromatic esters and has been utilised in investigations into solubility effects (Paruta 1969), and their interactions in binary/ternary mixtures both in the solution (with non-ionic surfactants) (Fukahori *et al.* 1996) and solid states (with trimethoprim) (Pedersen *et al.* 1994). The crystal structure of the methyl (Lin 1983), and ethyl (Lin 1986) esters have been reported and more recently, the structure of the propyl and butyl homologues have also been elucidated (Giordano *et al.* 1999). All four homologues of these have been shown to crystallise in the monoclinic system. In the present study, the series of paraben esters has been extended to include the pentyl (n = 5) to pentadecyl (n = 15) esters.

## 4.2 Materials characterisation and assessment of odd-even alternations

Analysis of the samples by DSC showed a single melting endotherm for each material. The mean melting temperature, from three consecutive DSC analyses, is plotted against the number of carbons in the ester unit as shown in Figure 4.1.

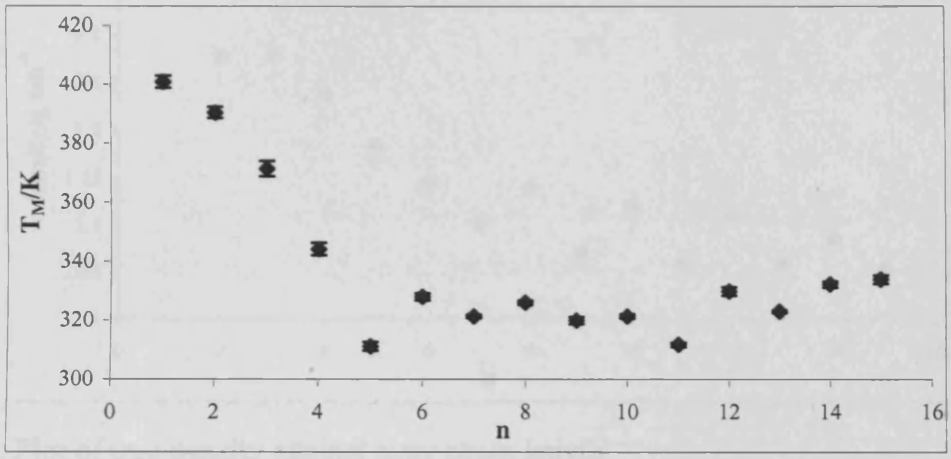



Figure 4.1: Plot of mean peak melting temperature against ester chain length. Error bars denote the standard deviation of replicate data

An almost linear relationship exists between ester chain length and melting temperature until n = 5, after which an apparent odd/even alternation is observed. The crystal structures of the n = 1 to n = 4 compounds are known and all exist in the monoclinic system (n = 2 and n = 3 are isomorphous), with ester to phenol hydroxy intermolecular hydrogen bonding (hereafter referred to as head to tail hydrogen bonding). It is anticipated that, in accordance with this linear relationship, the pentyl ester might also exist in the monoclinic system. Furthermore, in accordance with previously reported observations on homologous series, it is anticipated that the higher homologues may have different structures from the lower homologues, contributing to the non-linearity. Interestingly, all the ester homologues exhibit a single melting event with no phase transitions occurring prior to melting. A large number of long-chain flexible molecules such as n-alkanes undergo solid-to-solid transitions, which are accompanied by changes in the dynamics of the molecules in the crystals or the formation of mesophases, prior to the melt (Boese *et al.* 1999). All homologues were recrystallised from a range of different solvents of varying dielectric constants in order to induce polymorphism, but in all cases no polymorphs were found.

Figure 4.2 illustrates the chain length dependence of the packing or true density (ρ), obtained from helium pycnometry measurements of the materials at ambient temperature. The densities of the ethyl and propyl homologues are almost comparable due to the isomorphous structures.

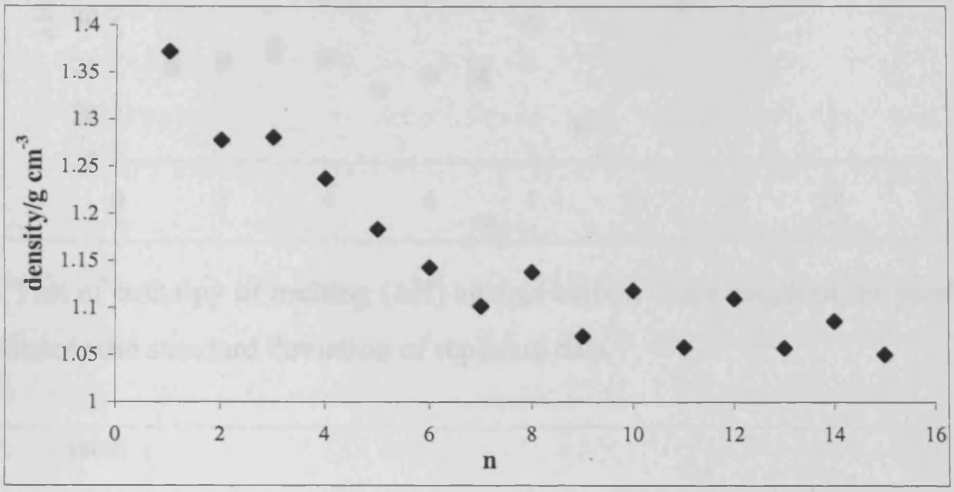



Figure 4.2: Plot of true density against ester chain length

The plot of the chain length dependence on density shows an almost perfect odd-even alternation for $n \geq 7$, and virtually mirrors the characteristics of the melting behaviour as shown in Figure 4.1. The initial decrease in the melting temperature is attributed to a rapid decrease in packing density.

Large molecules tend to have higher melting and boiling points, when compared with smaller molecules of the same type (Dunitz and Gavezotti 1999), and is attributed to the fact that larger molecules have a greater cohesive energy due to the increased attractive interactions with neighbouring molecules. For homologous series, the intermolecular forces are additive and as such, there is an increase in overall intermolecular attraction with increasing size, coupled also to a general increase in the packing coefficient (Dunitz and Gavezotti 1999). Furthermore, the fact that there is an inflection in the melting behaviour at $n = 5$ suggests that the aliphatic chain appears to exert more of an influence on the structure than the polar hydroxy benzoate head group when the chain length exceeds this point.

As expected, the odd-even alternation is also reflected in the variation of the enthalpy of melting and entropy of melting as a function of chain length, as illustrated in Figure 4.3 and 4.4 respectively.

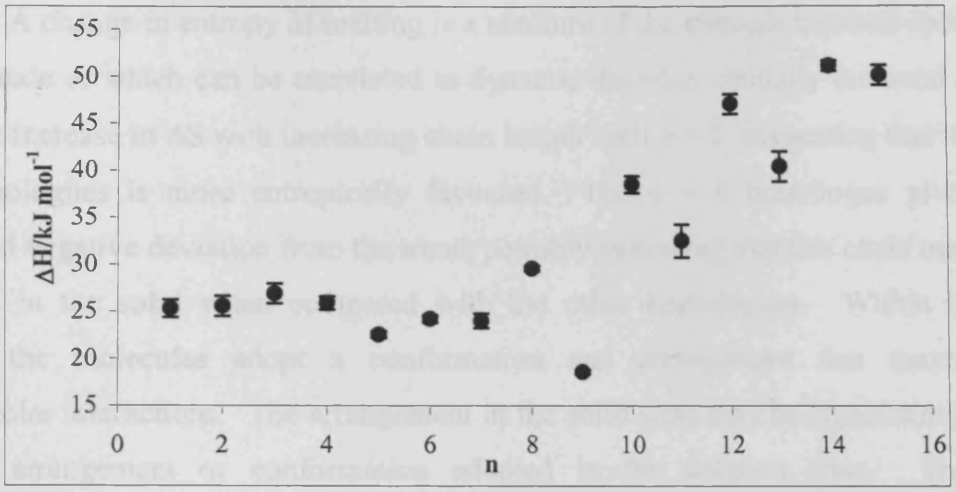



Figure 4.3: Plot of enthalpy of melting (ΔH) against carbon chain length of the paraben esters. Error bars denote the standard deviation of replicate data

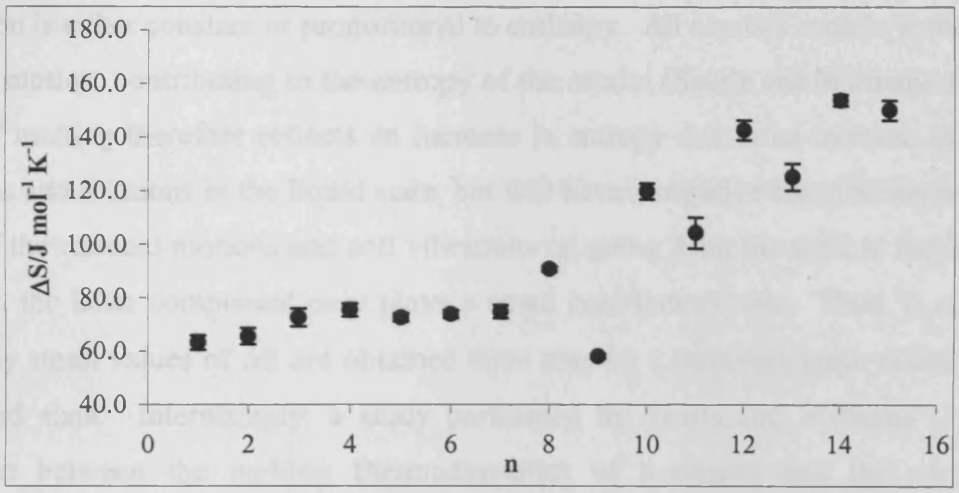



Figure 4. 4: Plot of entropy of melting (ΔS) against carbon chain length of the paraben esters. Error bars denote the standard deviation of replicate data

The enthalpy of fusion as a function of chain length initially shows a very slight linear increase, which is broken at n = 5 with an odd-even alternation. Interestingly, the esters containing an even number of carbons in the chain generally have a higher ΔH value than the adjacent odd esters that are one carbon unit shorter or longer. This behaviour is also reflected in the density pattern, where chains with an even number exhibit a more densely packed structure than chains with an odd number of carbons. Furthermore, the corresponding melting temperature and entropy of melting also show that the ester chains with an even number have a higher temperature and change in entropy on melting in comparison to the odd numbered homologues. This behaviour suggests that there are greater intermolecular attractive forces for the esters with even numbered chains.

A change in entropy of melting is a measure of the entropic increase upon melting, the magnitude of which can be correlated to dynamic disorder. Initially the trend is a slight monotonic increase in $\Delta S$ with increasing chain length until n = 7, suggesting that melting for these homologues is more entropically favoured. The n = 9 homologue gives a more pronounced negative deviation from the trend, possibly indicating that this chain may be more disordered in the solid when compared with the other homologues. Within the crystal structure, the molecules adopt a conformation and arrangement that maximises the intermolecular interactions. The arrangement in the solid state may be significantly different from the arrangement or conformation adopted in the solution state. The melting phenomenon generally has two sets of energetic contributions: differences in intermolecular forces between the solid and liquid state, and the variation in conformational energy between the solid and liquid states (Charton and Charton 2002). This assumes that any entropy contribution is either constant or proportional to enthalpy. All crystals contain some degree of molecular motion, contributing to the entropy of the crystal (Searle and Williams 1992). The entropy of melting therefore reflects an increase in entropy due to an increase in molecular translations and rotations in the liquid state, but will have a negative contribution arising from the loss of the residual motions and soft vibrations on going from the solid to the liquid state. Obviously, the latter component only plays a small contributory role. Thus, in cases where comparably small values of $\Delta S$ are obtained there may be a relatively large extent of motion in the solid state. Interestingly, a study performed by Searle and Williams (1992) on a comparison between the melting thermodynamics of n-alkanes and the corresponding cycloalkanes indicated that the cycloalkanes have higher melting temperatures and less positive $\Delta H$ and $\Delta S$ values of melting than their corresponding linear n-alkane counterparts. The higher melting temperature is attributed to the fact that there is no great entropic or enthalpic advantage to melting. Whilst the lower value of $\Delta S$ reflects the fact that there are no rotors (freely rotatable bonds) released upon melting of the cycloalkane.

Molecules in crystal structures are held together by the balance of attractive and repulsive forces (Dunitz and Gavezzotti 1999). The repulsive forces provide resistance to compression of the structure such that equilibrium positions (despite the occurrence of dynamic disorder) remain. A simplified representation of the attractive and repulsive forces may be ascribed to the enthalpy and entropy term. The balance between the enthalpic and entropic terms of any system forms the basis of the fundamental thermodynamics of a system (Dunitz 1995). A characteristic of weak interactions found in crystal structures involves

compensation of the enthalpy and entropy terms. Generally, as the enthalpy of melting becomes more positive (indicating a stronger interaction in the solid relative to the liquid phase) the entropic term also becomes more positive (due to loss of conformational flexibility in the solid phase) (Searle and Williams 1992, Dunitz 1995). Figure 4.5 shows an enthalpy-entropy compensation curve for the melting thermodynamics of all paraben esters.

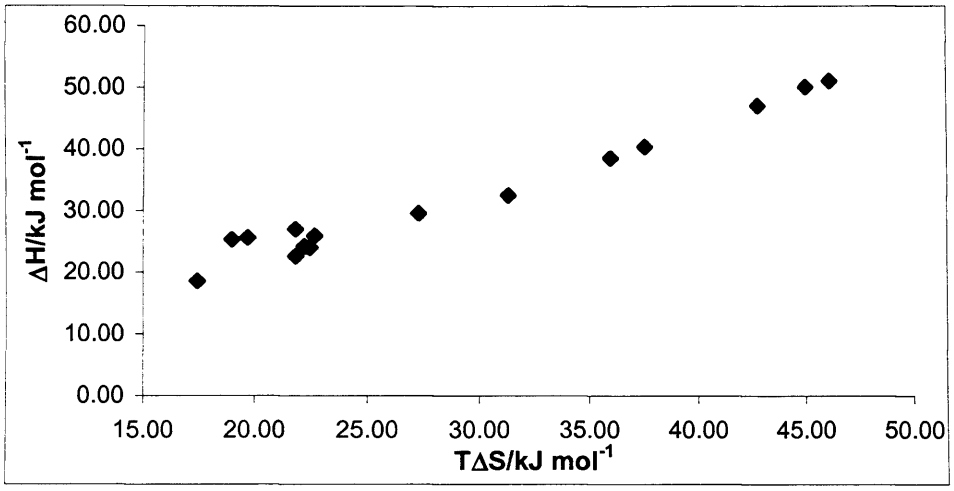



Figure 4.5: Enthalpy-entropy compensation curve for the paraben esters (at T = 300K)

The overall linear relationship seen in Figure 4.5 concords well with previous studies, in which a large population of both rigid and flexible structures showed an essentially linear relationship between the enthalpy and entropy of melting (Gilbert 1999). From these observations it can be inferred that for long chain flexible molecules (such as the paraben esters) the penalty for increased cohesive properties and perhaps conformational order is a loss in entropy. Conversely, a disruption in the conformational order affects enthalpy.

## 4.3 Spectroscopic assessment of the paraben esters

Assessment of the paraben ester homologues by FTIR spectroscopy showed some correlation between the stretching frequencies of the O—H bond and the C=O bond with the ester chain length, as illustrated in Figure 4.6.

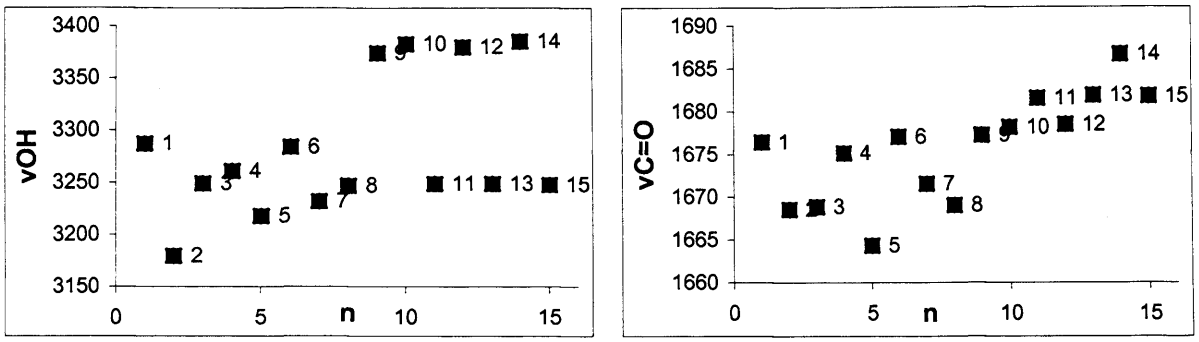



Figure 4.6: Chain length dependence of the O—H stretching frequency (νOH) and the carbonyl stretching frequency (νC=O). The points on the curves are labelled according to the number of carbons in the ester chain

In general, there is an overall trend of an increase in the carbonyl stretching frequency with increasing chain length. This trend is less clear with the O—H stretching frequency. In simplistic terms, an increase in the stretching frequency implies the formation of a stronger, shorter bond. In the case of the paraben esters, this implies that an increase in chain length is accompanied by a reduction in the strength of hydrogen bonding between adjacent 4-hydroxy benzoate groups. As discussed earlier, the lower four homologues (n = 1 to 4) have head-to-tail hydrogen bonds between neighbouring molecules, and it is anticipated that the higher homologues will also exhibit this type of interaction. The thermodynamic parameters of melting showed that there is an initial decrease in all parameters with increasing chain length. However, on increasing the chain length beyond a critical length (n = 8), a steady increase in the temperature, $\Delta H$ and $\Delta S$ of melting was observed. This observation is consistent with the 4-hydroxy benzoate moiety dominating the intermolecular interactions for the smaller chain esters, whilst the structures of the longer chain esters rely less on the polar group interactions to dominate the attractive forces of the structure.

FTIR spectroscopy is also an established technique for determining the chain conformation in n-alkanes and other molecules containing long chain aliphatic groups. A variation in localised wagging and rocking frequencies is associated with the methylene groups in either all trans or one or more gauche conformers. For example, FTIR spectroscopy was used to establish that phosphatidyl choline structures contain gauche-trans-gauche kinks as the dominant feature in the chain disorder (Casal and McElhaney 1990). Furthermore, the number of these kinks increases as chain length increases. As the frequencies of certain methylene vibrations are sensitive to the chain conformation, information on gauche-trans-gauche kinks, end chain gauche kinks, and gauche-gauche conformations can be obtained.

The use of FTIR spectroscopy alone, however, may not allow detection of single gauche conformers in the interior of the chain, and also does not provide information on the location of the conformational change along the chain.

A critical assessment was performed of the 1400 to 1300 $cm^{-1}$ region (frequency region of the conformationally sensitive methylene vibrations) of the FTIR spectra for the paraben esters, as shown in Figure 4.7. Table 4.1 summarises a list of peaks observed in this region. Typically, for a gauche-trans-gauche kink, peaks are observed at 1367 and 1306 $cm^{-1}$, for end gauche conformations peaks are seen at around 1341 $cm^{-1}$, and for gauche-gauche conformational kinks a peak is seen at around 1355 $cm^{-1}$ (Casal and McElhaney 1990).

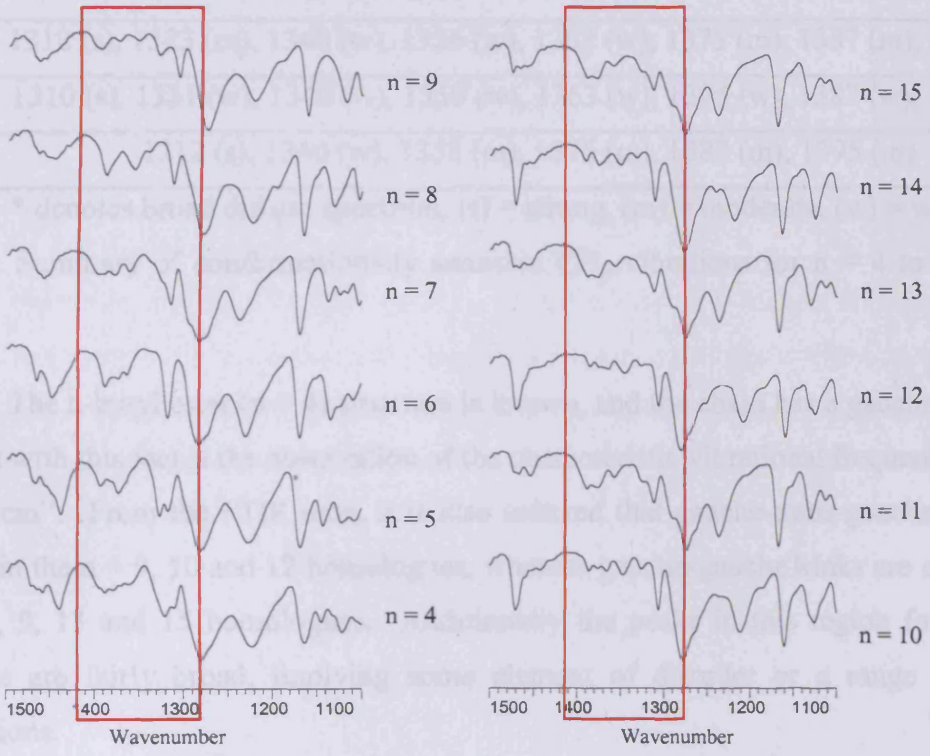



Figure 4.7: FTIR spectra of the n = 4 to 15 paraben esters showing the 1400 – 1300 $cm^{-1}$ frequency region (highlighted by the red box)

| n | Peaks in the 1300 –1400 $cm^{-1}$ range |
|---|---|
| 4 | 1308 (s), 1330 (w), 1354 (m), 1369 (m), 1388 (m) |
| 5* | 1317 (s), 1323 (s), 1374 (m), 1392 (m) |
| 6 | 1318 (s), 1340 (w), 1363 (m), 1375 (m), 1388 (s), 1395 (w) |
| 7 | 1313 (s), 1323 (m), 1360 (m), 1376 (m), 1388 (m), 1395 (m) |
| 8 | 1312 (s), 1341 (m), 1356 (m), 1388 (s), 1395 (w) |
| 9 | 1307 (s), 1330 (m), 1354 (w), 1375 (w), 1395 (m), 1399 (m) |
| 10 | 1306 (s), 1326 (m), 1348 (w), 1364 (m), 1375 (m), 1387(m) 1395 (m) |
| 11 | 1313 (s), 1340 (m), 1362 (m), 1375 (m), 1387 (s), 1395 (w) |
| 12 | 1307 (s), 1327 (w), 1346 (w), 1365 (m), 1375 (w), 1387 (m), 1395 (w) |
| 13 | 1312 (s), 1323 (m), 1340 (w), 1356 (w), 1363 (w), 1375 (m), 1387 (m), 1395 (m) |
| 14 | 1310 (s), 1331 (w), 1340 (w), 1350 (w), 1363 (w), 1375 (w), 1387 (w), 1395 (w) |
| 15 | 1312 (s), 1340 (w), 1358 (m), 1375 (m), 1387 (m), 1395 (m) |

* denotes broad diffuse spectrum, (s) = strong, (m) = moderate, (w) = weak

Table 4.1: Summary of conformationally sensitive $CH_2$ vibrations for n = 4 to 15 paraben esters

The n-butyl ester (n = 4) structure is known, and the chain has a gauche-trans kink. Consistent with this fact is the observation of the characteristic vibrational frequencies at 1308 and 1369 $cm^{-1}$. From the FTIR data, it is also inferred that gauche-trans-gauche kinks may also exist in the n = 9, 10 and 12 homologues, whereas gauche-gauche kinks are expected for the n = 8, 9, 13 and 15 homologues. Additionally the peaks in this region for the n = 5 homologue are fairly broad, implying some element of disorder or a range of possible conformations.

$^{13}C$ CP/MAS spectra were collected for the C5 to C15 paraben esters. The spectra, separated into odd and even sets, are illustrated in Figure 4.8. Spectra were also acquired with dipolar dephasing in order to afford non-quaternary suppression. Table 4.2 summarises the observations from both the full spectra and the NQS spectra.

| Homologue | Peaks per C environment in the molecule | NQC in NQS data |
|---|---|---|
| 5 | 2 | Some $CH_2$ |
| 6 | 1 | - |
| 7 | 2 | Some $CH_2$ |
| 8 | 1 | - |
| 9 | 1 | Most $CH_2$ |
| 10 | 1[a] | - |
| 11 | 2 | Some $CH_2$ |
| 12 | 1[b] | - |
| 13 | 1[b] | - |
| 14 | 1[b] | - |
| 15 | 1[b] | - |

[a] methylene peaks split

[b] peak overlap in aliphatic region

Table 4.2: Summary of $^{13}C$ CP/MAS spectra collected with and without dipolar dephasing. The peaks per carbon environment in the molecule represent all the carbons in the structure unless otherwise stated. NQC denotes non-quaternary carbons observed in addition to the expected quaternary peaks

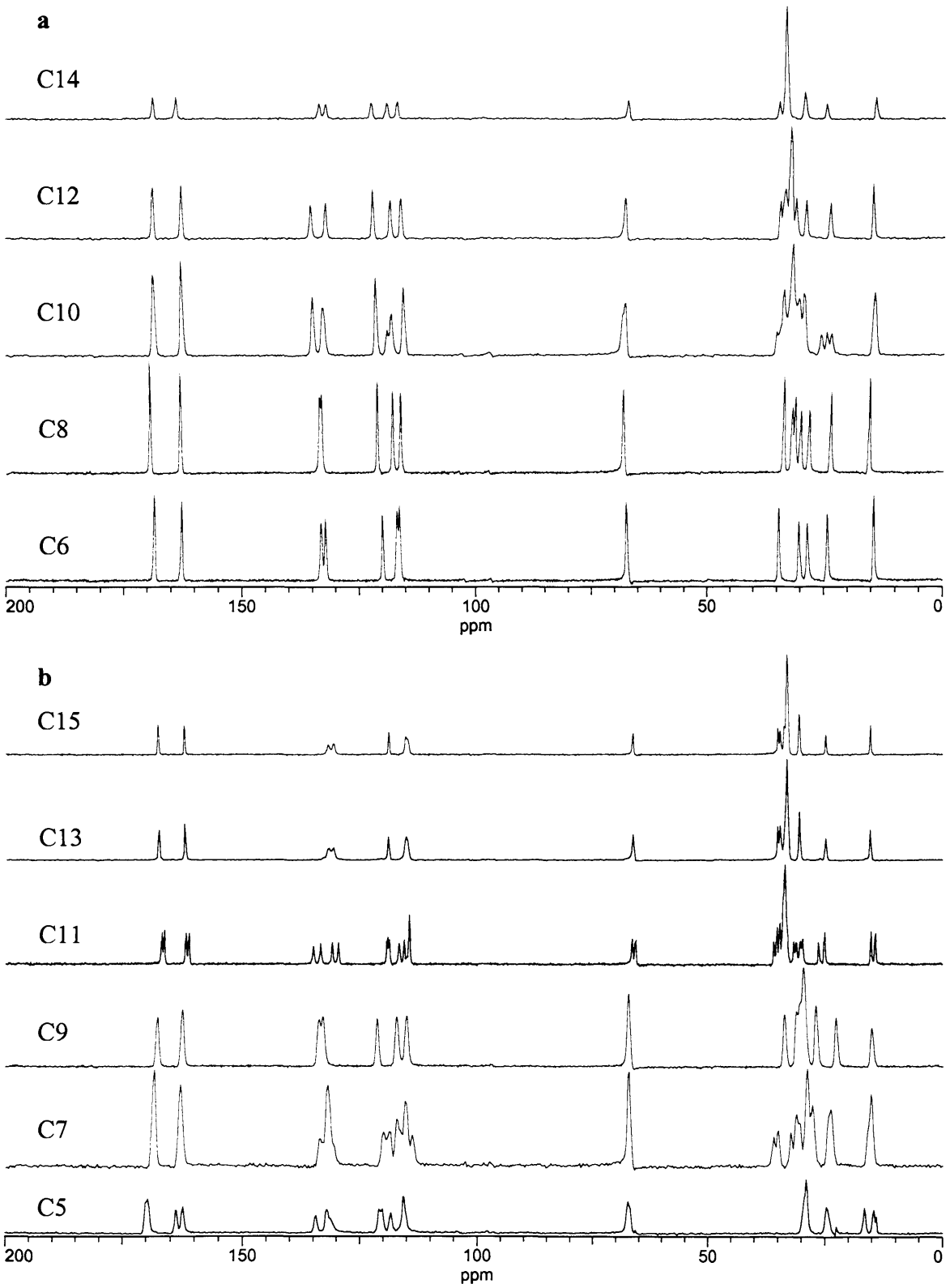



Figure 4.8: $^{13}C$ CP/MAS spectra of the paraben esters with (a) even chain length and (b) odd chain length

The spectra show that for some homologues there are doublet peaks indicating the possibility of either at least two crystallographically independent environments (giving a 1:1

peak intensity ratio) or possible disorder, although there is no odd-even trend to this observation. Furthermore, the presence of methylene resonances in the NQS data in odd numbered chain lengths up to n = 11 homologue would suggest that these esters exhibit some degree of dynamic disorder in the aliphatic chain. The NMR data will be used as an aid to spacegroup assignment for structure solution.

## 4.4 Structural assessment of the paraben esters

The structures of the four lower homologues have been reported previously (Giordano *et al.* 1999) and the structural parameters are summarised in Table 4.3.

| | **Methyl** | **Ethyl** | **Propyl** | **Butyl** |
|---|---|---|---|---|
| **Spacegroup** | Cc | $P2_1/c$ | $P2_1/c$ | C2/c |
| **a/Å** | 13.568(5) | 11.765(4) | 12.044(2) | 20.087(7) |
| **b/ Å** | 16.959(7) | 13.182(1) | 13.829(3) | 8.218(2) |
| **c/ Å** | 12.458(6) | 11.579(4) | 11.785(3) | 4.714(5) |
| **β/°** | 130.13(3) | 107.76(3) | 108.63(1) | 121.39(1) |
| **Z** | 12 | 8 | 8 | 8 |
| **V/ Å$^3$** | 2192.9 | 1710.2 | 1860.0 | 2073.4 |

Table 4.3: Summary of structural parameters for the n = 1 to 4 homologues

The methyl (n = 1) homologue contains three crystallographically independent molecules per unit cell, the butyl (n = 4) homologue contains one, and the ethyl (n = 2) and propyl (n = 3) homologues contain two. In all cases, a head-to-tail hydrogen-bonding motif is adopted, which leads to the formation of infinite chains. The n = 4 homologue was shown to have a gauche-trans kink in the chain (as discussed previously). Recrystallisation of the four lower homologues from single component solvent systems of varying polarity appeared not to afford polymorphic modifications, suggesting that the reported packing motif was the preferred and stable arrangement.

Attempts to grow single crystals of suitable quality and size, for single crystal XRD studies, of all the higher homologues were unsuccessful, often culminating in polycrystalline samples, oils or foams following solution-mediated growth. As such, high-resolution XRPD data were collected on all samples at ambient temperature, and attempts

were made to solve the structures from the powder data. Table 4.4 summarises the results from indexing and Pawley refinement of the powder data for the paraben esters (n = 5 to 15).

| n | Crystal system | Algorithm/ FOM | a/Å | b/Å | c/Å | α/° | β/° | γ/° | V/Å$^3$ |
|---|---|---|---|---|---|---|---|---|---|
| 5 | Monoclinic | Treor/18.0 | 13.65 | 22.59 | 7.65 | 90 | 93.9 | 90 | 2358 |
| 6 | Monoclinic | Dicvol/45.3 | 13.25 | 11.98 | 10.16 | 90 | 127.3 | 90 | 1289 |
| 7 | Triclinic | Dicvol/10.3 | 16.24 | 12.05 | 10.05 | 72.7 | 126.9 | 105.3 | 1493 |
| 8 | Orthorhombic | Dicvol/87.1 | 13.64 | 12.32 | 8.69 | 90 | 90 | 90 | 1460 |
| 9 | Monoclinic | Dicvol/14.8 | 32.59 | 8.23 | 6.05 | 90 | 92.6 | 90 | 1615 |
| 10 | Monoclinic | Dicvol/4.5 | 26.91 | 12.11 | 16.87 | 90 | 116.7 | 90 | 4917 |
| 11 | Triclinic | Treor/12.0 | 20.40 | 12.69 | 9.03 | 52.7 | 122.0 | 117.7 | 1541 |
| 12 | Monoclinic | Dicvol/20.0 | 28.65 | 8.47 | 16.96 | 90 | 117.8 | 90 | 3643 |
| 13 | Triclinic | Dicvol/12.6 | 21.18 | 12.00 | 5.68 | 112.6 | 96.2 | 116.0 | 1130 |
| 14 | Monoclinic | Xcell/*1 | 13.33 | 8.67 | 15.08 | 90 | 122.6 | 90 | 1467 |
| 15 | Triclinic | Xcell/*2 | 4.99 | 11.12 | 39.77 | 90 | 97.4 | 90 | 2188 |

*1 relative FOM = 0.27, FOM = 166

*2 relative FOM = 5.89, FOM = 3505

Table 4.4: Summary of indexed and Pawley refined cell parameters of the C5 to C15 homologues.

The results indicate that there does not appear to be a direct correlation between the odd even alternation and the crystal system, as observed for other long chain homologues discussed earlier. Furthermore, no obvious correlation exists between the chain length and the unit cell parameters, as illustrated graphically in Figure 4.9.

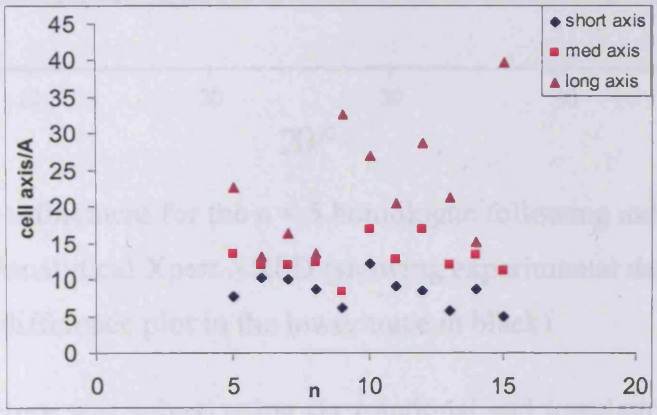



Figure 4.9: Plot of unit cell axes as a function of chain length

The next section relates to the structure solution of key paraben homologues.

### 4.4.1 Structure of the pentyl (n = 5) homologue

The XRPD data, acquired at 298 K, was indexed using the program TREOR (Werner *et al.* 1985) to give a monoclinic unit cell with parameters a = 13.65 Å, b = 22.59 Å, c = 7.65 Å, β = 93.9° and cell volume V = 2358 Å$^3$. Several datasets (from different sample preparations) were collected in an attempt to assess (and reduce) the influence of preferred orientation effects. The space group was assigned during the Pawley refinement procedure from consideration of systematic absences, giving $P2_1/c$ as the best spacegroup. Pawley refinement gave $R_{wp}$ = 4.10% (the profile fit is shown in Figure 4.10).

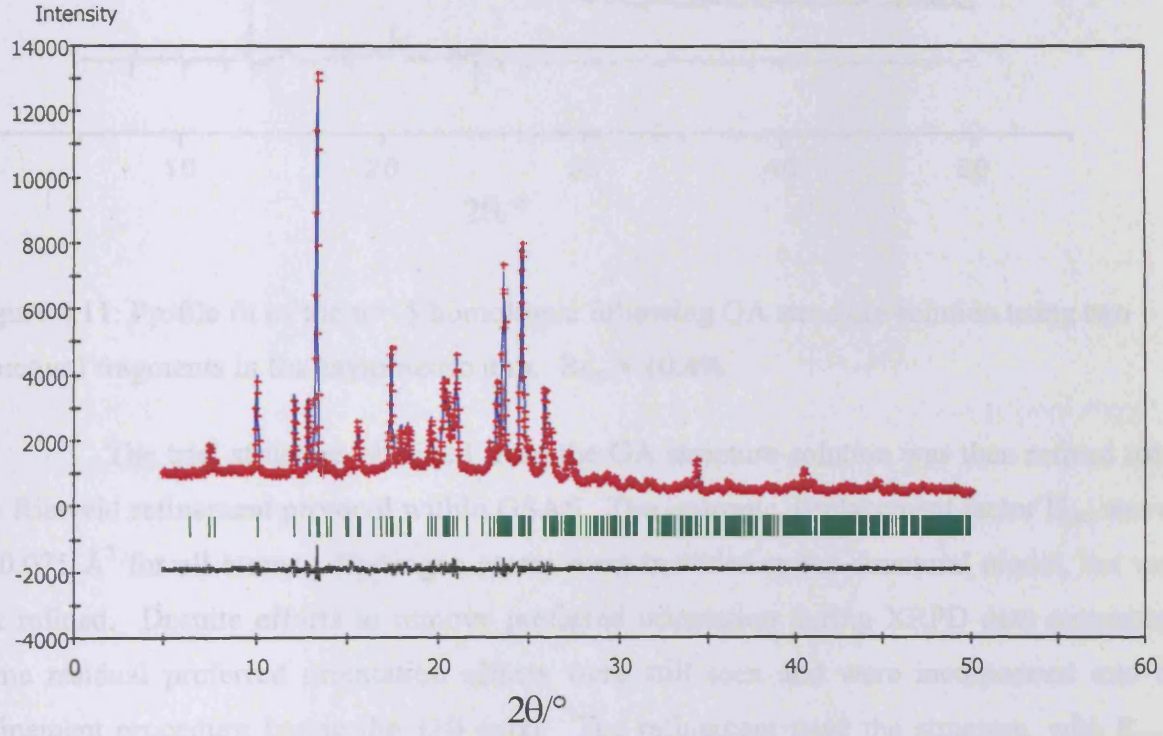



Figure 4.10: Pawley refinement for the n = 5 homologue following indexing by TREOR of data collected on a Panalytical Xpert XRPD (showing experimental data in red, calculated data in blue and the difference plot in the lower trace in black)

The structure was solved using six rotational and translational degrees of freedom for the molecule as a whole, and also allowing five internal rotations within the ester chain. Repeat attempts at structure solution using GA program EAGER, using Z = 8 (Z' = 2, i.e. two molecules in the asymmetric unit, in accordance with $^{13}$C CP/MAS studies and correct density values) gave reproducible results of $R_{wp}$ = 10.4%. A typical profile plot for the structure solution is shown in Figure 4.11.

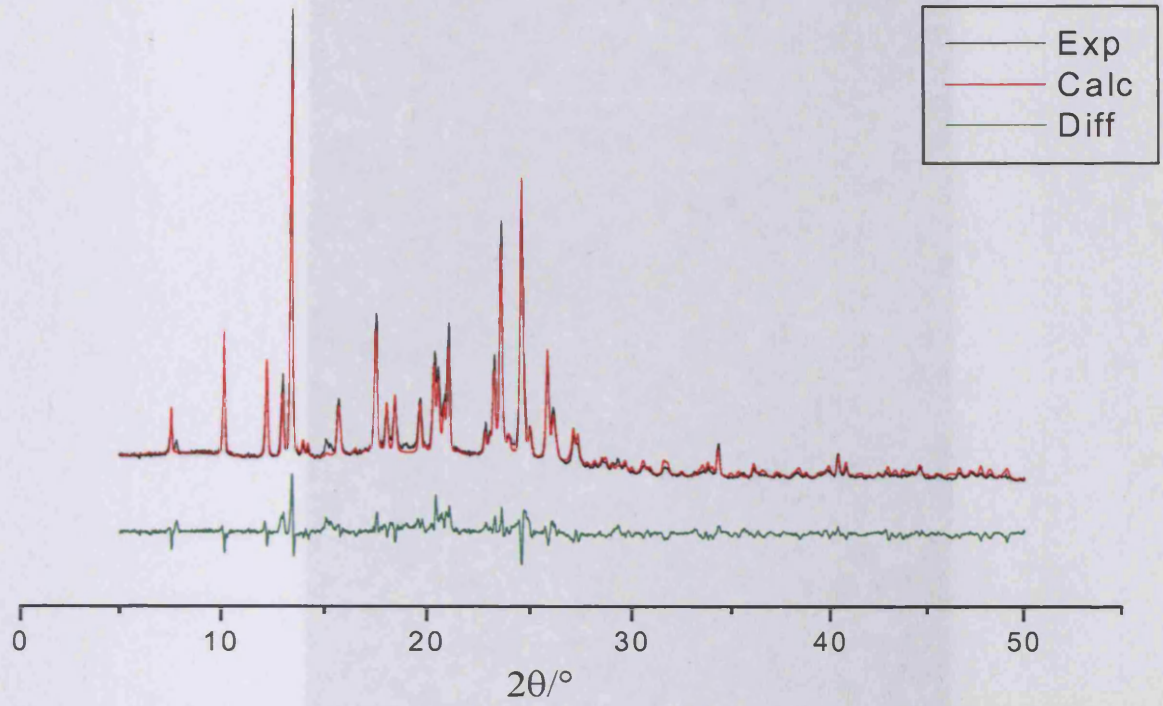



Figure 4.11: Profile fit of the n = 5 homologue following GA structure solution using two structural fragments in the asymmetric unit. $R_{wp}$ = 10.4%

The trial structure obtained from the GA structure solution was then refined using the Rietveld refinement protocol within GSAS. The isotropic displacement factor $U_{iso}$ was set at 0.075 $Å^2$ for all atoms. Hydrogen atoms were included in the structural model, but were not refined. Despite efforts to remove preferred orientation during XRPD data acquisition, some residual preferred orientation effects were still seen and were incorporated into the refinement procedure (using the 110 axis). The refinement gave the structure, with $R_{wp}$ = 6.55%, illustrated in Figure 4.12. Figure 4.13 shows the head to tail (ester to hydroxy) hydrogen bonding between the two molecules in the asymmetric unit creating a zigzag arrangement to the packing. The profile following the refinement is shown in Figure 4.14.

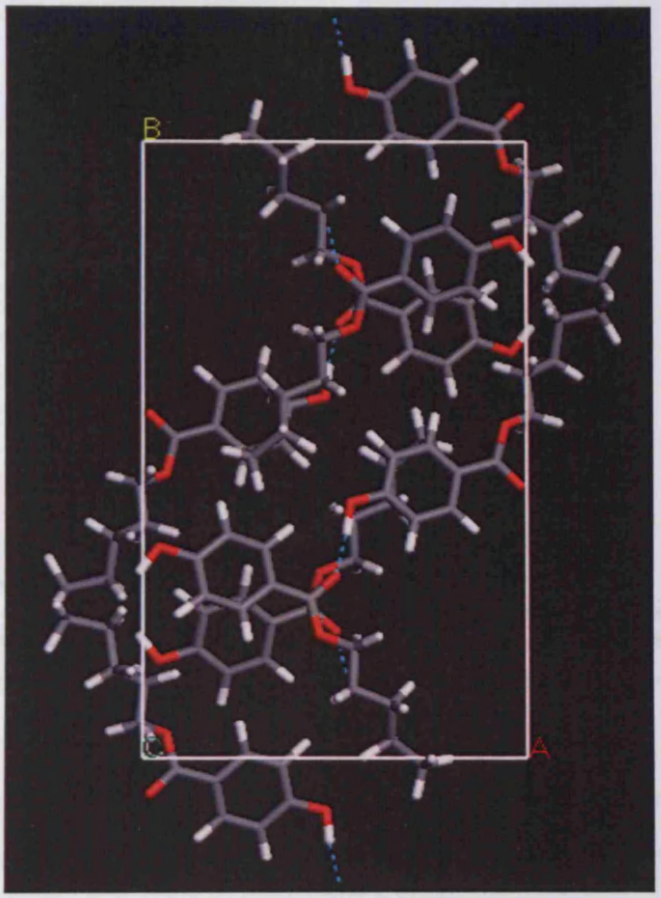



Figure 4.12: Rietveld refined structure of the n = 5 homologue showing the hydrogen bonding arrangement with the hydrogen bonds illustrated with the broken blue lines

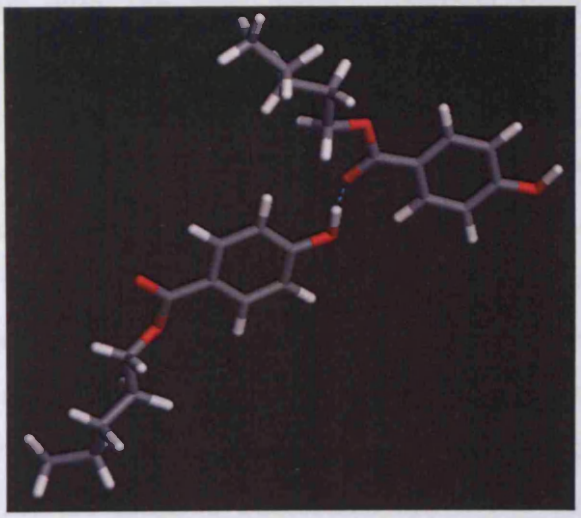


Figure 4.13: Arrangement of the two molecules in the asymmetric unit showing the head to tail hydrogen bonding that creates a zigzag arrangement in the structure

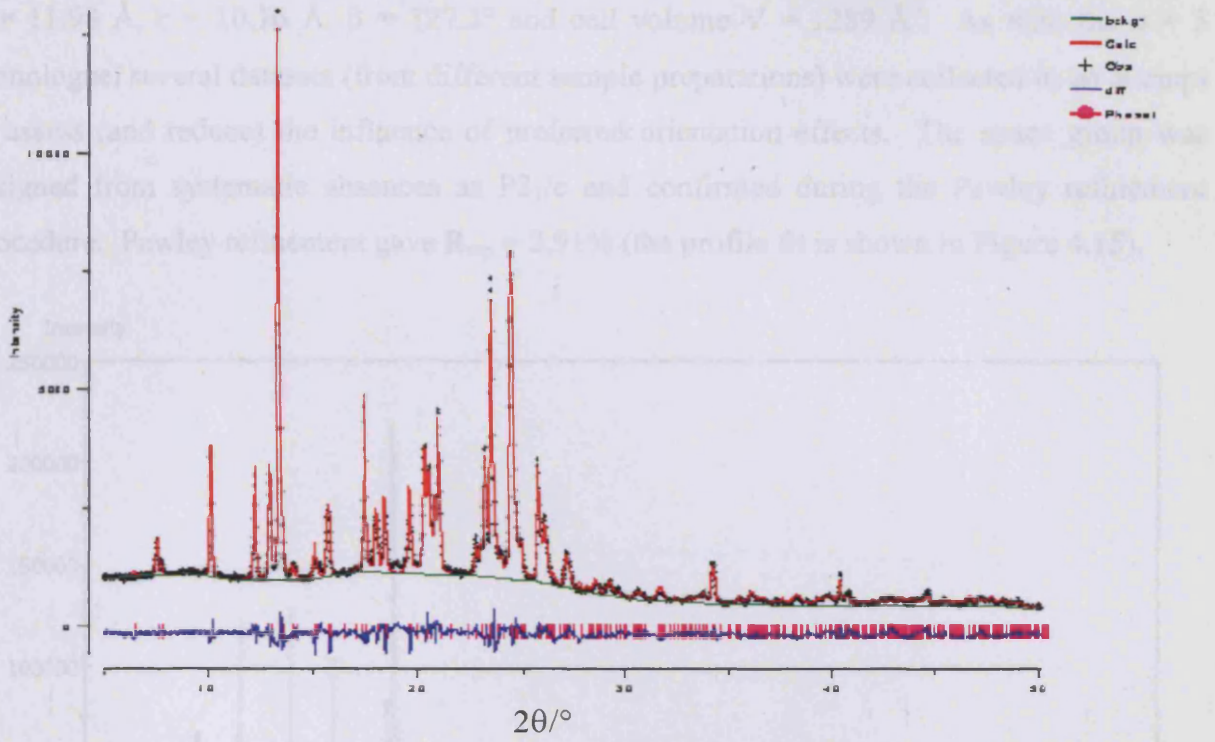



Figure 4.14: Profile fit with $R_{wp}$ = 6.55% ($R_p$ = 5.20%) following Rietveld refinement of the n = 5 homologue (showing the experimental data in black, the calculated data in red and the difference plot in blue in the lower trace)

The n = 5 density value obtained from the refined structure was shown to be 1.173 g $cm^{-3}$, consistent with the true density obtained from helium pycnometry at 1.184 g $cm^{-3}$. The hydrogen bonding arrangement is a head to tail motif (as anticipated) giving rise to a zigzag arrangement between the two molecules in the asymmetric unit along the direction of the b-axis. The chain conformation of both crystallographically independent molecules is largely all trans, with an end gauche arrangement for both molecules. However, in one of the molecules, the end gauche arrangement is staggered whilst the other is eclipsed arrangement. Assessment of this homologue by FTIR spectroscopy did not allude to the presence of end gauche conformers, however, the broad vibrational peaks intimated a range of conformational arrangements. The presence of a relatively high-energy conformational arrangement, especially for the eclipsed conformer, appears to be compensated for by the hydrogen bonding and the relatively densely packed structural arrangement.

### 4.4.2 Structure of the hexyl (n = 6) homologue

The XRPD data was acquired at 298 K and indexed using the program DICVOL (Boultif and Louer 1991) to give a monoclinic unit cell with parameters a = 13.25 Å,

b = 11.98 Å, c = 10.16 Å, β = 127.3° and cell volume V = 1289 Å$^3$. As with the n = 5 homologue, several datasets (from different sample preparations) were collected in an attempt to assess (and reduce) the influence of preferred orientation effects. The space group was assigned from systematic absences as $P2_1/c$ and confirmed during the Pawley refinement procedure. Pawley refinement gave $R_{wp}$ = 2.91% (the profile fit is shown in Figure 4.15).

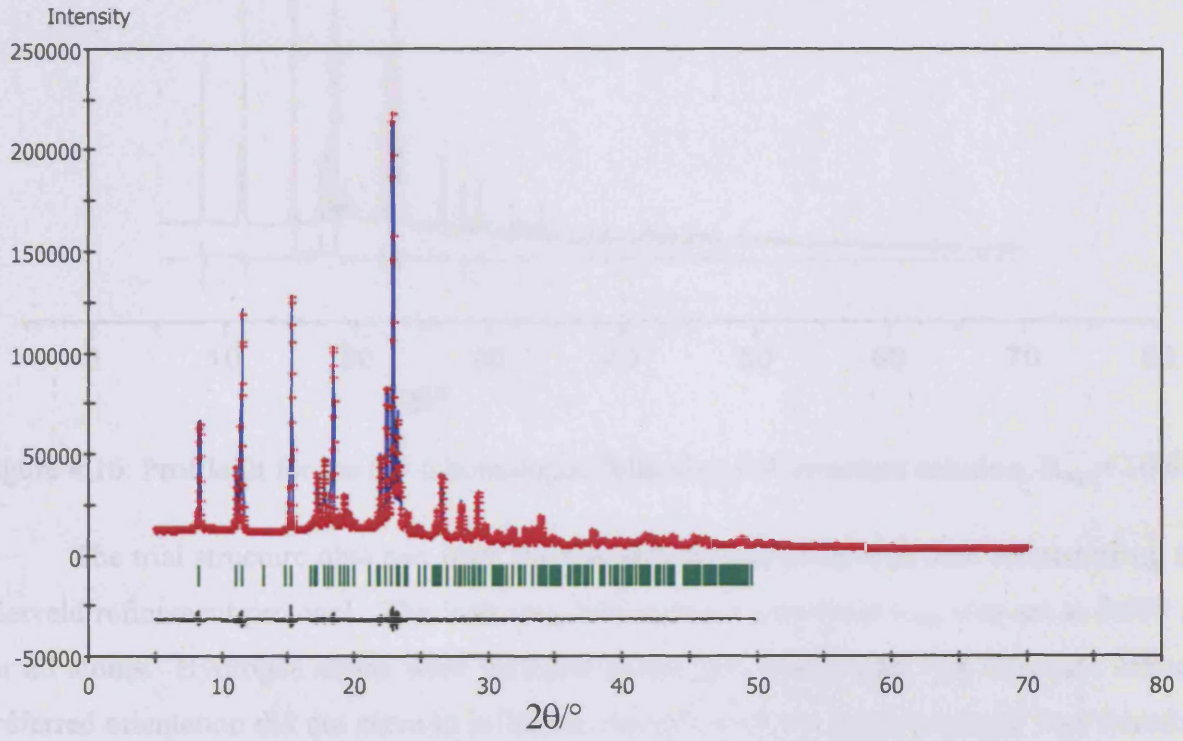



Figure 4.15: Pawley refinement for the n = 6 homologue following indexing by DICVOL of data collected on a Panalytical Xpert XRPD (showing experimental data in red, calculated data in blue and the difference plot in the lower trace in black)

Structure solution was performed using the GA program EAGER, allowing six rotational and translational degrees of freedom for the molecule as a whole, and also six internal rotations within the ester chain. Numerous repeat assessments were performed, using Z = 4, to give reproducible results of $R_{wp}$ = 10.6%. A typical profile plot for the structure solution is shown in Figure 4.16.

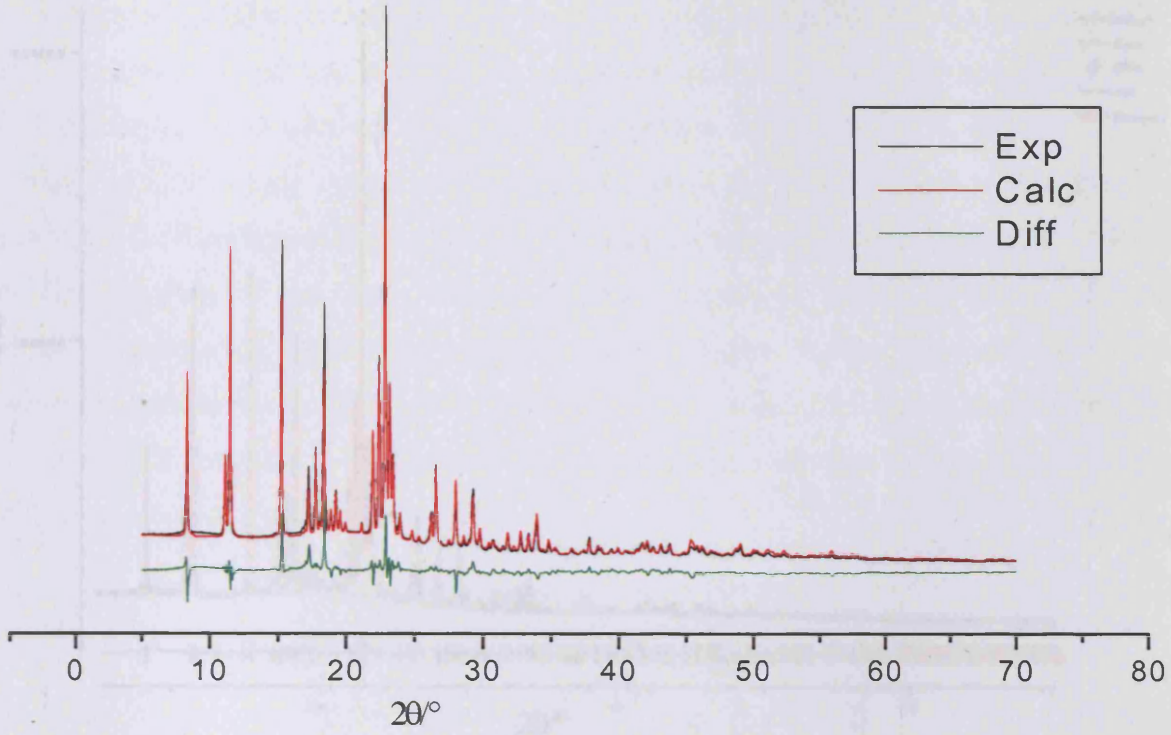



Figure 4.16: Profile fit for the n = 6 homologue following GA structure solution, $R_{wp}$ = 10.6%

The trial structure obtained from the GA structure solution was then refined using the Rietveld refinement protocol. The isotropic displacement parameter $U_{iso}$ was set at 0.075 $Å^2$ for all atoms. Hydrogen atoms were included in the structural model, but were not refined. Preferred orientation did not serve to influence the nature of the refinement and was therefore excluded from the final refined structure. The refinement gave the structure, with $R_{wp}$ = 5.10%, illustrated in Figure 4.17. Figure 4.18 shows the structural arrangement of the refined structure, also showing ester to hydroxy (head to tail) hydrogen bonding.

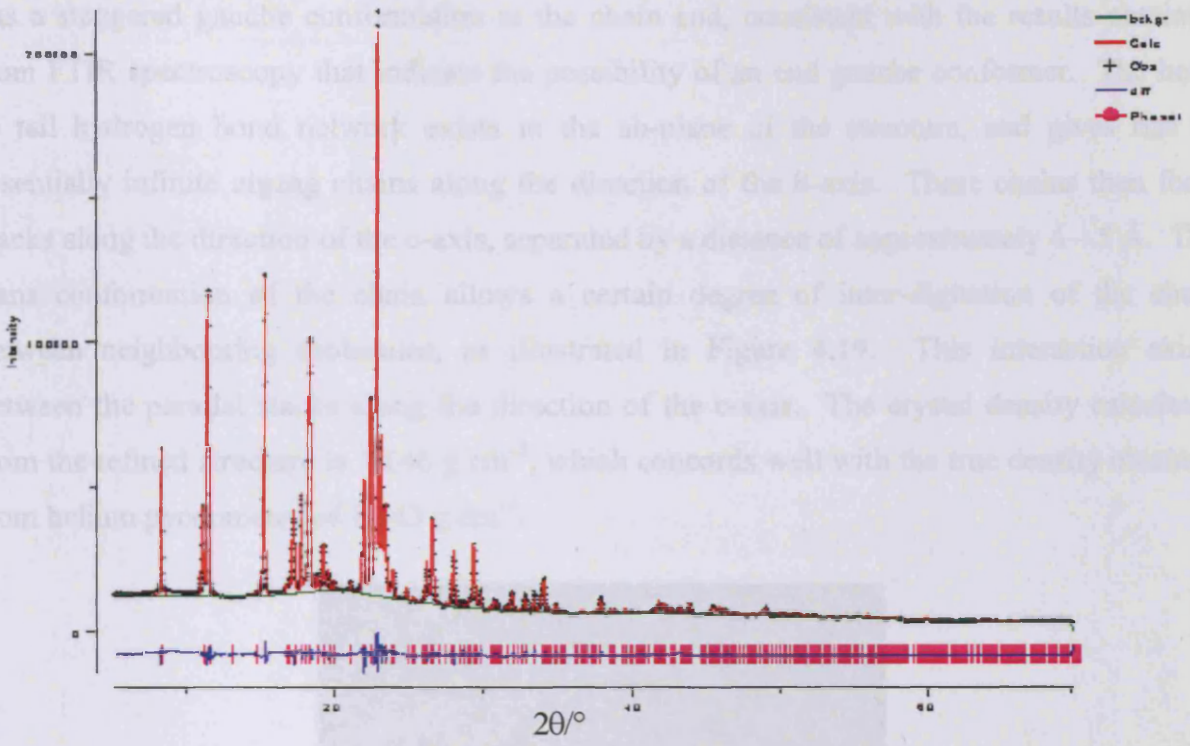



Figure 4.17: Profile fit with $R_{wp}$ = 5.10% ($R_p$ = 3.97%) following Rietveld refinement of the n = 6 homologue (showing the experimental data in black, the calculated data in red and the difference plot in blue in the lower trace)

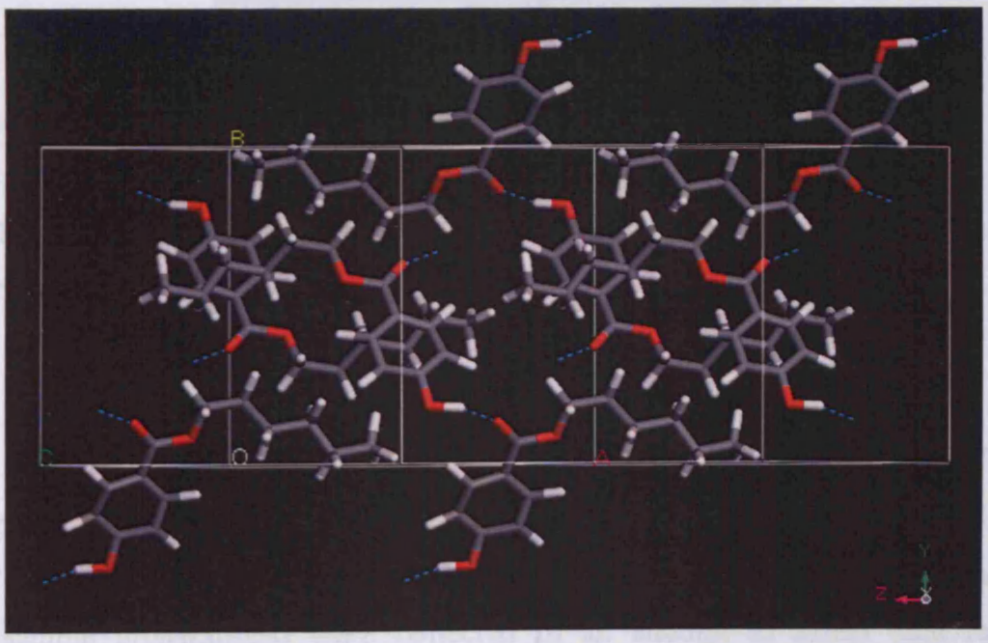


Figure 4.18: Rietveld refined structure for the n = 6 homologue showing the hydrogen bonding arrangement with the hydrogen bonds indicated by the broken blue lines

The structure is composed of a single molecule in the asymmetric unit. The alkyl chain is largely in an all trans configuration with the exception of the terminal region, which

has a staggered gauche conformation at the chain end, consistent with the results obtained from FTIR spectroscopy that indicate the possibility of an end gauche conformer. The head to tail hydrogen bond network exists in the ab-plane of the structure, and gives rise to essentially infinite zigzag chains along the direction of the b-axis. These chains then form stacks along the direction of the c-axis, separated by a distance of approximately 4 – 5 Å. The trans conformation of the chain allows a certain degree of inter-digitation of the chain between neighbouring molecules, as illustrated in Figure 4.19. This interaction exists between the parallel stacks along the direction of the c-axis. The crystal density calculated from the refined structure is 1.146 g $cm^{-3}$, which concords well with the true density obtained from helium pycnometry of 1.143 g $cm^{-3}$.

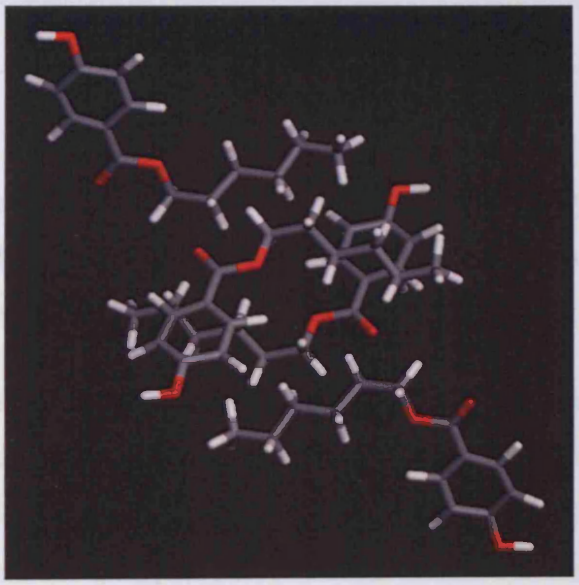


Figure 4.19: View of the n = 6 structure showing the Van der Waals interactions between neighbouring chains

### 4.4.3 Structure of the nonyl (n = 9) homologue

The XRPD data, acquired at 298 K, was indexed using the program ITO (Visser 1969) to give a monoclinic unit cell with parameters a = 32.59 Å, b = 8.23 Å, c = 6.05 Å, β = 92.6° and cell volume V = 1615 $Å^3$. As with the previous homologues, several datasets (from different sample preparations) were collected in an attempt to assess (and reduce) the influence of preferred orientation effects. The space group was assigned during the Pawley refinement procedure from consideration of systematic absences, giving $P2_1/a$ as the best spacegroup. Pawley refinement gave $R_{wp}$ = 6.80% (the profile fit is shown in Figure 4.20).

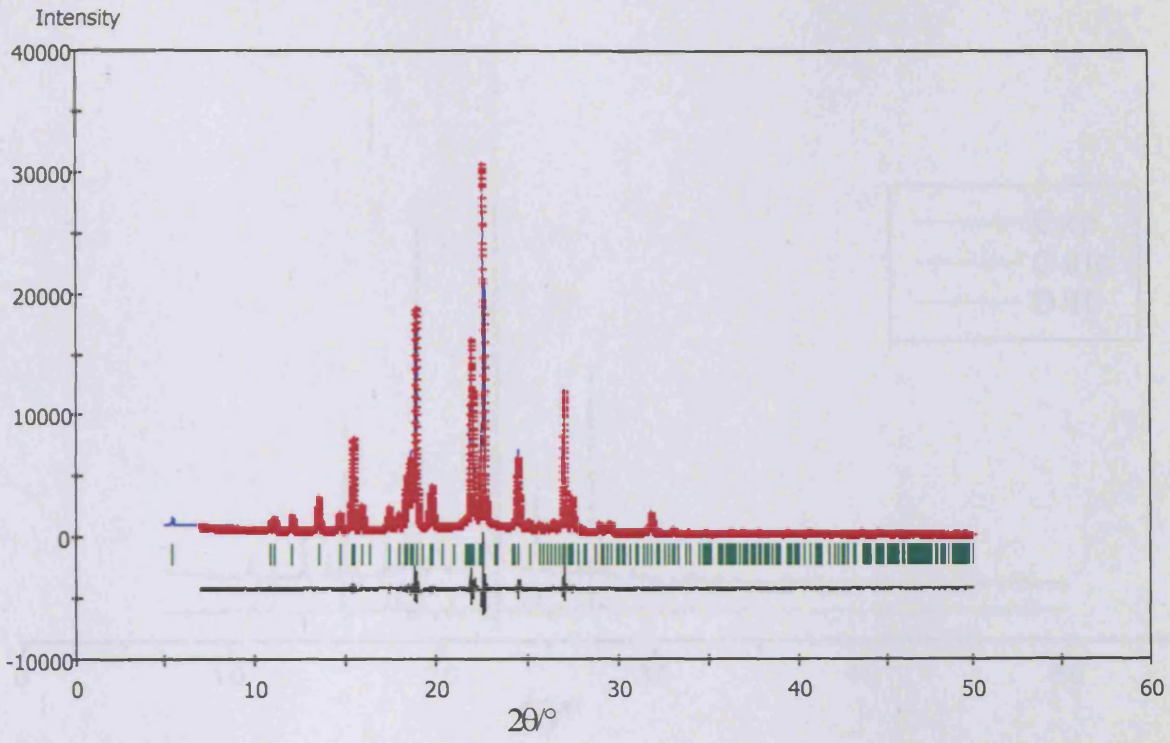



Figure 4.20: Pawley refinement for the n = 9 homologue following indexing by ITO of data collected on a Panalytical Xpert XRPD (showing experimental data in red, calculated data in blue and the difference plot in the lower trace in black)

Structure solution was performed using the GA program EAGER, allowing six rotational and translational degrees of freedom for the molecule as a whole, and also nine internal rotations within the ester chain. The use of one structural fragment for the structure solution reproducibly gave an arrangement with $R_{wp}$ = 25.0%. Furthermore, the use of two structural fragments (each allowing the variation of the degrees of freedom described above) with equal occupancies of 0.5, to mimic the possible disorder suggested by the NQS $^{13}C$ CP/MAS experiments, gave more reproducible structures with a lower value of $R_{wp}$ = 16.2%. A typical profile plot is shown in Figure 4.21.

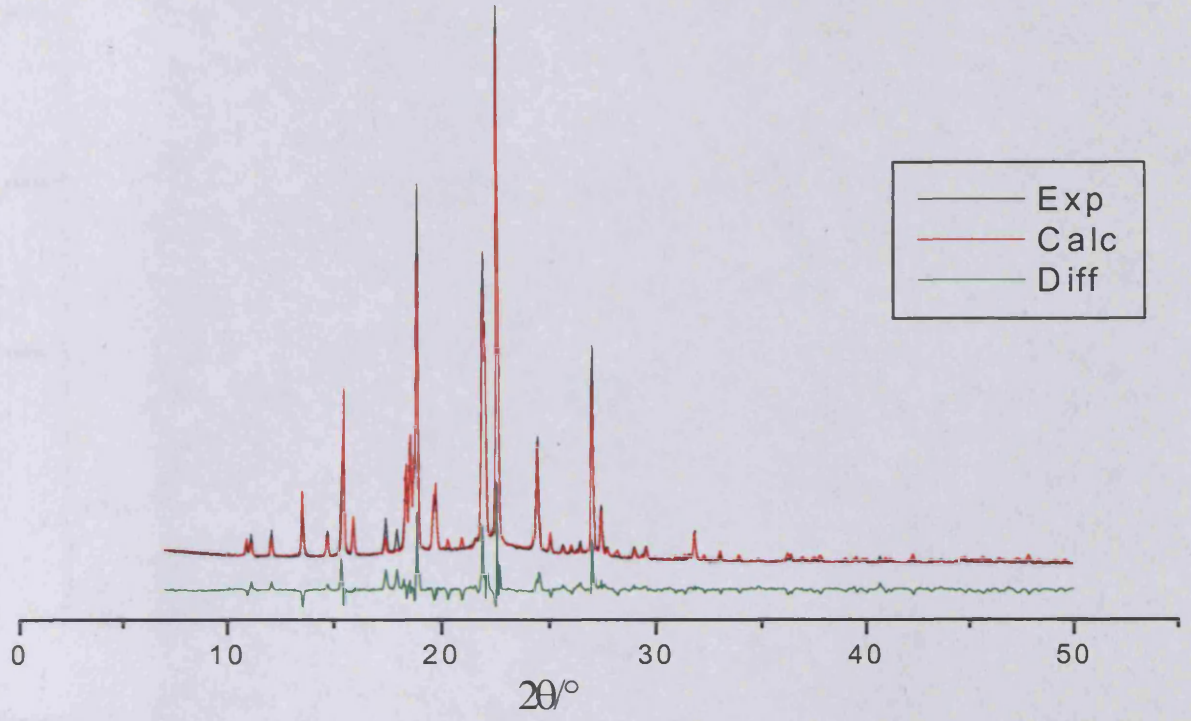



Figure 4.21: Profile fit for the n = 9 homologue following GA structure solution using two structural fragments in the asymmetric unit. $R_{wp}$ = 16.2%

The structure solution obtained from the GA calculations using two structural fragments was then refined using the Rietveld refinement protocol. As for the other homologues, the isotropic displacement parameter $U_{iso}$ was set at 0.075 $Å^2$ for all atoms. Hydrogen atoms were included in the structural model, but were not refined. Preferred orientation did not serve to influence the nature of the refinement and was therefore excluded from the final refined structure. The refinement gave $R_{wp}$ = 10.4% (Figure 4.22). Figure 4.23 shows the refined structure.

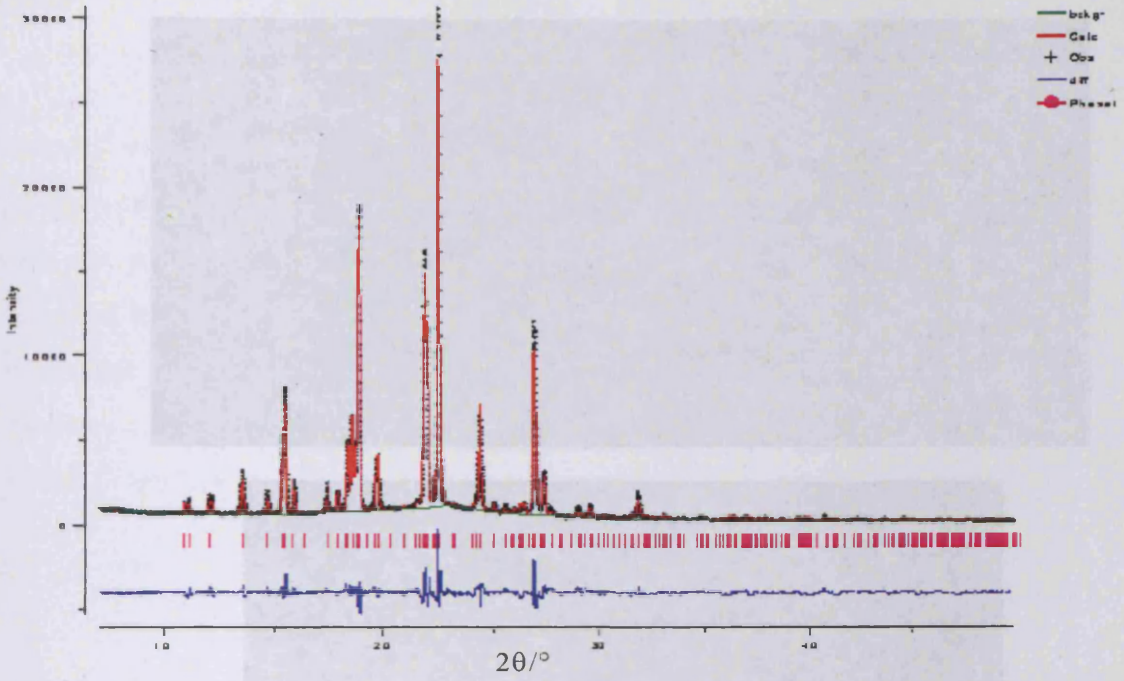



Figure 4.22: Profile fit with $R_{wp}$ = 10.41% ($R_p$ = 7.71%) following Rietveld refinement of the n = 9 homologue (showing the experimental data in black, the calculated data in red and the difference plot in blue in the lower trace)

The FTIR spectrum of this homologue has peaks characteristic of both gauche-trans-gauche, and gauche-gauche kinks. In the structure, one disorder component has almost all the chain torsions in gauche conformations, whilst the other disorder component contains both gauche-trans-gauche and gauche-gauche kinks. As discussed previously, the crystal structure represents a time and space averaged representation of the actual structure, and therefore it is difficult to ascertain whether any dynamic processes occur to interconvert the two disorder components (i.e. dynamic disorder). $^{13}C$ NQS NMR data on this homologue suggests the occurrence of dynamic disorder within the alkyl chains as evidenced by the presence of methylene peaks in the NQS spectrum, but that the remainder of the molecule is probably not subject to significant dynamics. Assessment of the structure shows that whilst the disorder representation provides a reasonable fit to the data, the interpretation is somewhat unclear and will require further work in order to ensure a comprehensive understanding of the nature of the disorder in this material.

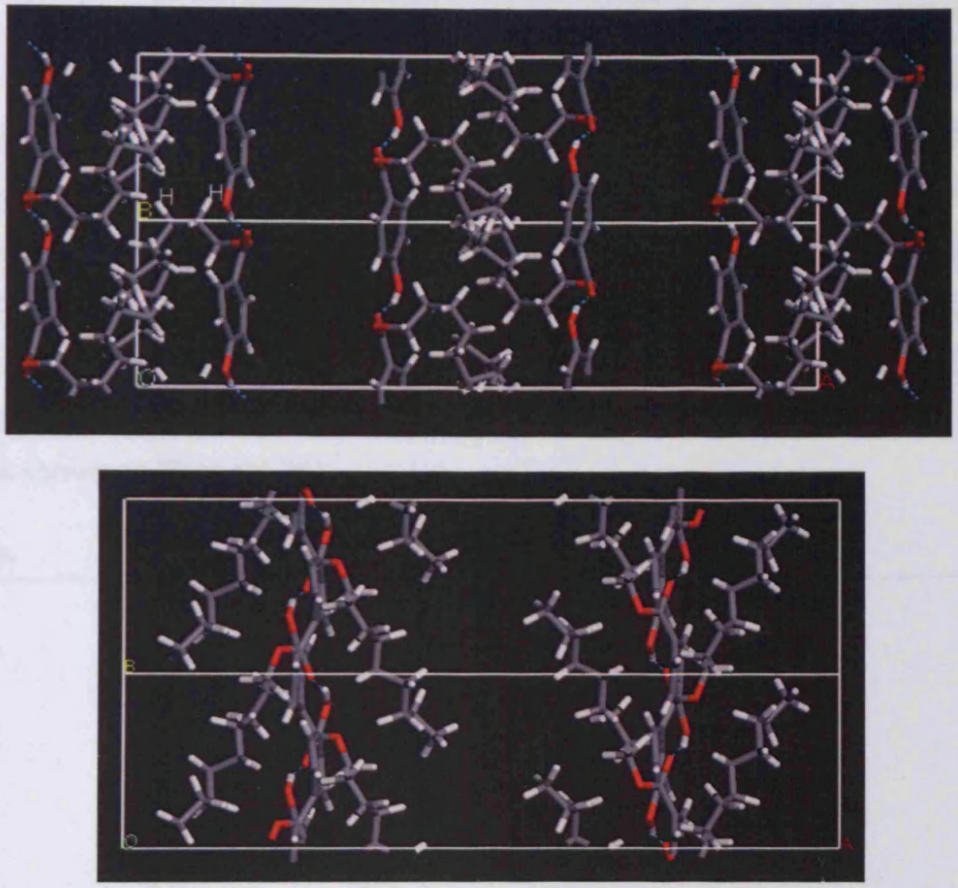


Figure 4.23: Rietveld refined structure of the n = 9 homologue showing each of the two disorder components. The broken blue lines illustrate hydrogen bonds

The entropy of melting associated with the n = 9 homologue shows a pronounced negative deviation from the observed chain length trend. This negative deviation is attributed to the fact that there is inherent disorder present within the structure. As discussed previously, the thermodynamic contributions to melting are associated with both differences in intermolecular forces and conformational arrangements on going from the solid to the liquid state. In the case of the n = 9 homologue, when compared to the structures of the lower homologues, there appears to be significant disorder in the solid state. Additionally, the NMR data indicate the presence of dynamic disorder associated with most of the methylene units in the chain. Data from both the structural arrangement and spectroscopic assessment support the observation of the lower melting parameters, and the marked deviation in the trend observed with the entropy of melting, when compared to the other homologues.

### 4.4.4 Structure of the pentadecyl (n = 15) homologue

As with the previous homologues, the XRPD data was acquired at 298 K and indexed using the program Xcell (Neumann 2003) to give a monoclinic unit cell with parameters a = 4.99 Å, b = 11.12 Å, c = 39.77 Å, β = 97.4° and cell volume V = 2188 Å$^3$. As with the previous homologues, several datasets (from different sample preparations) were collected in an attempt to assess (and reduce) the effects of preferred orientation. The space group was assigned during the Pawley refinement procedure from consideration of systematic absences, giving $P2_1/c$ as the best spacegroup. Pawley refinement gave $R_{wp}$ = 8.38% (the profile fit is shown in Figure 4.24).

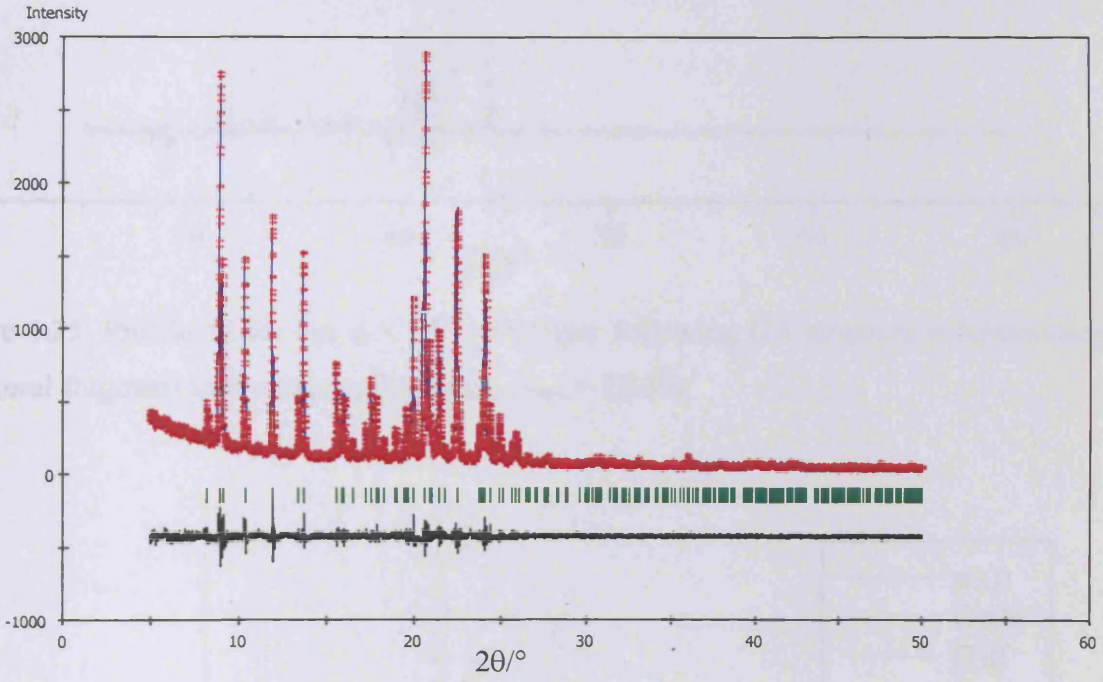



Figure 4.24: Pawley refinement for the n = 15 homologue following indexing by Xcell of data collected on a Panalytical Xpert XRPD (showing experimental data in red, calculated data in blue and the difference plot in the lower trace in black

Structure solution was performed using the GA program EAGER, and calculations were performed using both one and two structural fragments in the asymmetric unit. Several attempts were made in each case to ensure reproducibility. The use of one structural fragment, allowing six rotational and translational degrees of freedom for the molecule as a whole, and also fifteen internal rotations within the ester chain, with unit occupancy, reproducibly gave an arrangement with $R_{wp}$ = 22.3% (as shown in Figure 4.25). Furthermore, the use of two structural fragments (each allowing the variation of the degrees of freedom

described above) with equal occupancies of 0.5, to mimic any possible disorder, gave reproducible structures with $R_{wp}$ = 16.0%. A typical profile plot is shown in Figure 4.26.

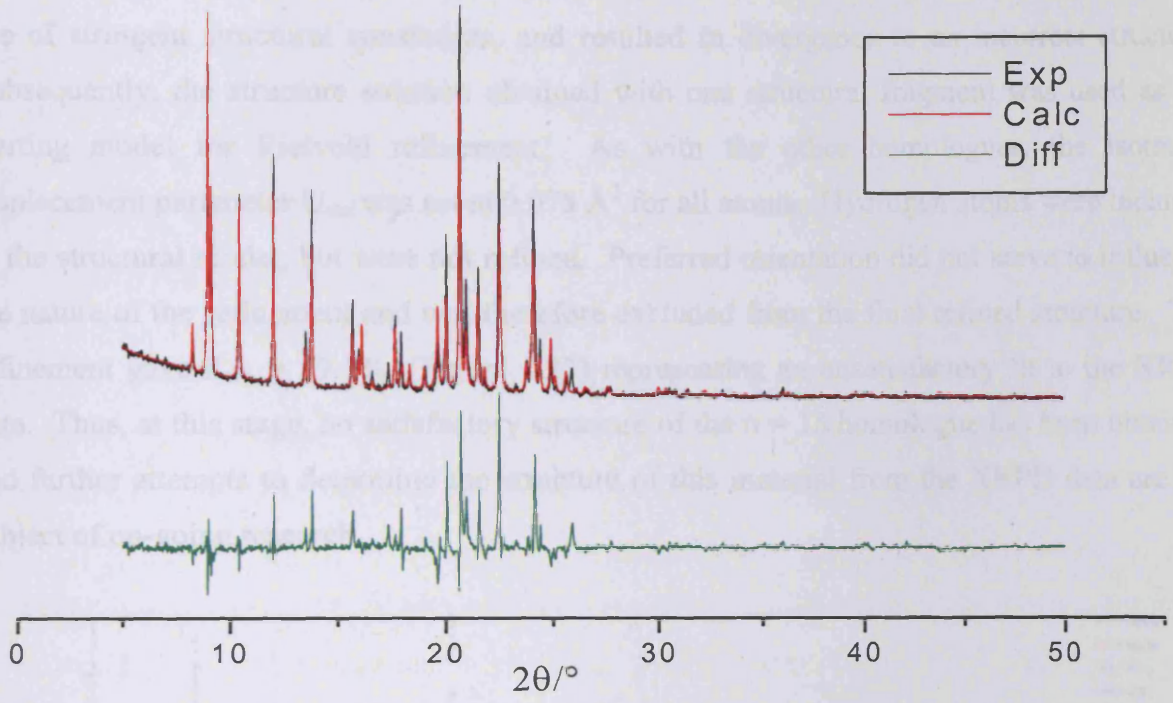



Figure 4.25: Profile fit for the n = 15 homologue following GA structure solution using one structural fragment in the asymmetric unit. $R_{wp}$ = 22.3%

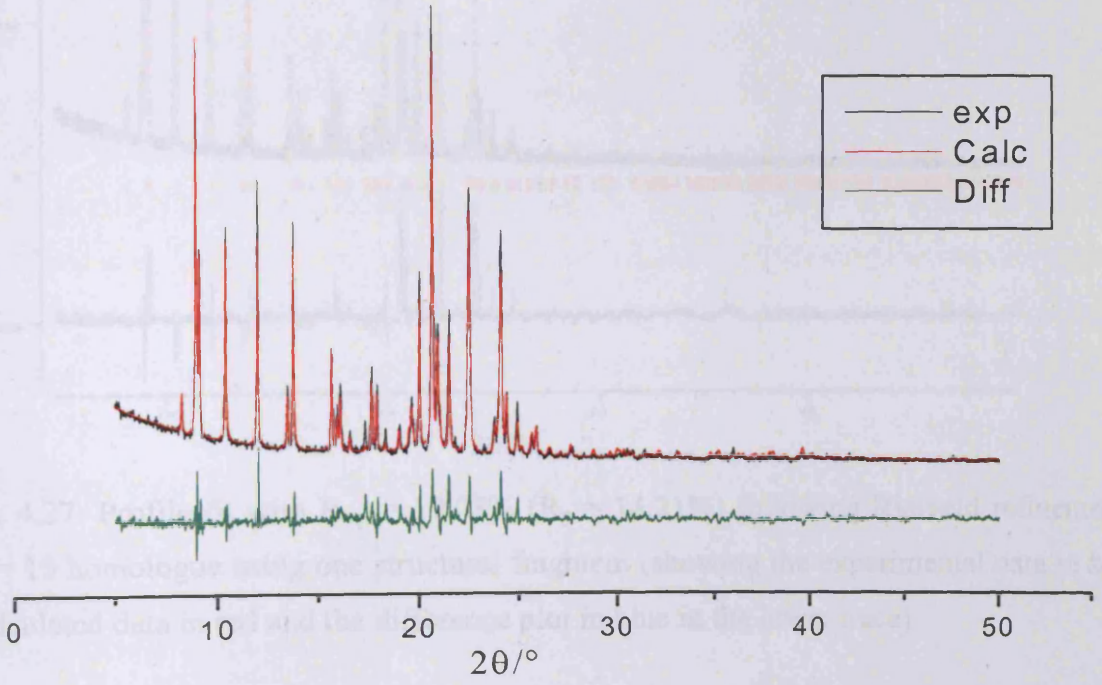



Figure 4.26: Profile fit for the n = 15 homologue following GA structure solution using two structural fragments in the asymmetric unit. $R_{wp}$ = 16.0%

Initially, the trial structure obtained from the GA structure solution using two structural fragments was then refined using the Rietveld refinement protocol. In all repeat assessments, the refinement process caused disruption of the molecular fragments, despite the use of stringent structural constraints, and resulted in divergence to an incorrect structure. Subsequently, the structure solution obtained with one structural fragment was used as the starting model for Rietveld refinement. As with the other homologues, the isotropic displacement parameter $U_{iso}$ was set at 0.075 $Å^2$ for all atoms. Hydrogen atoms were included in the structural model, but were not refined. Preferred orientation did not serve to influence the nature of the refinement and was therefore excluded from the final refined structure. The refinement gave $R_{wp}$ = 19.1%, (Figure 4.27) representing an unsatisfactory fit to the XRPD data. Thus, at this stage, no satisfactory structure of the n = 15 homologue has been obtained and further attempts to determine the structure of this material from the XRPD data are the subject of on-going research.

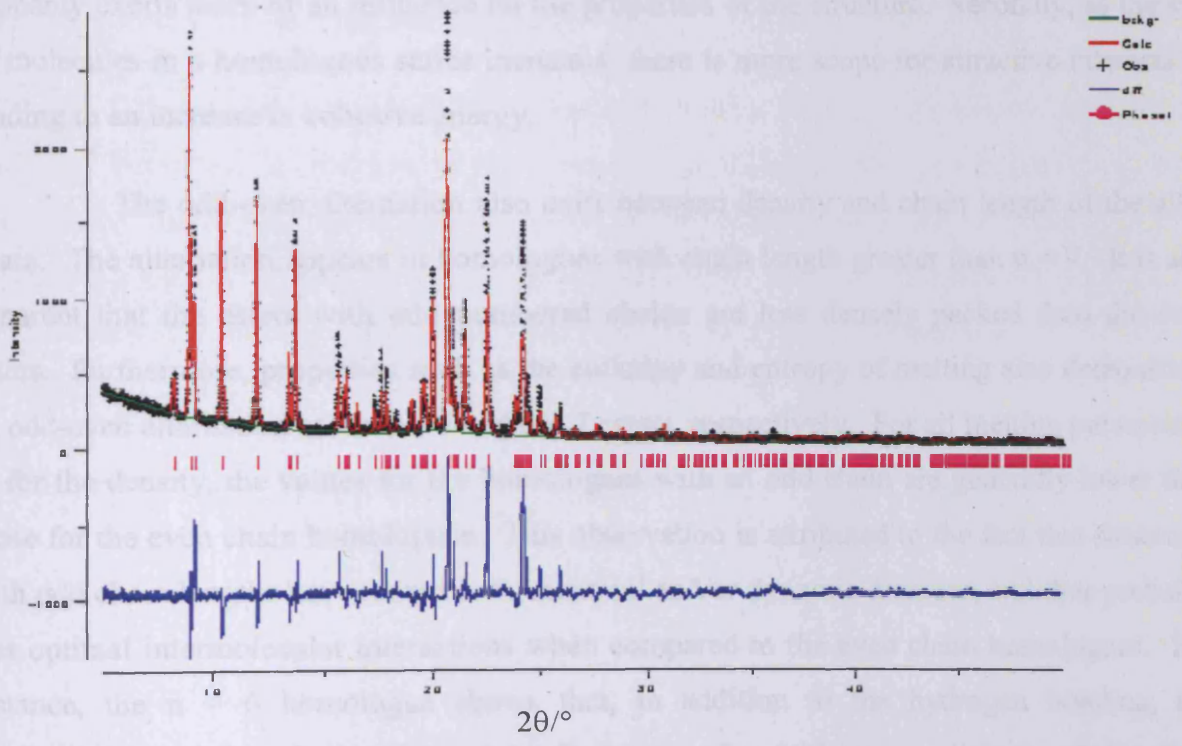



Figure 4.27: Profile fit with $R_{wp}$ = 19.08% ($R_p$ = 14.21%) following Rietveld refinement of the n = 15 homologue using one structural fragment (showing the experimental data in black, the calculated data in red and the difference plot in blue in the lower trace)

## 4.5 Discussion and conclusions

The long-range periodicity of a structure is brought about by the short-range interactions, both attractive and repulsive, that give rise to an efficiently packed array. These interactions are brought about by a balance or compensation between thermodynamic contributions, namely compensation between enthalpy and entropy constraints. For the paraben esters, as with other long chain homologous series, an odd-even variation exists in most physicochemical properties. For example, the melting temperature shows initially a linear decrease as a function of chain length until the n = 5 homologue, at which point there is an inflection in the trend accompanied by an odd-even alternation. The inflection corresponds to a general upward trend in the melting temperature as a function of chain length and is attributed to a change in the type and overall increase in intermolecular interactions. Firstly, after this point (the n = 5 homologue) the aliphatic chain, rather than the benzoate moiety, probably exerts more of an influence on the properties of the structure. Secondly, as the size of molecules in a homologous series increases, there is more scope for attractive interactions leading to an increase in cohesive energy.

The odd-even alternation also exits between density and chain length of the alkyl chain. The alternation appears in homologues with chain length greater than n = 7. It is also apparent that the esters with odd numbered chains are less densely packed than the even esters. Furthermore, properties such as the enthalpy and entropy of melting also demonstrate an odd-even alternation at the n = 5 and n = 7 esters, respectively. For all melting parameters, as for the density, the values for the homologues with an odd chain are generally lower than those for the even chain homologues. This observation is attributed to the fact that structures with odd chain lengths have more conformational and/or dynamic disorder, and thus probably less optimal intermolecular interactions when compared to the even chain homologues. For instance, the n = 6 homologue shows that, in addition to the hydrogen bonding, the conformation of the chain allows inter-digitation of neighbouring aliphatic chains thus maximising Van der Waals interactions. As a consequence, the n = 6 homologue has higher melting parameters and a more densely packed structure in comparison with the n = 5 homologue (which has a more disordered structure).

The n = 9 homologue, in addition, has thermodynamic properties ($\Delta H$ and $\Delta S$ of melting) that deviate markedly from the general trend observed as a function of chain length.

Whilst the final structure has yet to be confirmed, the tentative structure shows the presence of considerable disorder, which is consistent with the lower melting temperature and density.

There is no apparent odd-even alternation that could be correlated to the crystal system, as had been observed for other long chain homologous series such as n-alkanes (where odd chain length structures pack in orthorhombic cells and even structures in triclinic cells). Based on the current data, the odd-even alternation is ascribed to the variation in structural cooperativity to intermolecular interactions. For instance, the higher values of the melting thermodynamic parameters for the even chain length homologues are attributed to an enhanced ability for both hydrogen bonding and Van der Waals interactions to occur – leading to a more densely packed structure. Odd chain length homologues, conversely, possess geometrical constraints brought about by conformational or dynamic disorder that result in less densely packed structures due to the restricted scope for intermolecular interaction. These generalities need to be verified by solving more structures of homologues containing even chain lengths, but the general trend is evident.

The paraben esters (and in particular the higher homologues) possess considerable torsional flexibility in addition to the presence of disorder, and consequently represent particularly challenging structures to solve. The solution of some of these structures represents a significant step forward in the application of direct space methodologies to more complex systems, and is among the first attempts to tackle disordered structures by this method.

## 4.6 Experimental

### 4.6.1 Thermal analysis

Differential scanning calorimetry (DSC) experiments were carried out using a Perkin-Elmer DSC-7. Samples were analysed under a nitrogen atmosphere in 50 µl aluminium vented pans.

### 4.6.2 Synthesis of higher homologues

The decyl (C10), undecyl (C11), tridecyl (C13), tetradecyl (C14) and pentadecyl (C15) esters of 4-hydroxy benzoic acid were synthesised using the following procedure. Concentrated sulphuric acid (0.1 mL) was added to a stirred solution containing 4-hydroxy benzoic acid (1.38 g, 10 mmol) and the corresponding alcohol (11 mmol) in toluene (30 mL). The resultant solution was heated to 140°C for four hours with continual extraction of water using a Dean-Stark apparatus. The reaction mixture was then cooled to room temperature and silica gel (10 g) was added. The solvent was removed under reduced pressure and the resulting powder was loaded onto the top of a silica gel bed (4 x 15 cm). The product was eluted with 20% diethyl ether in iso-hexane to yield the ester as a white crystalline powder.

Decyl (C10) ester (1.67 g, 60.0%)

MS m/z: 277.4 ($M^-$) $C_{17}H_{26}O_3$ requires 278.4

HPLC purity: 98.2 %

$^1$H NMR δ($CDCl_3$): 7.98 – 7.95 (2H m), 7.00 – 6.84 (2H m)

Undecyl (C11) ester (1.43 g, 48.9%)

MS m/z: 291.4 ($M^-$) $C_{18}H_{28}O_3$ requires 292.4

HPLC purity: 99.8 %

$^1$H NMR δ($CDCl_3$): 7.98 – 7.95 (2H m), 7.00 – 6.84 (2H m)

Tridecyl (C13) ester (1.48 g, 46.1%)

MS m/z: 319.5 ($M^-$) $C_{20}H_{32}O_3$ requires 320.5

HPLC purity: 99.0 %

$^1$H NMR δ($CDCl_3$): 7.98 – 7.95 (2H m), 7.00 – 6.84 (2H m)

Tetradecyl (C14) ester (1.1 g, 32.8%)

MS m/z: 333.5 ($M^-$) $C_{21}H_{34}O_3$ requires 334.5

HPLC purity: 96.7 %

$^1$H NMR δ($CDCl_3$): 7.98 – 7.95 (2H m), 7.00 – 6.84 (2H m)

Pentadecyl (C15) ester (1.65 g, 47.3%)

MS m/z: 347.5 ($M^-$) $C_{22}H_{36}O_3$ requires 348.5

HPLC purity: 99.8 %

$^1$H NMR δ($CDCl_3$): 7.98 – 7.95 (2H m), 7.00 – 6.84 (2H m)

### 4.6.3 FTIR spectroscopy

FTIR spectroscopy was performed on pure samples using a Perkin-Elmer Spectrum-GX spectrometer fitted with a single bounce SpecAc variable temperature ATR attachment. Samples were analysed neat, and spectra were acquired over the range 4000 – 600 $cm^{-1}$ using a spectral resolution of 4 $cm^{-1}$ with 64 accumulations.

### 4.6.4 Solid state $^{13}$C CP/MAS NMR spectroscopy

High-resolution solid-state $^{13}$C NMR spectra were recorded at 75.5 MHz using a Chemagnetics Infinity 300 spectrometer and a Chemagnetics triple resonance probe. A standard CP and $^1$H decoupling sequence was used (as described in Chapter 2) and samples were packed in 4 mm zirconia rotors and spun at typical speeds of 5 kHz. NQS spectra were also acquired using the pulse-sequence described in Chapter 2.

### 4.6.5 Structure solution from X-ray powder diffraction data

Materials Studio Reflex module (version 1.2) from Accelerys was used for indexing and Pawley refinement. ITO (Visser 1969), TREOR (Werner *et al.* 1985), DICVOL (Boultif and Louer 1991) or X-cell (Neumann 2003) were used simultaneously to index the

XRPD patterns of the homologues. The best-fit cell from each analysis was used during the Pawley refinement procedure, which was used to refine the powder pattern with respect to the profile and background parameters (as described in Chapter 2). The measure of similarity used between the simulated and experimental profiles was $R_{wp}$. Structure solution was performed using the GA program EAGER, using the multi-population parallel genetic algorithm technique. Typically, the structure solution calculations involved 60 to 100 starting populations with 50 to 100 generations. The mating rate (i.e. the number of mating events per generation) was set at 15, and the mutation rate was typically set at 8 per generation.

Refinement was performed using the GSAS software as described in Chapter 2. Following an initial Le Bail refinement, the positions of the atoms were refined using the Rietveld protocol with standard geometrical restraints applied to bond lengths, bond angles and planar arrangements. For all non-hydrogen atoms, the isotropic displacement parameters were constrained to a common value. Hydrogen atoms were incorporated into the structure in the final stages of refinement but were not refined.

### 4.6.6 True density measurements

True density of powders was established using helium pycnometry, with a Micromeritics Accupyc 1330 and a 1 $cm^3$ sample cup. The headspace volume of the cup containing the sample relative to the calibration standard was used to determine density of the material.

## 4.7 References


- Boese R., Weiss H-C., Bläser D., (1999), *Angew. Chem. Int. Ed.*, **38**, 988
- Bond A., (2003), *Chem. Comm.*, 250
- Casal H. L., McElhaney R. N., (1990), *Biochemistry*, **29**, 5423
- Charton M., Charton B. I., (2002), *Advances in QSPR Volume 3*, JAI Press Inc.
- Dall'Acqua L., Della Gatta G., Nowicka N., Ferloni P., (2002) *J. Chem. Thermodynamics*, 34, 1
- Dorset D. L., Strauss H. L., Snyder R. G., (1991), *J. Phys. Chem,* **95**, 938
- Dunitz J. D., (1995), *Chem. & Biol.*, **2**, 709
- Dunitz J. D., Gavezotti A., (1999), *Acc. Chem. Res.*, **32**, 677
- Fukahori M., Takatsuji Y., Takahashi H., Sato H., Yotsuyanagi T., (1996), *Chem. Pharm. Bull.*, **44**, 1068
- Gilbert A. S., (1999), *Thermochim. Acta*, **339**, 131
- Giordano F., Bettini R., Donini C., Gazzaniga A., Caira M. R., Zhang G. G. Z., Grant D. J. W, (1999), *J. Pharm Sci.*, **88**, 1210
- Lin X., (1983), *J. Struct. Chem.*, **2**, 213
- Lin X., (1986), *J. Struct. Chem.*, **5**, 281
- McGann M. R., Lacks D. J., (1999), *J. Phys. Chem. B*, **103**, 2796
- Neumann M., (2003), *J. Appl. Crystallogr.*, **36**, 356
- Paruta A., (1969), *J. Pharm Sci.*, **58**, 216
- Pedersen S., Kristensen H. G., Cornett C., (1994), *S.T.P. Pharm. Sci.*, **4**, 292
- Searle M. S., Williams D. H., (1992), *J. Am. Chem. Soc.*, **114**, 10690

- Small D. M., (1986), *The Physical Chemistry of Lipids*, Plenum New York
- Visser J. W., (1969), *J. Appl. Crystallogr.*, **2**, 89
- Werner P.E., Eriksson L., Westdhal M., (1985), *J. Appl. Crystallogr.*, **18**, 367
- Yalkowsky S. H., Flynn G. L., Slunick T. G., (1972), *J. Pharm. Sci.*, **61**, 852

# 5 Understanding relative polymorph stability through structure and phase diagrams

## 5.1 Introduction

Polymorphism, as discussed previously in Chapter 1, arises from a variation of molecular packing in crystal structures, resulting from different intermolecular interactions. As would be expected, significant differences in the mode of intermolecular interactions give rise to dissimilarities in lattice energies, relative stability and other properties of each of the different polymorphs. In the area of pharmaceutical materials characterisation and selection, an understanding and control of polymorphism is crucial. The variation in the lattice energies of different polymorphs can profoundly affect the bioavailability of a solid dosage form. Furthermore, properties such as chemical and physical stability, and hygroscopicity, amongst other things, can also be governed by the nature of the crystal structure. In order to select an appropriate material for development, an understanding of the propensity for polymorphism is required. Once the occurrence of polymorphism has been established, the inter-relationships between the polymorphs (i.e. the ease of interconversion and the relative thermodynamic stability) and the respective properties of each polymorph need to be understood and addressed in order to rationalise the selection of a robust polymorph for development. Several methods can be used to establish the stability hierarchy and include solution-mediated processes, evaluation of the crystal structures with respect to hydrogen bonding and density, and assessment of energy or pressure temperature diagrams, as discussed in Chapters 1 and 2.

AZD7140, the molecular structure of which is shown in Figure 5.1, is a development compound for the treatment of inflammatory disorders. This compound exhibits both polymorphism and forms solvates, depending upon the conditions used during isolation or recrystallisation. The associated toxicity of the included solvent molecules in solvates renders these materials pharmaceutically undesirable. As such, the present study concentrates on the in-depth assessment of the two polymorphs (designated form I and form II) isolated for AZD7140, leading to the selection of the optimum form for development.

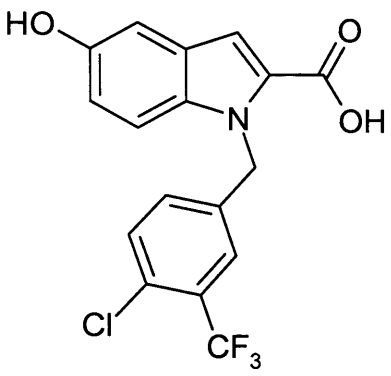



Figure 5.1: Molecular structure of AZD7140

## 5.2 Preparation and characterisation of the two polymorphs of AZD7140

AZD7140 polymorphs were prepared by recrystallisation from a hot saturated 3:2 v/v solution of methanol:water. Supersaturation effects governed the preferential formation of one polymorph over the other as a function of both the overall volume of recrystallisation solvent and the temperature range (or more specifically the cooling rate) during recrystallisation. Crystallisation, which is driven by the change in activity of the crystallising species between the solution state and crystalline state, is significantly influenced by both temperature and concentration. Form I was produced with slow cooling from a large volume of solvent, whilst form II was obtained from more rapid cooling with less solvent.

Figure 5.2 illustrates the DSC curves for both polymorphs of AZD7140, showing that at a moderate heating rate (~ 10 K $min^{-1}$) both polymorphs exhibit a single melting event. Assessment of the DSC profile at different heating rates showed the resolution of several events from the single melting event for form I. Figure 5.3 illustrates the DSC profile of form I at two different heating rates.

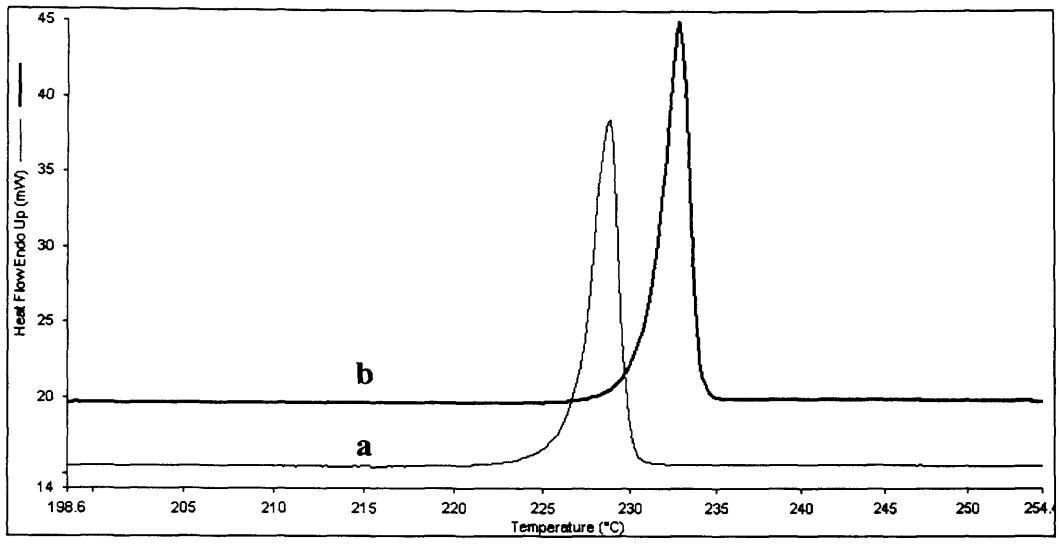



Figure 5.2: DSC profile of the two polymorphs of AZD7140 (a) form I and (b) form II

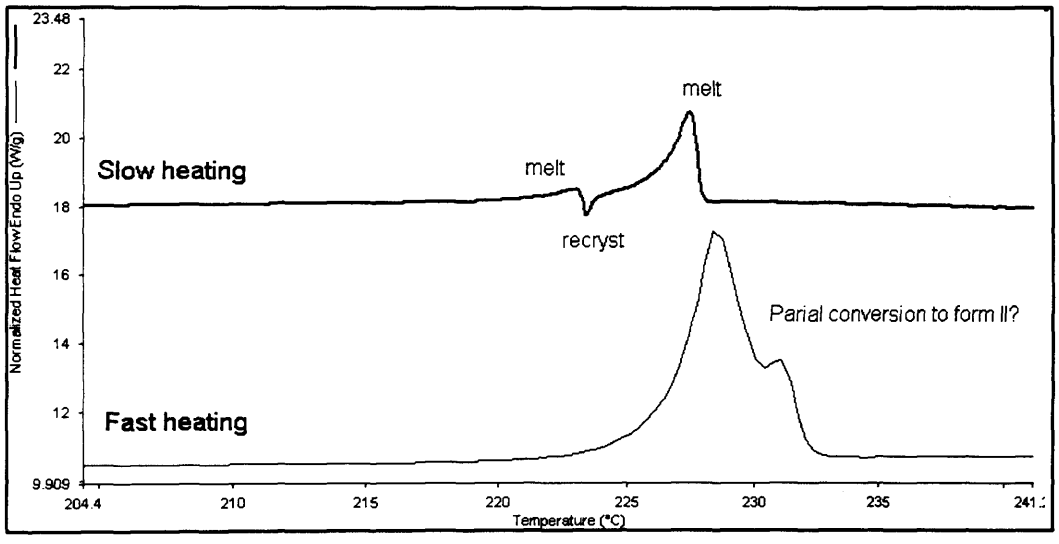



Figure 5.3: Variable heating rate DSC profiles of polymorph I of AZD7140

It is well known that experimental factors, such as heating rate and sample/particle size, can influence the nature of the melting behaviour (Giron 1995). By reducing the particle size (for instance by micronisation) or reducing the heating rate (typically 0.5 to 1 K $min^{-1}$) the single melting event for form I seen in the initial evaluation by DSC was resolved into either a fused doublet or three discrete events corresponding to the melting of form I followed by the recrystallisation of form II and subsequent melting of form II. To understand this behaviour further and to verify the sequence of events, hot stage microscopy was performed and indicated that the higher melting form (II) crystallises from the melt of form I as shown in Figure 5.4. It may be inferred that if the heating rate in the DSC experiments is sufficiently fast, there is no opportunity for the higher melting form II to crystallise from the melt of form I, and a single melting endotherm is observed. Conversely, if the heating is slowed or the

sample size is decreased sufficiently so as not to cause a large temperature differential across the sample, then a melt-recrystallisation-melt series of events is observed. Form II exhibits a single melting event irrespective of heating rate or particle size.

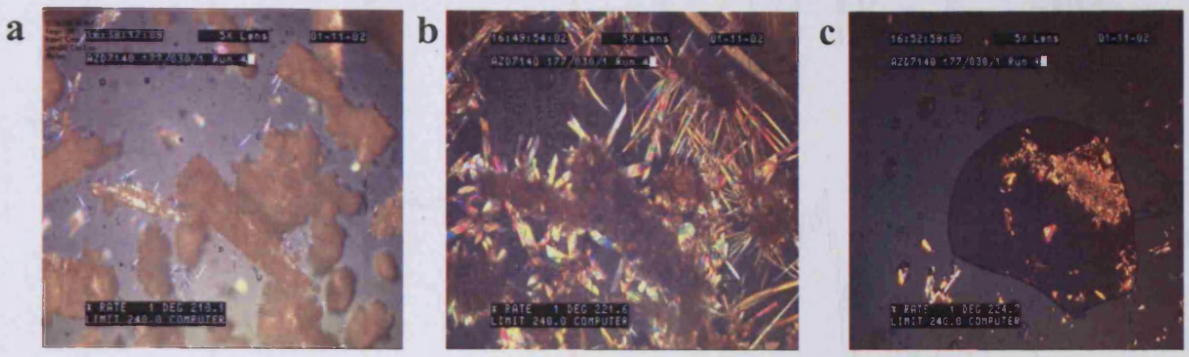



Figure 5.4: Hot stage micrographs of form I of AZD7140 showing (a) melt of form I, (b) recrystallisation of form II from the melt and (c) melt of form II

## 5.3 Spectroscopic assessments of the polymorphs of AZD7140

Spectroscopic assessment was performed on both polymorphs in order to establish spectroscopic differences between them and also in an attempt to assign hydrogen-bonding motifs. The FTIR-ATR spectra of the two polymorphs of AZD7140 are shown in Figure 5.5. The spectra highlight significant differences associated with the phenol indole acid moiety, and Figure 5.6 illustrates the carbonyl-stretching region, emphasising differences in the carboxylic acid groups between the two structures. The differences are also reflected in the Raman spectra as illustrated in Figure 5.7. The presence of two carbonyl groups in a structure with a centre of symmetry (such as an oxalate or carboxylic acid dimer arrangement) invokes an asymmetric and symmetric C=O stretching mode.

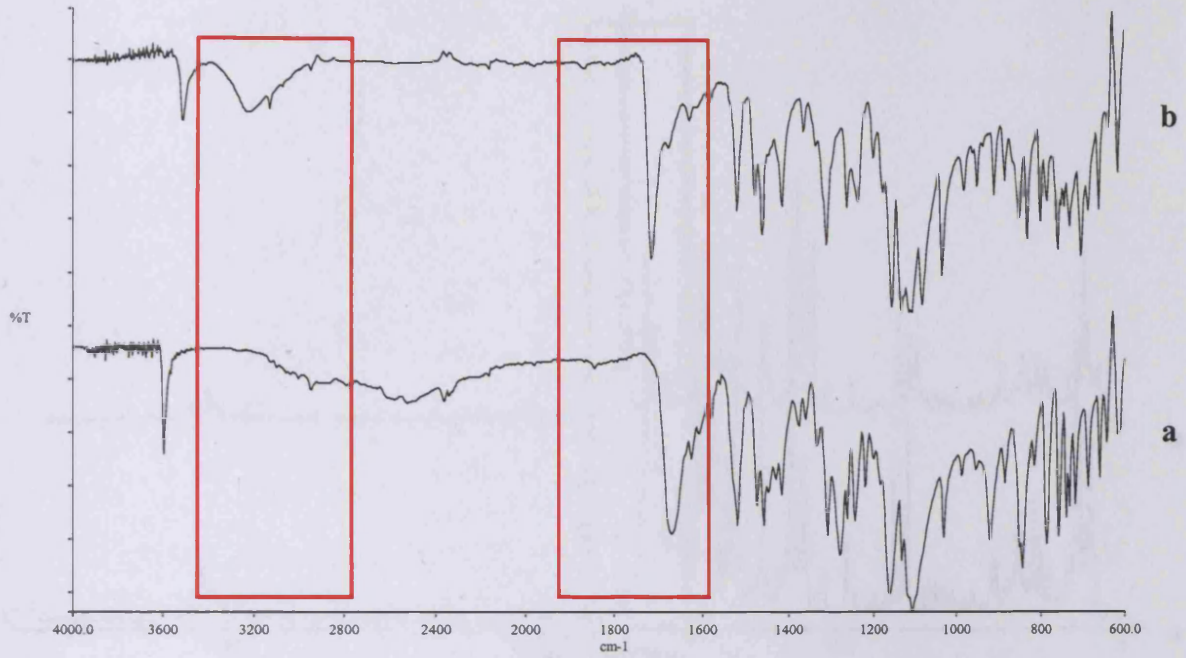



Figure 5.5: FTIR-ATR spectra of AZD7140 (a) form I and (b) form II, illustrating significant differentiation most notably in the phenol O—H and carbonyl stretching regions (as highlighted by the boxed regions in the Figure)

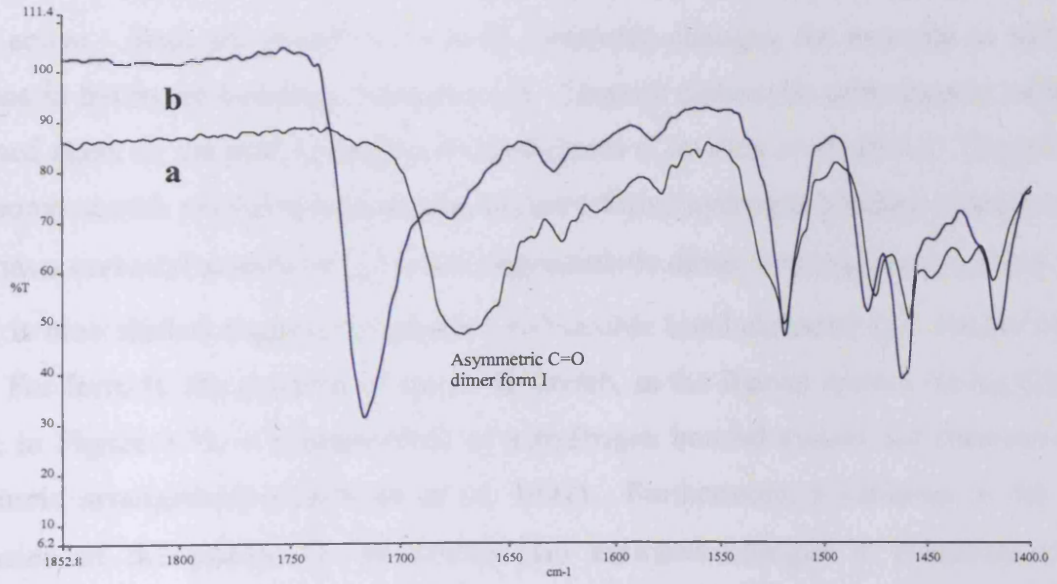



Figure 5.6: FTIR-ATR spectra of AZD7140 (a) form I and (b) form II, showing the C=O and aromatic ring stretching regions

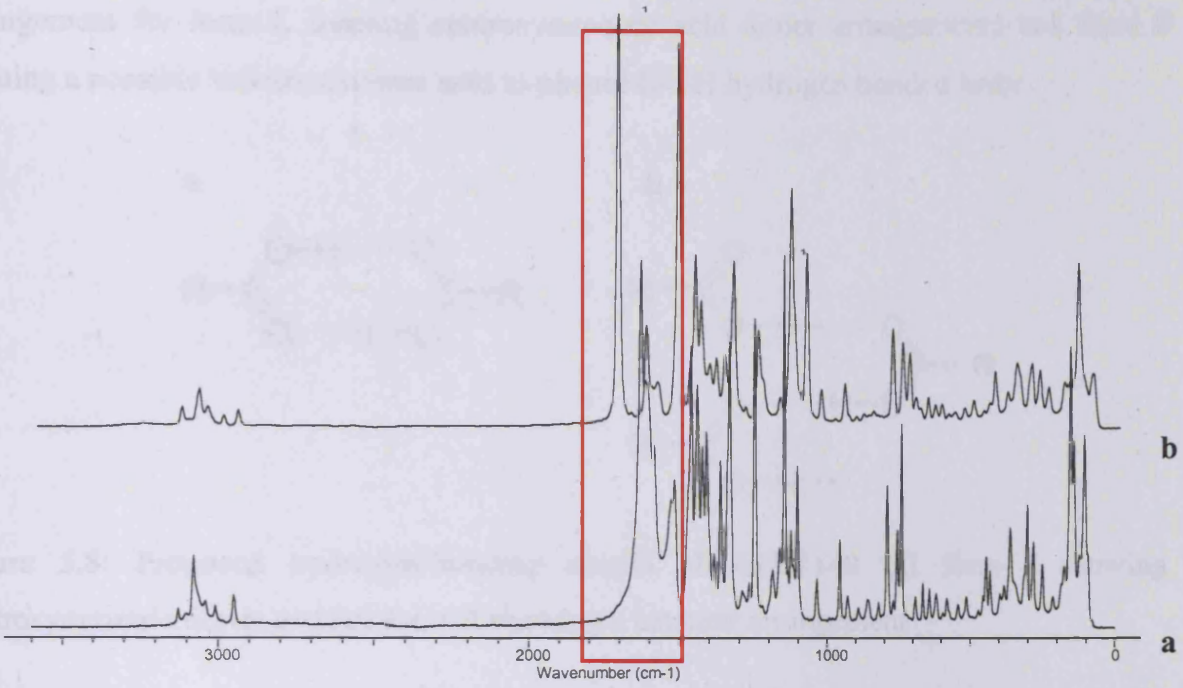



Figure 5.7: Raman spectra of AZD7140 (a) form I and (b) form II, the box highlights the most notable difference in the carbonyl symmetric stretching frequency

The asymmetric carbonyl stretch is IR active, whilst the symmetric stretch is Raman active. Both are sensitive to local electronic changes, for example as induced by variations in hydrogen bonding arrangements. Organic carboxylic acids usually exist, in the condensed state, as the acid hydrogen bonded dimer (Lin-Vien *et al.* 1991). The position of the carbonyl stretch provides information on the relative hydrogen bonding arrangement, and form I has a carbonyl stretch ( $\nu_{C=O}^{ass}$ ) in the characteristic dimer frequency range, whilst that for form II is blue shifted suggesting greater C=O double bond character (i.e. weaker hydrogen bond). For form II, the position of the C=O stretch, in the Raman spectra (as highlighted by the box in Figure 5.7), is characteristic of a hydrogen bonded system but consistent with a non-dimeric arrangement (Lin-Vien *et al.* 1991). Furthermore, a variation in the relative frequencies of the phenol O—H stretch also indicates changes in electronic character associated with this functional group. For instance, for form II the phenol O—H stretch is red-shifted whilst the acid O—H is blue shifted relative to those of form I suggesting that the phenol participates more intimately in hydrogen bonding in form II than in form I. Additionally, there are changes in the ring stretch frequencies (1600 – 1200 $cm^{-1}$). Whilst the individual peaks have not been ascribed to specific stretching modes, the variations allude to a clear differentiation between the two polymorphs by virtue of changes in the local environments of the functional groups. Figure 5.8 illustrates the proposed hydrogen bonding

arrangement for form I, forming centrosymmetric acid dimer arrangements and form II – forming a possible infinite catemer acid to phenol O—H hydrogen bonded links

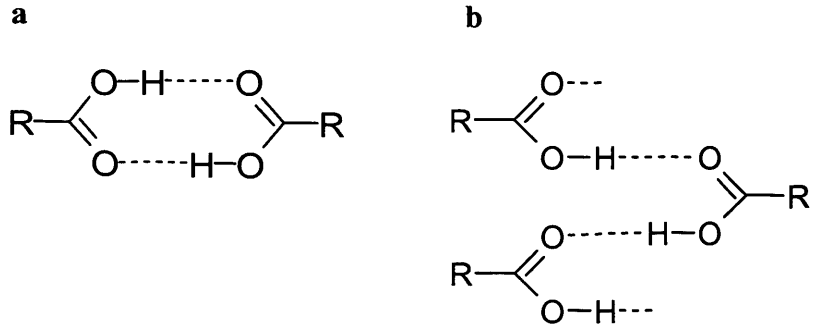



Figure 5.8: Proposed hydrogen-bonding motifs of AZD7140 (a) form I showing a centrosymmetric dimer and (b) form II showing a catemer arrangement

A structural search of the Cambridge Crystallographic Database using the phenol indole acid fragment showed no reported data. However an evaluation of organic carboxylic acids highlights the existence of both centrosymmetric dimers and hydrogen-bonded catemers, suggested by Leiserowitz (1976). For example, acetic acid is known to exist in the liquid and gaseous states as the centrosymmetric dimer but crystallises in a chiral space group forming a catemer arrangement (Boese *et al.* 1999). Structural assessment of a range of monocarboxylic acids reveal that approximately 95% of all reported structures have an acid dimer hydrogen bonded motif (Beyer and Price, 2000). However, when other functional groups with hydrogen bond donor/acceptor abilities are present (e.g. hydroxy benzoic acids), the occurrence of carboxylic acid dimers is somewhat reduced relative to the formation of other hydrogen bonding motifs, such as catemers. A study reported by Beyer and Price (2000) involving the assessment of lattice energy differences in structures of monocarboxylic acid structures exhibiting dimer and catemer arrangements showed that there is only a small energy difference between these two motifs. Furthermore, the presence of small substituents adjacent to the acid group can allow the formation of energetically quite similar catemer or dimer arrangements, whilst bulky groups hinder the ready formation of catemers. Overall, this study revealed that the prevalence of the dimer motif in crystal structures does not necessarily imply that it is energetically more favourable than other motifs, rather that its occurrence is often sterically driven. Furthermore, in addition to the acid hydrogen bonding motif itself, the crystal packing and other interactions will influence the energetics and hence stability of the crystal structure.

The high resolution XRPD data is shown in Figure 5.9. Helium pycnometry performed at 25°C, gave true density values of 1.63 and 1.64 g.cm$^{-3}$ for form I and form II respectively. This suggests that the crystal packing density of the two forms is essentially equivalent.

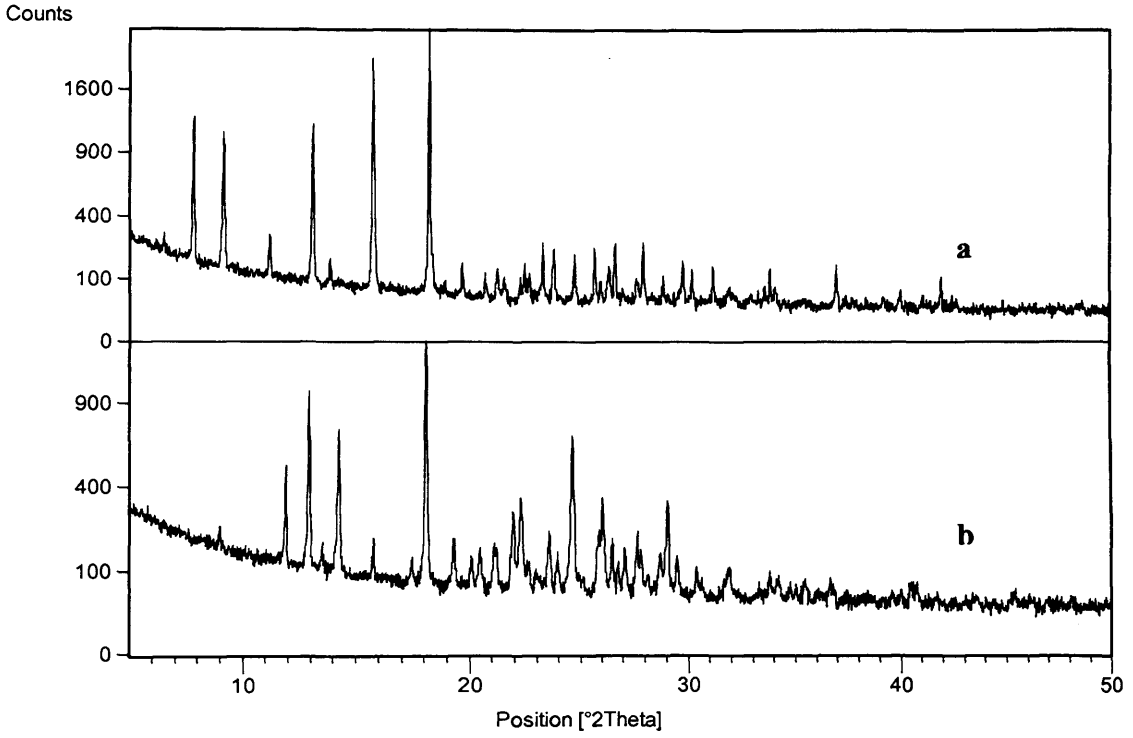



Figure 5.9: XRPD patterns of the polymorphs of AZD7140 (a) form I and (b) form II

The $^{13}C$ CP/MAS spectra of the two polymorphs were acquired in an attempt to corroborate the inference drawn from vibrational spectroscopy. Figure 5.10 illustrates the full $^{13}C$ CP/MAS spectra for both polymorphs of AZD7140. As highlighted by ATR spectroscopy, differences are seen in the phenol indole acid moiety. A variation in electron density is usually brought about by variations, for instance, in molecular conformation and/or hydrogen bonding arrangement. Thus the relative positions of the carbonyl resonance are suggestive of differences in hydrogen bonding. The changes associated with the phenol group established by ATR spectroscopy are reflected in the CP/MAS spectra. Furthermore, the spectra indicate that each carbon is represented by a single resonance signifying that for both polymorphs a single molecule is most likely to be present in the asymmetric unit. In addition, this observation would suggest that the proposed dimer arrangement in form I is consistent with a centrosymmetric structure.

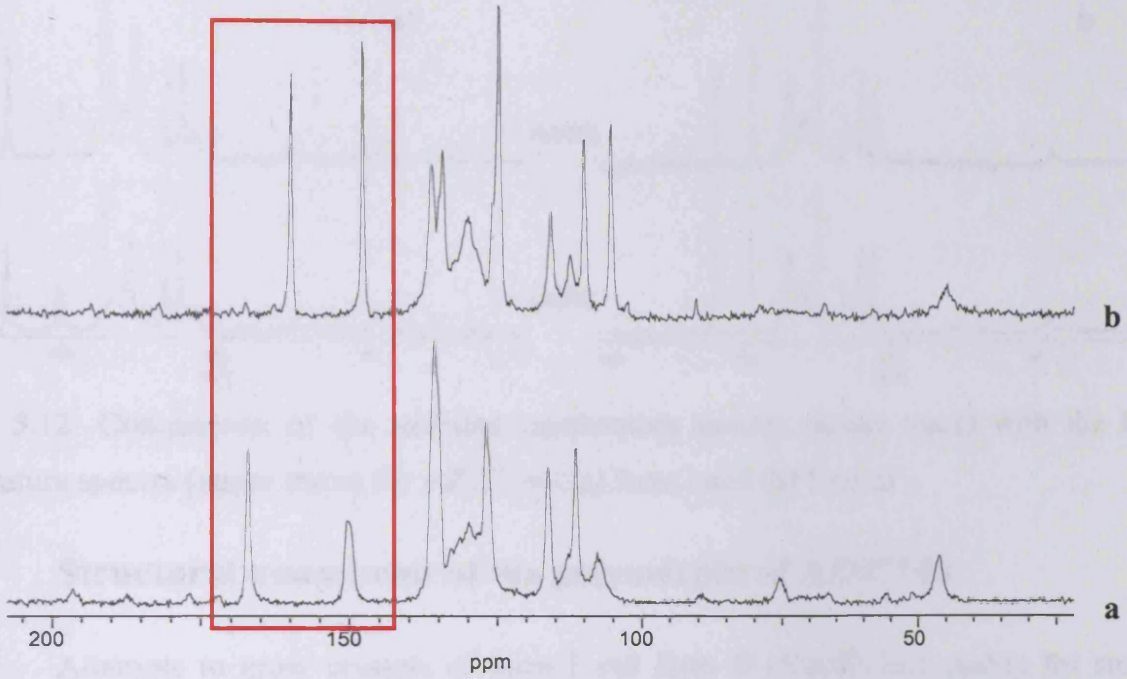



Figure 5.10: $^{13}C$ CP/MAS spectra for AZD7140 (a) form I and (b) form II with the changes in the carboxylic acid and phenol hydroxy resonances highlighted by the red box

For $^{13}C$ NMR analysis both the full spectra and the NQS data were obtained for samples at ambient temperature and -80°C, illustrated in Figures 5.11 and 5.12 respectively. The NQS data were acquired to aid spectral assignment and, along with the low temperature spectra, were also used to investigate the existence, if any, of dynamic disorder. The absence of any peaks associated with non-quaternary carbons in the NQS spectra, coupled to the similarity of both the ambient and low temperature spectra, suggest very little evidence of dynamic disorder.

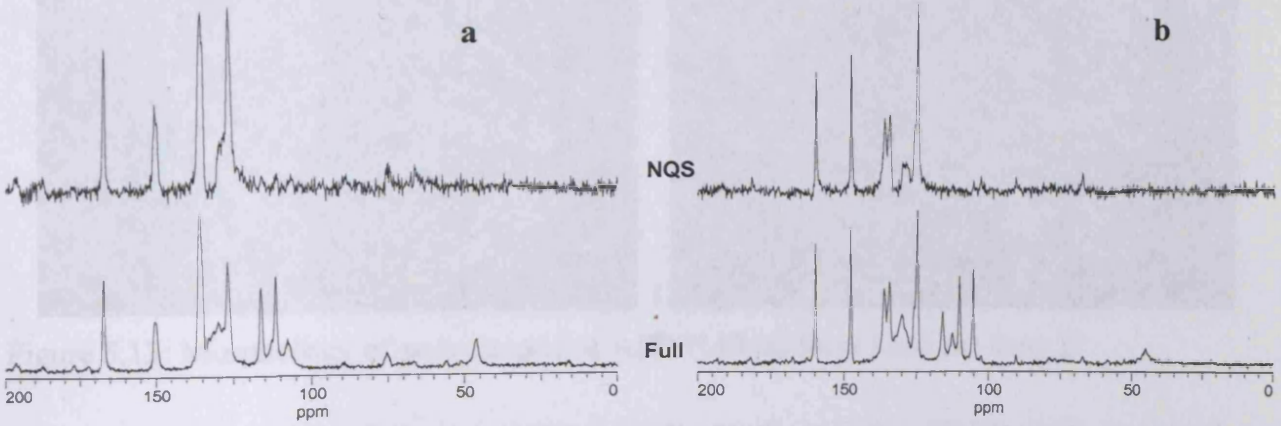



Figure 5.11: Comparison of the full spectra (lower trace) with the NQS spectra (upper trace) for AZD7140 (a) form I and (b) form II

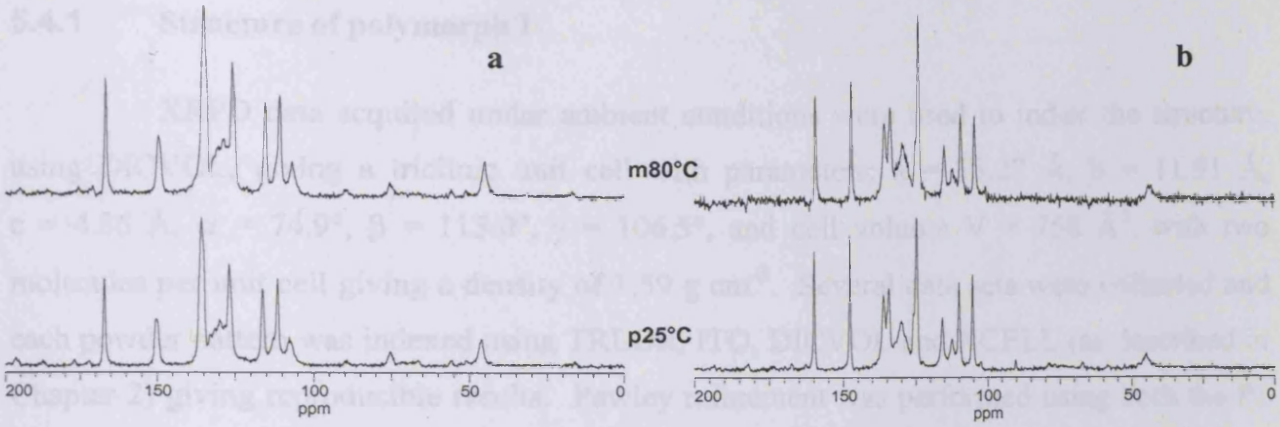



Figure 5.12: Comparison of the ambient temperature spectra (lower trace) with the low temperature spectra (upper trace) for AZD7140 (a) form I and (b) form II

## 5.4 Structural assessment of the polymorphs of AZD7140

Attempts to grow crystals of form I and form II of sufficient quality for single crystal X-ray diffraction studies always led to the formation of twinned and opaque crystals, or crystals of unsatisfactory proportions. Assessment of the crystal morphology indicates the difficulties in isolating suitable single crystals. Figure 5.13 shows the scanning electron micrographs of crystals of forms I and II obtained from attempts to grow suitable single crystals. In both cases very thin plates or needles were obtained with good proportions in two dimensions of the crystal but of insufficient thickness.

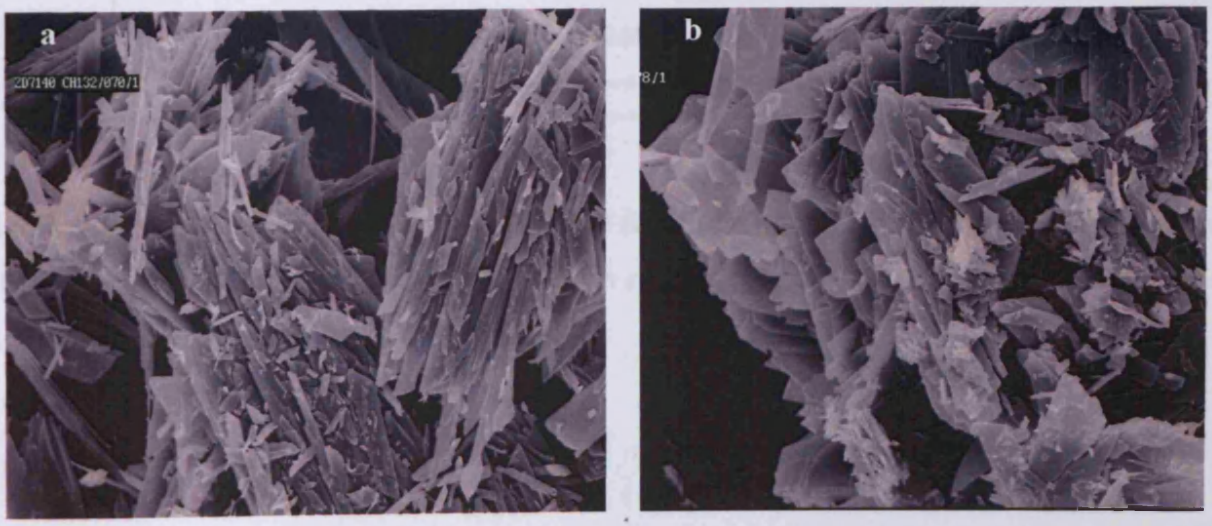



Figure 5.13: Morphology of polymorphs of AZD7140 (a) form I and (b) form II

As a result of unsuccessful crystal growth, powder XRPD data was used to determine the crystal structure of each polymorph. High-resolution XRPD data were collected at 298 K on both form I and form II and were used to solve their structures.

### 5.4.1 Structure of polymorph I

XRPD data acquired under ambient conditions were used to index the structure, using DICVOL, giving a triclinic unit cell with parameters; a = 15.27 Å, b = 11.91 Å, c = 4.86 Å, $\alpha$ = 74.9°, $\beta$ = 115.0°, $\gamma$ = 106.5°, and cell volume V = 758 Å$^3$, with two molecules per unit cell giving a density of 1.59 g cm$^{-3}$. Several data sets were collected and each powder pattern was indexed using TREOR, ITO, DICVOL and XCELL (as described in Chapter 2) giving reproducible results. Pawley refinement was performed using both the P1 and $P\bar{1}$ spacegroups, giving $R_{wp}$ = 1.81% in both cases. The Pawley fit profiles for both spacegroups are illustrated in Figures 5.14 and 5.15 respectively.

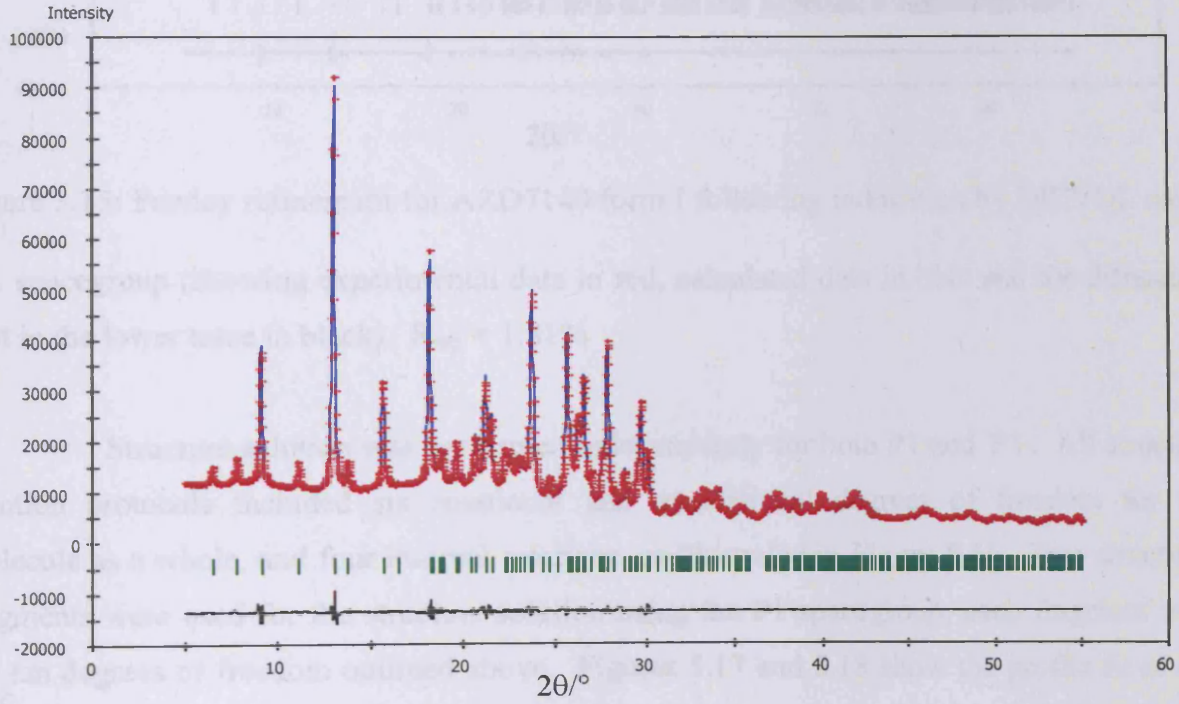



Figure 5.14: Pawley refinement for AZD7140 form I following indexation by DICVOL using P1 spacegroup (showing experimental data in red, calculated data in blue and the difference plot in the lower trace in black). $R_{wp}$ = 1.81%

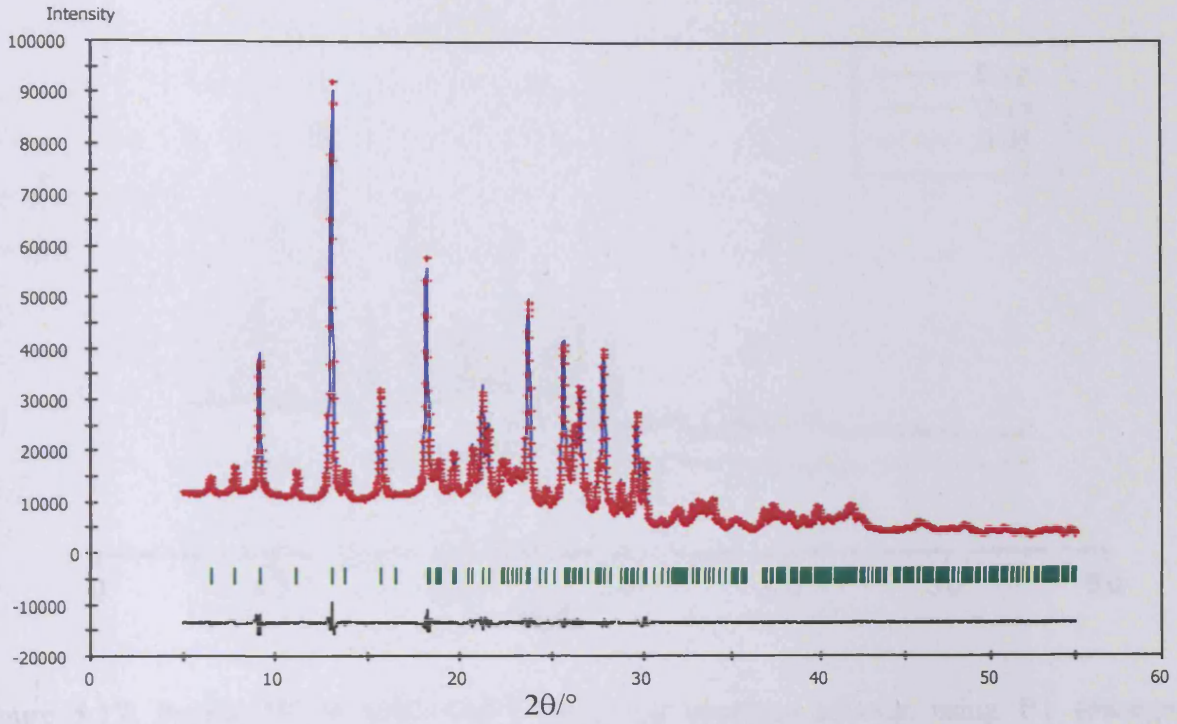



Figure 5.15: Pawley refinement for AZD7140 form I following indexation by DICVOL using $P\bar{1}$ spacegroup (showing experimental data in red, calculated data in blue and the difference plot in the lower trace in black). $R_{wp} = 1.81\%$

Structure solution was performed independently for both P1 and $P\bar{1}$. All structure solution protocols included six rotational and translational degrees of freedom for the molecule as a whole, and four internal rotations, as illustrated in Figure 5.16. Two structural fragments were used for the structure solution using the P1 spacegroup, each fragment with the ten degrees of freedom outlined above. Figures 5.17 and 5.18 show the profile fit of the structure solutions obtained using $P\bar{1}$ and P1 spacegroups respectively.

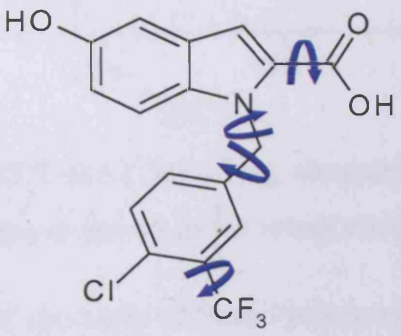



Figure 5.16: Molecular structure of AZD7140 showing, by means of the arrows, the internal rotations allowed during the structure solution process

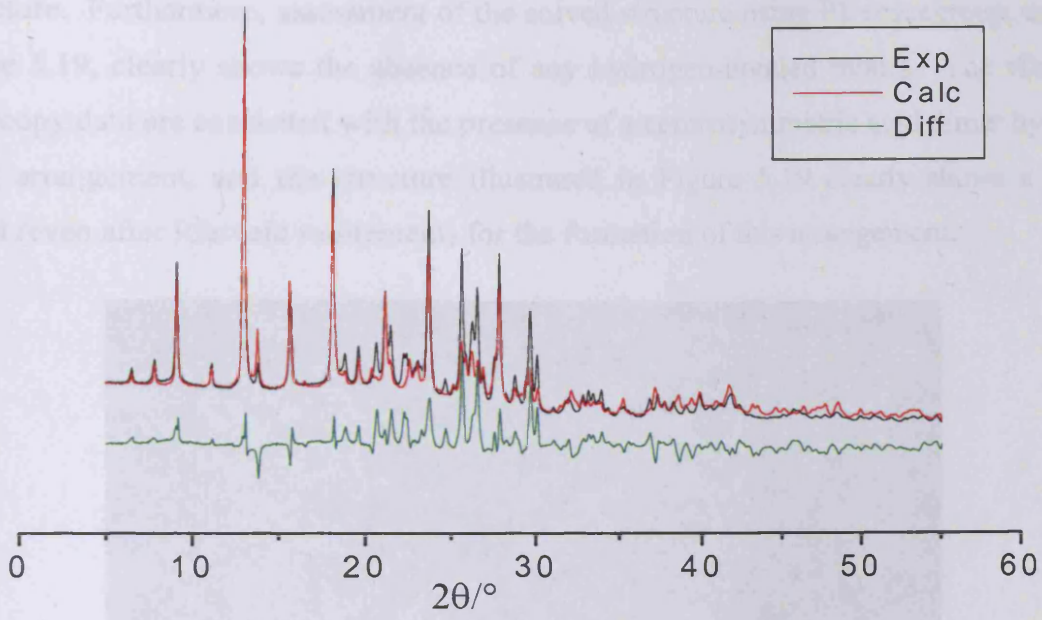



Figure 5.17: Profile fit of AZD7140 I following structure solution using $P\bar{1}$ spacegroup, $R_{wp}$ = 17.19% (the difference plot is shown in the lower trace)

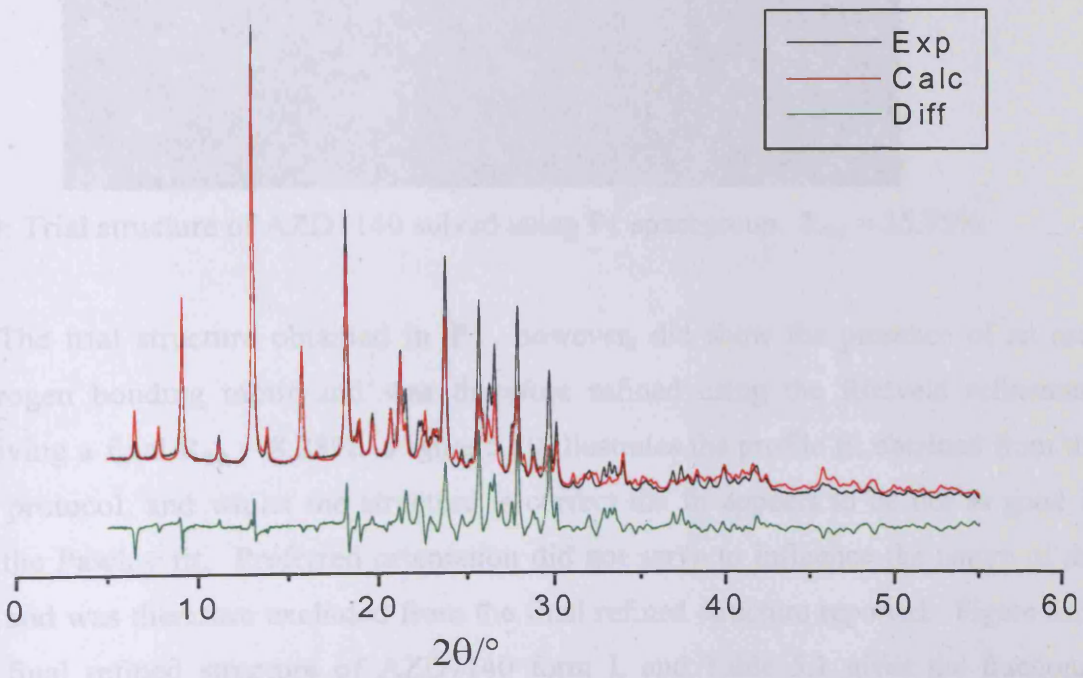



Figure 5.18: Profile fit of AZD7140 I following structure solution using P1 spacegroup, $R_{wp}$ = 15.75% (the difference plot is shown in the lower trace)

The solid-state NMR spectrum of form I indicates the presence of a single peak per carbon atom, usually consistent with the occurrence of one molecule per asymmetric unit, which would imply that the structure obtained using $P\bar{1}$ is a more correct representation of

the structure. Furthermore, assessment of the solved structure using P1 spacegroup, as shown in Figure 5.19, clearly shows the absence of any hydrogen-bonded motifs. The vibrational spectroscopy data are consistent with the presence of a centrosymmetric acid dimer hydrogen-bonding arrangement, and the structure illustrated in Figure 5.19 clearly shows a lack of potential (even after Rietveld refinement) for the formation of this arrangement.

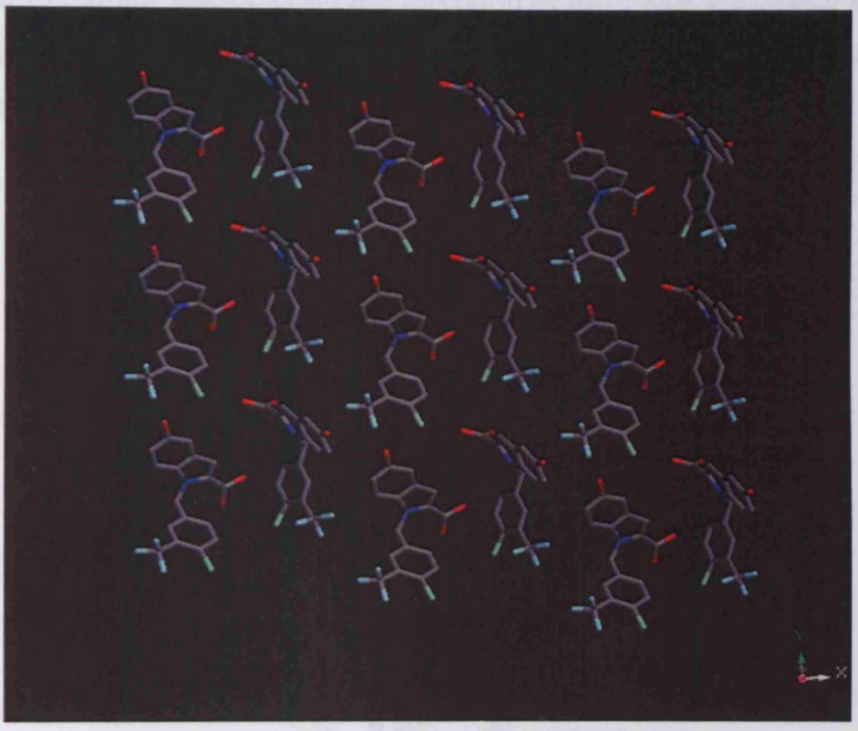


Figure 5.19: Trial structure of AZD7140 solved using P1 spacegroup. $R_{wp}$ = 15.75%

The trial structure obtained in $P\bar{1}$, however, did show the presence of an acid dimer hydrogen bonding motif and was therefore refined using the Rietveld refinement protocol, giving a final $R_{wp}$ = 8.28%. Figure 5.20 illustrates the profile fit obtained from the refinement protocol, and whilst the structure is correct the fit appears to be not as good in relation to the Pawley fit. Preferred orientation did not serve to influence the nature of the refinement and was therefore excluded from the final refined structure reported. Figure 5.21 shows the final refined structure of AZD7140 form I, and Table 5.1 gives the fractional coordinates for this refined structure.

The structure of form I shows the presence of a centrosymmetric acid dimer hydrogen bonded arrangement as predicted, located on a centre of inversion. The hydrogen bonded dimers form chains along the direction of the a-axis with each dimer related to the other by a unit cell translation along this axis. Furthermore, there is a unit cell translation of

this entire chain in the direction of the b-axis, giving rise parallel stack in the ab-plane. The parallel stacks interact with each other, in the direction of the b-axis, by means of a hydrogen bond between a fluorine atom in one stack and the phenol OH functionality in the adjacent stack. The molecules along the direction of the c-axis also form parallel stacks related to one another by inversion giving rise to the acid dimer hydrogen bonded arrangement. As seen along the a-axis, the dimer stacks along the direction of the c-axis are related to each other by a unit cell translation. The structural density of 1.613 g $cm^{-3}$ concords well with the true density value of 1.63 g $cm^{-3}$ obtained from helium pycnometry.

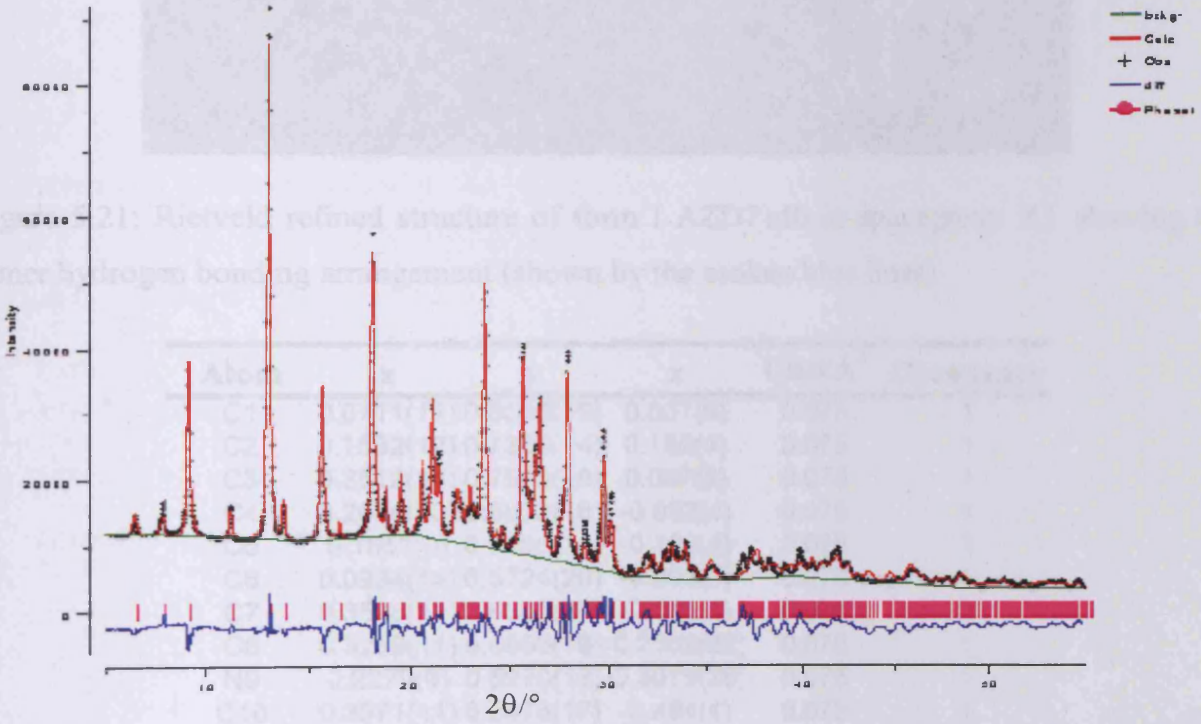



Figure 5.20: Profile fit with $R_{wp}$ = 8.28% ($R_p$ = 5.80%) following Rietveld refinement of form I of AZD7140 in spacegroup $P\bar{1}$ (showing the experimental data in black, the calculated data in red and the difference plot in blue in the lower trace)

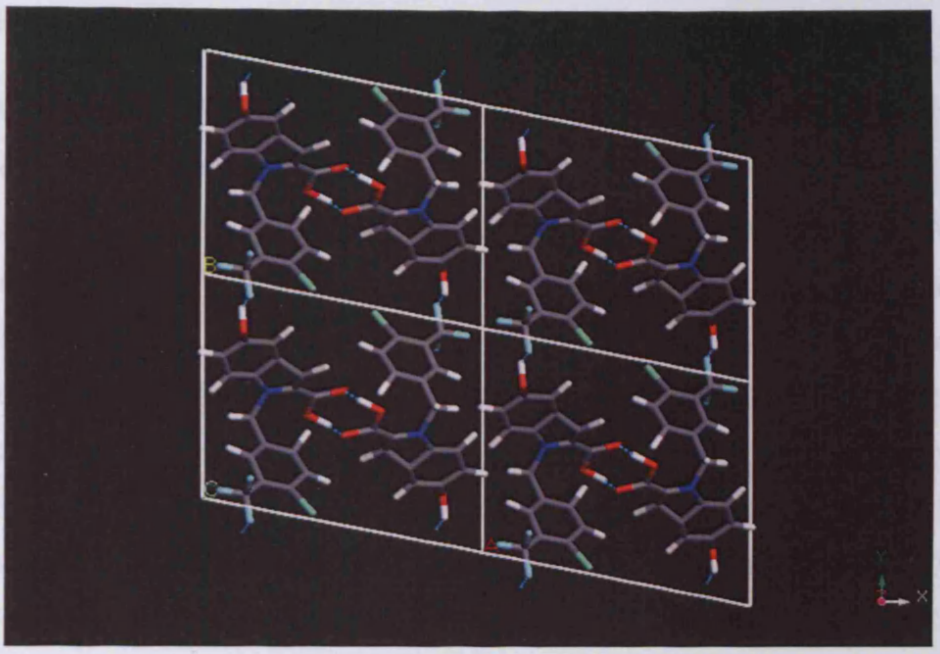


Figure 5.21: Rietveld refined structure of form I AZD7140 in spacegroup $P\bar{1}$ showing the dimer hydrogen bonding arrangement (shown by the broken blue lines)

| Atom | x | y | z | Uiso/Å$^2$ | Occupancy |
|---|---|---|---|---|---|
| C1 | 0.0711(11) | 0.6506(19) | 0.037(5) | 0.075 | 1 |
| C2 | 0.1532(12) | 0.7329(14) | 0.152(4) | 0.075 | 1 |
| C3 | 0.2512(14) | 0.7506(18) | 0.067(5) | 0.075 | 1 |
| C4 | 0.2656(10) | 0.6852(18) | -0.092(4) | 0.075 | 1 |
| C5 | 0.1851(9) | 0.5885(14) | -0.180(4) | 0.075 | 1 |
| C6 | 0.0934(14) | 0.5724(20) | -0.092(5) | 0.075 | 1 |
| C7 | 0.3538(12) | 0.6300(27) | -0.064(4) | 0.075 | 1 |
| C8 | 0.3259(11) | 0.5668(19) | -0.2602(32) | 0.075 | 1 |
| N9 | 0.2279(9) | 0.5270(13) | -0.3015(28) | 0.075 | 1 |
| C10 | 0.3971(11) | 0.5476(17) | -0.404(4) | 0.075 | 1 |
| O11 | 0.5021(21) | 0.6006(28) | -0.365(5) | 0.075 | 1 |
| O12 | 0.3679(14) | 0.4597(24) | -0.562(6) | 0.075 | 1 |
| C13 | 0.1724(12) | 0.3969(23) | -0.3459(32) | 0.075 | 1 |
| C14 | 0.2192(10) | 0.2866(16) | -0.104(4) | 0.075 | 1 |
| C15 | 0.3213(14) | 0.2799(18) | -0.094(6) | 0.075 | 1 |
| C16 | 0.3682(11) | 0.2000(23) | 0.143(5) | 0.075 | 1 |
| C17 | 0.3110(8) | 0.1205(12) | 0.333(4) | 0.075 | 1 |
| C18 | 0.2107(9) | 0.1222(12) | 0.3102(33) | 0.075 | 1 |
| C19 | 0.1714(10) | 0.2222(20) | 0.092(5) | 0.075 | 1 |
| Cl20 | 0.3901(9) | 0.0288(12) | 0.6721(24) | 0.075 | 1 |
| C21 | 0.1529(7) | 0.0552(10) | 0.5107(27) | 0.075 | 1 |
| F22 | 0.1689(12) | -0.0603(18) | 0.656(6) | 0.075 | 1 |
| F23 | 0.1695(14) | 0.1282(18) | 0.737(5) | 0.075 | 1 |
| F24 | 0.0603(15) | 0.0491(17) | 0.415(5) | 0.075 | 1 |
| O25 | 0.1468(16) | 0.7771(19) | 0.345(5) | 0.075 | 1 |

Table 5.1: Final refined atomic coordinates and isotropic displacement factors for AZD7140 form I $P\bar{1}$ structure (for non-hydrogen atoms only)

### 5.4.2 Structure of polymorph II

As for form I, high-resolution XRPD data were acquired under ambient conditions, and were used to index the structure, using XCELL, giving a monoclinic cell with parameters: a = 8.21 Å, b = 39.16 Å, c = 4.73 Å, β = 102.5, and cell volume V = 1484 $Å^3$, with four molecules per unit cell giving a density of approximately 1.65 g $cm^{-3}$, which is consistent with the true density value of 1.64 $cm^{-3}$ established from helium pycnometry. Furthermore, solid-state NMR data showed the presence of a single peak for each carbon atom in the molecular structure, usually consistent with there being a single molecule in the asymmetric unit, giving further confidence of the assigned cell with respect to density considerations. Several data sets were collected and each powder pattern was indexed using TREOR, ITO, DICVOL and XCELL (as described in Chapter 2) giving consistent and reproducible results. Consideration of systematic absences allowed the assignment of the spacegroup as $P2_1/c$, which was confirmed during the Pawley refinement procedure giving $R_{wp}$ = 9.23%. The Pawley fit profile is illustrated in Figure 5.22.

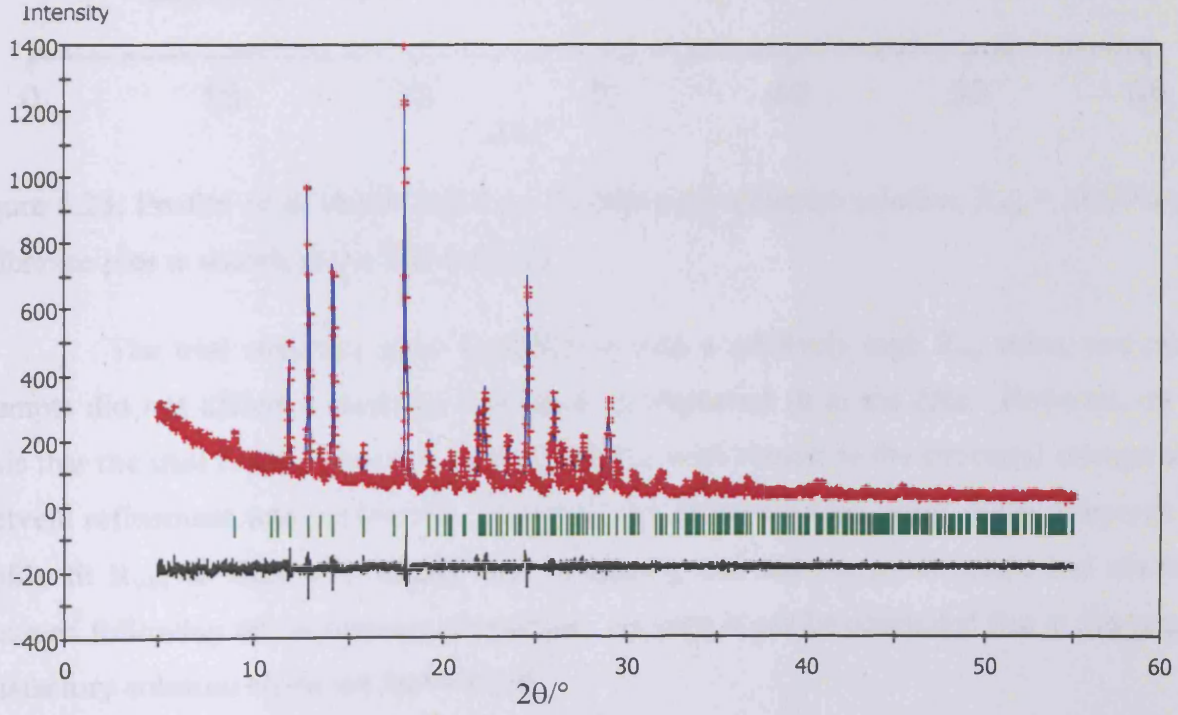



Figure 5.22: Pawley refinement for AZD7140 form II following indexing by XCELL using spacegroup $P2_1/c$ (showing experimental data in red, calculated data in blue and the difference plot in the lower trace in black). $R_{wp}$ = 9.23%

Structure solution was performed using six rotational and translational degrees of freedom for the molecule as a whole, with four internal rotations allowed, as illustrated

previously for form I (shown in Figure 5.16). Figure 5.23 shows the profile fit of the structure solution obtained for form II. Repeat structure solution calculations (six replicate determinations) were performed, starting from different random initial populations, giving rise to the same results in all cases.

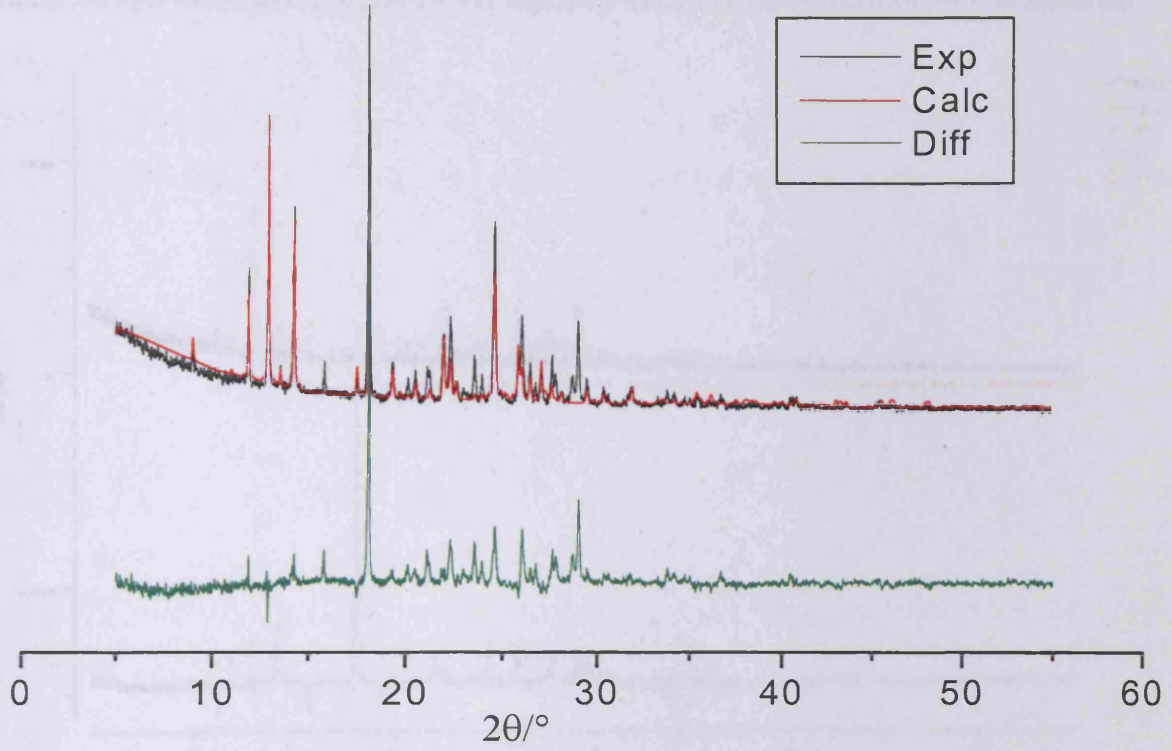



Figure 5.23: Profile fit of AZD7140 form II following structure solution, $R_{wp}$ = 28.20% (the difference plot is shown in the lower trace)

The trial structure gave a profile fit with a relatively high $R_{wp}$ value, and repeat attempts did not afford a structure that gave an improved fit to the data. However, on the basis that the trial crystal structure looked feasible with respect to the structural arrangement, Rietveld refinement was performed. However, the refinement procedure did not improve the profile fit $R_{wp}$, as shown in Figure 5.24, implying that the correct structure had not been achieved following the refinement procedure. As such it can be concluded that at this stage a satisfactory solution could not be obtained.

The profile fit obtained from the Rietveld refinement implicated the possibility of preferred orientation effects preventing the convergence to the correct structure. However, the use of preferred orientation refinement did not serve to influence the nature of the results. Additionally, attempts to minimise preferred orientation were made, by grinding the sample prior to XRPD data acquisition. The presence of disorder, either static or dynamic, was ruled

out from solid-state NMR data. Further investigations are required to establish the origin of the large peak at ~ 18° 2θ. At this stage, the unit cell and spacegroup are assigned with a high degree of confidence, but the unsatisfactory quality of the Rietveld refinements obtained so far are such that the structure itself must be considered as "undetermined". Future work will focus on further attempts to obtain a satisfactory structure determination of this material.

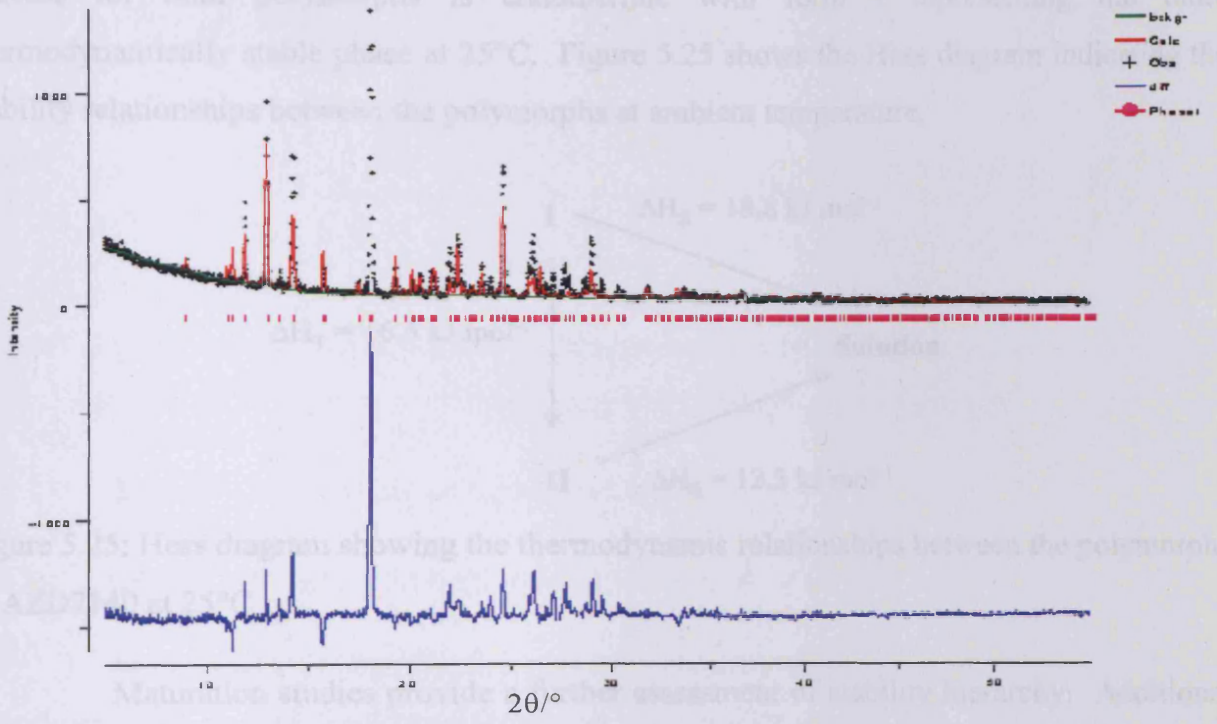



Figure 5.24: Profile fit with $R_{wp}$ = 28.39% ($R_p$ = 19.29%) following Rietveld refinement of AZD7140 II structure (showing the experimental data in black, the calculated data in red and the difference plot in blue in the lower trace)

## 5.5 Understanding the relative thermodynamic stability

### 5.5.1 Solution mediated assessments

Solution isothermal microcalorimetry provides a direct measure of the thermodynamics of dissolution. The cumulative enthalpy of solution encompasses a measure of the energy associated with wetting, breaking of intermolecular interactions in the crystal and solvation. Classification of thermodynamic stability by solution calorimetry relies upon the energetics of wetting and solvation across polymorphs to be constant, thereby providing a measure of lattice "strength" or energy. Whilst the component associated with heat of solvation may be correctly regarded as constant between polymorphs, the heat of wetting may

vary as a function of crystal habit and changes in surface characteristics, and thus could give rise to some variability. However, the enthalpy of dissolution ($\Delta_{sol}H$) was established from six replicate determinations at 25°C, leading to average values of $\Delta_{sol}H$ = 18.78 ± 3.37 kJ $mol^{-1}$ for form I, and 12.28 ± 1.65 kJ $mol^{-1}$ for form II, thus giving an enthalpy of transition between these two polymorphs of 6.5 kJ $mol^{-1}$. The data show that the dissolution process for both polymorphs is endothermic with form I representing the more thermodynamically stable phase at 25°C. Figure 5.25 shows the Hess diagram indicating the stability relationships between the polymorphs at ambient temperature.

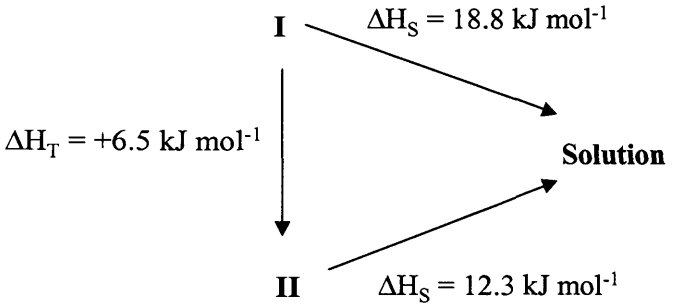



Figure 5.25: Hess diagram showing the thermodynamic relationships between the polymorphs of AZD7140 at 25°C

Maturation studies provide a further assessment of stability hierarchy. Additional confirmation for the stability order of the two polymorphs was shown experimentally by means of suspension mediated investigations (slurry experiments). The experiment is based upon solution mediated phase transformations of a 1:1 mixture of the two polymorphs with respect to time and solvent composition. At a given temperature, the stability order of the two polymorphs will be governed by the solubility order, such that the metastable form (at that temperature) will dissolve and recrystallise as the thermodynamically stable form. Experiments performed under ambient conditions, results from which are shown in Figure 5.26, revealed that form II readily converts to form I upon slurrying, confirming that I is the stable polymorph under ambient conditions. In addition, the conversion rate was facilitated by the presence of water in the 3:2 v/v methanol:water solvent mixtures, the solvent system from which the polymorphs were recrystallised.

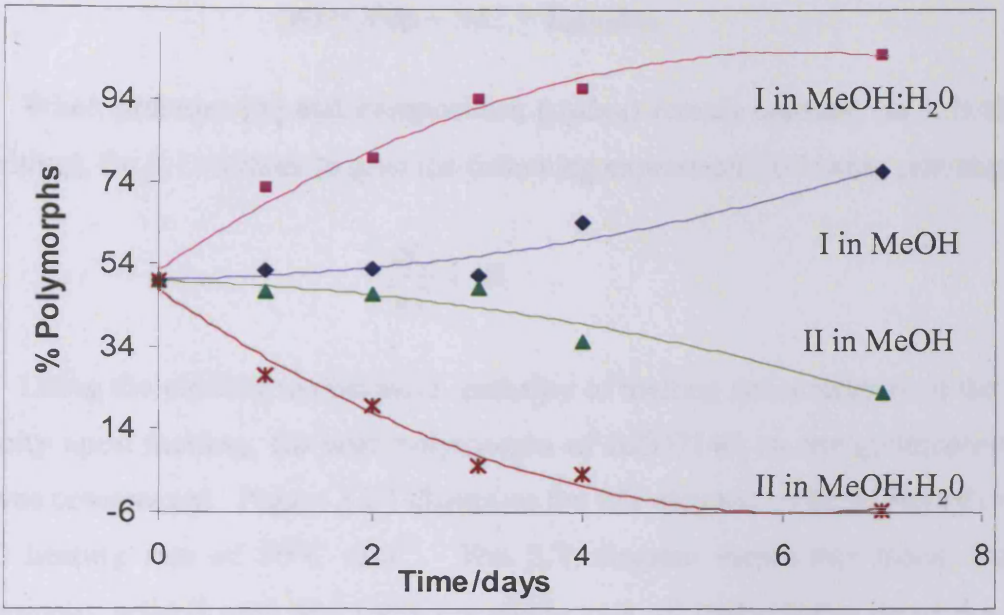



Figure 5.26: Plot showing the solution mediated conversion of a 1:1 mixture of form I and form II into form I, as a function of time, using two different solvent systems

### 5.5.2 Stability assignment using phase diagrams

#### 5.5.2.1 Stability assignment using an energy-temperature (E/T) phase diagram

Isothermal microcalorimetry was able to assign the stability relationship of the two polymorphs at a single temperature. However, the stability relationship of the polymorphs as a function of temperature needs to be established in order to assess whether pharmaceutical processing (such as storage or drying at elevated temperature, or milling) will cause an interconversion of one form to another. Further verification of the relative thermodynamic stability of the two forms was established from the extrapolation of the melting parameters as described by Yu (1995). The E/T diagram is composed of a series of both free energy (G) and enthalpy (H) isobars of all polymorphs positioned relative to one another and to those of the liquid phase. The relative positions are in accordance with thermodynamic stability (using the melting parameters) at the melting temperature and extrapolated to absolute zero. Construction of such curves relies on both statistical mechanics (based on ideality) and fundamental thermodynamics using the Gibbs function ($G = H - TS$). This procedure can only be applied to equilibrium first order phase transitions that do not involve a change in chemical composition (i.e. the melting event).

The Gibbs-Duhem equation (5.1) indicates the dependence of free energy of a system on pressure, temperature and composition.

$$dG = Vdp - SdT + \Sigma_B \mu_B dn_B \qquad \text{Eq 5.1}$$

When pressure (p) and composition ($\mu_B dn_B$) remain constant (as it is assumed to during melting), Eq 5.1 reduces to give the following expression (following rearrangement):

$$\frac{dG}{dT} = -S \qquad \text{Eq 5.2}$$

Using the melting temperature, enthalpy of melting and a measure of the change in heat capacity upon melting, for both polymorphs of AZD7140, an energy-temperature (E/T) diagram was constructed. Figure 5.27 illustrates the E/T diagram obtained from data acquired at a DSC heating rate of 50°C $min^{-1}$. The E/T diagram shows that forms I and II are enantiotropically related, and that form I is stable over all temperatures up until a transition point ($T_t$) is reached, estimated to be 158°C (431 K), beyond which form II becomes more stable.

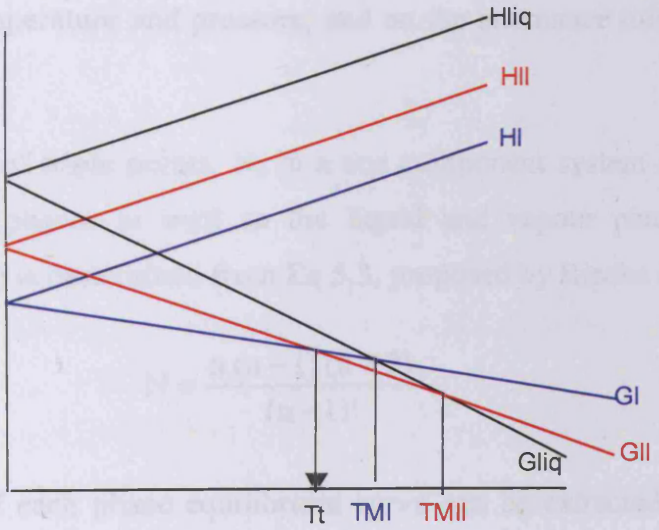



Figure 5.27: E/T diagram showing the stability relationships between the polymorphs of AZD7140 as a function of temperature (Tt represents the transition temperature, TMI and TMII represent the melting temperatures of form I and form II respectively, GI and GII represent the free energy curves for form I and form II respectively, Gliq is the free energy curve for the liquid phase and HI, HII and Hliq represent the enthalpy curves for form I, form II and the liquid phase respectively)

### 5.5.2.2 Stability assignments using a pressure-temperature (p/T) phase diagram

The p/T diagram is constructed using similar principles used for constructing the E/T diagram, i.e. it is based on the Gibbs-Duhem expression shown in Equation 5.1. However, in the case of the p/T diagram, both temperature and pressure are considered as variables, whilst the composition remains constant.

The free energy for the three states of matter (i.e. solid, liquid and vapour) is a function of pressure and temperature and can be represented by continuous surfaces in three-dimensional G-p-T space. These continuous surfaces are non-parallel and the slopes are described by the function dp/dT. The intersections of these surfaces occur at points where the three states of matter are in equilibrium, and are termed the triple points. The p/T diagrams are two-dimensional representations of the three-dimensional assessment of stability assignment, and the continuous surfaces are depicted as phase equilibrium curves that cross at the triple points. The stability hierarchy is assigned on the basis of Ostwald's criteria of positions relative to temperature and pressure, and on the alternance rule (Ceolin *et al.* 1992, Espeau *et al.* 2005).

The number of triple points, N, in a one component system of n-phases (where n encompasses all solid phases as well as the liquid and vapour phases, such that for a dimorphic system n = 4) is determined from Eq 5.3, proposed by Riecke (1890).

$$N = \frac{n\,(n-1)\,(n-2)}{(n-1)!} \qquad \text{Eq 5.3}$$

The slope of each phase equilibrium curve can be extracted from the Clapeyron equation, in the form shown in Eq 5.4, using the enthalpy and volume changes at temperature T where pressure $p \approx 0$. In Eq 5.4, $\Delta H$ represents the enthalpy of change of a given process, for example melting or sublimation, $\Delta V$ denotes the corresponding change in volume.

$$\frac{dp}{dT} = \frac{\Delta H}{T\Delta V} \qquad \text{Eq 5.4}$$

The vapour pressures of small organic materials, including polymorphs of pharmaceuticals are usually sufficiently small so as to allow the melting parameters from DSC experiments to provide a reasonable approximation to the solid-liquid-vapour triple point for each polymorph. Furthermore, the stability assignments and subsequent application

of the alternance rule is based on the order of the melting triple points (the point at which the solid, liquid and vapour states are in equilibrium as discussed in Chapters 1 and 2), i.e. the higher the melting triple point the more stable the phase (under those conditions). Additionally, the values of the exact vapour pressure are not necessary for the construction of the sublimation curves. Rather, the relative positions of the sublimation curves can be extrapolated and are sufficient to construct a topological p/T diagram suitable for assigning stability hierarchy.

The thermodynamic data obtained from the DSC measurements, shown in Table 5.2 were used to locate the solid-liquid-vapour triple points.

| **Form** | **ΔH/kJ mol$^{-1}$ dissolution at 298K** | **ΔH/kJ mol$^{-1}$ melting** |
|---|---|---|
| I | 18.9 | 53.1 at 500 K |
| II | 12.3 | 51.0 at 504 K |

Table 5.2: Summary of thermodynamic parameters obtained form solution calorimetry (dissolution) and DSC (melting) on AZD7140 polymorphs I and II (DSC melting parameters obtained using a heating rate of 50 K min$^{-1}$)

The liquid volume was determined at an intermediate temperature to the two melting points (at $T_{Mmean}$, representing the mean melting temperature, of 502 K) from experimental evaluation, performed by Prof. Ceolin as described in the literature (Espeau *et al.* 2005), of the density of the molten phase. The specific volume of the liquid phase was determined to be $V_{liq}$ = 0.7552 cm$^3$ g$^{-1}$ at 502 K. The solid specific volumes for each polymorph were obtained by extrapolation of a linear expression determined from a plot of the specific volume against temperature. These values were extracted from unit cell dilation experiments estimated from the indexing of XRPD data acquired as a function of temperature for both polymorphs. The specific volume of form I at 502 K was estimated to be $V_I$ = 0.63336 cm$^3$ g$^{-1}$, whilst that for form II was estimated to be $V_{II}$ = 0.61898 cm$^3$ g$^{-1}$. Hess diagrams representing the enthalpy and volume relationships between the two polymorphs were constructed as shown in Figure 5.28 and 5.29 respectively.

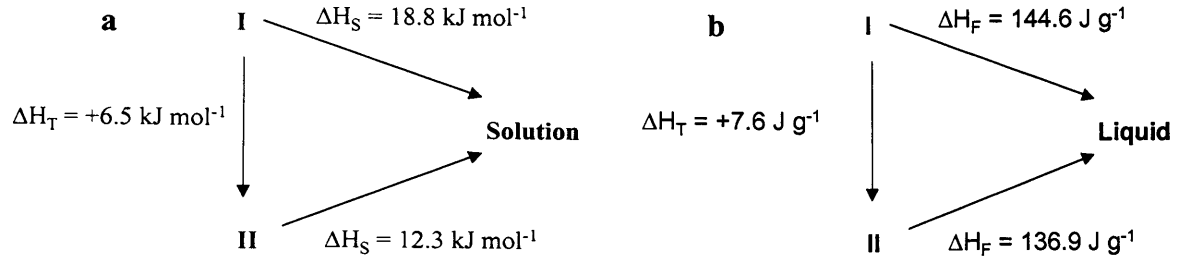



Figure 5.28: Diagrams showing the relationships between (a) the enthalpy of solution and (b) the enthalpy of melting of the polymorphs of AZD7140

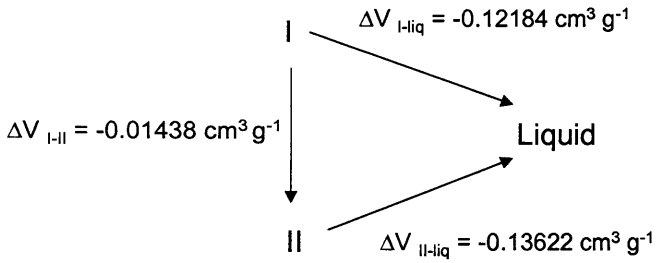



Figure 5.29: Diagram showing the specific volume relationships on going from the solid to liquid states for each polymorph of AZD7140

A plot of enthalpy (ΔH) values from DSC and solution calorimetry data against temperature gave an extrapolated linear expression of ΔH over temperature, allowing the enthalpy of melting to be estimated at 502 K. Taking into consideration the enthalpy and volume changes between each polymorph, the slope values for the various phase equilibrium curves can be estimated using the Clapeyron equation (Eq 5.4).

The slope for the I-II (solid to solid phase equilibrium) curve was estimated as shown in Eq 5.5

$$\text{Slope for I-II} = +7.62/(-0.01438 \times 502) = -1.055 \text{ MPa K}^{-1} \qquad \text{Eq 5.5}$$

The slope for the I-liquid (solid to liquid phase equilibrium) curve was estimated as shown in Eq 5.6

$$\text{Slope for I-liq} = +144.55/(0.12184 \times 502) = +2.363 \text{ MPa K}^{-1} \qquad \text{Eq 5.6}$$

The slope for the II-liquid (solid to liquid phase equilibrium) curve was estimated as shown in Eq 5.7

$$\text{Slope for II-liq} = +136.93/(0.13622 \times 502) = +2.002 \text{ MPa K}^{-1} \qquad \text{Eq 5.7}$$

Furthermore, the equations for the various curves for the p/T diagram were estimated using the expression shown in Eq 5.8, where m represents the slope discussed above.

$$p = m\,T + A \qquad \text{Eq 5.8}$$

On the basis of this expression, the equations for the I-liq and II-liq curves can be shown by ($p = 2.363\ T + A$) and ($p = 2.002\ T + A'$) respectively. It is assumed that p is negligible over the temperature range T = 0 – 500K (i.e. $p_{vap} \sim 0$). Therefore at T = 500 K and 504 K, for form I and form II respectively, p = 0. So the values of the constants A and A' can be calculated to be; A = -1181.5 MPa and A' = -1009.2 MPa. Thus, at the I-II-liq triple point the pressure of the system (p) is equal and therefore the temperature coordinate can be calculated as 2.363 x (T/K) – 1181.5 = 2.002 x (T/K) – 1009.21, T = 477.8 K. From this, the pressure coordinate is calculated from p/MPa = 2.363 x 477.8 – 1181.5, and thus p = -52.5 MPa (i.e. a virtual triple point).

The curve I-II crosses through the triple point I-II-liq, and from the Clapeyron equation has a negative slope (-1.055 $MPa.K^{-1}$). The equation for the I-II curve is represented by $p = -1.055\ T + A''$. At the triple point for this curve (I-II-liq) A'' can be calculated as follows; -52.5 = -1.055 x 477.8 + A''/MPa, and A'' = 451.58 MPa. Therefore the triple point I-II-v can be calculated as the I-II curve also crosses this point. If p = 0 at this point, then T can be calculated as 0 = -1.055 x (T/K) + 451.58, and thus T = 428K, which corresponds to the transition temperature. Tables 5.4 and 5.5 summarise the entire key data required to construct the p/T diagram.

| **Triple point** | **T/K** | **P/MPa** |
|---|---|---|
| I-liq-v | 500 | ~ 0 |
| II-liq-v | 504 | ~ 0 |
| I-II-liq | 478 | - 53 |
| I-II-v | 428 | 0 |

Table 5.3: Summary of the p and T coordinates of all the main triple points

| Curve | Slope |
|---|---|
| I-II | negative |
| I-liq | positive |
| II-liq | positive |

Table 5.4: Summary of the slopes of the phase equilibrium curves

The diagram can be constructed according to Ostwald's rules. The polymorph with the triple point at the highest temperature on the liquid-vapour curve is assigned as the most stable, and that at all other temperatures the most stable phase has the lowest vapour pressure. The p/T diagram is shown in Figure 5.30, and indicates that the melting curves diverge as the pressure increases, and that the two forms are enantiotropically related. Furthermore, the I-II-v triple point (or the transition temperature) occurs at 155°C. Form I is the most stable phase below this temperature, whilst form II is stable above it. These data concord well with the results obtained form the E/T diagram and the evaluation by solution mediated phase transition (maturation) studies.

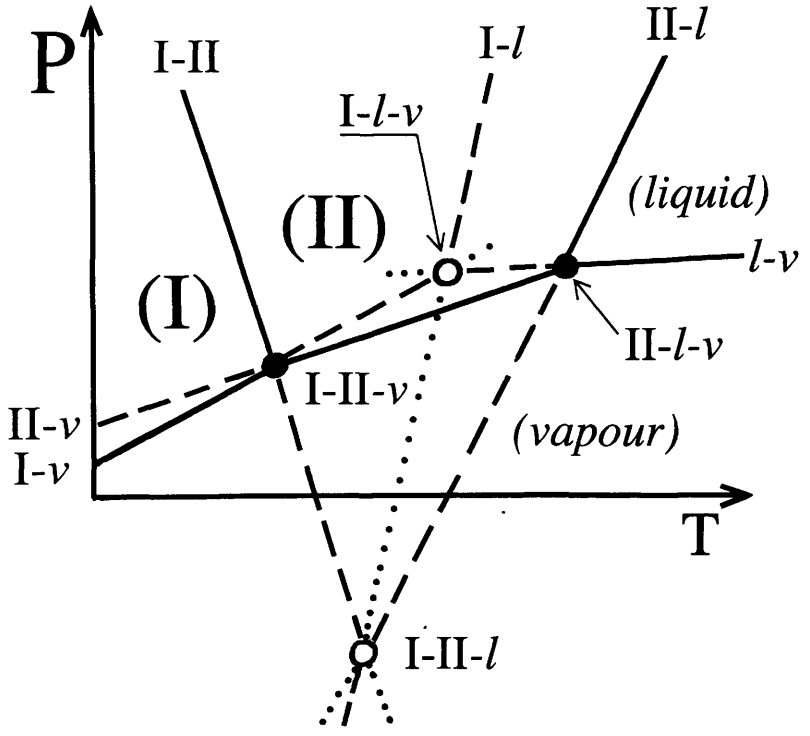



Figure 5.30: The topological p/T diagram for the polymorphs of AZD7140 showing the stable curve (—), metastable curves (----) and super-metastable curves (·····)

## 5.6 Discussion and conclusions

AZD7140 can be isolated in one of two polymorphic forms, along with a series of solvates. The two polymorphs, form I and form II, have melting parameters that are quite similar, and the higher melting form II recrystallises (under kinetic conditions) at high temperature from the melt of form I. The small difference in the melting temperature, around 4 K, may suggest that the lattice energies of the structures are largely similar with perhaps comparable hydrogen bonding motifs but small variations in conformational properties or packing arrangement. However, spectroscopic assessment of both polymorphs revealed that substantially different hydrogen bonding motifs exist between form I and form II. Form I was suggested to have a centrosymmetric carboxylic acid dimer motif, whereas form II was suggested to adopt a non-dimer hydrogen bonded arrangement, in which the phenol hydroxy group participates in hydrogen bonds more intimately than the acid functionality. That the melting temperatures are largely similar suggests that the different hydrogen bonding (and molecular packing) arrangements are energetically quite similar.

Solving the crystal structure of form I from XRPD data reveals that the structure does indeed contain the centrosymmetric acid dimer arrangement. Each molecule in the unit cell participates in three intermolecular hydrogen bonds, two forming part of the dimer sextet arrangement with each of the O—H ····· O=C hydrogen bond lengths being 1.36 Å, and the phenol O—H ····· F—C hydrogen bond length being 2.26 Å. Clearly the crystal structure and hydrogen bond arrangements for form I are in good agreement with the conclusions drawn from the vibrational spectroscopic assessment. Further investigations are required to obtain a satisfactory determination of the crystal structure of form II.

All thermodynamic stability assessments show that forms I and II are enantiotropically related, and that form I is stable below the transition temperature of 155 – 158°C, above which it becomes metastable with respect to form II, in good agreement with the observations from both thermal (DSC) and isothermal calorimetry (solution microcalorimetry) that indicate an endothermic transition on going from forms I to II. Calorimetric assessment gave an enthalpy of transition ($\Delta H_T$) on going from form I to form II of approximately 7 kJ $mol^{-1}$. Thus, whilst form I is thermodynamically more stable, the energetic difference between the two phases is relatively small, concurring well with the

structural interpretation of the relative stability of the two polymorphs. The enthalpy of solution is slightly more endothermic for form I than for form II.

Construction of both the E/T and p/T diagrams has allowed a robust assessment of the likely impact of temperature and pressure on the stability of the polymorphs. From both assessments, a transition temperature of 155 – 158°C was estimated. Form I was shown to have a stability domain below the transition temperature, and from the p/T diagram it is suggested that an increase in pressure might result in the conversion of form I into form II. This alerts the pharmaceutical formulator of solid oral dosage forms that secondary processing such as milling, micronisation or compaction of the tablet may need to be approached with caution. However, the present study conclusively shows that form I is the ideal polymorph to progress through development as it represents the thermodynamically stable form under ambient conditions, but form II may also be developable as it is energetically quite similar, and may thus be expected to show bioequivalence.

## 5.7 Experimental

### 5.7.1 Thermal analysis

Differential scanning calorimetry (DSC) experiments were carried out using a Perkin-Elmer DSC-7. Samples were analysed under a nitrogen atmosphere in 50 μl vented aluminium pans. Hot-stage microscopy (HSM) experiments were performed on a Zeiss Axioskope microscope fitted with a Linkham hot stage unit and controlled by Linksys software. Samples were heated from ambient temperature to 200°C at a rate of 10 °C $min^{-1}$ and 50 °C $min^{-1}$ for the phase diagram construction.

### 5.7.2 XRPD Analysis

High-resolution XRPD data were collected on a Philips Panalytical Xpert1 diffractometer in reflection mode (using Bragg-Brentano geometry). Typically, data were collected over the range 5 - 50° 2θ using a step size of 0.02 to 0.03° 2 θ and a counting time of 10 - 100 seconds per step.

### 5.7.3 Spectroscopy

#### 5.7.3.1 Vibrational spectroscopy

FTIR spectroscopy was performed on pure samples using a Perkin-Elmer Spectrum-GX spectrometer fitted with a single bounce SpecAc variable temperature ATR attachment. Samples were analysed neat and spectra were acquired over the range 4000 – 600 $cm^{-1}$ using a spectral resolution of 4 $cm^{-1}$ with 64 accumulations. Raman spectroscopy was performed using a Perkin-Elmer NIR FT-Raman spectrometer fitted with a NdYAG (1064 nm) laser. Data were collected over the range 4000 – 100 $cm^{-1}$ using 4 $cm^{-1}$ resolution and 64 accumulations. The samples were spun in a direction perpendicular to that of the incident radiation to minimise preferred orientation and laser heating effects. Data were collected using isothermal and non-isothermal conditions using a Ventacon non-ambient sample unit.

#### 5.7.3.2 $^{13}C$ CP/MAS experiments

High-resolution solid-state $^{13}C$ NMR spectra were recorded at 75.5 MHz using a Chemagnetics Infinity 300 spectrometer and a Chemagnetics triple resonance probe. A standard CP and $^{1}H$ decoupling sequence was used (as described in Chapter 2) and samples were packed in 4 mm zirconia rotors and spun at typical speeds of 5 kHz. NQS spectra were also acquired using the pulse-sequence described in Chapter 2.

### 5.7.4 Density measurements

#### 5.7.4.1 Density in the solid state

The true density, under laboratory conditions, was measured by helium pycnometry. A Micromeritics Accupyc 1330 helium pycnometer was used and the density measured on samples loaded into 1 $cm^3$ sample cups. The headspace volume of the cup containing the sample relative to the calibration standard was used to determine density of the material.

The change in solid-state density as a function of temperature (assessment of temperature induced dilation) was measured by indexing the XRPD patterns collected over the temperature range 25 - 100 °C. The variable temperature XRPD data were refined using Pawley refinement procedure and the corresponding specific volume was plotted against temperature.

#### 5.7.4.2 Density in the liquid state

Professor Rene Ceolin performed liquid density measurements in which 2.2115 g of the powder was introduced in a cylindrical silica tube (inner diameter = 7.797 ± 0.080 mm) and maintained overnight in a dynamic vacuum (approximately equating to $10^{-5}$ to $10^{-6}$ mmHg) in order to remove any adsorbed gases. Then the tube was sealed under vacuum. In order to circumvent any possible decomposition of the product, the furnace was heated to a temperature of 229.3 °C and held isothermally at this temperature during which melting occurred quickly. At this point the volume measurement was performed, which gave a specific volume of 0.7552 $cm^3\ g^{-1}$ (i.e. a density of 1.324 g $cm^{-3}$ for the liquid).

### 5.7.5 Scanning electron microscopy

Samples were deposited on sample holders and coated with a thin layer of gold and platinum to minimise sample charging during analysis. The prepared samples were then analysed by a LEO430 scanning electron microscope.

### 5.7.6 Structure solution from X-ray powder diffraction data

Materials Studio Reflex module (version 1.2) from Accelerys was used for indexing and Pawley refinement. ITO (Visser 1969), TREOR (Werner *et al.* 1985), DICVOL (Boultif and Louer 1991) or X-cell (Neumann 2003) were used simultaneously to index the powder patterns of both polymorphs. The best-fit cell was refined using the Pawley fit procedure, which incorporated refinement of the profile and background parameters (as described in Chapter 2). The measure of similarity used between the simulated and experimental profiles was $R_{wp}$. Structure solution was performed using the GA program EAGER, using the multi-population parallel genetic algorithm technique. Typically, the structure solution calculations involved 60 to 100 starting populations with 50 to 100 generations. The mating rate (i.e. the number of mating events per generation) was set at 15, and the mutation rate was typically set at 8 per generation.

Structure refinement was performed using the GSAS software as described in Chapter 2. Following an initial Le Bail refinement, the positions of the atoms were refined using the Rietveld protocol with standard geometrical restraints applied to bond lengths, bond angles and planar groups. For all non-hydrogen atoms, the isotropic displacement factors were constrained to a single value. Hydrogen atoms were incorporated into the structural model in the final stages of refinement but were not refined.

## 5.8 References


- Beyer T., Price S. L., (2000), *J. Phys. Chem. B*, **104**, 2647
- Boese R., Blaser D., Latz R., Baeumen A., (1999), *Acta Cryst.*, **C55**, IUC9900001
- Ceolin R., Toscani S., Agafonov V., Dugue J., (1992), *J. Solid State Chem.*, **98**, 366
- Espeau P., Ceolin R., Tamarit J-L., Perrin M-A., Leveiller F., (2005), *J. Pharm. Sci.*, **94**, 524
- Giron D., (1995), *Thermochim. Acta*, **248**, 1
- Leiserowitz L., (1976), *Acta Cryst.*, **B32**, 775
- Lin-Vien D., Colthup N. B., Fately W. G., Grasselli J. G., (1991), *The Handbook of Infrared and Raman Characteristics Frequencies of Organic Molecules*, Academic Press
- Riecke E., (1890), *Z. Phys. Chem.*, **6**, 411
- Yu L., (1995), *J. Pharm. Sci.*, **84**, 966

# 6 Models for structure solution of disordered systems

## 6.1 Introduction

### 6.1.1 Establishing a strategy to solve disordered structures

Structural disorder can occur whenever a crystal structure allows a molecule to adopt two, or more, different orientations (that are energetically similar). This type of disorder can often occur if the molecular shape is approximately symmetrical (Wright 1995). Furthermore, for non-centrosymmetric molecular structures to crystallise in spacegroups that require the molecules to occupy centrosymmetric sites, different molecular orientations should exist at each of these sites. This then fulfils the symmetry requirements in the average structure.

As discussed previously in Chapter 1, disorder can be either static or dynamic. The use of diffraction data does not readily allow a distinction to be made between the two forms of disorder, as the data generated is a time and spatial averaged representation of the structure. Analysis of single crystal x-ray diffraction data can be used to evaluate the existence and nature of disorder without any prior knowledge. This is based on variable temperature investigations and the assessment of anisotropic factors (as described in Chapters 1 and 2). Solving structures from XRPD data is less straightforward. For instance, a situation in which the powder data allows indexation and refinement of the cell, but does not allow direct-space structure solution because of the existence of disorder can be envisaged. This highlights the limitations of applying the current standard direct-space structure solution protocol to disordered systems. Furthermore, disorder may, in some instances, serve to exacerbate the limitations of peak overlap and broadening in the powder data.

The limitations of using direct space approaches (in the current format) to solve disordered structures have been reported recently (Tremayne *et al.* 2004). In this investigation, direct space methods were shown to be problematical in solving the structure of a seemingly tractable case of a substituted benzene molecule. A disorder model was used in which all the substituents occupied the various substitution patterns with varying occupancy. The structure was eventually solved using an exhaustive approach employed within the grid-search technique. Furthermore, structures were also solved by refinement based on hypothetical structures derived during polymorph prediction. Whilst this protocol represents

a good way forward in solving disordered structures, employing a strategy using direct space methods appears to be more tractable, more readily applicable, and less convoluted. The most significant limitation in relying on polymorph prediction to generate trial structures is the lack of universal applicability of the force fields employed to predict polymorphs, as well as the fact that such calculations are usually restricted to consider only a subset of all possible space groups. As such, a systematic approach using the direct space approach is seen as advantageous. Such an approach should ideally encompass the input of structural models that are time and space averaged representations of the possible disorder situation anticipated or understood to be present, based on spectroscopic assessment of the system.

### 6.1.2 1,3-dibromo-5-ethynyl benzene as a model compound

In this work, 1,3-dibromo-5-ethynyl benzene (DBEB) was used as a model compound. The crystal structure of this compound had been established from single crystal x-ray diffraction, and was shown to have orientational disorder of the molecule (Robinson *et al.* 1998). DBEB packs in an orthorhombic cell with spacegroup $P2_12_12_1$ and cell parameters a = 13.853 Å, b = 14.744 Å, c = 4.036 Å. The disorder is represented by the DBEB molecule adopting three orientations, such that the three non-hydrogen substituents were distributed over the three sites in the crystal. The carbon atoms of the benzene rings in each of the disorder components were indistinguishable. The relative occupancies of the three molecular orientations, in the reported crystal structure, are 65, 25 and 10%. This structure was selected as it represents a simple rigid molecule that exhibits a relatively non-complex disordered profile. Furthermore, it allows the possibility of different orientational disorder profiles to be simulated and studied.

## 6.2 Simulation of orientational disorder and defining strategies for solving the disordered structures

Models were constructed to mimic different scenarios of orientational disorder that may be exhibited by a structure such as DBEB, with the aim of "simulating" the procedure for tackling direct-space structure solution that may be followed if this example were being examined as an unknown structure. Seven models were generated to represent seven disordered crystal structures. The relative occupancies within each of the models are summarised in Table 6.1. Model 7 represents the actual disorder situation in the reported

crystal structure. The powder patterns of each of these disordered structures were then extracted using Material Studio Reflex package (2005), and are illustrated in Figure 6.1.

| Model | Name | Occupancies of the ethynyl substituent | | |
|---|---|---|---|---|
| | | Site 1 | Site 3 | Site 5 |
| 1 | Pm1 | 1 | 0 | 0 |
| 2 | Pm2 | 1/3 | 1/3 | 1/3 |
| 3 | Pm3 | 0.4 | 0.3 | 0.3 |
| 4 | Pm4 | 0.5 | 0.25 | 0.25 |
| 5 | Pm5 | 0.6 | 0.2 | 0.2 |
| 6 | Pm6 | 0.8 | 0.1 | 0.1 |
| 7 | Pm7 | 0.65 | 0.25 | 0.1 |

Table 6.1: Summary of disorder models used to generate seven disordered crystal structures, showing the occupancy of the ethynyl substituent at each site

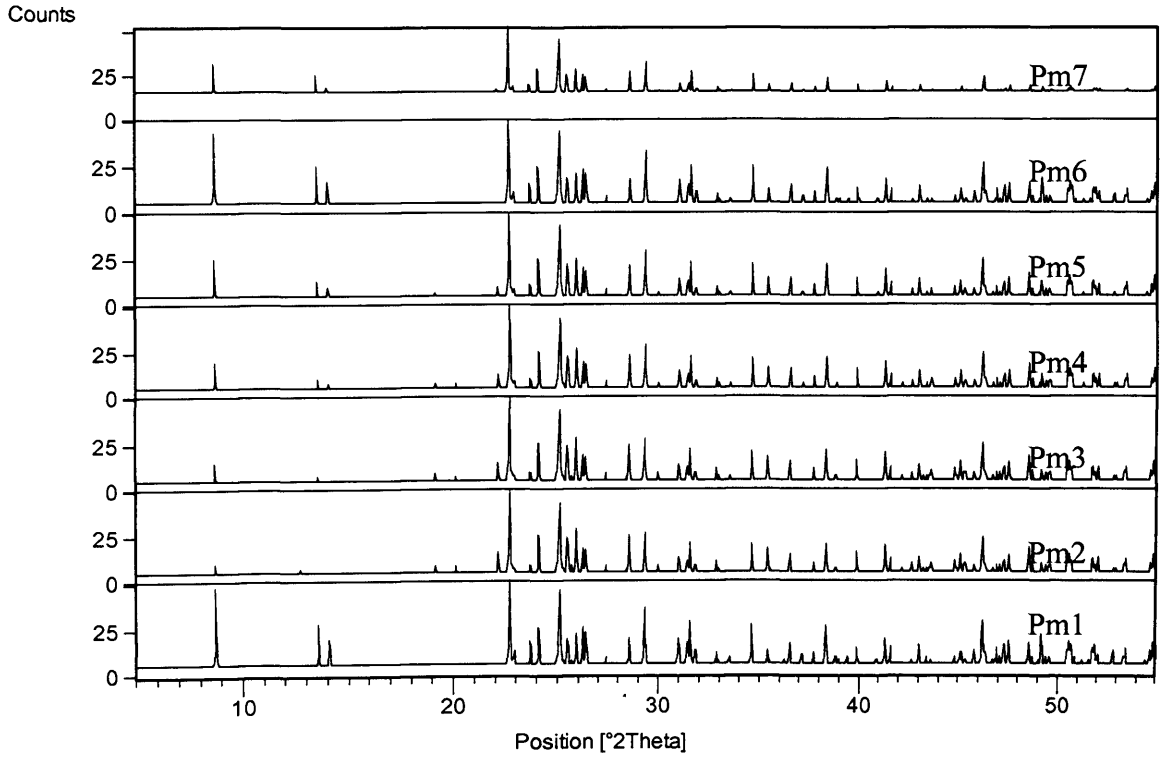



Figure 6.1: Simulated powder patterns from the seven disorder models

### 6.2.1 Defining input models to be used to describe disorder

Three test models were considered as the input for structure solution. These were selected to allow an evaluation of the minimum disorder representation that would be required in order to locate the correct structure solution using the GA. The details of the test models are summarised in the next sections.

#### 6.2.1.1 Test model f1: 1 structural fragment, composed of the ten atoms

This test model represents an ordered structure, as shown in Figure 6.2, with equal occupancies for all atoms in the structure. Consequently, this model is significantly constrained with respect to defining disorder.

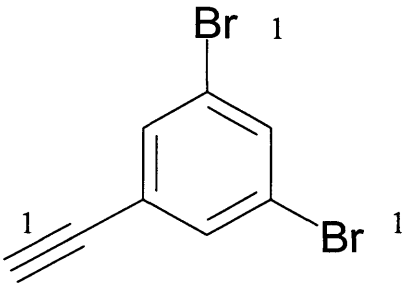



Figure 6.2: Schematic of the f1 model showing occupancies of the substituents

#### 6.2.1.2 Test model f3: 3 structural fragments

The f3 test model contains three independent fragments, with each fragment composed of all ten non-hydrogen atoms of DBEB. In each fragment, the occupancy of all of the atoms (including the ring carbon atoms as well as the substituents) were equally set to 1/3. A schematic of this model is illustrated in Figure 6.3. This disorder model represents a versatile and flexible model, capable of representing plausible trial structures for a large number of disorder regimes. The versatility of this model can be extended firstly to include more structural fragments in order to describe more complex disorder, and secondly by varying the relative occupancies of each fragment. The limitations of this model, however, are that the significant increase in the number of variables to be refined during the solution process (degrees of freedom) could result in structures that represent minima on the $R_{wp}$ hypersurface, but may not necessarily reflect a sensible and indeed correct structure.

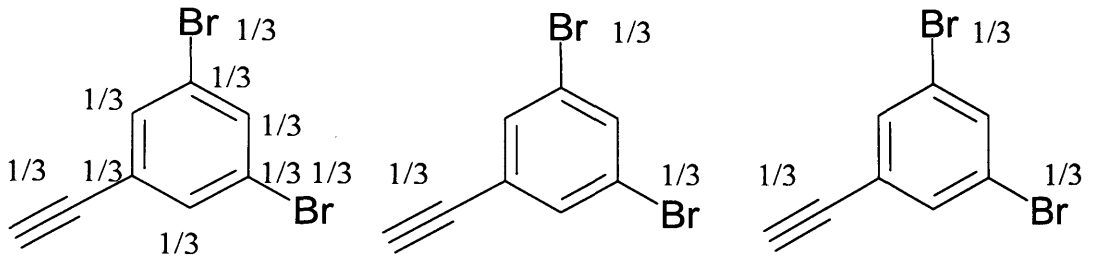



Figure 6.3: Schematic of the f3 model showing occupancies of the substituents and the phenyl ring carbons for all three fragments

### 6.2.1.3 Test model fd: 1 structural fragment, composed of 15 atoms

The fd test model contains one structural fragment with all ring carbon atoms set with an equal occupancy of 1. The bromine and ethynyl substituents are both located (with fractional occupancy) at each of the substituent sites (i.e. positions 1, 3 and 5 on the ring). The bromine substituents were all assigned an occupancy of 2/3, whilst the ethynyl carbon atoms were designated as 1/3 occupancy. Figure 6.4 represents a schematic of the model. This model can be used and adapted to represent different plausible orientational occupancies. Whilst it is heavily constrained to only represent orientational disorder with the occupancy defined by the model, i.e. in the present case substituents with a 1/3 occupancy in each of the three orientations, it could be generalised to represent other sets of occupancies.

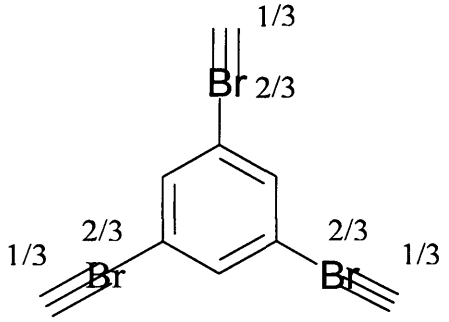



Figure 6.4: Schematic of the fd model showing 2/3 occupancy of the Br substituents, and 1/3 occupancy of the ethynyl substituents at each of the three sites

## 6.3 Structure solution using the test models to represent disorder

Each of the test models described in the preceding section was used as the molecular model in direct-space structure solution calculations using each of the powder x-ray diffraction patterns pm1 to pm7. For each test molecular model and data set, ten repeat structure solution calculations (using the GA) were performed from different random starting populations. Table 6.2 summarises the results obtained from the ten repeats from each data

set and model, showing that in some cases reproducible $R_{wp}$ values were obtained (showing convergence to a single structure), and others where a range is specified due to an incomplete convergence of the structure solution.

| Model | Name | $R_{wp}$ / % values from GA | | |
|---|---|---|---|---|
| | | f1 | fd | f3 |
| 1 | Pm1 | 9.7 | 39.8 | 9.7 - 16.9 |
| 2 | Pm2 | 43.5 | 7.0 | 7.1 |
| 3 | Pm3 | 43.4 | 8.5 | 8.6 |
| 4 | Pm4 | 38.9 | 13.9 | 12.3 – 15.1 |
| 5 | Pm5 | 31.2 | 20.0 | 15.1 – 17.0 |
| 6 | Pm6 | 10.2 | 36.6 | 10.3 – 14.0 |
| 7 | Pm7 | 25.9 | 24.4 | 8.3 – 12.2 |

Table 6.2: Summary of the results obtained from structure solution using each of the three test models described in section 6.1.3. The results shown are based on ten repeats. A single number denotes that no variation in the repeat analysis was observed

### 6.3.1 Assessment of the success rate for structure solution

Typically, a measure of whether a structure has been solved successfully is provided by the goodness of fit metric $R_{wp}$ (as described in Chapter 2). Furthermore, a visual assessment of the arrangement of the molecules in the structure coupled with an assessment of close contact interactions and scope for hydrogen bonding, further verifies the accuracy of the trial structure. In this study, an additional metric is used (particularly relevant for the trial structures generated using the f3 test model) in which the root mean square difference (rmsd) of the trial structure from the actual structure is calculated. The rmsd calculations are based on the positional differences between like carbon atoms in the aromatic rings of both the trial and actual structures.

#### 6.3.1.1 Assessment of the success rate for the pm1 structure

This structure represents a completely ordered phase. The success rate of achieving a correct trial structure following structure solution for the pm1 structure is

summarised in Table 6.3. Clearly, the f1 and f3 models are capable of representing this structure exactly, whilst the fd is not.

| Name | Success rate / % from GA | | |
|---|---|---|---|
| | f1 | fd | f3 |
| pm1 | 100 | 0 | 60 (11.6 Å) |

Table 6.3: Summary of the success rate for structure solution of the pm1 structure using the different input test models. The value in parentheses represents the rmsd between the trial and actual structures as described in the text

As expected, the f1 model provides the correct structure in each of the ten independent calculations, with the positions of the aromatic carbons and the substituents perfectly super imposable onto the correct structure. The fd model provides structures with high $R_{wp}$ values implying convergence to the incorrect structure. However, a visual assessment showed that the positions of the aromatic carbons were super-imposable onto the correct structure and that the substituents were correctly located in exactly the 1, 3 and 5 positions. The high $R_{wp}$ factor was attributed to the over-estimation of electron density in some positions, verifying that the fd model inadequately described this disorder. The fact that the positions of both the aromatic ring and the substituents were correctly located within the structure provides some confidence that the correct structure may be obtained following successful refinement by variation of the substituent occupancies.

The f3 model gives a 60% success rate in achieving the correct structures in the correct position and orientation (using a population of 60 generations), with all three fragments aligning orientationally with the expected occupancy, as illustrated in Figure 6.5. Of the remaining four incorrect structures using this input model, one of the fragments was not completely aligned in the correct position, indicating that complete convergence had not occurred. Table 6.3 summarises the rmsd for the deviation from convergence based on the alignment of like carbons in the ring. The magnitude of this deviation indicates that this fragment is a significant distance from converging. Figure 6.6 illustrates the unsuccessful trial structure. It was anticipated that the use of a larger population and generation size would lead to complete convergence to a good structure solution. Thus, increasing the number of

generations used during the GA procedure from 60 to 200 and 400 generations resulted in reproducible and successful convergence to give the correct structure.

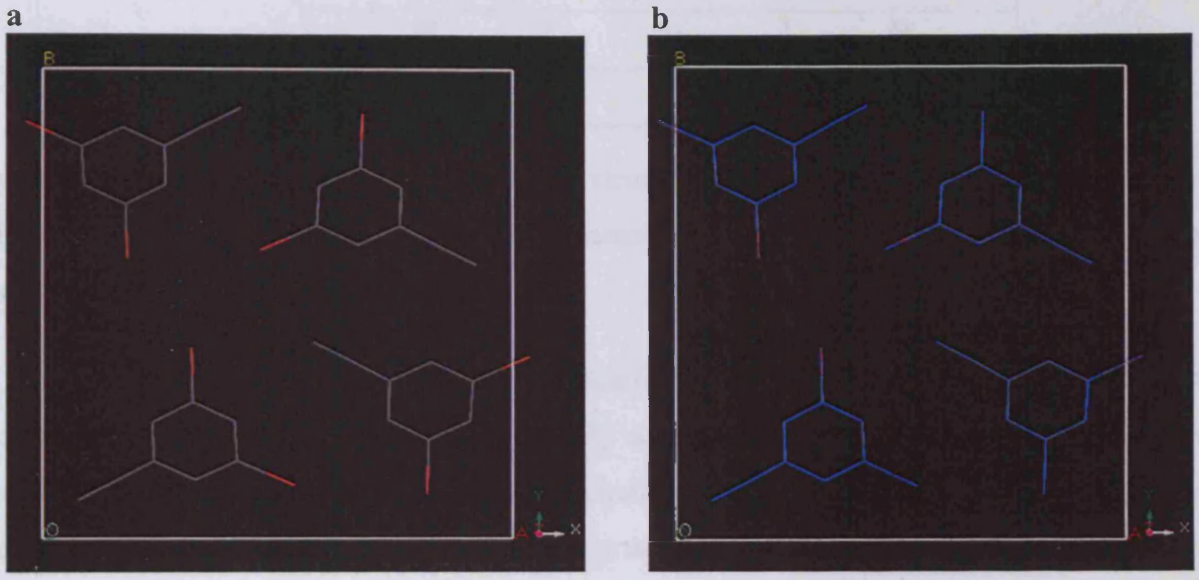



Figure 6.5: The successful pm1 trial structure obtained using the f3 test model, showing (a) the trial structure obtained and (b) an overlay of the trial and actual structure

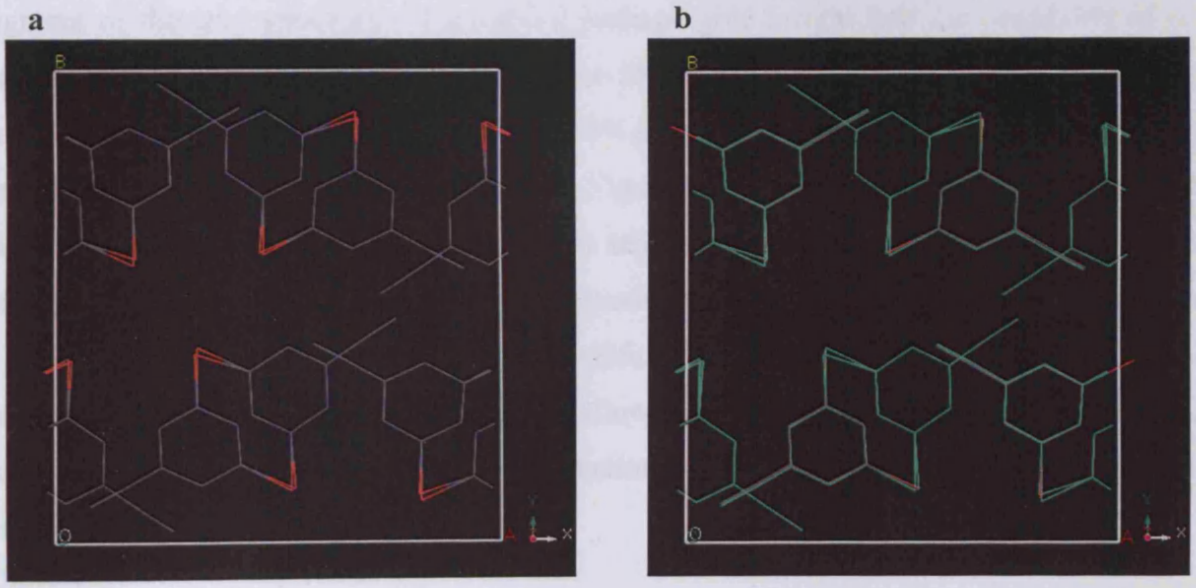



Figure 6.6: The unsuccessful pm1 trial structure obtained using the f3 test model, showing (a) the trial structure obtained and (b) an overlay of the trial and actual structure

### 6.3.1.2 Assessment of the success rate for the pm2 structure

This structure represents a relatively non-complex orientational disorder for each of the three substituents, each exhibiting 1/3 occupancy over the three substitution sites. It is anticipated that this mode of disorder can be represented exactly by the f3 and fd models, but not by the f1 model. The success rate of achieving a correct trial structure following structure solution for the pm2 structure is summarised in Table 6.4.

| Name | Success rate / % from GA | | |
|---|---|---|---|
| | f1 | fd | f3 |
| pm2 | 0 | 100 | 90 (1.92 Å) |

Table 6.4: Summary of the success rate for structure solution of the pm2 structure using the different input test models. The value in parentheses represents the rmsd between the trial and actual structure as described in the text

The f1 model gave the correct positions, with respect to the aromatic ring, and location of the substituents, but inadequately described the disorder in relation to electron density at each of the disordered site, as anticipated. Successful location of all components of the structure would suggest that the correct structure would be achieved upon refinement by variation of the substituent occupancies. Furthermore, of the ten repeats for the structure solution there was randomness in the orientation (over the 1, 3 and 5 positions) of the fragment of the trial structure. This could perhaps give insight into the possibility of equal distribution of the orientational disorder. The fd model gave 100% success rate as expected. The f3 model gave 9/10 correct structures with equal distribution of fragments in each of the correct orientations describing the disorder. Figure 6.7 illustrates the successful trial structure using the f3 input model. The one incorrect structure is due to one fragment, which, whilst being in the correct orientation, had not fully converged to the correct position with an rmsd of 1.92 Å. Figure 6.8 illustrates the unsuccessful trial structure. It was anticipated that more generations would be required in order to allow reproducible convergence of the structure. This was achieved upon using 200 generations in the GA solution process, where all structures were shown to be correct.

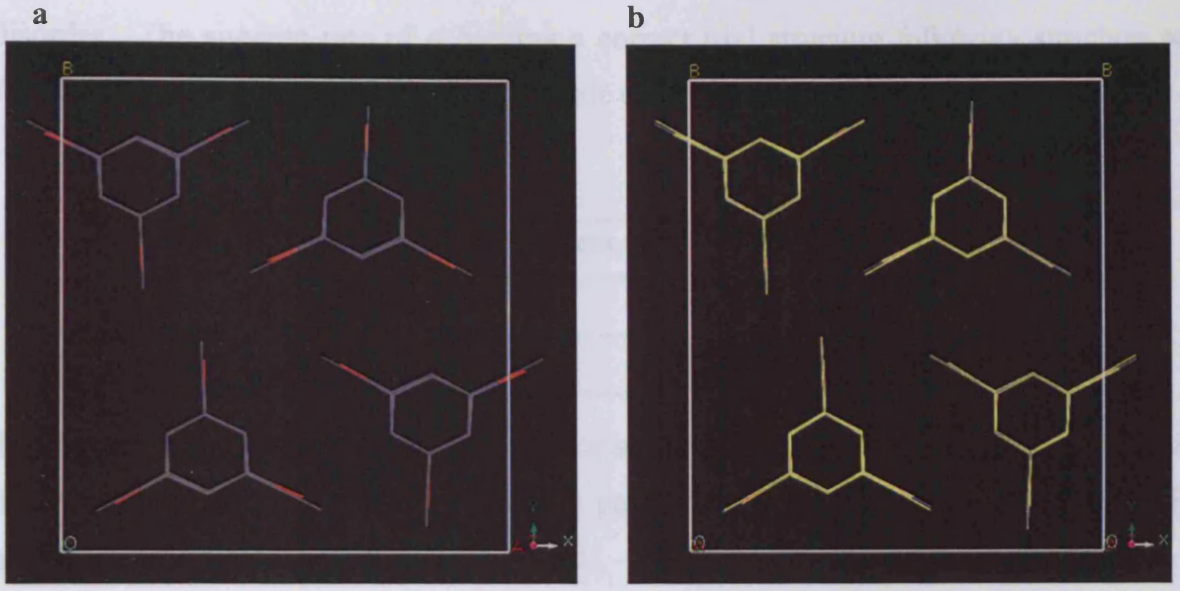



Figure 6.7: The successful pm2 trial structure obtained using the f3 test model, showing (a) the trial structure obtained and (b) an overlay of the trial and actual structure

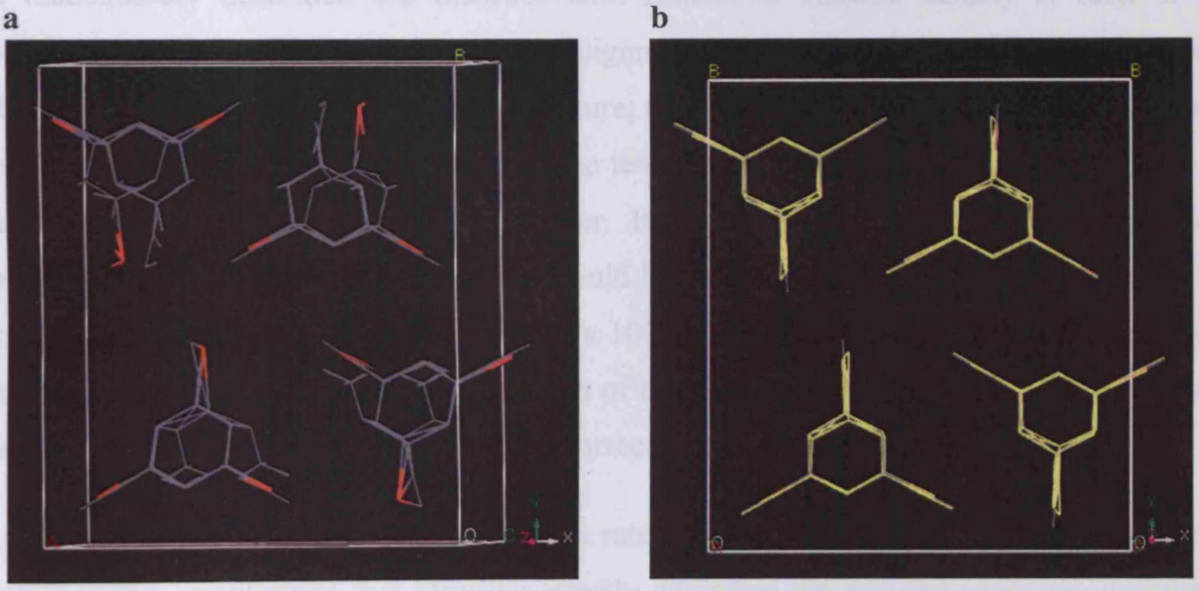



Figure 6.8: The unsuccessful pm2 trial structure obtained using the f3 test model, showing (a) the trial structure obtained (offset slightly to illustrate the incomplete positional convergence) and (b) an overlay of the trial and actual structure

### 6.3.1.3 Assessment of the success rate for the pm3 structure

This structure represents variable orientational disorder for each of the three substituents, with occupancies of 0.4, 0.3 and 0.3 at each of the three substitution sites. The mode of the positional disorder varies slightly from the pm2 model, and as such it is anticipated that both the fd and f3 models will provide approximate representations of the

disorder. The success rate of achieving a correct trial structure following structure solution for the pm3 structure is summarised in Table 6.5.

| Name | Success rate / % from GA | | |
|---|---|---|---|
| | f1 | fd | f3 |
| pm3 | 0 | 100 | 70 (7.67 Å) |

Table 6.5: Summary of the success rate for structure solution of the pm3 structure using the different input test models. The value in parentheses represents the rmsd between the trial and actual structure as described in the text

The f1 model gave the correct position of the aromatic ring in all of the ten repeats, but inadequately described the disorder with respect to electron density at each of the disordered sites, and showed no obvious alignment with the major occupancy orientation. Additionally, as observed for the pm2 structure, the structure solution showed randomness in the orientation of the fragment in each of the ten trial structures - implying the possibility of equal distribution of the orientational disorder. It is anticipated that the successful location of the substituents representing the disorder could be achieved upon refinement by varying the substituent occupancy. The fd model gave a 100% success rate, indicating that whilst it did not completely reflect the true representation of the disorder of the system it was fairly close. The occupancies could almost certainly be corrected during the refinement procedure.

The f3 model gave a 70% success rate, with an equal distribution of the fragments in each of the orientations that almost correctly described the disorder (correctly describing the disorder exhibited by the structure in pm2), as shown in Figure 6.9. The accurate location of all of the components of the structure increases the confidence that the structure representing the correct disorder may be achieved by refining the substituent occupancies. The three incorrect structures were due to the fact that one of the fragments had not fully converged, with an rmsd of 7.67Å, indicating that more generations would be required during the GA solution procedure in order to reproducibly converge to the correct trial structure. This was achieved with by increasing the number of generations used during the structure solution to 200. Figure 6.10 illustrates the unsuccessful trial structure.

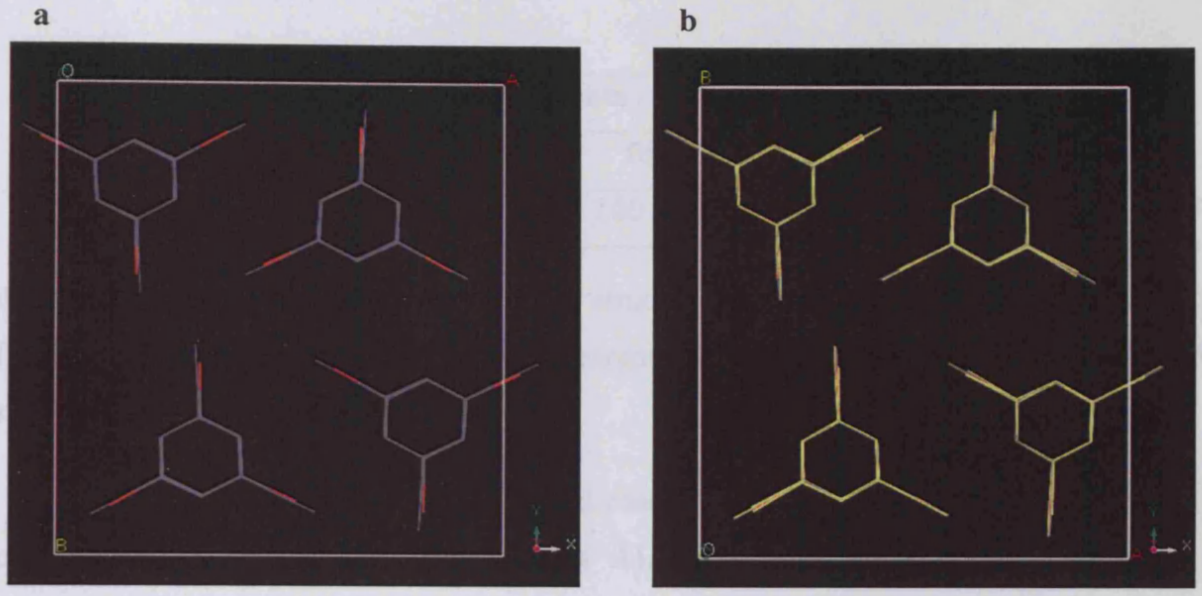



Figure 6.9: The successful pm3 trial structure obtained using the f3 test model, showing (a) the trial structure obtained and (b) an overlay of the trial and actual structure

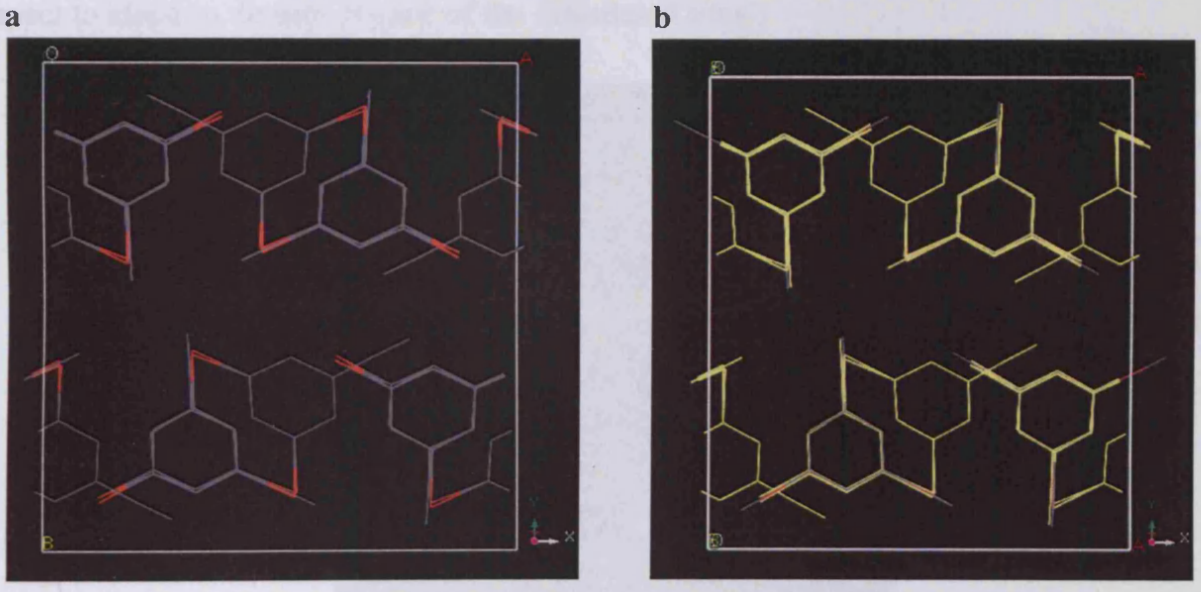



Figure 6.10: The unsuccessful pm3 trial structure obtained using the f3 test model, showing (a) the trial structure obtained and (b) an overlay of the trial and actual structure

### 6.3.1.4 Assessment of the success rate for the pm4 structure

This structure represents variable orientational disorder for each of the three substituents. Each of the substituents exhibits occupancies of 0.5, 0.25 and 0.25 at each of the three substitution sites. The success rate of achieving a correct trial structure following structure solution for the pm4 structure is summarised in Table 6.6.

| Name | Success rate / % from GA | | |
|---|---|---|---|
| | f1 | fd | f3 |
| pm4 | 0 (0.014 Å) | 100 / 0 | 0 |

Table 6.6: Summary of the success rate for structure solution of the pm4 structure using the different input test models. The value in parentheses represents the rmsd between the trial and actual structure as described in the text

As seen with the other disordered cases, the f1 model gave the correct position of the aromatic ring (with an rmsd of 0.014 Å), with the orientation of the fragment and substituents aligned to the major occupancy orientation (as shown in Figure 6.11) in all of the ten repeat calculations. The f1 model, however, inadequately described the disorder with respect to electron density at each of the disordered sites.

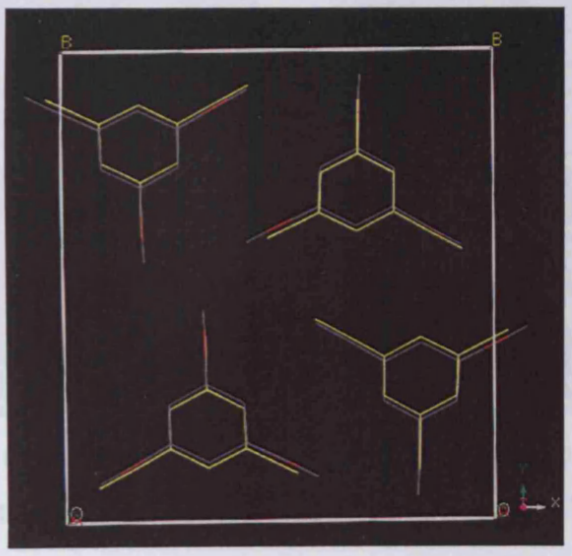


Figure 6.11: An overlay of the trial structure obtained using the f1 model with the actual pm4 structure showing alignment of the trial structure to the major occupancy (0.5) orientation

It was unclear whether the fd model represented 100% or 0% success rate. Verification of the success rate would only be achieved by refining the trial structure obtained using the fd model. The $R_{wp}$ value was still a little high, probably due to a larger discrepancy between actual and modelled disorder with respect to electron density at each of the disordered sites, despite the fact that the position of the aromatic ring and the substituents had been correctly located in all of the repeat analyses. The f3 model gave 0% success rate (with the use of 100 generations in each of the ten repeat calculations). In all of the unsuccessful

trial structures, the fragments were orientated such that each of the substituents were present at each of the disorder sites with a 1/3 occupancy (i.e. one ethynyl substituent was present at each of the disorder sites), suggesting a discrepancy with respect to electron density at each of the disordered sites. Two main different scenarios were seen in the failed structures designated as best and worst cases. In the best case, as shown in Figure 6.12, one of the three fragments had not fully converged, with an rmsd of 7.00 Å. In the worst case, illustrated in Figure 6.13, all three fragments had not fully converged to the correct position. The rmsd of all three fragments in this unsuccessful structure was found to be 0.59, 0.42 and 3.51 Å. Increasing the number of generations used in the GA calculations from 100 to 200 and then 400 did not improve the results, which still showed incomplete convergence of all three fragments.

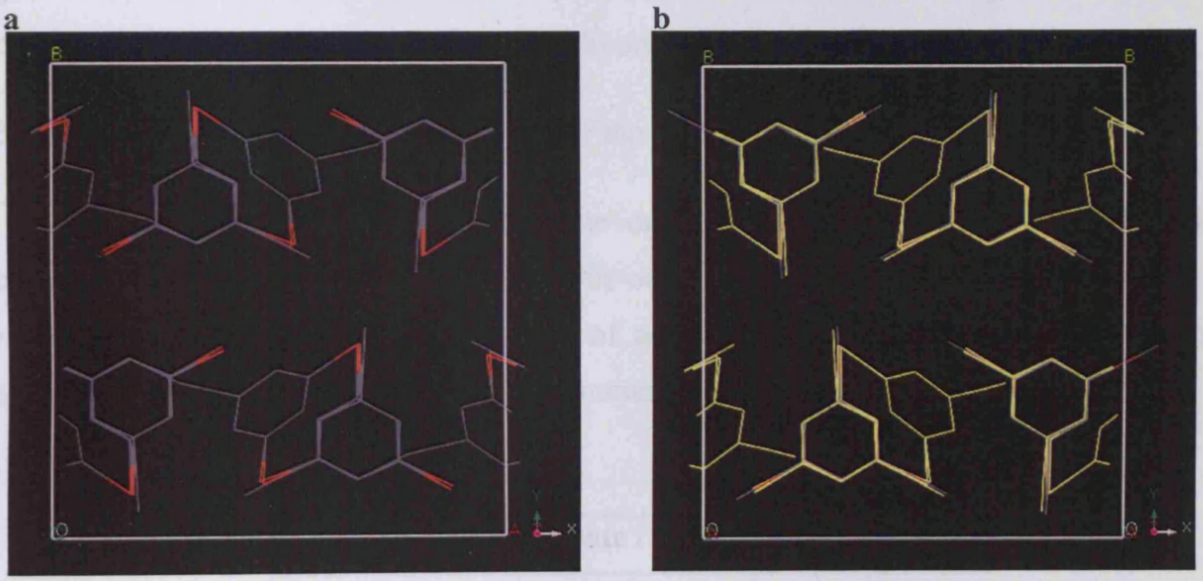



Figure 6.12: The best-case unsuccessful pm4 trial structure obtained using the f3 test model, showing (a) the trial structure obtained and (b) an overlay of the trial and actual structure

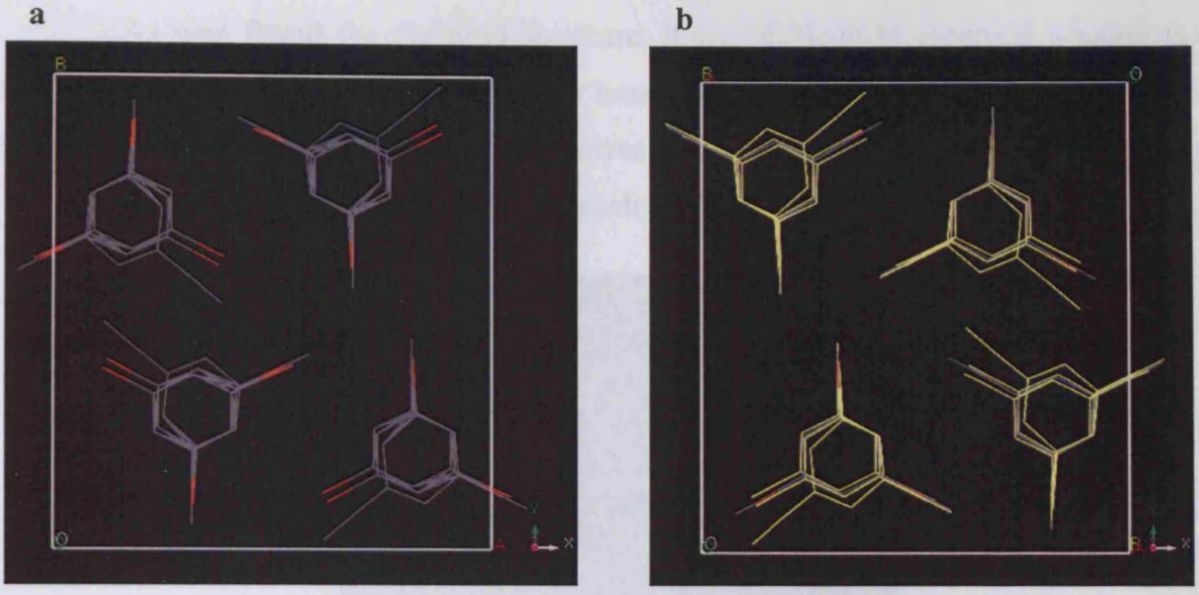



Figure 6.13: The worst-case unsuccessful pm4 trial structure obtained using the f3 test model, showing (a) the trial structure obtained and (b) an overlay of the trial and actual structure

### 6.3.1.5 Assessment of the success rate for the pm5 structure

This structure represents variable orientational disorder for each of the three substituents. Each of the substituents exhibits occupancies of 0.6, 0.2 and 0.2 at each of the three substitution sites. The success rate of achieving a correct trial structure following structure solution for the pm5 structure is summarised in Table 6.7.

| Name | Success rate / % from GA | | |
|---|---|---|---|
| | f1 | fd | f3 |
| pm5 | 0 | 100 / 0 | 0 |

Table 6.7: Summary of the success rate of structure solution of the pm5 structure using the different input test models.

As with the disorder model 4 (pm4), the f1 model gave the correct position of the aromatic ring, and with the orientation of the fragment aligned to the major occupancy of the correct structure (as shown in Figure 6.14). However, this model inadequately describes the disorder with respect to electron density at each of the disordered sites. That the positions of all the components of the structure were accurately located gave increased confidence of achieving the correct disorder structure by variation of the occupancies during the refinement procedure.

As was found for the pm4 structure, it was difficult to ascertain whether the fd model gave 100% or 0% success rate, solely based on the magnitude of the $R_{wp}$ value. Whilst the position of the aromatic ring had been correctly located, the high $R_{wp}$ value was attributed to the discrepancy in the electron density at each of the disordered sites.

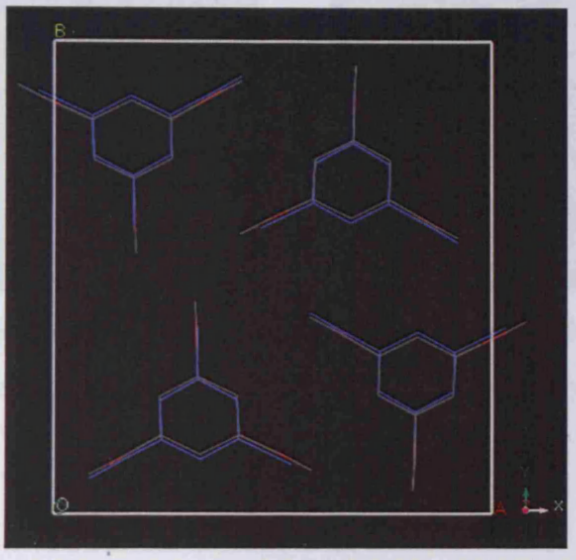


Figure 6.14: An overlay of the trial structure obtained using the f1 model with the actual pm5 structure, showing alignment of the trial structure to the major occupancy (0.6) orientation

The f3 model gave 0% success rate after 100 generations in each of the ten repeat calculations, with all three fragments still to converge to the correct position and orientation. The number of generations was increased to 200 and then 400, but with very little success in achieving complete convergence of the three fragments. Figure 6.15 shows the trial structure achieved, and Figure 6.16 illustrates the distribution of the substituents in the pm5 disorder regime over the three substitution sites. The rmsd values of the three fragments, relative to the correct structure, are 0.004, 0.042 and 7.09 Å. The fragments that had almost converged (rmsd of 0.004, 0.042 Å) had an arrangement in which each fragment has an ethynyl substituent in two of the three disordered orientations, as illustrated in Figure 6.17a. That no convergence of the third fragment was achieved was probably due to the discrepancy in describing the occupancy at each of the disordered sites. Figure 6.17b and c illustrate the two possible orientations for the third fragment to adopt during convergence. The rmsd value was calculated for the fragment to converge to either of these two scenarios and was found to be 7.09 Å for scenario 6.17b and 8.34 Å for scenario 6.17c, indicating that scenario b was perhaps more favourable with respect to the distribution of the substituents.

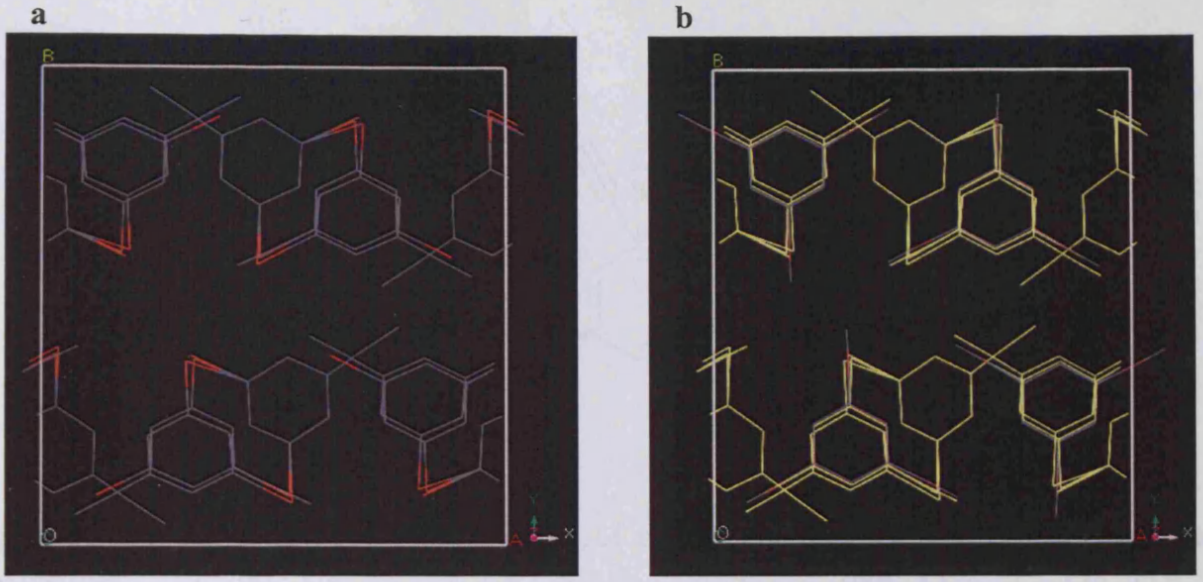



Figure 6.15: The unsuccessful pm5 trial structure obtained using the f3 test model, showing (a) the trial structure obtained and (b) an overlay of the trial (in yellow) and actual structure

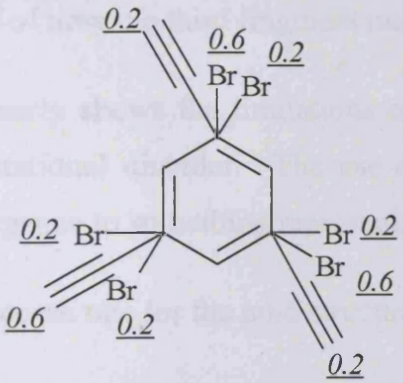



Figure 6.16: Schematic of the distribution of substituents and occupancies (in underlined italics) in the actual crystal structure of pm5, the substituent bonds at each of the three sites are skewed to provide clarity in visualising the substitution pattern and do not reflect the true conformation or spatial arrangement.

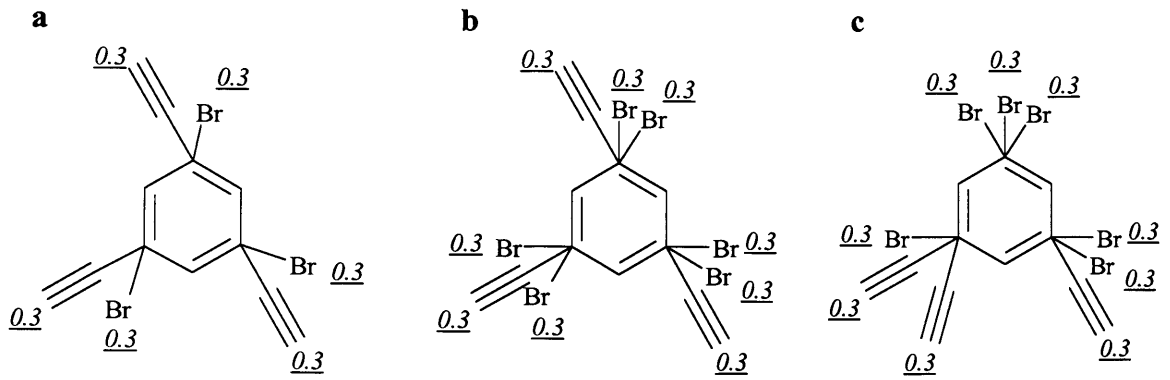



Figure 6.17: Schematic of the distribution of substituents and occupancies (in underlined italics). The substituent bonds at each of the three sites are skewed to provide clarity in visualising the substitution pattern and do not reflect the true conformation or spatial arrangement. Scheme (a) represents the status of the converged fragments and (b) and (c) represent two possible scenarios of how the third fragment may converge

This investigation clearly shows the limitations of using constrained models that inadequately describe the orientational disorder. The use of a more flexible or adaptable model may have allowed convergence to something representing the correct structure.

### 6.3.1.6 Assessment of the success rate for the pm6 structure

This structure represents an inequivalent orientational disorder for each of the three substituents. Each of the substituents exhibits an occupancy of 0.8, 0.1 and 0.1 at each of the three substitution sites. The success rate of achieving a correct trial structure following structure solution for the pm6 structure is summarised in Table 6.8.

| Name | Success rate / % from GA | | |
|---|---|---|---|
| | f1 | fd | f3 |
| Pm6 | 100 | 0 | 60 (8.21 Å) |

Table 6.8: Summary of the success rate for structure solution of the pm6 structure using the different input test models. The value in parentheses represents the rmsd between the trial and actual structure as described in the text

The f1 model gave 100% success rate of a structure that represented the actual structure based on the $R_{wp}$ assessment. Visual assessment of the fragment in the trial structure

indicated it to be perfectly aligned with the major occupancy orientation of the actual structure, as shown in Figure 6.18, and it is anticipated that refinement of the occupancies should give rise to the correct disorder structure.

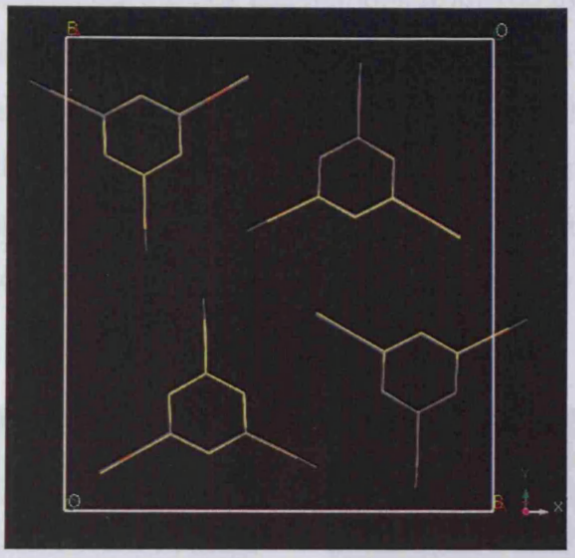


Figure 6.18: An overlay of the trial structure obtained using the f1 model with the actual pm6 structure, showing alignment of the trial structure to the major occupancy (0.8) orientation

The fd model gave the correct position of the aromatic ring and location of the substituents, but gave a relatively high $R_{wp}$ value due to the over-estimation of electron density in some positions, and under estimation in other positions. The f3 model gave 60% success rate, and of the correct structures all three fragments were aligned with the major occupancy, as illustrated in Figure 6.19.

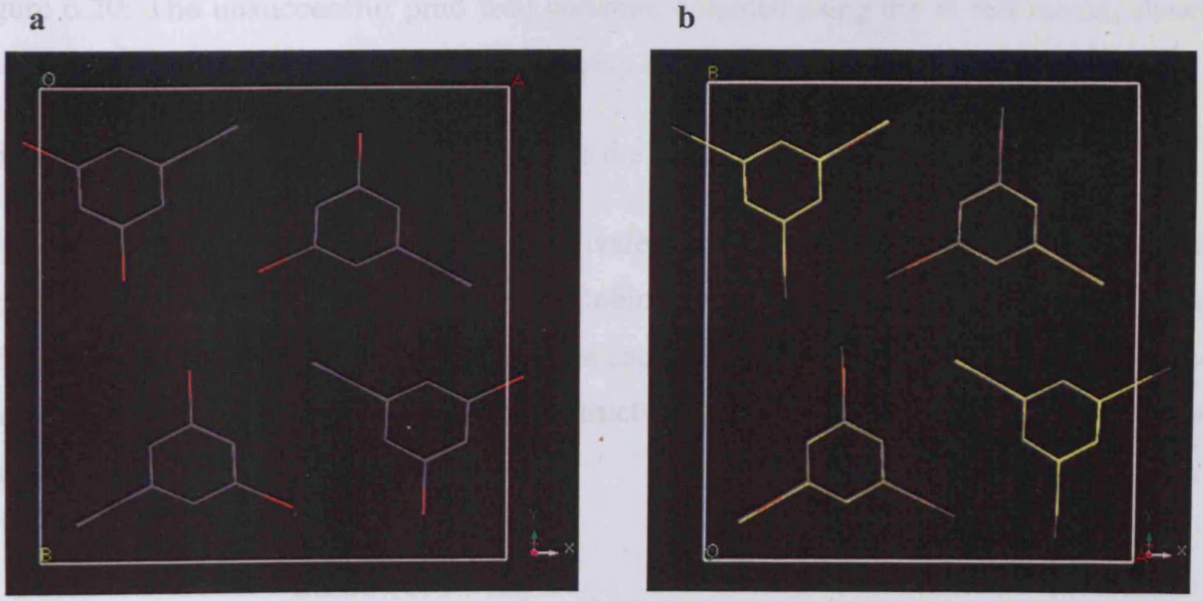



Figure 6.19: The successful pm6 trial structure obtained using the f3 test model, showing (a) the trial structure obtained and (b) an overlay of the trial and actual structure

The remaining four incorrect structures arose due to incomplete convergence of one of the three fragments in the f3 model, as illustrated in Figure 6.20. The rmsd deviation of the fragment to converge with the correct structural position was calculated as 8.21 Å. Increasing the number of generations used for the GA calculations from 100 to 200 or greater resulted in successful convergence of all three fragments. In all cases the orientation of the trial structure was aligned to the orientation adopted by the major occupancy in the actual crystal structure. However, this data could provide a falsely positive result in that it reflects a completely ordered system in which all structures used in the test models adopt a single orientation. It is anticipated that the description of the disorder would improve during the refinement procedure.

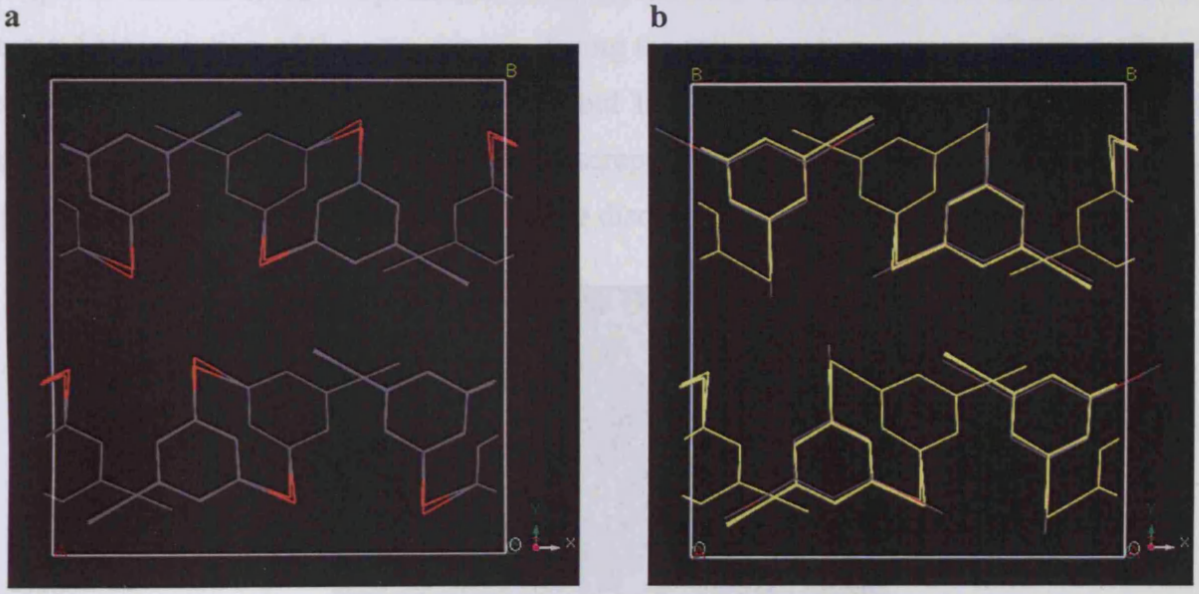



Figure 6.20: The unsuccessful pm6 trial structure obtained using the f3 test model, showing (a) the trial structure obtained and (b) an overlay of the trial and actual structure

### 6.3.1.7 Assessment of the success rate for the pm7 structure

This structure represents an inequivalent orientational disorder for each of the three substituents, as reported in the literature (Robinson *et al.* 1998). Each of the substituents exhibits an occupancy of 0.65, 0.25 and 0.1 at each of the three substitution sites. The success rate for the achievement of a correct trial structure following structure solution for the pm7 system is summarised in Table 6.9.

| Name | Success rate / % from GA | | |
|---|---|---|---|
| | f1 | fd | f3 |
| Pm7 | 100 | 100 / 0 | 30 |

Table 6.9: Summary of the success rate for structure solution of the pm7 structure using the different input test models.

The f1 model gave the correct structure on all occasions, with the structure aligned with the major occupancy orientation, as illustrated in Figure 6.21. As in previous cases, it is anticipated that the correct representation of the disorder exhibited by this structure would be achieved by variation of the occupancies during the refinement procedure. The fd model gave the correct position of the aromatic ring, but like models pm4 and 5, the $R_{wp}$ value was relatively high probably due to a larger discrepancy between actual and modelled disorder with respect to electron density at each of the disordered sites.

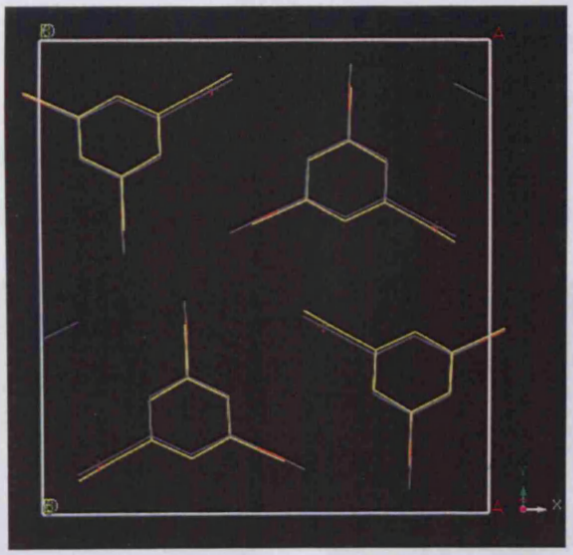


Figure 6.21: An overlay of the trial structure obtained using the f1 model with the actual pm7 structure, showing alignment of the trial structure to the major occupancy (0.65) orientation

The f3 model gave a 30% success rate. The correct structures, as illustrated in Figure 6.22, show two fragments aligning with the major occupancy orientation and the third fragment aligning with the next significant occupancy orientation. Complete convergence of all three fragments had not been achieved, with an rmsd deviation from the actual position of the aromatic ring calculated as 0.04, 0.03 and 2.19 Å. The solved structure shows an occupancy distribution of 0.666, 0.333, 0 for the three disordered sites. It is anticipated that

the correct representation of the disorder exhibited by this structure would be achieved following refinement of the occupancies.

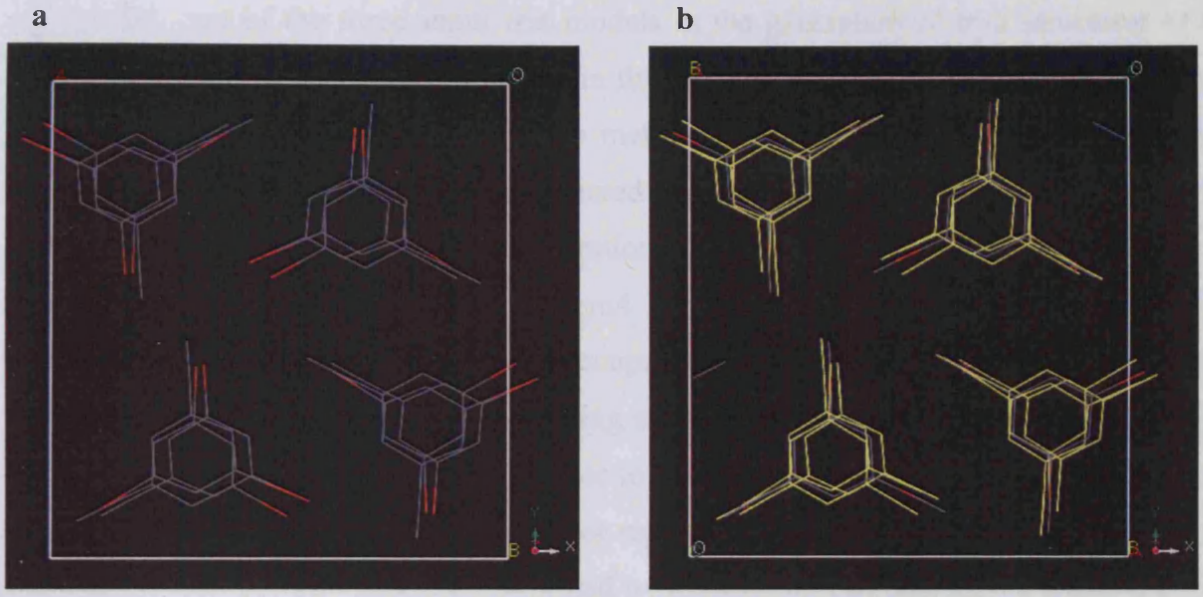



Figure 6.22: The successful pm7 trial structure obtained using the f3 test model, showing (a) the trial structure obtained and (b) an overlay of the trial and actual structure

The remaining six incorrect structures mainly arise from incomplete convergence of all three fragments (as shown in the Figure 6.23) with rmsd values calculated as 2.19, 2.06 and 2.14 Å. Increasing the number of generations used during the GA calculations from 100 to 200 resulted in complete convergence and 100% success rate of all repeats.


Figure 6.23: The unsuccessful pm7 trial structure obtained using the f3 test model, showing (a) the trial structure obtained and (b) an overlay of the trial and actual structure

## 6.4 Discussion and conclusions

The use of the three input test models in the generation of trial structures of the orientationally disordered systems assessed in this study clearly demonstrates that disordered structures can be solved using direct space methods in an entirely tractable and relatively straightforward process. Of the seven disordered structures that were evaluated, the use of the three input test models resulted in the generation of successful trial structures in all but two cases. These two exceptions were the pm4 and pm5 structures, for which the major orientation was approximately double the occupancy of the two minor orientations – i.e. the disorder was neither equally (or close to being equally) distributed across the three sites nor skewed heavily to one or two sites. In order to achieve a successful trial structure, in such cases, would require a more complex disorder representation in the input model. However, a correct representation of the disorder exhibited by these structures may be achieved following refinement of the trial structures obtained using the models discussed, with a variation of the site occupancies during the refinement procedure

The f3 test model correctly, and reproducibly, defined the disorder in a number of the structures evaluated in this study. The f3 model accurately describes the pm1 and pm2 structures and described (within error of the actual occupancy representation) the pm3, pm6 and pm7 structures. As such, this model shows versatility, and the presence of more variables (number of fragments) indicates that it could potentially represent a larger number of plausible structures, and could offer an insight as to whether disorder is indeed present. The fact that the variable space that it represents is significantly wide will, however, result in a significantly increased number of minima on the $R_{wp}$ hypersurface. Therefore, a large number of trial structures could be incorrectly defined as successful if the success rate is solely based on the $R_{wp}$ metric. Additionally, this representation of a wider variable space necessitates the use of a larger number of generations in order for the GA calculations to achieve the successful trial structure, indicating that the increased versatility of this model comes at the expense of increased computational time. Furthermore, this model could be adapted to include a greater or lesser number of fragments, with equal or inequivalent occupancy assignments, for each of the substituents in accordance with the type of disorder anticipated.

The fd test model is rather constrained and describes a very specific type of orientational disorder. This model successfully describes the disorder exhibited by the pm2 structure, as expected, and essentially describes the disorder seen in the pm3 structure. The

potential of this model to successfully locate the position of the aromatic ring and the substituents has also been demonstrated (for all the structures evaluated) and the measure of failure of this to achieve the correct structure was solely based on the high $R_{wp}$ value. However, there was some difficulty encountered in ascribing either complete success or complete failure based on the trial structures generated from this model. This difficulty arose from the fact that, whilst the ten repeats of the structure solution calculations were perfectly reproducible, the magnitude of the $R_{wp}$ value would suggest that the structure had not arrived at a suitable minimum, rather than the fact that it inaccurately describes the orientational disorder. The use of this model clearly illustrates the limitations of relying solely on a low $R_{wp}$ value as the best metric for a successful structure. That the repeat calculations produced indistinguishable trial structures provides weight to the fact that the atoms had been located at or close to the correct positions, and that the high $R_{wp}$ value was solely related to the inaccurate representation of substituent occupancy. The results show that the use of both the $R_{wp}$ value coupled with a measure of the location of the correct (and reproducible) atomic positions provides a better metric for assigning the success rate of a trial structure. The possibility of more accurately defining the orientational disorder during the refinement procedure (that is refine on substituent occupancy in each of the orientational sites) is yet to be explored, but is expected to be entirely achievable. Once this has been demonstrated, this would then illustrate the value of not rejecting spatially plausible and reproducible trial structures with high $R_{wp}$ values, and progressing either to refinement (in which particular attention to the substituent occupancy is assessed) or triggering the use of more complex orientational disorder models.

The f1 test model was used almost as a positive control to assess the presence/absence of any orientational disorder. That it correlates well with the disordered structure of pm6 illustrates the importance of not using this model in isolation if orientational disorder is suspected. In all cases, however, the position of the aromatic ring was located successfully, indicating that the measure of failure is based on the magnitude of the $R_{wp}$ metric arising from the inadequate representation of the substituents at each of the sites. Interestingly, the trial structures obtained using this model for the pm2 and pm3 structures show that the f1 fragment does not always align to the same orientation in each of the repeats, suggesting the possibility of equally distributed orientational disorder. Furthermore, for the remaining structures, the f1 fragment was shown to reproducibly align to the major occupancy

orientation of the actual structure. This indicates the utility of this model in obtaining an insight into the type of disorder that may be present.

This evaluation clearly shows the value of using more than one structural representation (i.e. more than one test model) in the input file during structure solution in cases where disorder is suspected. This study also emphasises the importance of using other parameters in addition to the $R_{wp}$ value as a measure of success.

## 6.5 Experimental

### 6.5.1 Extracting powder patterns

The disordered structures were generated by altering the occupancy of the crystallographic input file (.cif) and then extracting the powder pattern. This was achieved using the Material Studio Reflex module (2005).

### 6.5.2 Structure solution from X-ray powder diffraction data

Structure solution was performed using the GA program EAGER using parallel processing regime incorporated into the PEACSS program. Typically, the structure solution step involved 100 starting populations with 60 to 100 generations. The mating rate (i.e. mating per generation) was set at 15, and the mutation rate was typically set at 8 per generation.

## 6.6 References


- *Material Studio Reflex package v. 4.0.0.0*, Accelrys Inc., San Diego 2005
- Robinson J. M. A., Kariuki B. M., Harris K. D. M., Philp D., (1998), *J. Chem. Soc., Perkin 2 Trans*, 2459
- Tremayne M., Grice L., Pyatt J. C., Seaton C. C., Kariuki B. M., Tsui H. H. Y., Price S. L., Cherryman J. C., (2004), *J. Am. Chem. Soc.*, **126**, 7071
- Wright J.D., (1995), *Molecular Crystals*, Cambridge University Press