WARWICK
THE UNIVERSITY OF WARWICK

A Thesis Submitted for the Degree of PhD at the University of Warwick

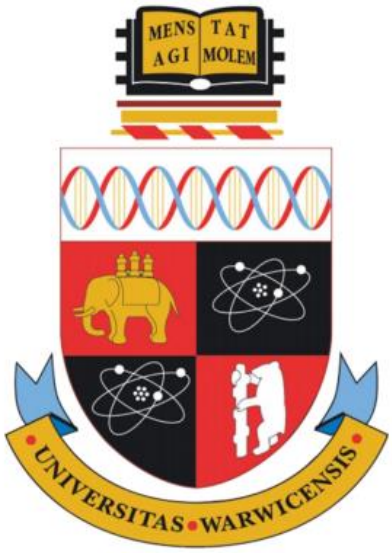
Permanent WRAP URL:

http://wrap.warwick.ac.uk/148414

Copyright and reuse:
This thesis is made available online and is protected by original copyright.
Please scroll down to view the document itself.
Please refer to the repository record for this item for information to help you to cite it.
Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwick.ac.uk/lib-publications



Combining experimental solid-state NMR and X-ray diffraction with density functional theory calcualtion to characterise multicomponent forms of organic molecules

by

Emily Kathleen Corlett

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philisophy

Molecular Analytical Sciences Centre for Doctoral Training

February 2020

Contents

Acknowledgements

I would like to thank my supervisor Professor Steven Brown, for all his support and advice throught my PhD, Professor Richard Walton, my second supervisor, and Dr David Walker for their guidance, particularly with XRD. I would also like to thank Dr Helen Blade and Dr Les Hughes at Astra Zeneca and Dr Philip Sidebottom at Syngenta for their help and expertise, as well as financial support from their respective organisations. Additional thanks go to Dr Guy Clarkson, Dr Steven Hubband and Dr Claire Murray. The opportunity and support provided by MAS and financial support from EPSRC is also gratefully acknowledged.

I am incredibly grateful to everyone, past and present, from the solid-state NMR groups who were always willing to help, problem solve, provide moral support, cake and cookies and generally make F16 a great office to work in. Finally, thank you to my family for all their love and encouragement throughout all these years.

Declarations

All work performed in this thesis is original researh carried out by this author and has not been submitted for a degree at another university. Any contributions from others have been stated accordingly. Chapters 3, 5 and 6 contain work published by the author:

E. K. Corlett, H. Blade, L. P. Hughes, P. J. Sidebottom, D. Walker, R. I. Walton and S. P. Brown, An XRD and NMR crystallographic investigation of the structure of 2,6-lutidinium hydrogen fumarate, *CrystEngComm*, 2019, **21**, 3502-3516.

E. K. Corlett, H. Blade, L. P. Hughes, P. J. Sidebottom, D. Walker, R. I. Walton and S. P. Brown, Investigating discrepancies between experimental solid-state NMR and GIPAW calculation: N=C-N ^{13}C and OH···O ^{1}H chemical shifts in pyridinium fumarates and their cocrystals, *Solid-State Nuclear Magn. Reson.*, 2020, **108**, 101662.

E. K. Corlett, H. Blade, L. P. Hughes, P. J. Sidebottom, D. Walker, R. I. Walton and S. P. Brown, 5-Amino-2-methylpyridinium hydrogen fumarate: an XRD and NMR crystallography analysis, *Magn. Reson. Chem.*, 2020, in press. DOI: 10.1002/mrc.5021.

Abstract

Structure analysis is an important step in the process of developing new solid forms of active pharmaceutical ingredients (APIs) and agrochemical ingredients (AIs) as it can provide a starting point for related structures to be designed and represents an intellectual property (IP) opportunity.

This thesis investigates the complementarity of XRD and solid-state magic-angle spinning (MAS) NMR, alongside density functional theory (DFT) calculations, for the characterisation of five related pyridine based molecules, each co-crystallised with fumaric acid. Structures and stabilities were investigated using both single crystal and powder XRD at a range of temperatures, ^{1}H MAS, ^{13}C cross polarisation (CP) MAS solid state NMR spectra and ^{1}H-^{1}H, ^{1}H-^{13}C and ^{14}N-^{1}H 2D correlation experiments and thermogravimetric analysis (TGA) and differential scanning calorimetry (DSC). DFT-based geometry optimisations were performed with the unit cell parameters both fixed and variable to investigate the convergence of structures recorded at different temperatures and calculations of the NMR parameters were conducted for both the full crystal structures and for isolated molecules from within the structures, highlighting the intermolecular interactions present.

Chapter 3 identifies the loss of the base molecule, through slow evaporation, from the system, resulting in formation of crystalline fumaric acid in the affected regions. Splitting of reflections in the powder XRD patterns was linked to the existence of a second 'polymorph', which shows noticeable variation in unit cell parameters while maintaining the overall molecular packing. Chapter 4 investigates the co-occurrence of at least three multicomponent forms from the same crystallisation media, isolates the transitions between the structures (from hydrated to anhydrous) and identifies methods both for improving the selectivity of their formation and allowing their co-identification by NMR. In Chapter 5 the multicomponent forms of three substitutional analogue bases are compared, one of which shows evidence of polymorphism. Chapter 6 presents the trend in the ^{14}N shift with crystal form for tertiary amine nitrogens as well as classifying common bonding patterns which, in a Cambridge Structural

Database (CSD) search, show significant differences in incidence depending on crystal form. Two chemical environments that show significant differences between experiment and calculation are also identified.

List of Figures

Abbreviations

1D/2D/3D	1 Dimensional/2 Dimensional/3 Dimensional	20
25AMP	2-Amino-5-MethylPyridine	14
25AMP:$F_{0.5}$:$FA_{0.5}$	2-Amino-5-MethylPyridinium hemi-Fumarate hemi-Fumaric Acid	15
25L	2,5-Lutidine	14
25L:$F_{0.5}$:FA	2,5-Lutidinium hemi-Fumarate Fumaric Acid	15
26AMP	2-Amino-6-MethylPyridine	15
$26AMP_2$:F:$FA_{0.5}$	Bis-(2-Amino-6-MethylPyridinium) Fumarate hemi-Fumaric Acid	16
$26AMP_2$:F:H_2O	Bis-(2-Amino-6-MethylPyridinium) Fumarate monoHydrate	16
26AMP:$F_{0.5}$:$(H_2O)_2$	2-Amino-6-MethylPyridinium hemi-Fumarate diHydrate	15
26L	2,6-Lutidine	14
26L:HF	2,6-Lutidinium Hydrogen Fumarate	15
52AMP	5-Amino-2-MethylPyridine	14
52AMP:HF	5-Amino-2-MethylPyridinium Hydrogen Fumarate	16
AI	Agrochemical Ingredient	1
API	Active Pharmaceutical Ingredient	1
BaBa	Back to Back	69
CCD	Charge-Coupled Device	33
CCDC	Cambridge Crystallographic Data Centre	15
CP	Cross Polarization	18
CSA	Chemical Shift Anisotropy	52
CSD	Cambridge Structural Database	5
CSP	Crystal Structure Prediction	7
DFT	Density Functional Theory	12
DLS	Diamond Light Source	34
DSC	Differential Scanning Calorimetry	18
DUMBO	Decoupling Using Mind Boggling Optimisation	67
DQ/SQ/ZQ	Double Quantum/Single Quantum/Zero Quantum	53
FA	Fumaric Acid	14
FID	Free Induction Decay	38

FT	Fourier Transform	25
FWHMH	Full Width Half Maximum Height	63
H-bond	Hydrogen bond	5
HETCOR	HETeronuclear CORrelation	68
HMQC	Heteronuclear Multiple Quantum Coherence	71
GGA	Generalised Gradient Approximation	78
GIPAW	Gauge Including Projector Augmented Wave	12
IP	Intellectual Property	3
MAC	Multi-Analysing Crystal	34
MAS	Magic Angle Spinning	12
MBD	Many Body Dispersion	79
NMR	Nuclear Magnetic Resonance	12
NOE	Nuclear Overhauser Effect	70
PAS	Principal Axis System	48
PBE	Perdew Burke Ernzerhof	78
PSD	Position Sensitive Detector	34
PXRD	Powder X-Ray Diffraction	11
RT	Room Temperature	5
SE	Schrödinger Equation	75
SA	Succinic Acid	187
SXRD	Single crystal X-Ray Diffraction	11
TGA	ThermoGravimetric Analysis	18
TMS	TetraMethylSilane	50
TPPI	Two Pulse Phase Incremented	65
VT	Variable Temperature	18
XRD	X-Ray Diffraction	6

Chapter 1. Solid-State Forms and the Structural Characterisation of Pharmaceuticals and Agrochemicals

The development of multicomponent forms, such as salts and cocrystals, is utilised by both the pharmaceutical and agrochemical industries to alter the physical properties of active pharmaceutical ingredients (APIs) and active ingredients (AIs). As a consequence there is a need for detailed structural characterisation of such systems. As the practice is far more widely researched and utilised within the pharmaceutical industry, due to the propensity for products to be administered in solid form, the use and development discussed here will largely be in the context of APIs. As AIs are generally dissolved to be sprayed onto crops, the solubility and dissolution profile are far less critical compared to in APIs, where their improvement is one of the key driving forces behind multicomponent development. Although it remains crucial to understand the solubility and dissolution of AIs, they require modification far less. There are, however, areas in which the development of multicomponent forms can be of use within the agrochemical industry and these will be commented on.

1.1. Multicomponent Forms

Attempting to change the solid-state chemistry of products is a standard requirement within the pharmaceutical industry and by varying the solid form, modified physical properties can be produced without changing the API pharmacology. To this end, salt formation has been common practice for more than 25 years[1] although in the last decade or so cocrystals have also been included in the search for the optimum solid form.[2-5] The promise of cocrystals has not yet been reflected by significant inclusion into products however, possibly due to regulatory uncertainty and limited proof of their effectiveness *in vivo*. Recent regulatory changes introduced by the U. S. Food and Drug Administration and reviews from the European Medical Agency may improve things, however, as they suggest a shift towards

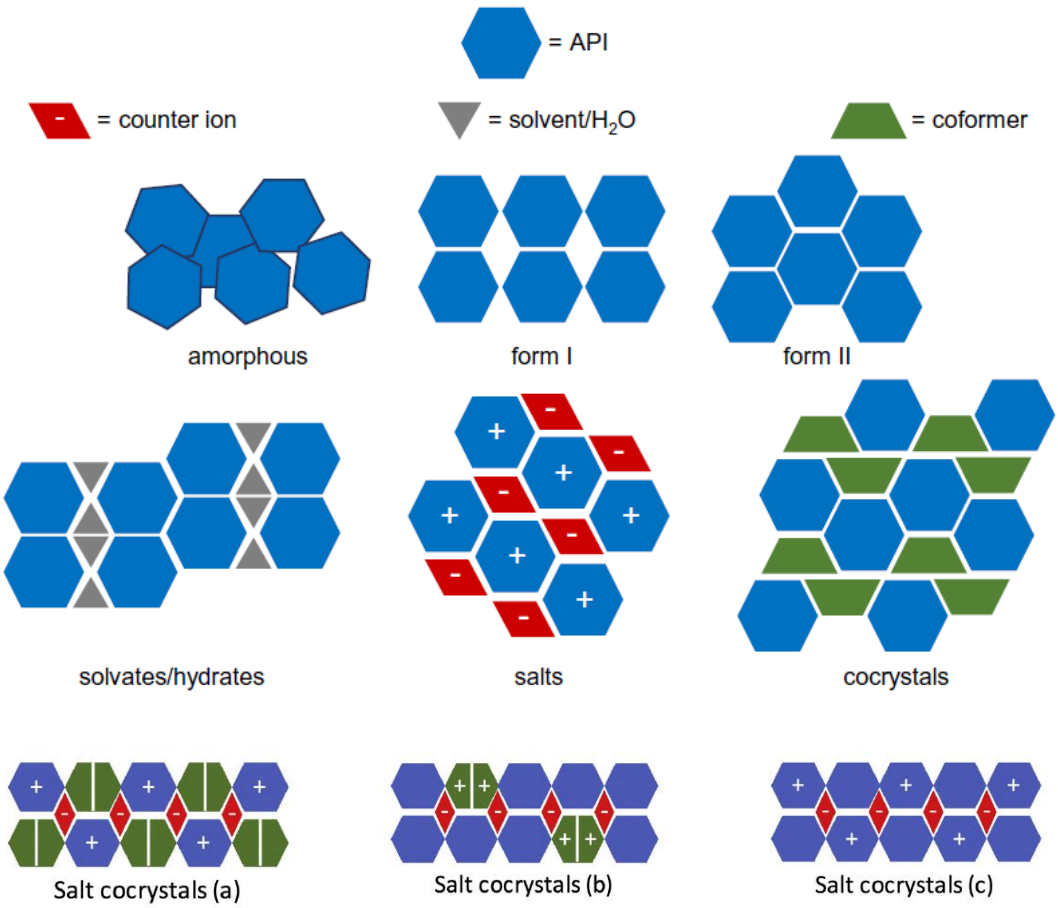


Figure 1.1: The different possible solid-state crystal forms for a generic API (modified from Vioglio et al.).[9] Salt cocrystals of type (a) and (b) are also referred to as cocrystal of a salt and salt of a cocrystal, respectively.

classifying systems by the final properties and uses of each form, rather than more prescriptive definitions (salt vs. cocrystal).[6-8]

A major driving force in the pharmaceutical industry is to improve the solubility, bioavailability and thermal stability of products along with numerous other physical characteristics, such as manufacturing properties like flow characteristics, hygroscopicity, particle size, filterability and compressibility. 65-70 % of APIs are developed as solid form products and crystalline forms are preferred for ease of synthesis, rejection of impurities and as they generally have better stability than amorphous forms.[9] The need for effective property alteration strategies is of increasing importance as it is estimated that in excess of 60% of new APIs under development are poorly soluble.[10, 11] Finding the optimal solid form involves a broad screening process to ensure that the crystallisation landscape is thoroughly explored, discovering the extent of polymorphism and potential multicomponent forms available (Fig.

1.1). To allow for ease of manufacture and the best stability, generally the most thermodynamically stable form should be chosen.[12] Newly determined forms represent an intellectual property (IP) possibility,[13] through either new patents or patent extensions.[14] It is accepted that the bonding behaviour of salts and cocrystals is similar but there is tension added to classifications within industry by the regulatory requirements and IP potential. A detailed definition of the new phase is required for patenting and the correct characterisation is therefore considered crucial.[6] As increasing dissolution and absorption of the API[15, 16] also changes its bioavailability, thorough investigations are required for new forms as there is a chance of adverse effects.

1.1.1. Salts

Salts have been used for ~ 60 years and account for more than 50 % of drugs on the market.[9] They are defined as being a 'chemical compound consisting of an assembly of cations and anions' by IUPAC.[17] To ensure charge balance to the supramolecular structure, the stoichiometry of salts must be exact.[9]

1.1.2. Cocrystals

Cocrystals have been the subject of much debate at both a chemical and regulatory level. The most general definition includes all non-covalently bound multi-component forms, particularly solvates and some metal complexes, although the addition of excluding salts and solvates narrows it to the more generally used definition: a system consisting of multiple components that are neutral and independently solid at room temperature.[9] The effect of cocrystal formation on the properties is similar to that seen for salts as the two forms represent the extremes of a continuum of proton transfer, which results in cases where the definition of a system as a cocrystal or a salt can be based on the position of a single proton.[6] In some cases proton transfer can vary depending on temperature, creating additional challenges for classification. There was initial reticence to view salts and cocrystals as part of a continuum[18]

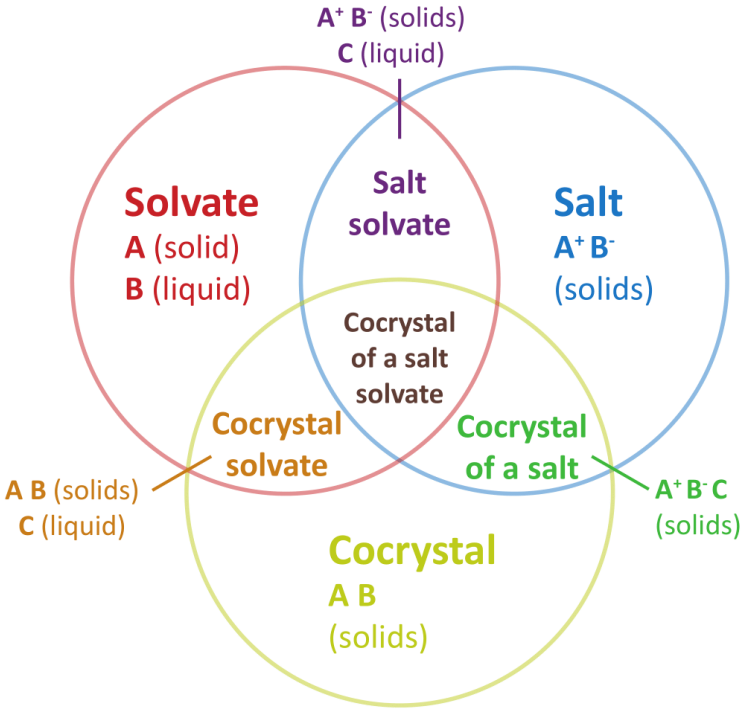


Figure 1.2: Venn diagram approach to the classification of multicomponent crystal forms. Names of forms are chosen to be consistent with their usage in this work (adapted from Grothe et al.2016).[19]

but this has changed as more evidence has been generated, and many now use more of a Venn diagram of forms[19] to classify systems (Fig. 1.2).

1.1.3. Cocrystals of Salts

Cocrystals of salts contain a mixture of ionic and neutral forms of one of the components[20] (some groups also refer to salts of cocrystals, depending on whether the interaction between hetero-molecules is ionic or neutral).[19] It is possible to define a category of salt cocrystals as having a mix of ionic and neutral APIs and co-formers/counter ions within the unit cell (which both of the above would fall into) as done by Voiglio *et al* in Fig. 1.1.[9] Cases where the API is present in both neutral and ionised forms are attractive as both have the same pharmaceutical activity, allowing the potential to maintain drug loading and gain the improved manufacturing performance of a salt without increasing hygroscopicity, as seen for sodium valproate valproic acid.[21, 22]

1.1.4. Solvates

Solvates are systems that contain solvent molecules within the crystal lattice, with the term hydrate used where the solvent in question is water. The choice is generally made to distinguish solvates by whether or not a component is liquid or solid at 293.15 K, room temperature (RT), which is a practical if somewhat arbitrary definition.[6, 9, 20] If a crystalline system has the same structure following de-solvation, it is termed pseudo-polymorphic, whereas instances where the structure changes are polymorphic solvates.[20]

Solvent association can take a range of forms: adsorption onto the surface; adsorption and absorption into disordered regions or defects; inclusion as liquid pockets during crystal growth; and inclusion as part of the crystal packing within the unit cell. Amorphous and disordered systems generally have larger volumes with less tightly packed molecules and weaker intermolecular interactions so they tend to show greater solvent association than crystalline systems.[23, 24] The introduction of solvent can lead to significant problems by introducing disorder and creating metastable systems[25-27] but it can also have a positive effect. Solvent molecules within the lattice can form strong interactions, *e.g.* hydrogen bonds (H-bonds), with APIs, potentially improving the stability of metastable systems[28] and sometimes leading to the formation of flexible clusters.[29] Instances of solvent molecules taking the role of space fillers to stabilise the structure without the presence of strong interactions have also been seen.[30]

Water is exceptionally abundant on earth and has a unique chemistry which allows it to form interactions through a variety of mechanisms, including H-bonding and Van der Waals interactions through dispersion forces and induced dipoles.[24, 31] The orientation of the molecule's two H-bond acceptors (the oxygen lone pairs) and two H-bond donors allows them to form both tetrahedral networks as well as multiple interactions with the components of the crystal system.[20] Charged ions interact strongly with polar molecules such as water,[32] as clearly evidenced by the proportion of hydrates in the Cambridge Structural Database (CSD) that involve highly polar molecules being found to account for more than 90 % of structures.[33]

In the case of salts that have a single strong ionic interaction, where the rest of the molecular surface is left exposed and weakly ionic, hydrate formation is promoted as the remaining molecular surface can be stabilised by association with water.[29] This can lead to instabilities within the crystal lattice so is not generally sought after. The polarity association also explains why hydrate formation is less common within cocrystal systems.[6]

1.1.5. Amorphous Solids

Amorphous solids represent an alternative development route to the crystalline forms discussed above. They are classified as solids that have a random structure without long range order, containing only short range order.[20, 34] They show diffuse X-ray diffraction (XRD), no melting endotherm and no symmetry.[34] Mesophase systems (such as ionic liquids) show intermediate symmetry.[35]

1.1.6. pK_a

pK_a represents the pH at which the solubilised population is 50% charged and 50% neutral. For a difference in pK_a between two components of more than 2-3, proton transfer is considered essentially complete, with the system generally forming a salt, whereas for a difference < 2 it is difficult to predict which form will occur.[36] However, the use of the empirical pK_a rule to determine crystal form is highly debated as the parameters are not determined in the solid state and so their applicability is questionable. It is considered to hold well for salt formation where $\Delta pK_a > 3$, but for components where ΔpK_a is lower, the classification is less certain and detailed characterisation to determine the proton position is required as it impacts upon the interaction energies and can have a significant effect.[37-40] It has also been noted that there is some dependence on the base : acid ratio and the pressure under which formation occurs.[41, 42]

1.2. Molecular Interactions

Non-covalent and supramolecular interactions within crystals can be difficult to characterise and understand due to the close packing of molecules in three dimensions. Only a full understanding of the contribution of each interaction to the overall stability would make it possible to confidently predict which structure is the most stable and even then such a structure might not be experimentally accessible as the intricacies of the energy landscapes for nucleation and crystal growth are poorly understood. Although, judging by the recent Cambridge Blind Tests,[43] crystal structure prediction (CSP) methods have made great progress in accurately predicting energetically feasible structures, the ability to determine the most probable for a given set of crystallisation conditions remains impractical, let alone deriving the conditions for a given structure.[44] Multicomponent systems are even more challenging for CSP, although not impossible.[45] Despite the increased number of degrees of freedom, some success has been possible from the addition of more empirical constraints – for example, ternary cocrystals (with three components) have been both designed and synthesised by Aakeroy *et al*. and Adsmond *et al*., utilising an understanding of supramolecular synthons (discussed below) and, in the latter case, pK_a values and H-bond basicity.[46, 47]

The need in the above examples of additional constraints suggests that, development of increasingly powerful CSP methods aside, a thorough understanding of the key interactions that maintain crystal structures is essential to the further development of crystal prediction and engineering, particularly for the next step of accurately predicting the physical properties for each potential structure.[48] Many different intermolecular interactions can affect crystal packing and it is important to understand the role and characteristics of each. Some relevant examples are introduced below.

Hydrogen bonding This is one of the most important interactions as it's so ubiquitous. They can range in strength but typically have energies between 4-60 kJ mol^{-1},[49] with most of those discussed in this work considered 'strong' with energies of ~17 kJ mol^{-1} and above.[50] They are

	directional and can persist into solution giving it an important role in the association of the API and conformer prior to crystallisation.[6]
π-π	These interactions can be face to face or face to edge and play an important role in stabilising systems with aromatic rings.[51, 52]
Ion-π	They can be equivalent in strength to a H-bond so they tend to form in systems whenever suitable components are present.[53]
Halogen bond	A closed shell interaction between an electron deficient σ-hole and a covalently bonded halogen that lies opposite it. They are not as strong as H-bonds but often occur simultaneously to them in halogen containing systems.[54]
Closed shell	The broader set of interactions under which halogen bonds fall. Other forms, which are less common, include aurophilic and argentophilic interactions[55] (in gold and silver complexes, respectively) and secondary bonding.[56]
Van der Waals	The most important interaction but hard to understand or control. Molecules like to be close packed with as little empty space within the structure as possible. Stabilisation by van der Waals forces is proportional to contact surface area. Non-self complementary shapes are likely to incorporate other molecules, possibly leading to a propensity for salt and cocrystal formation[57] but also to the generation of solvates (particularly where API and conformer are not complementary).[6]

A balance which optimises all the potential interactions and minimises the overall energy leads to the formation of the final crystal structure. When a favoured close packing is at odds with preferred H-bonding, the likelihood of polymorphism increases because of the trade-off

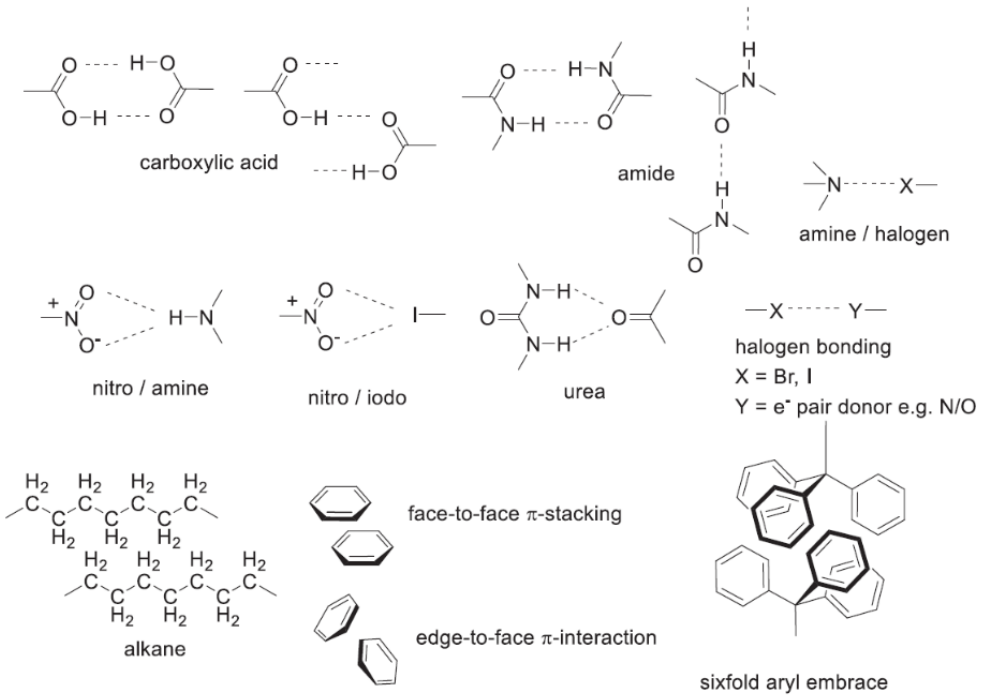


Figure 1.3: Examples of common supramolecular synthons (taken from Berry et al.).[6]

between the two. Alternatively, the number of independent units in the crystallographic unit cell, Z', will be greater than 1.[58]

Supramolecular synthons[59] are patterns of intermolecular interactions that occur frequently and are therefore predictable and reproducible (Fig. 1.3), with the robust motifs conserved across a family of structures.[60] For example, the formation of carboxylic acid dimers, amide NH···O motifs, π-stacking and halogen bonding are all regularly exhibited by molecular crystals, recognition of which has led to their use in the rational design of numerous cocrystals.[2, 4, 6, 19, 61-66] A part of their success within crystal engineering is the hierarchy between different synthons, with analysis of the CSD providing likelihoods of certain synthons forming given the presence of other structural moieties. For example, formation of the pyridine-hydroxyl synthon appears as near certain in the absence of other H-bonding moieties.[67] This is particularly useful for determining whether co-crystallisation of the components is more or less likely than a mixture containing phases of the isolated forms.

The carboxylic acid group is particularly versatile in crystal engineering due to its ability to form both a homo-synthon as well as wide range of hetero-synthons, one of which in particular has been extensively studied, that of carboxylic acid with a heterocyclic base.[60,

[61, 68, 69] It is also generally accepted that they allow easy distinction of crystal form, as the C-O bond lengths of carboxylate have been found to be very similar, while they differ in the neutral molecule (C-OH ~1.2 Å, C=O ~1.3 Å).[9] Both the homo- and hetero-synthons of carboxylic acid are integral to the structures discussed in this work.

A more general method of defining both local patterns of H-bonding and longer range structures has also been developed based on graph notation.[70, 71] Within this notation, different types of hydrogen bonds and patterns can be distinguished simply. Each motif is described by $\mathbf{X}_{\mathbf{a}}^{\mathbf{d}}(\mathbf{n})$, where **X** denotes the type of pattern (**C** – chain, **R** – ring, **D** – dimer and **S** – intramolecular), the superscript and subscript refer to the number of H-bond donors and acceptors involved, respectively, and **n** is the number of atoms involved in the entire repeat unit. This convenient shorthand is utilised for crystal structure descriptions throughout this work.

1.3. Key Properties of Pharmaceutical and Agrochemical Solids

Although most widely associated with changing dissolution, the formation of multicomponent systems can affect a wide range of manufacturing and biopharmaceutical properties that are crucial to successful development and use. Some of those most commonly altered are outlined here:

Dissolution	Increased rates are generally desired and have been reported for both salts and cocrystals *in vitro* (and *in vivo* for salts).[1, 60, 72] Conversely, some systems show delayed dissolution compared to the native API/AI, which notably has potential applications within the agrochemical industry (*e.g.* by slowing the rate at which the AI is washed off leaves).[73]
Melting point	Crucial for determination of physical stability and appropriate design of a manufacturing route since processes such as milling

	and roller compaction will increase the temperature whereas others, like hot melt extrusion, require melting.[74, 75] An analysis shows that the melting point of the multicomponent system is between that of the API/AI and the co-former in ~51 % of cases but it can be either above or below.[76]
Hygroscopicity	Cocrystals are more resistant to hydration than salts (although not necessarily more resistant than the isolated API). As hydration can impact the utility and stability of the product, this is one of the driving factors behind cocrystal investigation.[77]
Compression	Roller compaction and tabletting need good compression characteristics. They can be modified with excipients during formulation but it is sometimes necessary to consider the API's property when a high drug loading is required.[78-81] Co-crystallisation has been shown to have an impact, but it is not straightforward.[82-84]
Pharmacokinetics and -dynamics	The effect of body on drug and drug on body, respectively. Impact of salt or cocrystal formation is complex and still unclear.[76]

It is clear that as well as improving manufacturability, altering the dissolution and pharmaco-kinetics or pharmacodynamics will have a significant impact on the API/AI's release profile and dosage efficiency.

1.4. Characterisation of Pharmaceutical and Agrochemical Solids

Structure determination of any crystalline compound is most commonly achieved by XRD with single crystal XRD (SXRD) the cornerstone of structure solution, particularly for organics, while powder XRD (PXRD) is utilised as a powerful fingerprinting tool, signalling

the presence of or changes to already determined structures. Knowing the crystal structure of a solid form allows modelling and prediction of properties as well as providing a starting point for related structures to be designed. This is therefore an important step in the characterisation of a product. However, XRD, like all techniques, has several limitations and the use of complementary methods can therefore be of great benefit. When dealing with organic molecules, the most obvious of these is that single crystals of sufficient size and quality are not necessarily available at the development stage. This problem has been exacerbated in recent years by the move towards greener chemistry, with an increasing use of solvent free methods such as kneading, grinding and co-melting, all of which produce microcrystalline powders.[85-87] The use of dehydration or de-solvation of solvate forms also tends to cause crystals to crack.[88] It is possible to solve the crystal structure from PXRD data but the compression of the data into a single dimension means that advanced data analysis techniques are required, such as direct space methods and genetic algorithms.[89-92] As such, complementary information can greatly improve success rates, with solid-state nuclear magnetic resonance (NMR) able to provide key data.[93-95]

As expanded upon in Chapter 2, nuclear spins are very sensitive to their local environments, allowing a wealth of information on the local to intermediate length scales to be probed. Many of the interactions involved are anisotropic (orientation dependent) which leads to broad lineshapes in the solid-state, whereas in solution NMR these are averaged by the rapid molecular tumbling, producing narrow resonances influenced by the isotropic components. The development of magic angle spinning (MAS) in the late 1950's was vital to improving solid-state resolution as it achieves averaging of the anisotropic interactions.[96-98] Ongoing advances in hardware and pulse sequences, alongside the development of computational methods, has led to the emergence of the field of NMR crystallography.[99-103] The information inherent to NMR spectroscopy, notably chemical shift, dipolar interactions, spin diffusion, chemical shielding tensors and quadrupolar couplings, provides insight into the crystal symmetry, Z', atomic proximities and bonding arrangements and molecular orientation as well as dynamics and the presence of different phases within a system. Used alongside

molecular modelling and/or quantum mechanical calculations, this information can aid in the development, characterisation and validation of structures.[9, 104, 105] The use of density functional theory (DFT) calculations to both geometry optimise and determine NMR parameters (by employing the gauge-including projector-augmented wave (GIPAW) method) is particularly powerful.[103, 106, 107]

As XRD is inherently insensitive to low atomic number elements, hydrogen atoms are essentially invisible because they have so few electrons. Most hydrogens will therefore be placed at chemically sensible positions, rather than being found in the electron density map, limiting the detail that can be provided on the hydrogen moderated intermolecular interactions by XRD. In particular, determining the crystal form with confidence can be challenging as the overall geometry of the molecules may not differ noticeably despite the change in ionisation.[108, 109] For example, the geometry of the pyridine-carboxylic synthon is expected to be similar for both salts and cocrystals despite the sensitivity to proton position of the surrounding energy landscape.[38] Neutron diffraction can be utilised to provide accurate hydrogen positions, as the neutrons interact with the nuclei of atoms rather than the electron shell and scattering does not increase as a function of atomic mass, unlike the case for X-rays.[110] Like NMR, it is often used as a complementary technique to XRD as the weaker diffraction of neutrons means it requires larger sample sizes that often make obtaining single crystals unfeasible, with the presence of positive and negative scattering lengths adding further complications.[111-113] Analysis of hydrogen bonding networks and hydrogen dense systems is difficult, however, as hydrogens produce a very large inelastic scattering which means samples generally need to be deuterated. Another drawback is that neutron diffraction requires a nuclear reactor or spallation source to generate neutrons so it is only available at a limited number of sites and therefore cannot be implemented as a standard technique. The wider availability and access to solid-state NMR make the development of its use as a complementary crystallographic method highly desirable. NMR crystallography is particularly useful when determining crystal form (salt/cocrystal/*etc.*) or investigating H-bonding motifs and networks.[114-118] The same phenomena in NMR which can facilitate

structure determination also allows both weak and strong intermolecular interactions to be interrogated, with the ^{1}H chemical shifts in particular inherently sensitive to local interactions.[119]

When performing calculations of the NMR parameters relating to a crystal structure, a comparison of the chemical shifts calculated for the full crystal structure to those calculated for individual isolated molecules, as extracted from the geometry optimised crystal structure, can provide additional insight into the crystal packing.[118, 120-124] The difference between the crystal and isolated molecule chemical shifts indicates the presence of intermolecular interactions, with changes of more than 1 ppm being considered significant. Positive values of $\Delta\delta_{\text{Crystal - Molecule}}$ are attributed to the presence of hydrogen bonding, while negative values arise due to ring currents from C–H···π interactions and π- π stacking.[115, 120, 121, 125]

1.5. Thesis Outline: Systems Investigated and Their Characterisation

1.5.1. Systems

The primary systems studied within this work result from the co-crystallisation of a set of methyl or amino substituted pyridines with fumaric acid. Fumarate (F), or hydrogen fumarate (HF), is a pharmaceutically acceptable counterion for salt formation and is already employed to improve the properties of various APIs, including Fe^{3+} ions, bisoprolol and tenofovir disoproxil.[126] Fumaric acid (FA) is a di-carboxylic acid that has three possible conformations which are energetically similar, with either both carboxylic groups in s-*cis* configurations (two s-*cis*), one in s-*cis* and one in s-*trans* (s-*cis*/s-*trans*) or both in s-*trans* configurations (two s-*trans*) (Fig. 1.4, left to right). Computational studies have found that they differ in energy by ~3-5 kJ mol^{-1}, with the two s-*cis* conformation having the lowest energy and being the form present in the room temperature crystalline solid, although experimentally they are observed to be even closer in energy, ~1 kJ mol^{-1}, with approximately equal amounts of each conformer

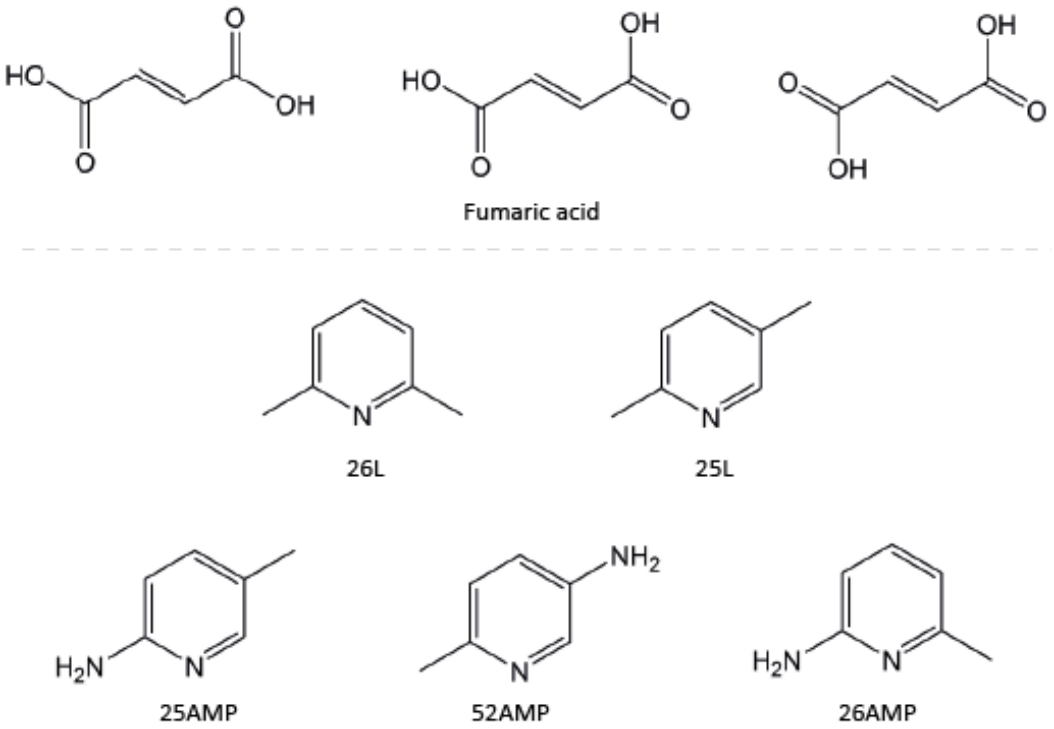


Figure 1.4: Component molecules used to make the multicomponent systems studied in this work. The three possible conformations of fumaric acid (FA) are shown at the top, two s-cis, s-cis/s-trans and two s-trans (left to right). The five bases used are 2,6-lutidine (26L), 2,5-lutidine (25L), 2-amino-5-methylpyridine (25AMP), 5-amino-2-methylpyridine (52AMP) and 2-amino-6-methylpyridine (26AMP).

present.[127, 128] Each of the five pyridine molecules studied in this work are related by a change in substitution position and/or a change of the substituent between a methyl group or an amino group, specifically two dimethylpyridine (lutidine) isomers and three amino-methylpyridine isomers. As shown in Fig. 1.4, the chosen molecules are 2,6-lutidine (26L), 2,5-lutidine (25L), 2-amino-5-methylpyridine (25AMP), 5-amino-2-methylpyridine (52AMP) and 2-amino-6-methylpyridine (26AMP). These bases show a wide range of physical characteristics, with melting points ranging between −15 °C and 287 °C and the known pK_a values varying between 6.18 and 7.41 (Table 1.1).

Table 1.1: Physical properties for each component in Figure 1.4.

Component	Melting (°C)	Boiling (°C)	pK_a	ΔpK_a[a]	H-bond donors	H-bond acceptors
FA	287[129]	290[129]	3.02 & 4.54[130]	-	2	4
26AMP	40-44[131]	208-209[131]	7.41[132]	4.38	1	2
25AMP	76-77[133]	227[133]	7.22[134]	4.19	1	2
52AMP	95-99[135]	-	-	-	1	2
25L	-15[136]	157[136]	6.40[130]	3.47	0	1
26L	-6[137]	143-145[137]	6.65[130]	3.62	0	1

[a] Compared to the value for removing the first carboxylic acid proton from FA.

The resulting multicomponent solid-state forms reported in this thesis include four previously published structures and three newly determined structures. The four published structures are found in the CSD,[138] and are available from the Cambridge Crystallographic Data Centre (CCDC) with the CCDC references and deposition numbers given in Table 1.2.

Table 1.2: Previously published structures downloaded from the CCDC with relevant references and multicomponent form.

Name	Thesis code	Lit. reference	CSD reference	Deposition no.	Form and base
2,6-lutidinium hydrogen fumarate	**26L:HF**	Pan *et al.*[139]	MIBYEB	181445	Salt with 26L
2,5-lutidine hemi-fumarate fumaric acid	**25L:F$_{0.5}$:FA**	Haynes *et al.*[140]	RESGEC	615314	Cocrystal of a salt with 25L
2-amino-5-methylpyridinium hemi-fumarate hemi-fumaric acid	**25AMP:F$_{0.5}$:FA$_{0.5}$**	Hemamalini *et al.*[141]	DUTNUC	788456	Cocrystal of a salt with 25AMP
2-amino-6-methylpyridinium hemi-fumarate dihydrate	**26AMP:F$_{0.5}$:(H$_2$O)$_2$**	Selyani *et al.*[142]	COGCIN	1521964	Salt hydrate with 26AMP

No crystal structures have previously been published for the other three solid state forms (Table 1.3), which have been determined within the course of this study. They have been deposited with the CSD.

Table 1.3: New structures whose structure determination is presented in this thesis, with the CCDC deposition number and multicomponent form.

Name	Thesis code	Deposition no.	Form and base
5-amino-2-methylpyridinium hydrogen fumarate	**52AMP:HF**	1952142-3	Salt with 52AMP
Bis-(2-amino-6-methylpyridinium) fumarate monohydrate	**26AMP$_2$:F:H$_2$O**	1952132-40	Salt hydrate with 26AMP

Bis-(2-amino-6-methylpyridinium) fumarate hemi-fumaric acid	**26AMP$_2$:F:FA$_{0.5}$**	1952129-31	Cocrystal of a salt with 26AMP

As shown in Tables 1.2 and 1.3 and Fig. 1.5, there are three cocrystal of salt systems, two salts and two salt hydrates. The absence of a cocrystal is not surprising considering the pK_a of fumaric acid (3.02 for the first proton and 4.38 for the second) and, therefore, the resulting difference in pK_a (ΔpK_a) between the acid and any given base is less likely to be smaller than 3.[37-40] As discussed above, the boundary between cocrystal and salt is a continuum, however, with borderline cases and those falling around the boundary less easily predicted. For some combinations, the same components are able to crystallise in both neutral and ionic forms as well as more disordered systems (*e.g.* cocrystals of salts, *etc.*). In some cases a range of crystal forms can be produced from the same crystallisation media.[140]

As can be seen in Table 1.1, all of the components have a ΔpK_a of more than three compared to fumaric acid (although this has not been calculated for 52AMP as its pK_a is not reported in the literature). The expectation is therefore that they will all definitely form salts. The smallest difference, ΔpK_a = 3.15, corresponds to 25L which forms a cocrystal of a salt with fumaric acid but so does 26AMP, which has the largest, ΔpK_a = 4.38. It is therefore noted that, within this set, there appears to be no correlation between ΔpK_a and the likelihood of forming a salt as opposed to a cocrystal of a salt, suggesting that the inclusion of both neutral and ionic forms of the same molecule is related instead to the packing of the molecules. Both 25AMP and 26AMP, which have the largest ΔpK_a values, form cocrystals of salts, with 26AMP only forming a classic salt with fumarate when also hydrated. There is also no evident correlation between the number of donors and acceptors (see Table 1.1) and the adopted solid state forms within this small set, with a mixture of forms seen for both lutidines and among the amino-methylpyridines. The combination, and relative positions/orientations of substitution, of both donors and acceptors and of groups that introduce steric hindrance (the methyl groups) seems to have a bigger impact, as will be explored further in the results chapters.

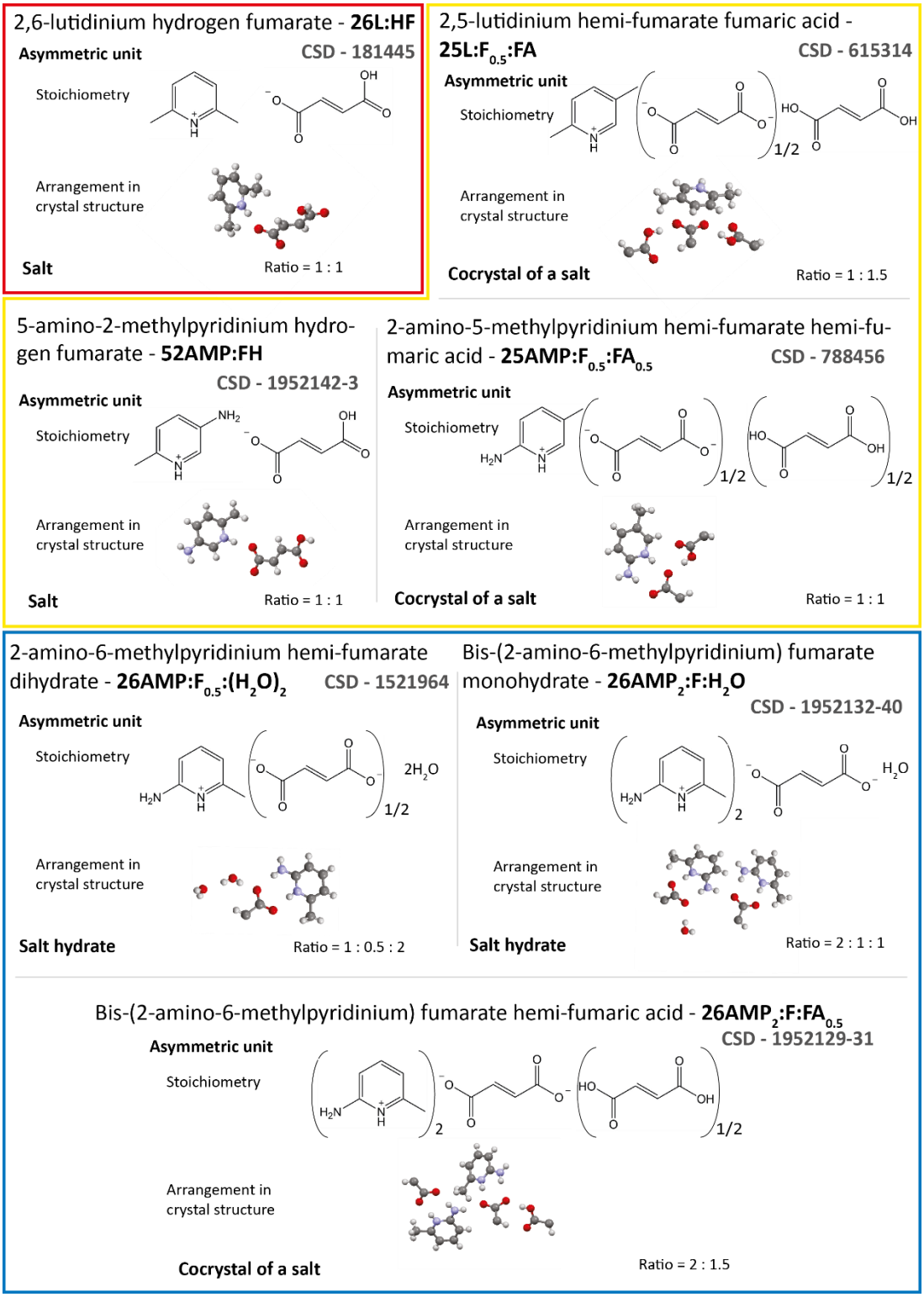


Figure 1.5: Summary of the multicomponent crystal systems studied in this work. All names are given relative to a single base molecule whereas the stoichiometric ratio is here given in terms of the asymmetric unit. Coloured boxes indicate the groups in which the structures are discussed in Chapter 3 (red), Chapter 4 (blue) and Chapter 5 (yellow). Note this Figure is reproduced at the end of this thesis to provide a system reference.

1.5.2. Aims of Research

This thesis investigates the complementarity of XRD and solid-state NMR, alongside DFT calculations, for the characterisation of the multicomponent crystals of small organic molecules introduced in Section 1.5.1. The combined approach is used to probe crystal form, stability and H-bonding networks. Both SXRD and PXRD are employed to determine the initial structures of each compound and investigate their stability on grinding, with variable temperature (VT) experiments conducted for some systems to explore the thermal phase transitions and/or the expansion of the unit cell with temperature. DFT-based geometry optimisations and GIPAW calculations of the NMR parameters were conducted for both the full crystal structure, allowing a comparison with experimental MAS NMR spectra, and for isolated molecules from within the structure, to highlight the intermolecular interactions present. ^{1}H MAS, ^{13}C cross polarisation (CP) MAS solid state NMR spectra and ^{1}H-^{1}H (DQ-SQ and SQ-SQ), ^{1}H-^{13}C and ^{14}N-^{1}H 2D correlation experiments were conducted for each system. Chapter 2 outlines the relevant theory of the key techniques as well as the experimental approaches employed for each. Thermogravimetric analysis (TGA) and differential scanning calorimetry (DSC) were also performed for each system to investigate their thermal properties.

The first three experimental Chapters 3-5 focus on the thorough characterisation of each system by the general approach outlined above (although some system dependence to the approach chosen will be evident). Chapter 3 describes the structure and stability of 26L:HF, including room temperature degradation due to the volatility of the base, 26L. Chapter 4 studies the three structures based on 26AMP and the phase transitions that link each of the pseudo-polymorphs. Chapter 5 explores the dramatic effect due to changes in a single functional group by comparing the structures of 25AMP:$F_{0.5}$:$FA_{0.5}$, 25L:$F_{0.5}$:FA and 52AMP:HF.

The final chapter, Chapter 6, serves to review and compare the above structures, identifying trends and intermolecular structural patterns, and also develops more general

schemes for prediction and analysis of future systems. A possible solid-state NMR approach based on the ^{14}N shift is outlined for determining crystal form for pyridine based systems where the crystal structure is unknown. This is based on DFT calculations of the expected range of ^{14}N shifts depending on the presence of ionic or neutral H-bonding. Finally, CSD searches for both fumarate/fumaric acid and succinate/succinic acid multicomponent crystals reveal trends in both crystal form, that reflect their relative pK_a, and in the observed intermolecular structural patterns.

Chapter 2. Theory and Experimental Considerations

2.1. XRD

Unless otherwise stated, the theory contained within this Section is based upon that presented in: *X-ray Crystallography*. 2nd ed.; Oxford University Press: Oxford, 2015;[143] *Fundamentals of Crystallography*. Second ed.; Oxford University Press: Oxford, 1992;[144] and *Structure from Diffraction Methods. Wiley: Chichester, 2014*;[145]

2.1.1. Crystals

2.1.1.1. Unit cell and crystal lattice

Crystalline solids are those that possess long-range order. Theoretically, they are infinite but, due to finite size, defects and impurities, they are more practically viewed to consist of structures with periodicity over 10^3-10^{20} atomic dimensions. The periodic structure is defined by the unit cell and the crystal lattice (Fig. 2.1).

Unit cell — Parallelepiped whose shape and contents can be repeated (without transformation) to describe the entire crystal. There are generally multiple possible choices but the smallest is chosen by convention. In three dimensions (3D), the side lengths are denoted $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ and the internal angles are $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$.

Lattice — Describes the crystal's periodic structure and the placement of unit cell repeats to build the crystal via translational symmetry, defined by lattice points with directions described by lattice vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$.

The possible relationships between the six 3D unit cell parameters lead to the definition of seven crystal classes, specified in Table 2.1.

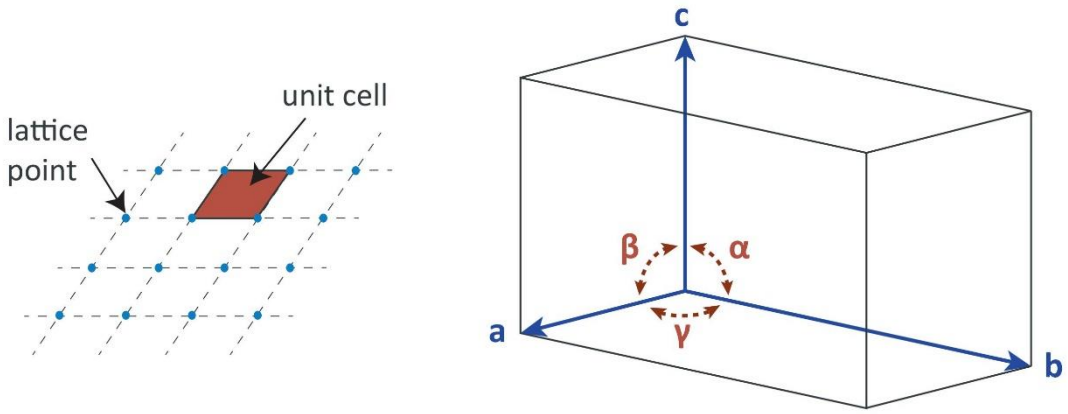


Figure 2.1: Diagram showing the selection of a 2D unit cell and the lattice describing the wider crystal structure (left) and a 3D unit cell with side lengths and internal angles labelled (right).

Table 2.1: The seven crystal classes.

Unit cell parameters		**Crystal class**
a = b = c	$\alpha = \beta = \gamma = 90°$	Cubic
a = b ≠ c	$\alpha = \beta = \gamma = 90°$	Tetragonal
$a \neq b \neq c$	$\alpha = \beta = \gamma = 90°$	Orthorhombic
$a \neq b \neq c$	$\alpha = \gamma = 90°$ $\beta \neq 90°$	Monoclinic
a ≠ b ≠ c	$\alpha \neq \beta \neq \gamma \neq 90°$	Triclinic
a = b ≠ c	$\alpha = \beta = 90°$ $\gamma = 120°$	Hexagonal
a = b = c	$\alpha = \beta = \gamma \neq 90°$	Trigonal/Rhombohedral

There are also four possible lattices for 3D structures: primitive (P), body centred (I) and two forms of face centred (F and C). As can be seen in Fig. 2.2, in primitive lattices, the lattice points are generally located at corners of the unit cell so that each cell contains a single lattice point, meaning that the cells contents are not linked by translational symmetry, whereas in a body centred lattice, there is an additional point in the centre of each cell with fractional coordinates (½, ½, ½), *i.e.* an atom placed at (x, y, z) will be identical to the atom at (x + ½, y + ½, z + ½). The two face centred lattices have an additional lattice point at either the centre of all (F) or one (C) of the unit cell faces. These lattice points therefore sit at the fractional coordinates (0, ½, ½), (½, 0, ½) and (½, ½, 0). By convention, when there is only one additional face lattice point it lies on the *ab* face, hence being denoted C. As with the body

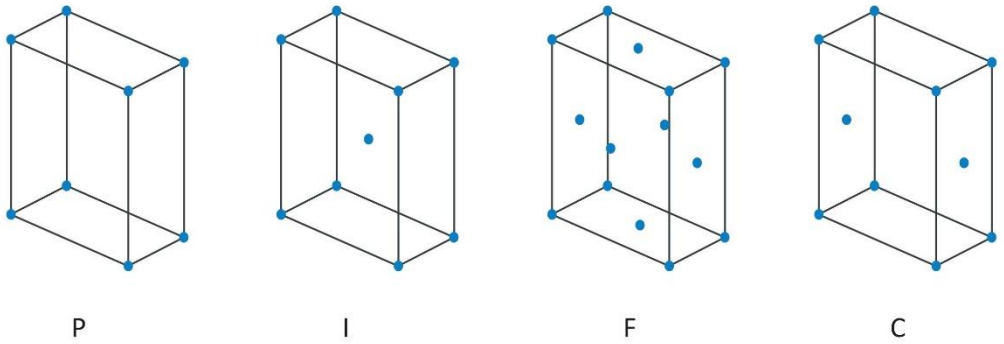


Figure 2.2: Possible 3D crystal lattices. From left to right: primitive (P), body centred (I) and the two face centred lattices (F and C).

centred lattice, each additional lattice point (three or one for F and C, respectively) means that any given atom will have an identical one at the distance of the lattice point's fractional coordinates. Combining the possible crystal classes and lattice types leads to fourteen unique lattices, known as the Bravais lattices (Table 2.2).

Table 2.2: The fourteen Bravais lattices.

Crystal classes	**Bravais lattices**
Cubic	P I F
Tetragonal	P I
Orthorhombic	P I F C
Monoclinic	P C
Triclinic	P
Hexagonal	P
Trigonal/Rhombahedral	P

The contents of the unit cell are described by the asymmetric unit and the symmetry operations of the *space group*. The asymmetric unit is the smallest group of atoms or molecules that can be used to describe the crystal structure. The unit cell usually contains multiple copies of the asymmetric unit that are related to each other via symmetry operations, *e.g.* rotations, inversions, reflections and translations (Fig. 2.3). Simple operations can be combined to form more complex symmetry elements, such as:

- Inversion axis (rotation and inversion)
- Screw axis (rotation and translation)
- Glide planes (reflection and translation)

The parameter Z is used to denote the number of formula units within the unit cell, while Z' denotes the number of atoms or molecules within each asymmetric unit.

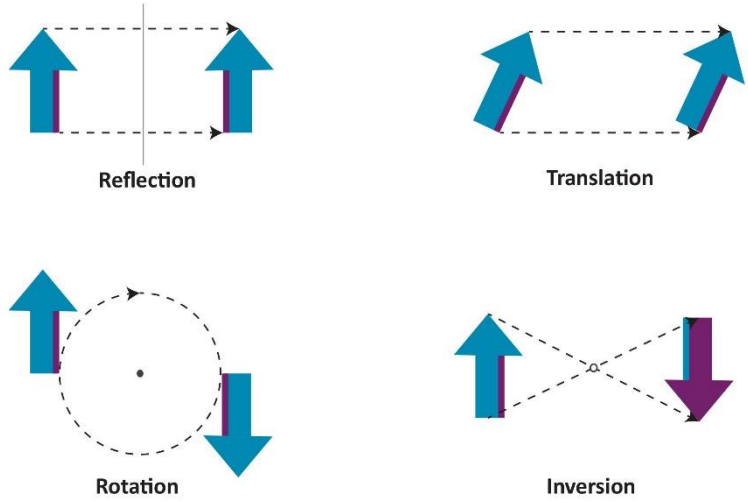


Figure 2.3: Simple symmetry operations

Applying these additional symmetry operations to the Bravais lattices leads to the definition of 230 distinct space groups. The key systems considered within this work fall into two space groups: $P\bar{1}$ and $P2_1/c$.

- $P\bar{1}$ is one of the two triclinic space groups. It is the centrosymmetric form of $P1$. The only symmetry elements that correspond to $P1$ are lattice translations along ***a***, ***b*** and ***c*** whereas $P\bar{1}$ also possesses a point of inversion, generally defined to lie at the origin.
- $P2_1/c$ is a monoclinic space group with a primitive lattice. It has a two-fold screw axis along *b* (a rotation by 180° followed by translation of ½ the lattice vector) and a glide plane along *c* (a reflection followed by translation), resulting in an inversion centre.

2.1.1.2. Crystal planes

Crystal planes are geometrical constructs to help illustrate the diffraction process (Fig. 2.4). Each family of planes intersect all the lattice points and are parallel and equally spaced:

Miller indices **(hkl)** indices that indicate the family of planes that intersect the lattice vectors at *a/h*, *b/k* and *c/l*. They are always integers, they can be positive, negative (commonly denoted by addition of a bar) or zero.

***d*-spacing** d_{hkl}, the distance between neighbouring planes. Its relationship to the lattice parameters is dependent on the crystal class (see Table 2.3).

Table.2.3: Definition of the d-spacing for each crystal class.

Crystal class	d_{hkl} in terms of *abc* and *hkl*
Cubic	$\frac{1}{d^2} = \frac{h^2 + k^2 + l^2}{a^2}$
Tetragonal	$\frac{1}{d^2} = \frac{h^2 + k^2}{a^2} + \frac{l^2}{c^2}$
Orthorhombic	$\frac{1}{d^2} = \frac{h^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2}$
Monoclinic	$\frac{1}{d^2} = \frac{1}{\sin^2\beta}\left(\frac{h^2}{a^2} + \frac{k^2\sin^2\beta}{b^2} + \frac{l^2}{c^2} - \frac{2hl\cos\beta}{ac}\right)$
Triclinic	$\frac{1}{d^2} = \frac{\left(\begin{array}{l}\frac{h^2}{a^2\sin^2\alpha} + \frac{2kl}{bc}(\cos\beta\cos\gamma - \cos\alpha) \\ + \frac{k^2}{b^2\sin^2\beta} + \frac{2hl}{ac}(\cos\alpha\cos\gamma - \cos\beta) \\ + \frac{l^2}{c^2\sin^2\gamma} + \frac{2hk}{ab}(\cos\alpha\cos\beta - \cos\gamma)\end{array}\right)}{1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma}$
Hexagonal	$\frac{1}{d^2} = \frac{4}{3}\left(\frac{h^2 + hk + k^2}{a^2}\right) + \frac{l^2}{c^2}$
Trigonal/ Rhombahedral	$\frac{1}{d^2} = \frac{(h^2 + k^2 + l^2)\sin^2\alpha + 2(hk + kl + hl)(\cos^2\alpha - \cos\alpha)}{a^2(1 - 3\cos^2\alpha + 2\cos^3\alpha)}$

2.1.2. Bragg Diffraction

When radiation is scattered by the crystal lattice, the incoming and outgoing waves are related by the Laue equations. As periodic repeat distances in crystalline solids are comparable to the wavelength of X-rays, the scattered waves can interfere with each other. In the ordered array

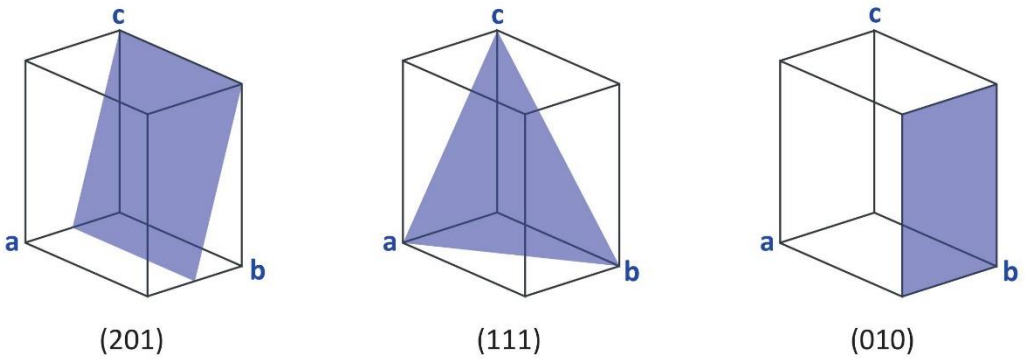


Figure 2.4: Three examples of crystal planes described by different Miller indices.

of a crystal lattice, scattering from the subsequent lattice planes can produce constructive interference if the difference between the path lengths of the scattered waves is equal to an integer multiple of their wavelength, λ. The path difference is determined by the interplanar spacing, d, and the angle of incidence of the radiation, θ (Fig. 2.5). Constructive interference occurs when:

$$2d \sin\theta = n\lambda \tag{1}$$

where n is an integer. First derived by W. H. and W. L. Bragg in 1913, Eqn. 1 is commonly referred to as the Bragg condition. When it is met, a diffraction peak is produced, known as a Bragg peak. The net scattering angle, with respect to the incident beam, is 2θ.

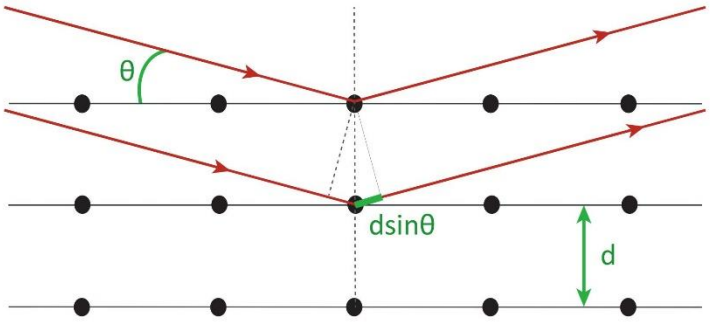


Figure 2.5: Schematic of Bragg diffraction

2.1.2.1. Reciprocal lattice

The reciprocal lattice is the Fourier transform (FT) of the direct crystal lattice which exists in momentum space, k-space. A set which contains an infinite number of crystallographic planes in the direct lattice can be represented by a single vector, or the point at the end of a vector, in the reciprocal lattice. As the momentum difference between incoming and diffracted X-rays of a crystal is a reciprocal lattice vector, the diffraction pattern of a crystal can be used to determine the reciprocal vectors of the lattice. The symmetry is maintained in k-space but diffraction peaks are represented only by reciprocal lattice points with varying intensity (whereas all unit cells are the same in direct space).

2.1.2.2. Fourier transforms

There are two numerical values associated with each reflection in the diffraction pattern, the amplitude $|F|$ and the phase φ of the resultant diffracted wave (only the amplitudes are obtained experimentally). Each resultant wave is labelled by its Miller indices and can be written in a coordinate representation as:

$$F(hkl) = |F(hkl)|e^{i\varphi(hkl)} \tag{2}$$

F(hkl), the structure factor, therefore represents the amplitude and phase for reflection (*hkl*) and is given by the FT of the electron density ρ with respect to that reflection:

$$F(hkl) = \int \rho(xyz)e^{2\pi i(hx+ky+lz)}dV \tag{3}$$

As Eqn. 3 is an integral over the entire unit cell and contains a continuous function, *ρ(xyz)*, it is inconvenient to actually calculate. Expressing the electron density in terms of individual atoms that can be summed over is more practical. As the incident beam is scattered off the electrons in each atom, the number of electrons present has an effect on the variation of intensity with angle (with maximum intensity at $2\theta = 0$, where all the X-rays scattered off the electrons are in phase). This variation is captured by the atomic scattering factor, *f(θ)*, a function of $\sin\theta/\lambda$. For atoms within a crystalline solid, this is modified by a further term to

capture the effects of isotropic vibration with an isotropic displacement parameter, *U*. Eqn. 3, therefore, becomes:

$$F(hkl) = \sum f_j(\theta) e^{\frac{-8\pi^2 U_j \sin^2\theta}{\lambda^2}} e^{2\pi i(hx_j + ky_j + lz_j)} \quad (4)$$

The sum over all the atoms within the unit cell represents the sum over multiple waves to give the resulting reflection for each direction, allowing the expected diffraction pattern to be calculated for a given structure.

The reverse FT of the diffraction pattern, which for crystals is a sum over the discrete reflections, gives the electron density:

$$\rho(xyz) = \frac{1}{V} \sum_{h,k,l} |F(hkl)| e^{i\varphi(hkl)} e^{-2\pi i(hx+ky+lz)} \quad (5)$$

Eqn. 5 constitutes a sum over the structure factors of all diffracted beams, including their amplitudes and phases, with a phase shift for each geometrical position in the resulting image relative to the origin of the unit cell. As mentioned above, *φ(hkl)* is unknown for experimental reflections, making it impossible to calculate Eqn. 5. This is referred to as the 'phase problem'.

2.1.3. Measurement and Data Analysis

XRD is studied using an X-ray diffractometer, consisting of a radiation source, a goniometer and a detector. The measured property is the intensity of the diffracted X-rays as a function of the detector angle, 2θ.

Source *X-ray tube*: where a high energy beam of electrons is accelerated towards a metal target leading to its excitation and subsequent emission of energy in the X-ray range.

	Synchrotron: the radial acceleration of relativistic charged particles causes them to lose energy by emitting electromagnetic radiation, including X-rays, with high flux, brilliance and stability.
Detector	*Gas-filled*: where X-ray photons ionise (e. g. xenon) and these ions are then accelerated into a detector which converts their electrical signal into voltage. *Scintillation*: X-rays hit a phosphor screen which emits visible range photons which strike the detector to be converted to voltage pulses. *Solid state*: X-rays photons hit a semi-conductor and are converted to an electron-hole pair.
Goniometer	Allows the precise measurement of the angle at which Bragg diffraction occurs.

SXRD is performed on a single unfractured crystal, ideally between 30-300 μm in size. This produces clear diffraction spots but, for any given orientation of the crystal, the Bragg condition will be satisfied for very few reflections. Rotation of the crystal is necessary to allow reflection from more lattice planes to occur. Different portions of the diffraction pattern are also recorded separately, to prevent loss of information due to overlap reflections (the result of compressing 3D information into two dimensions).

Assignment of the correct *(hkl)* indices to the observed reflections allows the six unit cell parameters to be determined via the Bragg equation. It may also be possible to determine the space group from systematic absences and comparison of symmetry equivalent intensities. As this can be performed with a small subset of the complete diffraction pattern, it is commonly implemented as part of a preliminary screening scan, as the quality of the crystal can also be determined before proceeding to a full experiment.

2.1.3.1. Structure solution/refinement

Corrections are often required following data collection to account for variation in X-ray intensity and absorption as well as crystal deterioration and non-uniformity in the detector response. Corrected intensities are proportional to $|F|^2$. Some reflections may also be rejected from consideration if the standard uncertainty, σ, in the intensity (and therefore in $|F|$) is too high.

The 'phase problem' can be overcome in a variety of ways, chosen depending on the properties of the system in question.

Patterson Synthesis	Based on the FT of the reflection intensities, $|F|^2$, rather than $|F|$, a map of vectors between atoms in the structure is produced. Common for systems containing some heavy atoms as vectors involving them produce significant peaks in the map allowing their positions to be determined.
Direct Methods	Utilise probability relationships among the reflection phases, which depend on their relative intensities. They are commonly used for systems with similar weight atoms and consist of trial and error of phases from a considered starting point (the probable phase relationships between the most important reflections).
Dual-space Methods	Exploit the limited information available in direct space (from the crystal structure) as well as in reciprocal space (the diffraction pattern). The model is repeatedly interconverted between direct and reciprocal space between iterations of modification.

The resulting structural model is refined through a least-squares approach by varying the numerical parameters that describe the structure and comparing calculated structure factor amplitudes, $|F_c|$, to those observed experimentally, $|F_o|$, until they match as closely as possible. To aid the refinement process, constraints may be applied to fix certain parameters to

particular values or restraints, introducing penalties within the least squares analysis to prevent a parameter varying too far from a target value.

2.1.3.2. R-factors

At each stage of the process the quality of the data and structure is monitored using *R*-factors. During data reduction, the self-consistency of the data can be determined from the merging error, R_{int}, and a measure of the signal to noise ratio, R_{sigma}:

$$R_{int} = \frac{\sum\left|{F_o}^2 - {F_o}^2(mean)\right|}{\sum {F_o}^2} \tag{6}$$

$$R_{sigma} = \frac{\sum \sigma\left({F_o}^2\right)}{\sum {F_o}^2} \tag{7}$$

During structure solution and refinement of the structural model, a commonly used *R*-factor is *R*1:

$$R1 = \frac{\sum\left||F_o| - |F_c|\right|}{\sum|F_o|} \tag{8}$$

Alternatively, F^2 may be used, as for the merging *R*-factors, to produce *R*2 and weights may be incorporated to account for the variation in reliability between reflections, for example in *wR*2:

$$wR2 = \sqrt{\frac{\sum w\left({F_o}^2 - {F_c}^2\right)^2}{\sum w\left({F_o}^2\right)^2}} \tag{9}$$

2.1.4. PXRD

While diffraction from a single crystal in a given orientation (SXRD) produces discrete Laue spots corresponding to its reflections (Fig. 2.6a), a set of crystals (of the same material) in different orientations will produce a set of overlapping diffraction patterns which will also have different orientations. On an area detector, a given reflection will therefore appear as identical spots around a circle (for which the scattering angle is 2θ) corresponding to each

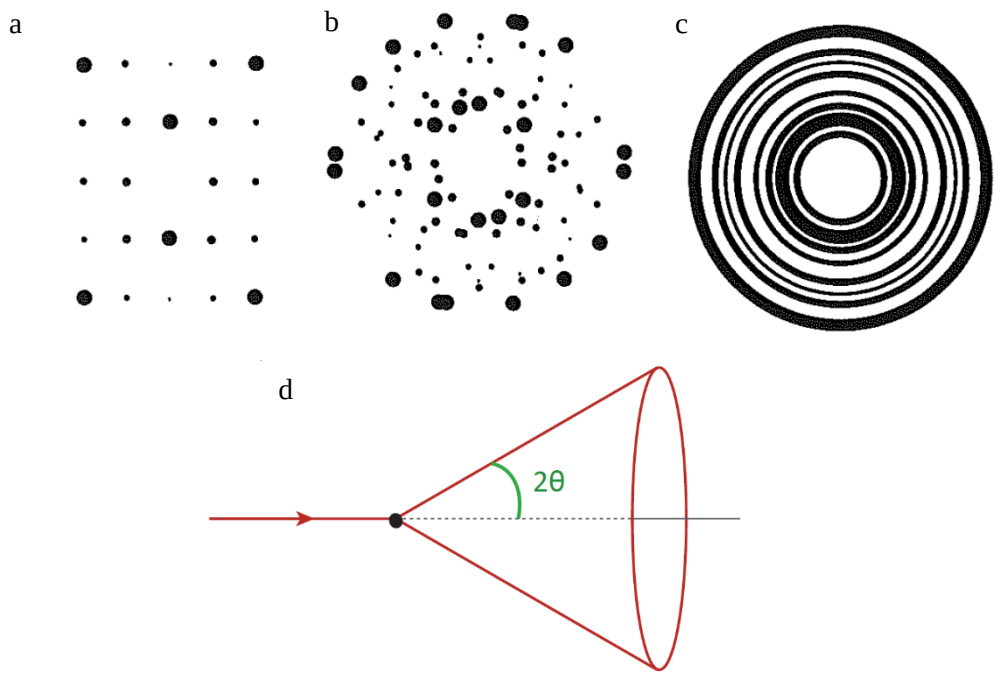


Figure 2.6: The expected diffraction pattern produced on an area detector by (a) a single crystal, (b) four crystals with a random relative orientations and (c) a microcrystalline powder (reproduced from X-ray Crystallography 2nd ed.; Oxford University Press: Oxford, 2015). Also shown is (d) a cone of diffracted X-rays produced by each reflection of a microcrystalline powder sample.

crystal (Fig. 2.6b). A microcrystalline powder theoretically contains an isotropic distribution of every possible crystal orientation randomly throughout, producing a full circle as the resultant cone of diffracted X-rays reaches the detector (Fig. 2.6c and d). Although area detectors are used for PXRD, it is common practice to record only a radial strip of the diffraction cone, recording intensity as a function of angle.

Some crystal morphologies will naturally diverge from an isotropic distribution of crystallite orientations when the powder is packed into a sample holder. Plate and rod shaped crystals in particular tend to align with each other, introducing preferred orientation to the powder and resulting in deviations from the expected reflection intensities. Ensuring samples are fine grained and spinning them during data collection can help to minimise this effect.

Although the compression of the 3D structural information into one dimension makes it difficult, structure solution from powder data is possible. The overlap in reflections for complex structures, with larger unit cells and lower symmetry, make patterns difficult to index and the effects of preferred orientation more significant. Although high resolution data and

advanced techniques have increased the use PXRD for structure solution, it has always been an important tool for material identification and refinement.

2.1.4.1. PXRD refinement

Pawley and le Bail are both profile fitting methods which do not require a structural model, merely constraints on the unit cell parameters, peak shape and background parameters. The size and symmetry of the unit cell are used to constrain the 2θ values, allowing these to be refined against the experimental pattern, but the intensities of each peak are unconstrained to find the optimal fit to the data. These fitting methods are therefore routinely used prior to the Rietveld method as they can be viewed as the best possible fit and *R*-factors against which the success of the Rietveld fitting to the structural model can then be judged. Unlike in the le Bail method, Pawley fitting also generates a co-variance matrix correlating the extracted experimental peak intensities that can theoretically then be used for structure solution alongside these intensities. Within the Rietveld method, the intensities of every data point are treated as refinable parameters, not only the maximum intensity of each experimental reflection, preventing overlaps in experimental reflections inhibiting successful refinement. A full structural model is required for the experimental powder pattern to be refined against, to allow determination of the relative peak intensities in the calculated powder pattern. As in SXRD, least-squares methods are used to find the most optimal fit between experimental data and the structural model. The *R*-factors used to judge the quality of the fit are also analogous.

The most straight-forward is the weighted profile *R*-factor, R_{wp}, roughly equivalent to *wR*2 (similarly *R*2 is ~ equivalent to R_p) although, as it works with the intensities (of every data point for Rietveld refinements, rather than just each reflection) and not structure factors, it is written in terms of *I* rather than F^2:

$$R_{wp} = \sqrt{\frac{\sum w(I_o - I_c)^2}{\sum w(I_o)^2}} \tag{10}$$

For an ideal model, where the average discrepancy between observed and calculated is the same as σ, the 'ideal' R_{wp} can be determined, referred to as the expected *R*-factor, R_{exp}. For a number of data points *N*:

$$R_{exp} = \sqrt{\frac{N}{\sum w(I_o)^2}} \tag{11}$$

Within the Rietveld method, due to the refinement of at all data points, a mechanism was developed to apportion the intensity of peaks containing multiple overlapping reflections on the basis of the contribution to the intensity of the calculated reflections. This allows values of F_o to be estimated and therefore a powder equivalent of *R*1 to be determined, R_{Bragg}.[146]

2.1.5. Experimental Details

Crystals were initially selected for SXRD using polarised light microscopy with an Olympus SZ61 Stereomicroscope. Those that appeared by shape and birefringence to be single crystals were chosen. SXRD was carried out under Cu K_α (1.5406 Å) on either a Rigaku Oxford Diffraction SuperNova diffractometer with an Atlas S2 charge-coupled device (CCD) detector equipped with an Oxford Cryosystems N-Helix cooling system (for 26L:HF and 52AMP:HF) or a Gemini R diffractometer with a Ruby CCD detector equipped with an Oxford Cryosystems Cobra for cooling ($26AMP_2$:F:$FA_{0.5}$, 26AMP:$F_{0.5}$:$(H_2O)_2$, $26AMP_2$:F:H_2O, 25L:$F_{0.5}$:FA and 25AMP:$F_{0.5}$:$FA_{0.5}$). Crystal screening was conducted at room temperature. CrysAlisPro[147] data-collection and processing software was used, allowing crystals to be checked for quality and giving a preliminary unit cell determination by using a short pre-experiment prior to full data collection. This pre-experiment was used to screen a large number of crystals from each crystallisation, with full data collection run for unknown structures or if a discrepancy was identified between the experimental unit cell parameters and those found in the CCDC. Following full data collection, ShelXL[148] was used for structure solution and a least-squares refinement was run, using the Olex2[149] software. Temperatures down to 100 K

were employed to monitor the contraction of the unit cell and both 26AMP:F hydrates were also recorded at temperatures up to 340 K.

Following screening by SXRD, the most crystalline components of each crystallisation were ground to a fine powder and the structure was checked by PXRD to determine bulk purity and ensure no changes had occurred under grinding, by comparing the experimental powder pattern to the pattern predicted from the crystal structure. Standard PXRD was performed on a Panalytical X'Pert Pro MPD equipped with a curved Ge Johansson monochromator, giving pure Cu $K_{\alpha 1}$ radiation and a solid state PiXcel detector. The powder samples were mounted on a zero-background offcut-Si holder, spinning at 30 rpm. Each sample was run with a step size of 0.013° and time per step ranged from 750 to 3500 ms, depending on the sample volume available. VT-XRD of 26AMP:$F_{0.5}$:$(H_2O)_2$ was performed on the same instrument with an Anton Paar HTK1200N spinner chamber and a time per step of 200 s. Temperatures were increased from room temperature (RT), ~25 °C, to 150 °C in steps of 10 °C from 30 °C onwards, and then decreased back to RT.

For 26L:HF, static transmission PXRD experiments were performed on a Xenocs Xeuss 2.0 SAXS diffractometer under Cu K_{α} radiation and a Pilatus 300 K area detector. The powders were loaded into 1.0 mm diameter borosilicate capillaries. High resolution PXRD patterns were collected on beamline I11 at the Diamond Light Source (DLS), UK, using a wavelength of 0.8249 Å. Powders were loaded into 0.7 mm diameter borosilicate capillaries, mounted on the beamline on a spinning brass holder. Diffraction patterns were recorded using both the position sensitive detector (PSD) and multi-analysing crystals (MAC). Half-way through the MAC scans (20 min), the position of the capillary in the beam was moved to prevent sample degradation from the beam. PSD scans (~ 2 min) were recorded at the starting position on the capillary before and after the MAC scan to check for any change. A comprehensive analysis of the powder patterns was undertaken using TOPAS Academic v6,[150] which was used for Le Bail[151] and Rietveld[152] refinements.

2.2. NMR

Unless otherwise stated, the theory contained within this Section is based upon that presented in: *NMR: The Toolkit*. Oxford University Press: Oxford, 2000;[153] *Introduction to Solid-State NMR Spectroscopy*. Blackwell: Oxford, 2004;[154] *Spin Dynamics*. Wiley: Chichester, 2001;[155]. and *Understanding NMR Spectroscopy*. Wiley: Chichester, 2010.[156]

2.2.1. Introduction to Basic Principles and Terminology

Nuclei have an intrinsic property called spin which is a form of angular momentum. This is characterised by a nuclear spin quantum number, *I*. The nuclear spins are formed by combining the spins of their protons and neutrons. The rules by which spins are summed mean that isotopes with even mass numbers have integer or zero spins and those with odd mass numbers have half-integer spins.

A nuclear state with spin *I* is (2*I* + 1) degenerate but application of a magnetic field breaks this degeneracy, as with ordinary angular momentum. The splitting of the nuclear spin levels is called nuclear Zeeman splitting. The energy of a single nucleus in an external field, $\boldsymbol{B_0} = B_0\boldsymbol{k}$ (where $\boldsymbol{k}$ is a unit vector), is given by the Zeeman Hamiltonian:

$$\hat{H} = \boldsymbol{\mu} \cdot \boldsymbol{B_0} \tag{12}$$

where $\boldsymbol{\mu}$ is the magnetic moment and is defined as:

$$\boldsymbol{\mu} = \gamma \boldsymbol{I} \tag{13}$$

γ is the gyromagnetic ratio, a proportionality constant for each nuclear species. The *z*-axis, by convention taken to lie along the external magnetic field, is considered the longitudinal direction. The projection of the spin angular momentum onto *z* is:

$$I_z = m\hbar \tag{14}$$

The magnetic quantum number, m, can take any value at integer steps between $-I$ and $+I$. When the spin angular momentum is aligned along the same direction as $\boldsymbol{B_0}$, Eqn. 12 can be written as:

$$\hat{H} = -\gamma \hat{I}_z B_0 = \omega_0 \hat{I}_z \tag{15}$$

ω_0 is the Larmor frequency, corresponding to the splitting between energy states under the applied magnetic field, as shown in Fig. 2.7 for a spin-½ nucleus. It is an angular frequency proportional to γ and B_0:

$$\omega_0 = -\gamma B_0 \tag{16}$$

As will be explained below, this splitting of the energy levels due to quantisation and is sensitive to subtle changes in local environments since the experience magnetic field varies slightly from B_0. It is the foundation of NMR, producing a measurable property containing a wealth of information on chemical environments and dynamics. Nuclei with $I = 0$ have no magnetic moment and therefore no splitting, so are not observable.

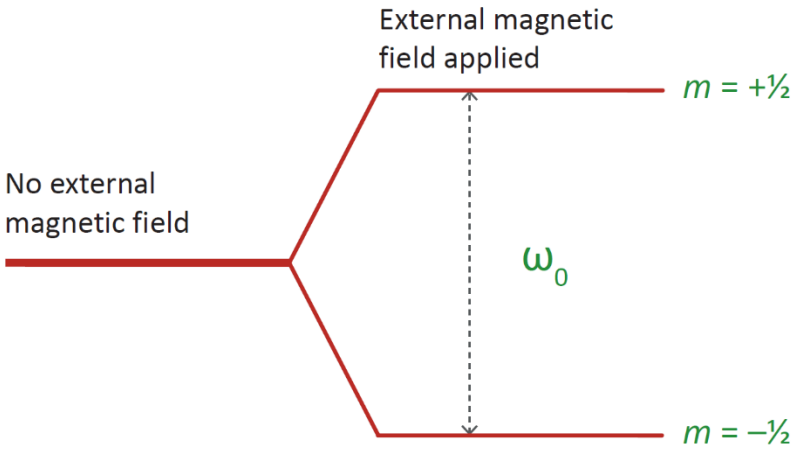


Figure 2.7: Zeeman splitting of a spin-½ nucleus upon application of an external magnetic field, corresponding to an energy difference between states of ω_0.

In a spin-½ system, the two states are commonly labelled α and β, corresponding to the z-component of the spin angular momentum being aligned parallel and anti-parallel to the magnetic field, respectively. At thermal equilibrium, statistical physics denotes that the population, p, of states at different energy levels is governed by a Boltzmann distribution:

$$\frac{p_\alpha}{p_\beta} = e^{\frac{\omega_0}{k_b T}} \tag{17}$$

k_bT, Boltzmann's constant multiplied by temperature, corresponds to the thermal energy. This results in a small bias in favour of the lower energy state and net magnetisation, that is the vector sum of the individual magnetic moments in a bulk sample, of a sample along the same axis as B_0.

To push the system out of thermal equilibrium, and produce a signal that can be observed, the magnetisation must be rotated into the transverse, *xy*, plane. A radio frequency, *rf*, pulse with an oscillation frequency, ω_{rf}, is applied perpendicular to B_0. It constitutes a weak linearly oscillating magnetic field, B_1, with amplitude $|B_1|$, duration t and phase φ, described by:

$$B_1 = |B_1| \cos(\omega_{rf} t + \varphi) \tag{18}$$

When ω_{rf} is close to ω_0, the magnetisation nutates about B_1 at the nutation frequency, ω_1 (Fig. 2.8):

$$\omega_1 = -\gamma B_1 \tag{19}$$

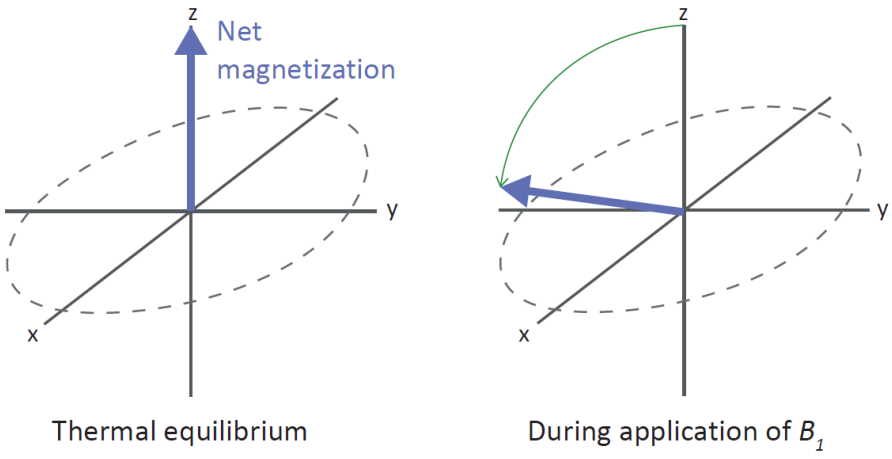


Figure 2.8: The effect of applying an rf pulse along the x-axis, causing the net magnetisation to nutate about B_1.

Following nutation into the transverse plane, the individual spins, and hence the bulk magnetization, precess about the *z*-axis at ω_0, producing a changing magnetic field. This generates a current in a coil through Faraday's law of electromagnetic induction, the NMR signal. As most samples studied by NMR generate spectra with multiple resonances, ω_{rf} cannot be set to exactly ω_0 for all of them. The difference between these two frequencies is the resonance offset, Ω, and is often non zero. A rotating frame of reference is often used to make the interpretation of the evolution of the system with time more comprehensible. By treating B_1 as comprising two counter rotating fields, with frequencies $+\omega_{rf}$ and $-\omega_{rf}$, the oscillating pulse may be treated as static as only the component rotating in the same sense as ω_0 will be retained. The other component is hundreds of megahertz (MHz) off-resonance and can therefore be ignored. The rotating frame rotates at ω_{rf}. This means that, in the rotating frame, precession occurs under a residual field determined by Ω.

The application of an *rf* pulse displaces the system from thermal equilibrium but, once the pulse has been removed, the system will return to this state. The spins dissipate energy through two relaxation mechanisms in order to return to their low energy alignment with B_0 along the *z*-axis. Transverse relaxation, T_2, describes the loss of coherence of magnetisation in the transverse plane as energy is lost to other spins within the sample. Longitudinal relaxation, T_1, describes the return of magnetisation to lie along the *z*-axis as energy is dissipated to the lattice. T_1 determines the rate at which experiments can be repeated.

Signals are measured relative to the same rotating frame. As the time domain signal (as recorded) is sensitive to the sign of Ω, it is necessary to record two sets, when phases are offset from each other by 90°, which constitute the real and imaginary parts of the time domain signal. This is called quadrature detection and is achieved by mixing down with a reference signal whose phase φ differs by 90°. Therefore it is possible to sign discriminate the frequency spectrum following FT of the recorded free induction decays (FIDs).

2.2.2. Quantum-Mechanical Notation

To understand the quantum mechanical description of NMR, as presented from section 2.2.3 onwards, it is useful to first introduce the notation that will be employed. Also, depending on the situation, different levels of theory may be required to analyse the system so a range of approaches are outlined.

2.2.2.1. Bra and ket

It is convenient to describe systems with multiple states with the Dirac bra and ket notation. Any arbitrary state, ket $|\psi\rangle$, is the superposition of orthogonal basis states, for example $|1\rangle$ and $|2\rangle$ with coefficients c_1 and c_2. Alternatively, $|\psi\rangle$ can be written in vector notation as a column vector, $\boldsymbol{\Psi}$:

$$|\psi\rangle = c_1|1\rangle + c_2|2\rangle \qquad \text{or} \qquad \boldsymbol{\Psi} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \tag{20}$$

Each ket has a corresponding bra, $\langle\psi|$, which can also be written as a row vector – the complex conjugate transpose of $\boldsymbol{\Psi}$:

$$\langle\psi| = \langle 1|{c_1}^* + \langle 2|{c_2}^* \qquad \text{or} \qquad \boldsymbol{\Psi}^\dagger = \begin{pmatrix} {c_1}^* & {c_2}^* \end{pmatrix} \tag{21}$$

The orthogonality of $|1\rangle$ and $|2\rangle$ means the scalar product of $\langle\psi|$ and $|\psi\rangle$ is given by:

$$\langle\psi|\psi\rangle = {c_1}^* c_1 + {c_2}^* c_2 = |c_1|^2 + |c_2|^2 \tag{22}$$

The vector representation also derives this result directly from matrix multiplication:

$$\langle\psi|\psi\rangle = \boldsymbol{\Psi}^\dagger\boldsymbol{\Psi} = \begin{pmatrix} {c_1}^* & {c_2}^* \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = {c_1}^* c_1 + {c_2}^* c_2 \tag{23}$$

An operator, $\hat{A}$, acting upon a function yields another function. For the case where $|\psi\rangle$ is an eigenvector of $\hat{A}$, application of the operator scales the *ket* to give $a|\psi\rangle$, where a is the corresponding eigenvalue. For any given physical quantity, an operaor can be defined which acts upon the system to yield a measurement from the quantum system. The average value of

a property from an experimental measurement, referred to as the expectation value, is given for a two level system by:

$$\begin{aligned}\langle\hat{A}\rangle = \langle\psi|\hat{A}|\psi\rangle &= [\langle 1|{c_1}^* + \langle 2|{c_2}^*]\hat{A}[c_1|1\rangle + c_2|2\rangle] \\ &= {c_1}^* c_1\langle 1|\hat{A}|1\rangle + {c_1}^* c_2\langle 1|\hat{A}|2\rangle + {c_2}^* c_1\langle 2|\hat{A}|1\rangle \\ &+ {c_2}^* c_2\langle 2|\hat{A}|2\rangle\end{aligned} \tag{24}$$

Where the basis vectors are eigenvectors of $\hat{A}$, this simplifies to:

$$\langle\hat{A}\rangle = a_1|c_1|^2 + a_2|c_2|^2 \tag{25}$$

For this two-level system, the matrix representation of $\hat{A}$ is:

$$\boldsymbol{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} \langle 1|\hat{A}|1\rangle & \langle 1|\hat{A}|2\rangle \\ \langle 2|\hat{A}|1\rangle & \langle 2|\hat{A}|2\rangle \end{pmatrix} \tag{26}$$

This allows Eqn. 24 to be rewritten using the vector notation:

$$\begin{aligned}\langle\hat{A}\rangle = \boldsymbol{\Psi}^\dagger\boldsymbol{A}\boldsymbol{\Psi} &= \begin{pmatrix} {c_1}^* & {c_2}^* \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\ &= {c_1}^* c_1 A_{11} + {c_1}^* c_2 A_{12} + {c_2}^* c_1 A_{21} + {c_2}^* c_2 A_{22}\end{aligned} \tag{27}$$

2.2.2.2. The density operator

For systems involving numerous spins, rather than an isolated spin-½, the situation is significantly more complicated. As the expectation value of an operator depends on the products of its coefficients, it is convenient to define a density operator:

$$\hat{\rho} = |\psi\rangle\langle\psi| \tag{28}$$

For a spin-½ system, the matrix representation of this is:

$$\boldsymbol{\rho} = \begin{pmatrix} c_\alpha \\ c_\beta \end{pmatrix}\begin{pmatrix} {c_\alpha}^* & {c_\beta}^* \end{pmatrix} = \begin{pmatrix} c_\alpha {c_\alpha}^* & c_\alpha {c_\beta}^* \\ c_\beta {c_\alpha}^* & c_\beta {c_\beta}^* \end{pmatrix} = |\psi\rangle\langle\psi| = \boldsymbol{\Psi}\boldsymbol{\Psi}^\dagger \tag{29}$$

Inspection of the matrix elements of $\boldsymbol{\rho}$ and the final form of the expectation value seen in Eqn. 27 (section 2.2.2.1) shows that the expectation value can be rewritten as the trace of the product of $\boldsymbol{\rho}$ and $\boldsymbol{A}$:

$$\langle \hat{A} \rangle = \mathrm{Tr}(\boldsymbol{\rho A}) \tag{30}$$

To better highlight the effect of the elements in $\boldsymbol{\rho}$, a system is defined with coefficients that can be written as the product of a real coefficient a and a phase, with phase constant φ, such that Eqn. 20 can be rewritten as:

$$|\psi\rangle = a_\alpha e^{i\varphi_\alpha}|1\rangle + a_\beta e^{i\varphi_\beta}|2\rangle \tag{31}$$

Substituting Eqn. 31 into Eqn. 29 yields:

$$\boldsymbol{\rho} = \begin{pmatrix} {a_\alpha}^2 & a_\alpha a_\beta e^{i(\varphi_\alpha - \varphi_\beta)} \\ a_\alpha a_\beta e^{i(\varphi_\beta - \varphi_\alpha)} & {a_\beta}^2 \end{pmatrix} = \boldsymbol{\psi\psi}^\dagger \tag{32}$$

Eqn. 32 describes a system containing multiple isolated spins. If they all occupy slightly different states, the off-diagonal terms will average to zero. This situation – where only on-diagonal terms remain in Eqn. 37 – is called the population state and occurs when the system is in thermal equilibrium. This is analogous to longitudinal spin angular momentum, I_z. The alternative case, where the off-diagonal terms are non-zero, only occurs when there is phase coherence - the phase terms do not cancel as the ensemble of spins are in the same state. This is analogous to magnetisation in the transverse plane, corresponding to transverse spin angular momentum. A spin system needs to be pushed out of thermal equilibrium in NMR experiments as only the phase coherence states are observable (with only single quantum coherences observable directly via the induction of a current in the coil – see section 2.2.1).

2.2.2.3. Product operators

Although density operator theory is necessary when strong couplings are present in an ensemble of spins, a simpler approach using product operators can be applied when only weak couplings are present (*e.g.* where evolution under a resonance offset during a *rf* pulse can be neglected and the offset and *J*-coupling Hamiltonians commute). The change in net

magnetization caused by a *rf* pulse, for example along the *y* axis in the rotating frame with a flip angle of β, is given by:

$$I_x \xrightarrow{\beta_y} I_x \cos\beta - I_z \sin\beta \tag{33}$$

$$I_y \xrightarrow{\beta_y} I_y \tag{34}$$

$$I_z \xrightarrow{\beta_y} I_z \cos\beta + I_x \sin\beta \tag{35}$$

Evolution under a resonance offset, Ω, for a time, *t*, can be described as:

$$I_x \xrightarrow{\Omega t} I_x \cos\Omega t + I_y \sin\Omega t \tag{36}$$

$$I_y \xrightarrow{\Omega t} I_y \cos\Omega t - I_x \sin\Omega t \tag{37}$$

$$I_z \xrightarrow{\Omega t} I_z \tag{38}$$

The product operator formalism can be used to describe the time evolution of coupled systems, with magnetisation initially on *I* evolving with time to a coupled state which is the product of two single-spin operators. It is therefore possible to describe evolution between two spins *I* and *S*, which can be the same or different nuclear species, due to a weak *J* coupling, an internal interaction between chemically bonded nuclei (see section 2.2.5.4):

$$I_x \xrightarrow{\pi J_{IS} t} I_x \cos\pi J_{IS} t + 2I_y S_z \sin\pi J_{IS} t \tag{39}$$

Evolution under both a *J* coupling and Ω can be dealt with sequentially despite them actually occurring simultaneously.

2.2.3. Spin Angular Momentum Operators

Within Dirac notation, where $|\psi\rangle$ is an eigenvector of $\hat{I}^2$ and $\hat{I}_z$, the eigenvalues are given by:

$$\hat{I}_z|\psi\rangle = m|\psi\rangle \tag{40}$$

Here *m* is the *z*-component of the spin-angular momentum. As stated in section 2.2.1, this component can take values of $-I, -I+1,\ldots, +I$. For nuclei where $I = \frac{1}{2}$, the Zeeman eigenstates, $|\alpha\rangle$ and $|\beta\rangle$, correspond to eigenvalues of:

$$\hat{I}_z|\alpha\rangle = +\frac{1}{2}|\alpha\rangle \qquad \hat{I}_z|\beta\rangle = -\frac{1}{2}|\beta\rangle \tag{41}$$

$$\hat{H}|\alpha\rangle = +\frac{1}{2}\omega_0|\alpha\rangle \qquad \hat{H}|\beta\rangle = -\frac{1}{2}\omega_0|\beta\rangle \tag{42}$$

By convention, $|\alpha\rangle$ and $|\beta\rangle$ are the eigenstates corresponding to spin up and down, respectively.

Only a single spin angular momentum component commutes with the total spin angular momentum and the individual $\hat{I}_x$, $\hat{I}_y$ and $\hat{I}_z$ components do not commute with each other, implying that only one is observable at any given time. By convention this is taken to be $\hat{I}_z$. The *x* and *y*-components, $\hat{I}_y$ and $\hat{I}_y$, lie in the plane transverse to B_0 and do not commute with $\hat{I}^2$, meaning that the individual components cannot be observed simultaneously. The above relationships can be written as:

$$\left[\hat{I}^2, \hat{I}_z\right] = 0 \tag{43}$$

$$\left[\hat{I}_x, \hat{I}_y\right] = i\hat{I}_z \tag{44}$$

The total magnitude of the spin-angular momentum operator squared is given by:

$$\hat{I}^2 = \hat{I}_x^2 + \hat{I}_y^2 + \hat{I}_z^2 \tag{45}$$

$|\alpha\rangle$ and $|\beta\rangle$ are not eigenstates of $\hat{I}_x$ and $\hat{I}_y$ but are instead interconverted:

$$\hat{I}_x|\alpha\rangle = +\frac{1}{2}|\beta\rangle \qquad \hat{I}_x|\beta\rangle = +\frac{1}{2}|\alpha\rangle \tag{46}$$

$$\hat{I}_z|\alpha\rangle = +\frac{1}{2}i|\beta\rangle \qquad \hat{I}_z|\beta\rangle = -\frac{1}{2}i|\alpha\rangle \tag{47}$$

The matrix form of each of the *x*, *y* and *z*-components is:

$$\hat{\boldsymbol{I}}_{\boldsymbol{x}} = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}, \quad \hat{\boldsymbol{I}}_{\boldsymbol{y}} = \begin{pmatrix} 0 & -\frac{1}{2}i \\ \frac{1}{2}i & 0 \end{pmatrix}, \quad \hat{\boldsymbol{I}}_{\boldsymbol{z}} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} \tag{48}$$

As the wavefunction describing a spin-½ nucleus can be considered a superposition of $|\alpha\rangle$ and $|\beta\rangle$, such that, according to the notation in section 2.2.2.1, $|\psi\rangle = c_\alpha|\alpha\rangle + c_\beta|\beta\rangle$, the corresponding expectation values for these components are:

$$\langle \hat{I}_x \rangle = \frac{1}{2}\left(c_\alpha {c_\beta}^* + {c_\alpha}^* c_\beta\right) = \mathrm{Re}\left(c_\alpha {c_\beta}^*\right) \tag{49}$$

$$\langle \hat{I}_y \rangle = \frac{1}{2} i\left(c_\alpha {c_\beta}^* - {c_\alpha}^* c_\beta\right) = -\mathrm{Im}\left(c_\alpha {c_\beta}^*\right) \tag{50}$$

$$\langle \hat{I}_z \rangle = \frac{1}{2}|c_\alpha|^2 - \frac{1}{2}\left|c_\beta\right|^2 \tag{51}$$

To consider the behaviour of a system during an NMR experiment, the time dependence of the system must be taken into account. The time-dependent Schrödinger equation for $\langle\psi|$ and $|\psi\rangle$ is given by:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi| = i\langle\psi|\hat{H} \qquad\qquad \frac{\mathrm{d}}{\mathrm{d}t}|\psi\rangle = -i\hat{H}|\psi\rangle \tag{52}$$

Therefore, differentiating $\hat{\rho}$ (see Eqn. 28) with respect to time gives:

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}(t) = \frac{\mathrm{d}}{\mathrm{d}t}|\psi\rangle\langle\psi| = \left(\frac{\mathrm{d}}{\mathrm{d}t}|\psi\rangle\right)\langle\psi| + |\psi\rangle\left(\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi|\right) = -i\left[\hat{H}\,,\hat{\rho}(t)\right]$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}(t) = -i\left[\hat{H}\,,\hat{\rho}(t)\right] \tag{53}$$

This equation (Eqn. 53) relates the evolution of $\hat{\rho}$ to the Hamiltonian. It is known as the Liouville von-Neumann equation and its solution is most conveniently written in terms of operator exponentials, $e^{-i\hat{H}t}$. It is more complicated when the Hamiltonian also varies with time but so long as the Hamiltonian can be considered constant during each particular period of time that contributes to the evolution period, then a simple extension is possible:

$$\hat{\rho}(t) = e^{-i\hat{H}t}\hat{\rho}(0)e^{+i\hat{H}t} \tag{54}$$

$$\hat{\rho}(0) \xrightarrow{\hat{H}_1 t_1} e^{-i\hat{H}_1 t_1}\hat{\rho}(0)e^{+i\hat{H}_1 t_1} \xrightarrow{\hat{H}_2 t_2} e^{-i\hat{H}_2 t_2}e^{-i\hat{H}_1 t_1}\hat{\rho}(0)e^{+i\hat{H}_1 t_1}e^{+i\hat{H}_2 t_2} \tag{55}$$

The full Hamiltonian operator (used in all previous Eqn.) consists of the sum of the Hamiltonians for all the relevant interactions acting upon the system, both external and

internal. The external interactions relevant in NMR experiments are the Zeeman interaction and radio frequency pulses (represented with subscripts *Z* and *rf*, respectively). The internal interactions relevant to this thesis are chemical shielding, dipolar coupling, quadrupolar coupling and *J*-coupling (represented by with subscripts *σ*, *D*, *Q* and *J*, respectively).

$$\hat{H} = \hat{H}_Z + \hat{H}_{rf} + \hat{H}_\sigma + \hat{H}_D + \hat{H}_Q + \hat{H}_J \tag{56}$$

The individual Hamiltonians can be expressed in Cartesian coordinates, for spin operator $\hat{I}$, interaction $\tilde{A}$ and additional interacting spin operator $\hat{S}$, as:

$$\hat{H}_A = \hat{I}\tilde{A}\hat{S} = \begin{pmatrix} \hat{I}_x & \hat{I}_y & \hat{I}_z \end{pmatrix} \begin{pmatrix} A_{xx} & A_{xy} & A_{xz} \\ A_{yx} & A_{yy} & A_{yz} \\ A_{zx} & A_{zy} & A_{zz} \end{pmatrix} \begin{pmatrix} \hat{S}_x \\ \hat{S}_y \\ \hat{S}_z \end{pmatrix} \tag{57}$$

2.2.4. External Interactions

2.2.4.1. Thermal equilibrium

The Hamiltonian describing the Zeeman interaction between B_0 and *I* is as stated in Eqn. 15, above:

$$\hat{H}_Z = -\gamma \hat{I}_z B_0 = \omega_0 \hat{I}_z \tag{58}$$

For NMR experiments (particularly in the solid-state), B_0 is on the order of tesla (T), with magnetic field strengths between 11.7 T and 16.4 T for results presented in this thesis. These are, however, commonly reported in terms of ω_0 for ^{1}H.

2.2.4.2. **rf *pulses***

As discussed in section 2.2.1, only transverse magnetisation is observable so it is necessary to displace the magnetisation from thermal equilibrium along the *z* direction defined by B_0 by applying an *rf* pulse. This corresponds (see section 2.2.2.2) to the excitation of coherence

states. From Eqn. 18, the application of a B_1 field on the *x*-axis (at two counter rotating frequencies) can be written as

$$\begin{aligned} \boldsymbol{B_1}(\boldsymbol{t}) &= 2B_1 \cos\left(\omega_{rf} t + \varphi\right) \mathbf{i} \\ &= B_1[e^{+\omega_{rf} t} + e^{-\omega_{rf} t}]\mathbf{i} \qquad \text{if } \varphi = 0 \end{aligned} \tag{59}$$

Here, **i** is a unit vector aligned along the *x*-axis. As discussed above, the $-\omega_{rf}$ term is usually hundreds of MHz off-resonance and can therefore be ignored, simplifying the Hamiltonian for an *rf* pulse. In the rotating frame, it is simplified even further as the *rf* field appears static so can be treated as time independent:

$$\hat{H}_{rf} = \hat{I}\tilde{Z}\hat{B}_1 = -\gamma B_1\left[\hat{I}_x \cos\left(\omega_{rf} t + \varphi\right) + \hat{I}_y \sin\left(\omega_{rf} t + \varphi\right)\right] \tag{60}$$

$$\begin{aligned} \hat{H}_{rf}^{rot} &= -\gamma B_1\left[\hat{I}_x \cos\varphi + \hat{I}_y \sin\varphi\right] \\ &= \omega_1 \hat{I}_x \mathbf{i} \qquad \text{if } \varphi = 0 \end{aligned} \tag{61}$$

where the nutation frequency ω_1 was introduced in section 2.2.1, Eqn. 19. By convention, $\varphi = 0$ corresponds to a pulse on the *x*-axis and $\varphi \neq 0$ to a pulse applied on a different axis in the transverse plane.

The time-dependent density operator for this *rf* pulse follows from the solution of the Liouville von Neumann equation (Eqn. 54):

$$\hat{\rho}(t) = e^{-i\omega_1 \hat{I}_x t} \hat{\rho}(0) e^{+i\omega_1 \hat{I}_x t} \tag{62}$$

If the pulse is applied to a spin-½ system at thermal equilibrium, where $\hat{\rho}(0) = \hat{I}_z$, Eqn. 62 can be expressed as:

$$\hat{\boldsymbol{\rho}}(\boldsymbol{t}) = \frac{1}{2}\begin{pmatrix} \cos\omega_1 t & i\sin\omega_1 t \\ -i\sin\omega_1 t & \cos\omega_1 t \end{pmatrix} \tag{63}$$

There are both diagonal elements (population) and off-diagonal elements (coherences). Applying Eqn. 30 and Eqn. 48, the expectation values are therefore:

$$\langle \hat{I}_x \rangle = \mathrm{Tr}(\boldsymbol{\rho I_x}) = 0 \tag{64}$$

$$\langle \hat{I}_y \rangle = \mathrm{Tr}\left(\boldsymbol{\rho I_y}\right) = \frac{1}{2}\sin\omega_1 t \tag{65}$$

$$\langle \hat{I}_z \rangle = \mathrm{Tr}(\boldsymbol{\rho I_z}) = \frac{1}{2}\cos\omega_1 t \tag{66}$$

2.2.4.3. Resonance off-set

As introduced in section 2.2.1, in the rotating frame, the residual field about which the transverse plane magnetisation will precess is determined by the resonance offset, $\Omega = \omega_0 - \omega_{rf}$. The Hamiltonian for the Zeeman interaction can therefore be written as:

$$\hat{H}_Z^{rot} = (\omega_0 - \omega_{rf})\hat{I}_z = \Omega\hat{I}_z \tag{67}$$

The solution to the Liouville von Neuman equation (Eqn. 54) determines the time evolution of $\hat{\rho}$ under Ω for, *e.g.*, transverse magnetisation $\hat{I}_x$:

$$\hat{\rho}(t) = e^{-i\Omega\hat{I}_z t}\hat{I}_x e^{+i\Omega\hat{I}_z t} \tag{68}$$

$$\hat{\boldsymbol{\rho}}(\boldsymbol{t}) = \frac{1}{2}\begin{pmatrix} 0 & e^{-i\Omega t} \\ e^{+i\Omega t} & 0 \end{pmatrix} \tag{69}$$

Quadrature detection is described by a raising operator, $\hat{\boldsymbol{I}}_+$ ($\equiv \hat{\boldsymbol{I}}_-^{\dagger}$), which corresponds to two components that are 90° out of phase with respect to each other (see section 2.2.1) that are the real and imaginary components of the FID (this will be discussed in more detail in section 2.2.6):

$$\hat{\boldsymbol{I}}_+ = \hat{\boldsymbol{I}}_{\boldsymbol{x}} + i\hat{\boldsymbol{I}}_{\boldsymbol{y}} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{70}$$

The evolution of the NMR signal, $s(t)$, under Ω is calculated by taking the trace of the product of $\hat{\boldsymbol{\rho}}(\boldsymbol{t})$ and $\hat{\boldsymbol{I}}_+$:

$$\begin{aligned} s(t) &= \mathrm{Tr}\left[\hat{\boldsymbol{I}}_+\hat{\boldsymbol{\rho}}\right] \\ &= \mathrm{Tr}\left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & e^{-i\Omega t} \\ e^{+i\Omega t} & 0 \end{pmatrix}\right) \\ &= e^{+i\Omega t} = \cos\Omega t + i\sin\Omega t \end{aligned} \tag{71}$$

This corresponds to precession of the real and imaginary components of magnetisation in the transverse plane. This rotating magnetisation, with a 90° separation in phase between the two components, induces a current in the coil, producing the NMR signal.

2.2.5. Internal Interactions

Unlike the external interactions outlined in section 2.2.4, internal interactions are first described with respect to their own frame of reference, a Principal Axis System (PAS), rather than using the Cartesian coordinate axis of either the laboratory frame or its rotating frame. In each PAS, the interaction matrix of the Hamiltonian is diagonalized (off-diagonal components are zero).

As the Zeeman interaction is so dominant, all the internal interactions must be rotated to this frame of reference, which being aligned with B_0 is the laboratory frame (B_0 along the z-axis). Rotations can be broken into three stages, each corresponding to a rotation about a specific axis by one of the three characterising angles – the Euler angles: α, β and γ (Fig. 2.9).

$$(X, Y, Z) \longrightarrow (X^a, Y^a, Z^a)$$ The system is rotated by α about Z

$$(X^a, Y^a, Z^a) \longrightarrow (X^{ab}, Y^{ab}, Z^{ab})$$ The system is rotated by β about Y^a

$$(X^{ab}, Y^{ab}, Z^{ab}) \longrightarrow (X^{abc}, Y^{abc}, Z^{abc})$$ The system is rotated by γ about Z^{ab}

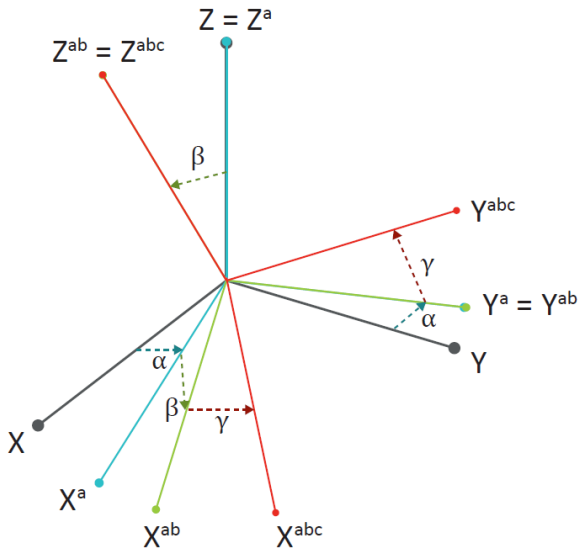


Figure 2.9:Euler angles as defined in the Cartesian coordinate system.

Spherical tensor coordinates are used instead of Cartesian coordinates to simplify the process. Written with spherical tensors, the generic form of the interaction Hamiltonian (Eqn. 57) becomes:

$$\hat{H}_A = \sum_{j=0}^{2} \sum_{m=-j}^{+j} -1^m A_{jm} \hat{T}_{j-m} \tag{72}$$

A_{jm} corresponds to the spatial component, representing the magnitude of the internal interaction, and $\hat{T}_{j-m}$ the spin component, expressed in terms of the required quantum mechanical angular momentum operators, in terms of the spherical tensor with order *j* and rank *m*. Transformations only affect the spatial component of spherical tensors, which is why they are simpler to perform. Given the aforementioned diagonalization in the PAS, Eqn. 72 simplifies to:

$$\hat{H}_A^{PAS} = A_{00}^{PAS}\hat{T}_{00} + A_{20}^{PAS}\hat{T}_{20} + A_{22}^{PAS}\hat{T}_{2-2} + A_{2-2}^{PAS}\hat{T}_{22} \tag{73}$$

The rotation transformation gives:

$$R(A_{jm}) = \sum_{m=-j}^{j} D_{m\prime m}^{j}(\alpha, \beta, \gamma)\, A_{jm\prime} \tag{74}$$

$D_{m\prime m}^{j}(\alpha, \beta, \gamma)$ represents the Wigner matrix elements, with reduced Wigner rotation matrices $d_{kl}^{j}(\beta)$:

$$D_{kl}^{j}(\alpha, \beta, \gamma) = e^{-ik\alpha} d_{kl}^{j}(\beta) e^{-il\gamma} \tag{75}$$

$$d_{kl}^{j}(\beta) = \langle jk | e^{-i\beta \hat{J}_y} | jl \rangle \tag{76}$$

A rotation from the PAS (P) to the laboratory frame (L) under the Euler angles is, therefore:

$$\hat{A}_{jm\prime}^{L} = \sum_{m=-j}^{j} D_{mm\prime}^{j}(\alpha_{PL}, \beta_{PL}, \gamma_{PL})\, A_{jm\prime}^{PAS} \tag{77}$$

As noted above, the Zeeman interaction is generally the dominant interaction. It is therefore possible to apply the secular approximation: all other interactions are considered first-order perturbations to the Zeeman Hamiltonian. This means that Eqn. 56 can be written in terms of the zero order perturbation (equal to $\widehat{H}_Z$) and this first order perturbation $\widehat{H}_1$:

$$\widehat{H} = \widehat{H}_0 + \widehat{H}_1 \tag{78}$$

Only spin terms which commute with $\hat{I}_Z$ are considered in this approach:

$$[\hat{I}_Z, \hat{T}_{jm}] = m\hat{T}_{jm} \tag{79}$$

The commutator in Eqn. 79 is only zero when $m = 0$. The secular approximation is only valid when B_0 is sufficiently high-field to ensure the dominance of the Zeeman interaction.

2.2.5.1. Chemical shielding

The ability of NMR to distinguish different chemical environments for the same nuclear species stems from an effect called chemical shielding, σ. This arises from the electronic configuration as, under B_0, the electrons orbit the nucleus. This induces a small local magnetic field so the each nucleus is shielded with respect to B_0 in a site specific manner and its resonance frequency, ω, changes by a small but measurable amount:

$$\omega = \frac{\gamma B_0(1-\sigma)}{2\pi} \tag{80}$$

As the absorption frequency scales linearly (to a good approximation) with magnetic field, it is convenient to quote the chemical shift relative to a known reference compound (*e.g.* tetramethylsilane (TMS) for protons) rather than giving the frequency in Hz as this allows a direct comparison of data collected at different magnetic fields:

$$\delta = \frac{\omega - \omega_{ref}}{\omega_{ref}} \times 10^6 \tag{81}$$

As the chemical shift is generally very small, the factor of 10^6 is included, converting the scale into parts per million (ppm).

The Cartesian form of the Hamiltonian for the chemical shielding can be written as:

$$\hat{H}_{CS} = \gamma \hat{I} \tilde{\sigma} \hat{B}_0 \tag{82}$$

The second rank shielding tensor, $\tilde{\sigma}$, has both symmetric and asymmetric components but only the symmetric part has a significant impact on the NMR experiment:

$$\tilde{\sigma}_s = \begin{pmatrix} \sigma_{xx} & \frac{1}{2}(\sigma_{xy} + \sigma_{yx}) & \frac{1}{2}(\sigma_{xz} + \sigma_{zx}) \\ \frac{1}{2}(\sigma_{xy} + \sigma_{yx}) & \sigma_{yy} & \frac{1}{2}(\sigma_{yz} + \sigma_{zy}) \\ \frac{1}{2}(\sigma_{xz} + \sigma_{zx}) & \frac{1}{2}(\sigma_{yz} + \sigma_{zy}) & \sigma_{zz} \end{pmatrix} \tag{83}$$

To describe the effect on the NMR spectrum, the spherical tensor form of the Hamiltonian in PAS (Eqn. 73) needs to be transformed to the laboratory frame. The A_{00}^{PAS} term is isotropic and thus invariant under the rotation. Under the secular approximation:

$$\begin{aligned} A_{20}^{L} &= \sum_{m=-2}^{2} A_{2m}^{PAS} d_{m0}^{2}(\beta_{PL}) e^{-im\alpha_{PL}} \\ &= A_{20}^{PAS} d_{00}^{2}(\beta_{PL}) + A_{22}^{PAS} d_{20}^{2}(\beta_{PL}) e^{-2i\alpha_{PL}} \\ &+ A_{2-2}^{PAS} d_{-20}^{2}(\beta_{PL}) e^{2i\alpha_{PL}} \\ &= A_{20}^{PAS} \frac{1}{2}(3\cos^2\beta_{PL} - 1) + A_{2\pm2}^{PAS} \sqrt{\frac{3}{2}} \sin^2\beta_{PL} \cos 2\alpha_{PL} \end{aligned} \tag{84}$$

The spatial terms are given by:

$$A_{00}^{PAS} = \gamma \sqrt{\frac{1}{3}} \left(\sigma_{xx}^{PAS} + \sigma_{yy}^{PAS} + \sigma_{zz}^{PAS}\right) \tag{85}$$

$$A_{20}^{PAS} = \gamma \sqrt{\frac{1}{6}} \left(2\sigma_{xx}^{PAS} - \sigma_{yy}^{PAS} - \sigma_{zz}^{PAS}\right) \tag{86}$$

$$A_{2\pm2}^{PAS} = \gamma \frac{1}{2} \left(\sigma_{xx}^{PAS} - \sigma_{yy}^{PAS}\right) \tag{87}$$

In terms of isotropic chemical shielding, σ_{iso}, anisotropic chemical shielding, σ_{aniso}, and the asymmetry parameter, η, the chemical shielding Hamiltonian in the laboratory frame is given by:

$$\begin{aligned}\hat{H}_{\sigma}^{L} = -\omega_0\sigma_{iso}\hat{T}_{00} \\ -\frac{1}{2}\omega_0\sigma_{aniso}(3\cos^2\beta_{PL} - 1 \\ +\eta\sin^2\beta_{PL}\cos 2\alpha_{PL})\hat{T}_{20}\end{aligned} \tag{88}$$

Conversion from chemical shielding to the chemical shift tensor gives:

$$\delta = \frac{\sigma_{iso}(\text{ref}) - \sigma_{iso}}{1 - \sigma_{iso}(\text{ref})} \times 10^6 \tag{89}$$

The definitions for δ_{iso}, δ_{aniso} and η are:

$$\delta_{iso} = \frac{1}{3}\left(\delta_{xx}^{PAS} + \delta_{yy}^{PAS} + \delta_{zz}^{PAS}\right) \tag{90}$$

$$\delta_{aniso} = \delta_{zz}^{PAS} - \delta_{iso} \tag{91}$$

$$\eta = \frac{\delta_{xx}^{PAS} - \delta_{yy}^{PAS}}{\delta_{aniso}} \tag{92}$$

The Haeberlen convention[157] gives the relative magnitudes of the three chemical shielding principal components as:

$$|\delta_{zz}^{PAS} - \delta_{iso}| \geq |\delta_{xx}^{PAS} - \delta_{iso}| \geq \left|\delta_{yy}^{PAS} - \delta_{iso}\right| \tag{93}$$

The rapid tumbling of molecules in solution means that, on the NMR timescale, the anisotropic component averages to zero and only the isotropic component is observed. In a static powder sample, this does not occur and the anisotropic component of all the orientations have an effect, producing a continuous range of overlapping chemical shifts. An example of the simulated static powder line-shape, for a spin-½ nucleus with chemical shift anisotropy (CSA) is shown in Fig. 2.10. Magic angle spinning (MAS) is employed to combat this anisotropic line broadening (see section 2.2.5.5).

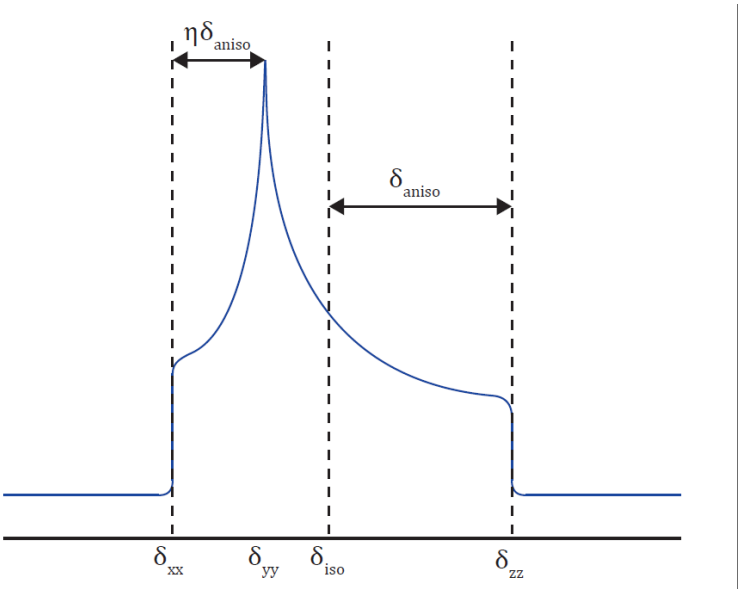


Figure 2.10: A simulated pattern for a static powder sample with a CSA line-shape, where η = 0.5. The components of the chemical shift in PAS are labelled.

2.2.5.2. Dipolar coupling

Dipolar coupling is due to the interaction between the nuclear magnetic moments of different nuclear spins. For two spins *I* and *S*, spin *S* will 'feel' the magnetic field produced by spin I and vice versa if they are close enough in space (< 10 Å) with the strength of the interaction directly related to the internuclear distance. In the quantum mechanical description for two spin-½ nuclei, there are four possible Zeeman transition states (Fig. 2.11). A transition between $|\alpha\beta\rangle$ and $|\beta\alpha\rangle$ energy levels corresponds to a zero quantum (ZQ) transition, whereas moving between $|\alpha\alpha\rangle$ and $|\beta\beta\rangle$ is a double quantum (DQ) coherence. The remaining possible transitions are single quantum (SQ) coherences. The energy levels of $|\alpha\beta\rangle$ and $|\beta\alpha\rangle$ are practically degenerate for homonuclear dipole couplings whereas there can be a significant difference between them in the heteronuclear case. The Cartesian form of the Hamiltonian for the dipolar coupling can be written as:

$$\hat{H}_D = 2\sum_{i<j} \hat{I}_i \tilde{D} \hat{S}_j \tag{94}$$

The dipolar coupling is generally of the order of 10s of kHz and is much smaller than the Zeeman interaction. The PAS of the dipolar Hamiltonian, which is aligned with the

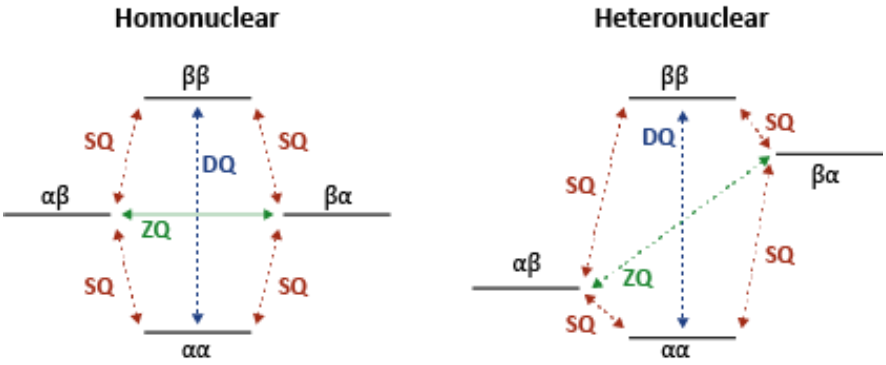


Figure 2.11: Energy level diagram for two coupled spin-½ nuclei.

internuclear vector, must be rotated into the laboratory frame. In the PAS Hamiltonian for dipolar coupling, the only non-zero term in Eqn. 73 is the A_{20}^{PAS} term, with:

$$A_{20}^{PAS} = \sqrt{6} d_{IS} \tag{95}$$

The dipolar coupling constant, d_{IS}, is given (in rad s^{-1}) by:

$$d_{IS} = -\frac{\mu_0}{4\pi}\frac{\hbar \gamma_I \gamma_S}{r_{IS}^3} \tag{96}$$

Under the secular approximation:

$$\begin{aligned} A_{20}^{L} &= A_{20}^{PAS} D_{00}^{2} = \sqrt{6} d_{IS} e^{-2i\alpha_{PL}0} d_{00}^{2} e^{2i\alpha_{PL}0} \\ &= \sqrt{6} d_{IS} \frac{1}{2}(3\cos^2\beta_{PL} - 1) \end{aligned} \tag{97}$$

The corresponding spin term, $\hat{T}_{20}$, is written as:

$$\hat{T}_{20} = \frac{1}{\sqrt{6}}\left(\hat{I}_z\hat{S}_z - \frac{1}{2}\left(\hat{I}_x\hat{S}_x + \hat{I}_y\hat{S}_y\right)\right) \tag{98}$$

The laboratory frame dipolar Hamiltonian can therefore be written as:

$$\hat{H}_{D,het} = d_{IS}\frac{1}{2}(3\cos^2\beta_{PL} - 1)\left(2\hat{I}_z\hat{S}_z\right) \tag{99}$$

$$\hat{H}_{D,homo} = d_{IS}\frac{1}{2}(3\cos^2\beta_{PL} - 1)\left(2\hat{I}_z\hat{S}_z - \left(\hat{I}_x\hat{S}_x + \hat{I}_y\hat{S}_y\right)\right) \tag{100}$$

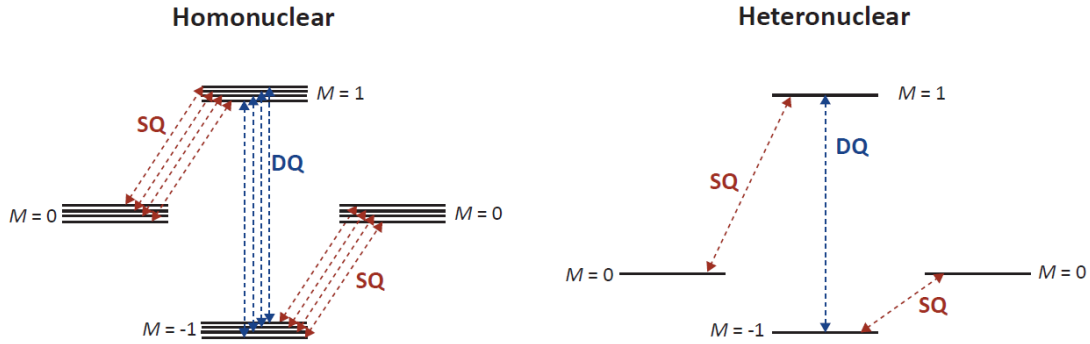


Figure 2.12: Energy level transitions for a pair of spin-½ dipolar coupled nuclei. Energy levels correspond to a linear combination of Zeeman eigenstates for a homonuclear pair (left) and to the Zeeman product eigenstates for a heteronuclear pair (right).

The matrix form of the spin operators are:

$$2\hat{I}_z\hat{S}_z = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix}, \qquad \hat{I}_x\hat{S}_x + \hat{I}_y\hat{S}_y = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{101}$$

For two spin-½ nuclei, the spin eigenstates correspond to the Zeeman transition states in Fig. 2.11, which produce a first order shift in the energy of the Zeeman interaction to a limited range of Zeeman product states (Fig. 2.12). The line broadening from the heteronuclear dipolar coupling is dependent on orientation. The static line-shape for a system with heteronuclear dipolar coupling is known as a Pake doublet, shown in Fig. 2.13. The splitting between the two horns is equal to $|d_{IS}|/2\pi$. In solution, rapid tumbling of the molecules removes the

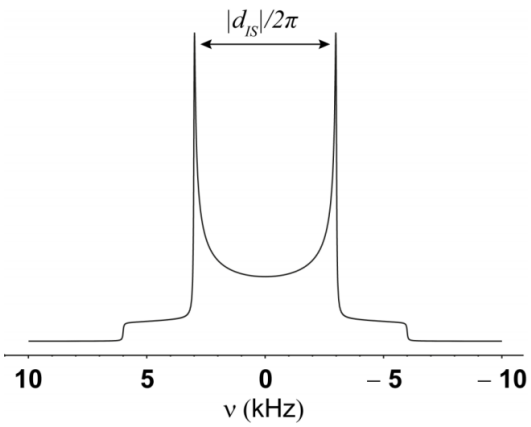


Figure 2.13: Simulated NMR line shape for a heteronuclear dipolar coupling between a pair of spin-

anisotropic broadening due to dipolar coupling as this is averaged out to zero. It can be completely removed under MAS.

Due to the $\left(\hat{I}_x\hat{S}_x + \hat{I}_y\hat{S}_y\right)$ term in Eqn. 99, the homonuclear dipolar coupling is less straightforward as the spin operator (Eqn. 101) contains non-zero off-diagonal terms: $\left(\hat{I}_x\hat{S}_x + \hat{I}_y\hat{S}_y\right)$ can be conveniently expressed in terms of raising and lowering operators as $\left(\hat{I}_-\hat{S}_+ + \hat{I}_+\hat{S}_-\right)$, these are referred to as the flip-flop term. The spin eigenstates correspond to linear combinations of degenerate Zeeman levels. In systems containing numerous nuclei, the presence of multiple degenerate eigenstates, with different transition frequencies, means that the Hamiltonians do not commute at different time points. The result is that the subsequent line broadening from homonuclear dipolar coupling is only partially removed under MAS.

2.2.5.3. Quadrupolar interaction

The quadrupolar interaction occurs if the nuclear spin is ≥ 1. In these cases, nuclei possess an electric quadrupolar moment as well as well as a magnetic dipolar moment. This quadrupolar moment is caused by the charge distribution in the nucleus and results in a perturbation to the Zeeman interaction. The Cartesian form of the Hamiltonian for the quadrupolar interaction can be written as:

$$\hat{H}_Q = \frac{eQ}{2I(2I-1)\hbar}\hat{I}\tilde{V}\hat{I} \tag{102}$$

$\tilde{V}$ describes the components of the electric field gradient in Cartesian coordinates:

$$\tilde{V} = \begin{pmatrix} V_{xx} & V_{xy} & V_{xz} \\ V_{yx} & V_{yy} & V_{yz} \\ V_{zx} & V_{zy} & V_{zz} \end{pmatrix} \tag{103}$$

In the PAS, the components of $\tilde{V}$are described by the quadrupolar coupling constant, C_Q, which describes the strength of the quadrupolar moment and the asymmetry parameter, η_Q, which describes the relative strength of the electric field gradient in three orthogonal directions:

$$C_Q = \frac{e^2 qQ}{h} = \frac{eQV_{zz}}{h} \tag{104}$$

$$\eta_Q = \frac{V_{xx} - V_{yy}}{V_{zz}} \quad \text{where} \quad |V_{xx}| \leq |V_{yy}| \leq |V_{zz}| \tag{105}$$

Q is the isotope specific nuclear quadrupolar moment, q the electric field gradient at the nucleus and e the electric charge. A high C_Q corresponds to a large electric field gradient across the nucleus.

Quadrupolar interactions are typically on the order of MHz, and are thus significantly larger than the other internal interactions. However, the Zeeman interaction remains sufficiently dominant for the nuclei considered in this work that it is necessary to rotate the PAS to the laboratory frame. From Eqn. 73 and Eqn. 102, the PAS Hamiltonian for the quadrupolar interaction is:

$$\hat{H}_Q^{PAS} = \frac{2\pi}{2I(2I-I)}\left(A_{20}^{PAS}\hat{T}_{20} + A_{22}^{PAS}\hat{T}_{2-2} + A_{2-2}^{PAS}\hat{T}_{22}\right) \tag{106}$$

The spatial terms are given by:

$$A_{20}^{PAS} = \sqrt{\frac{2}{3}}C_Q \tag{107}$$

$$A_{22}^{PAS} = A_{2-2}^{PAS} = \frac{1}{2}\eta_Q C_Q \tag{108}$$

The full Hamiltonian in the laboratory frame is:

$$\hat{H}_Q^L = \frac{2\pi}{2I(2I-I)}\left(A_{20}^L\hat{T}_{20} - A_{21}^L\hat{T}_{2-1} - A_{2-1}^L\hat{T}_{21} + A_{22}^L\hat{T}_{2-2} + A_{2-2}^L\hat{T}_{22}\right) \tag{109}$$

When considered as a first-order perturbation, under the secular approximation, only the contribution of A_{20}^L term is needed:

$$\begin{aligned} A_{20}^L &= A_{20}^{PAS}D_{00}^2 + A_{22}^{PAS}D_{20}^2 + A_{2-2}^{PAS}D_{-20}^2 \\ &= \sqrt{\frac{3}{2}}\frac{C_Q}{2I(2I-I)}\frac{1}{2}\left((3\cos^2\beta_{PL} - 1)\right. \\ &\left. + \eta\sin^2\beta_{PL}\cos 2\alpha_{PL}\right) \end{aligned} \tag{110}$$

The first-order quadrupolar Hamiltonian is therefore:

$$\hat{H}_Q^L = \sqrt{\frac{3}{2}}\frac{C_Q}{2I(2I-I)}\frac{1}{2}\left((3\cos^2\beta_{PL}-1)+\eta\sin^2\beta_{PL}\cos 2\alpha_{PL}\right)\hat{T}_{20} \tag{111}$$

The first order perturbation can be written as:

$$E_m^1 = \langle m|\hat{H}^1|m\rangle \tag{112}$$

Substitution of Eqn. 111 yields:

$$E_m^1 = \sqrt{\frac{3}{2}}\frac{C_Q}{2I(2I-I)}\frac{1}{2}\left(3m^2-I(I+1)\right)\Big((3\cos^2\beta_{PL}-1) + \eta_Q\sin^2\beta_{PL}\cos 2\alpha_{PL}\Big) \tag{113}$$

For ^{14}N, a spin-1 nucleus and the only quadrupolar nucleus considered in this work, there are three possible energy levels ($m=-1,0,1$) and thus two possible transitions as shown in Fig. 2.14. For the first-order perturbation, the energy splitting for the total transition between $m=-1$ and $m=1$ is the same as for the unperturbed Zeeman energy levels but the separate transition energies are different. This is not the case for the second-order perturbation.

The second-order perturbation to the Zeeman interaction needs to be considered when the quadrupolar coupling is of the order of MHz, which is often the case. This perturbation is given by:

$$E_m^2 = \sum_{m\neq n}\frac{\langle n|\hat{H}^1|m\rangle\langle m|\hat{H}^1|n\rangle}{E_n^0-E_m^0} \tag{114}$$

The secular approximation is not valid when calculating the quadrupolar Hamiltonian to the second order so all terms in Eqn. 109 need to be considered.

$$E_m^2 = -\left(\frac{C_Q}{4I(2I-I)}\right)^2\frac{2}{\omega_0}m\begin{bmatrix}(I(I+I)-3m^2)D_{00}^{0(Q)}\\ +(8I(I-I)-3m^2-3)D_{20}^{2(Q)}\\ +(18I(I-I)-34m^2-5)D_{40}^{4(Q)}\end{bmatrix} \tag{115}$$

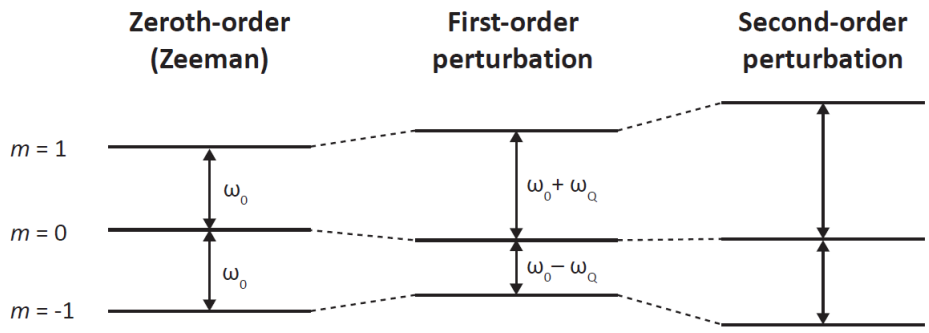


Figure 2.14: First- and second-order perturbation to the Zeeman energy levels due to the quadrupolar interaction for a spin-1 nucleus.

$$D_{00}^{0(Q)} = -\frac{1}{5}\left(3 + \eta_Q^2\right) \tag{116}$$

$$D_{20}^{2(Q)} = \frac{1}{28}\left[\left(\eta_Q^2 - 3\right)\left(3\cos^2\beta_{PL} - 1\right) + 6\eta_Q^2 \sin^2\beta_{PL}\cos 2\alpha_{PL}\right] \tag{117}$$

$$D_{40}^{4(Q)} = \frac{1}{8}\left[\begin{array}{c}\left(\frac{1}{140}\left(18 + \eta_Q^2\right)\left(35\cos^4\beta_{PL} - 30\cos^2\beta_{PL} + 3\right)\right) \\ + \left(\frac{3}{7}\eta_Q \sin^2\beta_{PL}\left(7\cos^2\beta_{PL} - 1\right)\cos 2\alpha_{PL}\right) \\ + \left(\frac{1}{4}\eta_Q^2 \sin^4\beta_{PL}\cos 4\alpha_{PL}\right)\end{array}\right] \tag{118}$$

For the second-order perturbation, the energy splitting of the total transition between $m = -1$ and $m = 1$ is altered relative to the zeroth order. As the second-order perturbation in Eqn. 115 is inversely proportional to ω_0, the effects decrease with increasing magnetic field strength.

The zero-rank term in Eqn. 115 is isotropic. The isotropic contribution to a resonance frequency in a NMR spectrum for a quadrupolar nucleus contains an additional quadrupolar isotropic shift, δ_{iso}^{Q}, along with the isotropic chemical shift. δ_{iso}^{Q} is determined by the quadrupolar parameters:

$$\delta_{iso}^{Q} = -\left(\frac{3}{40}\right)\left(\frac{P_Q}{\nu_0}\right)^2 \frac{[I(I+1) - 9m(m-1) - 3]}{[I^2(2I-1)^2]} \times 10^6 \tag{119}$$

When $I = 1$ and $m = 0{,}1$, Eqn. 119 simplifies to:

$$\delta_{iso}^{Q} = -\left(\frac{3}{40}\right)\left(\frac{P_Q}{\nu_0}\right)^2 \times 10^6 \tag{120}$$

The quadrupolar product, P_Q, is given by:

$$P_Q = C_Q \sqrt{1 + \frac{\eta_Q^2}{3}} \tag{121}$$

The second-rank term is removed under MAS (see section 2.2.5.5). However, residual broadening from the quadrupolar interaction will remain under MAS as the fourth-rank term is also anisotropic and both second- and fourth-rank terms cannot be averaged by rotation about the same angle simultaneously (as for MAS).

2.2.5.4. J-Coupling

J-coupling is mediated by the sharing of electrons so only occurs between covalently bonded electrons. It is the weakest of the internal interactions, with maximum values of ~100 Hz for light atoms. The anisotropic component of the *J*-coupling, which has the same orientation dependence as the dipolar coupling. is small enough for the nuclei studied in this work that it can be ignored. The isotropic component is the trace of the interaction tensor:

$$J_{iso} = \frac{1}{3}\,\mathrm{Tr}[\tilde{J}] = \frac{1}{3}\left(J_{xx} + J_{yy} + J_{zz}\right) \tag{122}$$

The Cartesian form of the Hamiltonian for the *J*-coupling interaction can be written as:

$$\hat{H}_J = \sum_{i<j} \hat{I}_i \tilde{J} \hat{S}_j \tag{123}$$

Although the splittings caused by the *J*-coupling can be seen in the solution state, due to the narrow lineshapes, the anisotropic line broadening in solid-state NMR means there is usually insufficient resolution to observe a *J*-splitting. *J*-couplings can, however, be employed as a method of polarisation transfer, identifying chemically bonded nuclei in correlation experiments, though this method is not utilised within this work.

2.2.5.5. Magic Angle Spinning

Magic-angle spinning (MAS) is capable of averaging many anisotropic interactions present in solid-state NMR. This is the case when the orientation dependence of these anisotropic contributions is $3\cos^2\theta - 1$, with θ corresponding to the angle between the interaction tensor and B_0. MAS utilises this dependence as when θ = 54.74° (the so-called magic angle) $3\cos^2\theta - 1 = 0$. Averaging of the anisotropic interactions (over one full rotor period) is therefore efficiently achieved by the physical rotation of a powder sample about an axis that sits at the magic angle relative to B_0, ensuring the interactions along the axis of rotation are also removed as they are zero.

The mathematical description of MAS requires the introduction of an intermediate rotation to the rotor frame (R) into the rotation between the PAS and the laboratory frame. The Euler angles used are therefore:

$$PAS \longrightarrow Rotor\ frame \qquad R(\alpha_{PR}, \beta_{PR}, \gamma_{PR})$$

$$Rotor\ frame \longrightarrow Laboratory \qquad R(\alpha_{RL}, \beta_{RL}, \gamma_{RL})$$

These transformations are illustrated in Fig. 2.15, where it can also be seen that α_{PR} is subtended by the physical rotation of the rotor, at a spinning frequency ω_R. The angle γ_{RL} defines the rotor phase.

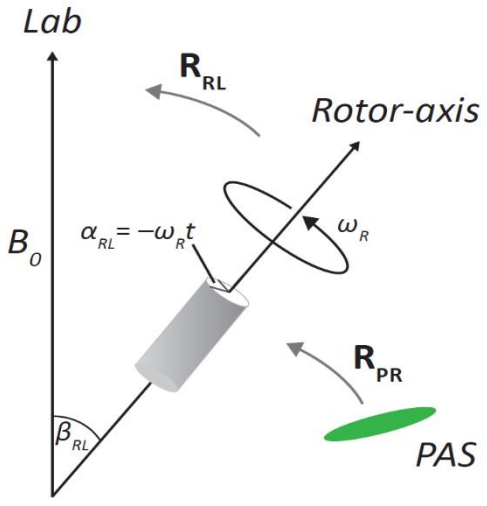


Figure 2.15: A rotor aligned at β_{RL} relative to B_0. The subsequent rotations from PAS-rotor and rotor-laboratory frames are illustrated.

The A_{20}^{PAS} term is non-zero for both the chemical shielding and dipolar Hamiltonians. Rotation of this term from the PAS frame via the rotor frame gives (using the secular approximation) the laboratory frame term:

$$A_{20}^{L} = A_{20}^{PAS} \sum_{m'=-2}^{2} D_{0m'}^{2}(\alpha_{PR}, \beta_{PR}, \gamma_{PR}) d_{m'0}^{2}(\beta_{RL}) e^{im'\omega_R t} \tag{124}$$

Terms with non-zero values of *m'* will, over a 360° rotation of the rotor (the time taken for this is one rotor period, τ_R), average to zero:

$$\int_0^{\tau_R} e^{im'\omega_R t} dt = 0 \qquad \text{if} \quad m' \neq 0 \tag{125}$$

Therefore, provided acquisition is over exactly an integer multiple of τ_R, applying rotation matrices means Eqn. 124 becomes:

$$\begin{aligned} \langle A_{20}^{L} \rangle &= A_{20}^{PAS} d_{00}^{2}(\beta_{PR}) d_{00}^{2}(\beta_{RL}) \\ &= A_{20}^{PAS} \left[\frac{1}{4} (3\cos^2 \beta_{PR} - 1)(3\cos^2 \beta_{RL} - 1) \right] \end{aligned} \tag{126}$$

Hence if β_{RL} is set to 54.74°, the aforementioned magic angle, the anisotropy averaged over one rotor period will cancel to zero. If acquisition is not over an integer multiple of τ_R this is not the case, the $m' \neq 0$ terms will not average to zero and additional relevant spatial components must be considered. In this case, Eqn. 126 becomes:

$$A_{20}^{L} = A_{20}^{PAS} \begin{bmatrix} \frac{1}{4} (3\cos^2 \beta_{PR} - 1)(3\cos^2 \beta_{RL} - 1) \\ -\frac{3}{4} \sin 2\beta_{PR} \sin 2\beta_{RL} \cos(\gamma_{PR} + \omega_R t) \\ +\frac{3}{4} \sin^2 \beta_{PR} \sin^2 \beta_{RL} \cos(2\gamma_{PR} + 2\omega_R t) \end{bmatrix} \tag{127}$$

The first term is the same as for Eqn. 126 so, if β_{RL} is set to 54.74°, this will cancel to zero. The final two terms are seen experimentally as spinning sidebands when acquisition is not over an integer multiple of τ_R, such spinning sidebands are found at increments of ω_R.

2.2.6. Experimental Methods

2.2.6.1. Lineshapes

In NMR, the time domain signal is generated by induction of a current in a coil and detected using quadrature detection (as mentioned in section 2.2.1). By mixing the recorded signal with two reference frequencies, that have a phase difference of 90°, two orthogonal components of the magnetisation, that constitute the real and imaginary parts of the complex time-domain signal, are measured. In the absence of internal interactions, the time-domain signal for evolution under a resonance offset, Ω, is given by:

$$s(t) = \cos \Omega t + i \sin \Omega t \tag{128}$$

A damping term characterised by the transverse relaxation, T_2, is typically included to capture the loss of coherence between spins. Eqn. 128 can therefore be written in exponential form as:

$$s(t) = e^{i\Omega t} e^{-\frac{t}{T_2}} \tag{129}$$

The complex time domain signal, the FID, is FT to give the frequency domain spectrum, which is more easily interpreted. For simplicity, we define a parameter $R = 1/T_2$.

$$S(\omega) = \frac{R}{R^2 + (\omega - \Omega)^2} - i\frac{\omega - \Omega}{R^2 + (\omega - \Omega)^2} = A(\omega) - iD(\omega) \tag{130}$$

The real part of the spectrum, $A(\omega)$, generates an absorption lineshape while the imaginary part, $D(\omega)$, has a dispersion lineshape (a phase adjustment is often necessary to ensure each component is purely absorptive or dispersive). $A(\omega)$ is at a maximum intensity at Ω, while $D(\omega)$ is has zero intensity. The linewidth, when measured as the full width at half maximum height (FWHMH), of $A(\omega)$ is given by $1/\pi T_2$ (in Hz). Therefore a short T_2 is associated with marked line broadening.

2D experiments are described by two evolution periods: t_2, which is analogous to the FID acquisition time in 1D, and t_1, which describes a second evolution time in 2D experiments. To produce a 2D dataset, t_1 is incremented (Δt_1) over the collection of a series of FIDs. The

direct, t_2, dimension produces the frequency ω_2 on the axis F_2 and the indirect, t_1, dimension produces the frequency ω_1 on the axis F_1. A generic 2D experiment is described by four periods of time: *preparation*, in which the desired nuclear spin coherences are excited; *evolution* of the spins under the interactions present during the incremented t_1; *mixing*, in which the spin coherences are converted into an observable form; and *detection* during t_2 as an in-phase SQ coherence. The general form of the 2D time domain signal, for evolution under only Ω, is given by:

$$s(t_1, t_2) = e^{-ip\Omega t_1} e^{-\frac{t_1}{T_2^{(1)}}} e^{i\Omega t_2} e^{-\frac{t_2}{T_2^{(2)}}} \tag{131}$$

In an experiment that is phase modulated (pm) with respect to t_1, then only a single selected coherence order, p, is present during t_1. For the case where $p = 1$, Eqn. 131 becomes:

$$s_{pm}(t_1, t_2) = e^{-i\Omega t_1} e^{-\frac{t_1}{T_2^{(1)}}} e^{i\Omega t_2} e^{-\frac{t_2}{T_2^{(2)}}} \tag{132}$$

FT of Eqn. 132, with respect first to t_2 and then to t_1, gives:

$$\begin{aligned} s_{pm}(\omega_1, \omega_2) &= (A_1^- - iD_1^-)(A_2^+ - iD_2^+) \\ &= (A_1^- A_2^+ - D_1^- D_2^+) - i(A_1^- D_2^+ + D_1^- A_2^+) \end{aligned} \tag{133}$$

As the superscript describes the sign of the precession frequency, it can be easily seen that the F_1 components only have negative frequencies, indicating that sign discrimination has been achieved. However, there are both absorptive and dispersive elements in both the real and imaginary components. This results in a broad, phase-twisted lineshape. Therefore, NMR experiments are collected using amplitude modulation (am), where both signs of coherence order are present during t_1 ($p = \pm 1, \pm 2, ...$). In this case, Eqn. 132 becomes:

$$\begin{aligned} s_{am}(t_1, t_2) &= \left(e^{-i\Omega t_1} + e^{i\Omega t_1}\right) e^{-\frac{t_1}{T_2^{(1)}}} e^{i\Omega t_2} e^{-\frac{t_2}{T_2^{(2)}}} \\ &= 2\cos\Omega t_1 \, e^{-\frac{t_1}{T_2^{(1)}}} e^{i\Omega t_2} e^{-\frac{t_2}{T_2^{(2)}}} \end{aligned} \tag{134}$$

FT with respect to t_2 proceeds as before but FT for t_1 is performed separately for the real and imaginary components:

$$s_{am}^{Re}(\omega_1, \omega_2) = (A_1^- + A_1^+)A_2^+ - i(D_1^- + D_1^+)A_2^+ \quad (135)$$

$$s_{am}^{Im}(\omega_1, \omega_2) = (A_1^+ + A_1^-)D_2^+ - i(D_1^+ + D_1^-)D_2^+ \quad (136)$$

Amplitude modulation ensures that the real part of the spectrum is purely absorptive but does not provide sign discrimination. Additional acquisition methods (section 2.2.6.2) are required for this.

2.2.6.2. 2D sign discrimination

A common approach to provide sign discrimination when acquiring 2D datasets is to employ the States method.[158] The FID for each value of t_1 is recorded twice, with a difference in phase of $\pi/2|p|$ between the preparation pulses used for each. This results in the two signals being 90° out of phase with respect to each other, with the second sine-modulated:

$$s_{am,sin}(t_1, t_2) = 2i \sin \Omega t_1 \, e^{-\frac{t_1}{T_2^{(1)}}} e^{i\Omega t_2} e^{-\frac{t_2}{T_2^{(2)}}} \quad (137)$$

The FT of Eqn. 137, with the real and imaginary t_1 components transformed separately as general amplitude modulation, produces the sine modulated equivalent of Eqn. 135 as:

$$s_{am,sin}^{Re}(\omega_1, \omega_2) = (A_1^- - A_1^+)A_2^+ - i(D_1^- - D_1^+)A_2^+ \quad (138)$$

The difference between the real, absorptive components of the two signals (the difference between Eqn. 135 and Eqn. 138) is purely absorptive and sign-discriminated:

$$(A_1^- + A_1^+)A_2^+ - (A_1^- - A_1^+)A_2^+ = 2A_1^+A_2^+ \quad (139)$$

An alternative approach is the two pulse phase incremented (TPPI) method, whereby a single FID is recorded for each t_1 but Δt_1 is halved with respect to the spectral width, sω_1. A $\pi/2|p|$ phase change is introduced for each subsequent FID, modulating the coherence by a t_1 specific phase and thus ensuring sign discrimination.

Within this work, the combined approach, States-TPPI,[159] is employed. Two FIDs are recorded for each value of t_1 but the phase is modulated as 0°-90°-180°-270° rather than 0°-90°-0°-90° for SQ t_1 evolution.

2.2.6.3. Phase cycling

Coherences are excited by applying *rf* pulses but a single pulse introduces a range of coherences relating to the possible eigenstates and often only specific ones are desired, corresponding to filtering with respect to certain interactions. Selection of the desired coherence pathways is achieved by phase cycling. The pulse sequence is repeated *N* times and the phases of certain pulses and the receiver are altered between repeats. Addition over all FIDs removes unwanted coherence pathways through destructive interference. The 'Golden Rules' of phase cycling that must be applied are:

1. Changing the phase of a pulse or group of pulse by $\Delta\varphi$ results in a phase shift to any coherence that is changing coherence order $\longrightarrow -\Delta\varphi\Delta p$.
2. For a phase cycle step of 360°/*N*, in addition to the desired coherence change, Δp, pathways corresponding to $\Delta p \pm nN$ will also be selected. All other pathways will be blocked.

The initial and final coherences are always fixed as the initial state is always thermal equilibrium, $p = 0$, and only in-phase SQ coherences are observable directly (generally $p = -1$ is chosen). For the phase cycle to correctly select and remove the desired pulses, the total number of FIDs must be an integer multiple of *N*, otherwise interference between the signal of subsequent scans will be incomplete. The second rule results in selection of coherences of higher quantum order as well as the desired one but these can generally be ignored as excitation efficiency decreases with increasing quantum order. For complex experiments, it is possible to employ nested phase cycles to allow for multiple coherence changes. These can select symmetric positive and negative coherence pathways and so permit amplitude

modulated acquisition. The receiver phase is the sum of the individual receiver phases for each coherence change.

2.2.6.4. Decoupling

Organic materials typically have dense proton environments resulting in significant hetero- and homo-nuclear dipolar couplings that contribute to the broadness of the NMR linewidths (see Fig. 2.11 in section 2.2.5.2). Decoupling techniques are applied to reduce the NMR linewidths.

When observing 'dilute' nuclei, with low natural abundance, it is assumed that the H-X heteronuclear dipolar coupling is the most significant interaction remaining under MAS. Heteronuclear dipolar decoupling is generally applied as a continuous, high *rf* frequency to the abundant nuclei, exciting continuous transitions between its $|\alpha\rangle$ and $|\beta\rangle$ states. If the *rf* nutation frequency, $\omega_1/2\pi$, is fast enough, *i.e.* more than three times larger than the heteronuclear dipolar coupling, the rapid transitions aid in averaging the coupling. For ^{1}H-^{13}C dipolar coupling an $\omega_1/2\pi$ of 100 kHz is sufficient. To avoid a periodicity being introduced when decoupling is employed alongside MAS, it is implemented in blocks with phase switching between pulses. The small phase incremental alternation with 64 steps (SPINAL-64)[160] approach to this is employed within this work.

^{1}H-^{1}H homonuclear couplings present even more of a problem in proton dense systems as they can only partially be removed under MAS, although increasing the spinning frequency can produce significant improvements. It is therefore common to employ homonuclear decoupling sequences. The method employed in this work is Decoupling Using Mind Boggling Optimisation (DUMBO), specifically eDUMBO$_{22}$.[161, 162] The DUMBO sequences apply on-resonance pulses, at a constant nutation frequency, with rapid phase switching between the discrete pulses (on the order of 100 ns). The original sequence DUMBO-1 was optimised computationally for static conditions while eDUMBO-1$_{22}$ was further optimised experimentally at 22 kHz MAS.

2.2.6.5. Cross-polarization

Cross-polarisation, CP, enhances the magnetisation of dilute, low-γ nuclei, by exploiting heteronuclear dipolar couplings to transfer polarization from an abundant, high-γ nucleus, like ^{1}H. This is particularly useful for ^{13}C as its natural abundance is only 1.1 % (^{12}C is the most abundant carbon isotope but is NMR invisible as $I = 0$) and its relatively low γ also typically results in slow spin-lattice relaxation compared to ^{1}H. The CP MAS pulse sequence is shown in Fig. 2.16, below.

Cross-polarisation occurs when the Hartmann-Hahn matching condition[163] is satisfied:

$$\gamma_I B_{1I} = \gamma_S B_{1S} \tag{140}$$

In this work, spin *I* always corresponds to ^{1}H and spin *S* to ^{13}C. The Hartmann-Hahn condition is valid as written in Eqn.140 when under static conditions. When the system is under MAS, it can be rewritten as:

$$\gamma_I B_{1I} = \gamma_S B_{1S} + n\omega_R \quad \text{for} \quad n = 1, 2, \ldots \tag{141}$$

CP transfer occurs via heteronuclear dipolar couplings so magnetisation transfers most quickly between nuclei that are near in space, with longer contact times required to ensure transfer for longer range couplings. The peak intensities are therefore indicative of the transfer distance/efficiency rather than the number of chemical environments for that resonance. The chosen contact time also needs to take longitudinal relaxation during the *rf* pulse, $T_{1\rho}$, into account as it will make lengthy contact times inefficient. Fast MAS also decreases the efficiency of transfer, although ramped contact pulses have been found to partially compensate for this.[164]

The 2D ^{1}H-^{13}C heteronuclear correlation (HETCOR) experiment employed in this work is based on CP MAS transfer with eDUMBO-1_{22} homonuclear-decoupling (2.2.6.4) applied during t_1 to reduce the ^{1}H linewidths.

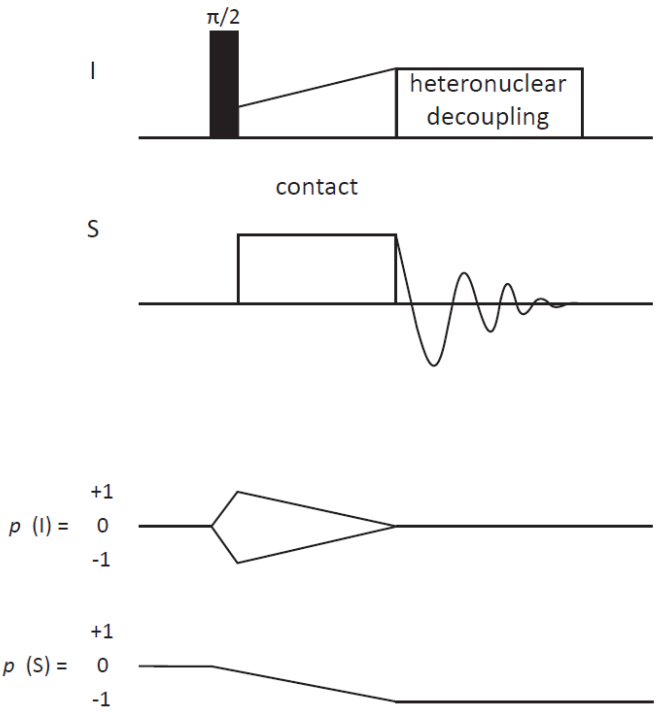


Figure 2.16: A ramped CP pulse sequence alongside the coherence transfer pathway for spins I and S. Heteronuclear decoupling (in this work SPINAL-64) is applied during acquisition to remove strong heteronuclear dipolar couplings.

2.2.6.6. ^{1}H-^{1}H DQ experiment

The ^{1}H-^{1}H DQ experiments conducted in this work employ the Back to Back (BaBa) method[165, 166] for dipolar recoupling. This consists of four $\pi/2$ pulses per rotor period, τ_R, although this can be repeated over multiples of τ_R to produce longer dipolar recoupling times. BaBa recoupling is a non-γ encoded DQ recoupling sequence and has an inherent dependence on the amplitude of the rotor phase (Fig. 2.17). 2D experiments therefore need to be recorded in a rotor synchronised manner to prevent rotor-encoded spinning sidebands in the DQ dimension. The sidebands are folded back onto the centreband under rotor synchronisation which leads to an increase sensitivity. The number of t_1 increments needed is also reduced, making it possible to cut down experimental times.

The dependence of the dipolar coupling on the interatomic distance ($d_{IS} \propto 1/r_{IS}^3$, in Eqn. 96) results in the maximum intensities, I, and their build-up with recoupling time being indicative of relative ^{1}H-^{1}H distance.[106, 167, 168] Furthermore, the relative intensities of DQ

peaks for a specific SQ resonance relate, to a good approximation, to the ratio of squares of the relevant dipolar coupling constants, allowing an experimental determination of the relative distances from a single experiment:

$$\frac{d_{IS_1}^2}{d_{IS_2}^2} = \frac{r_{IS_2}^6}{r_{IS_1}^6} \tag{142}$$

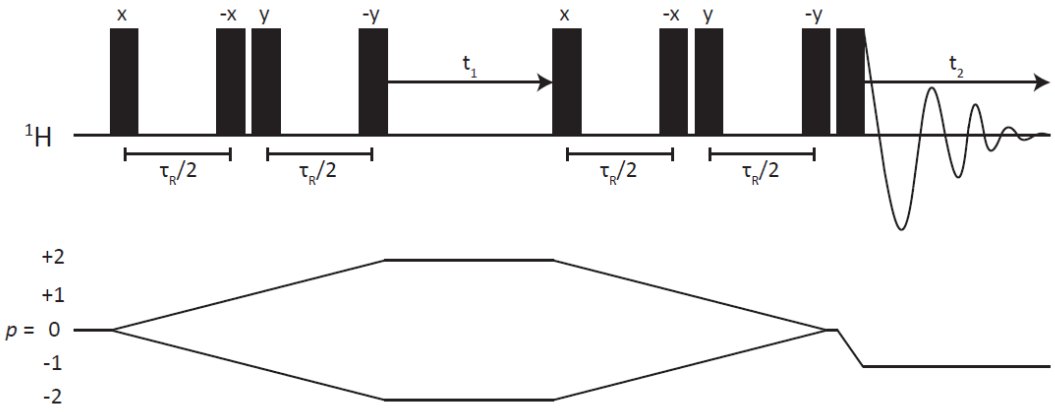


Figure 2.17: A ^{1}H-^{1}H DQ/SQ pulse sequence with BaBa recoupling alongside the coherence transfer pathway. An initial block of BaBa recoupling is employed for DQ excitation and then a second is employed after t_1 for DQ reconversion.

2.2.6.7. ^{1}H-^{1}H spin diffusion

A ^{1}H-^{1}H SQ spin diffusion experiment, where *z* magnetisation is transferred through homonuclear dipolar couplings, can also be employed to probe structural features of systems. The same pulse sequence (Fig. 2.18) is used in both solid-sate and solution state NMR, where,

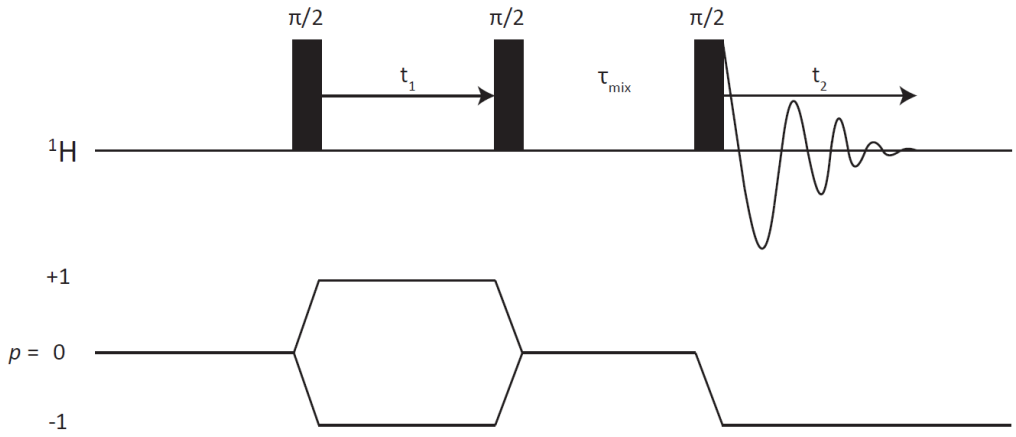


Figure 2.18: A ^{1}H-^{1}H SQ/SQ spin diffusion (NOESY like) pulse sequence alongside the coherence transfer pathway.

in the latter case, polarisation transfer is *via* the Nuclear Overhauser Effect (NOE) in the NOESY experiment. Although these experiments are often used to probe dynamics or, over a range of mixing times, τ_{mix}, to investigate the relative distances, in this work the ^{1}H-^{1}H SQ spin diffusion experiment is used to determine the number of separate phases within a sample. For a sufficiently long τ_{mix}, all resonances within the same phase will transfer magnetisation between all others, allowing the identification of more than one phase by noting the absence of correlations.

2.2.6.8. HMQC

Heteronuclear Multiple Quantum Coherence (HMQC)[169-172] experiments can be utilised to determine correlations between different nuclear species. Within this work they are utilized for ^{14}N-^{1}H correlations, with direct detection on the ^{1}H channel to provide a sensitivity enhancement relative to detection on the ^{14}N channel (Fig. 2.19). This is due to ^{1}H being high-γ and ^{14}N low-γ, resulting in a larger magnetic moment and Larmor frequency for ^{1}H (as both of these properties are proportional to γ).

Rotary resonance recoupling (R^3)[173, 174] is employed to reintroduce the heteronuclear dipolar couplings that are removed under MAS (alternative dipolar recoupling schemes or magnetisation transfer *via J*-couplings are possible). For the rotary resonance condition, $\omega_1 = n\omega_R$, at $n = 2$ there is efficient heteronuclear dipolar (and CSA) recoupling with simultaneous ^{1}H-^{1}H homonuclear dipolar decoupling. The *rf* pulses for R^3 are only applied to the ^{1}H channel as the *rf* nutation frequencies are too small to manipulate the large ^{14}N quadrupolar interaction. Additionally, there is also an increase in ^{1}H coherence lifetimes as a result of the recoupling, increasing the signal.

Application of a 4-step nested phase cycle (2-steps on the first ^{1}H pulse and 2-steps on the last ^{14}N pulse) selects two SQ changes of coherence order $\Delta p = \pm 1$.

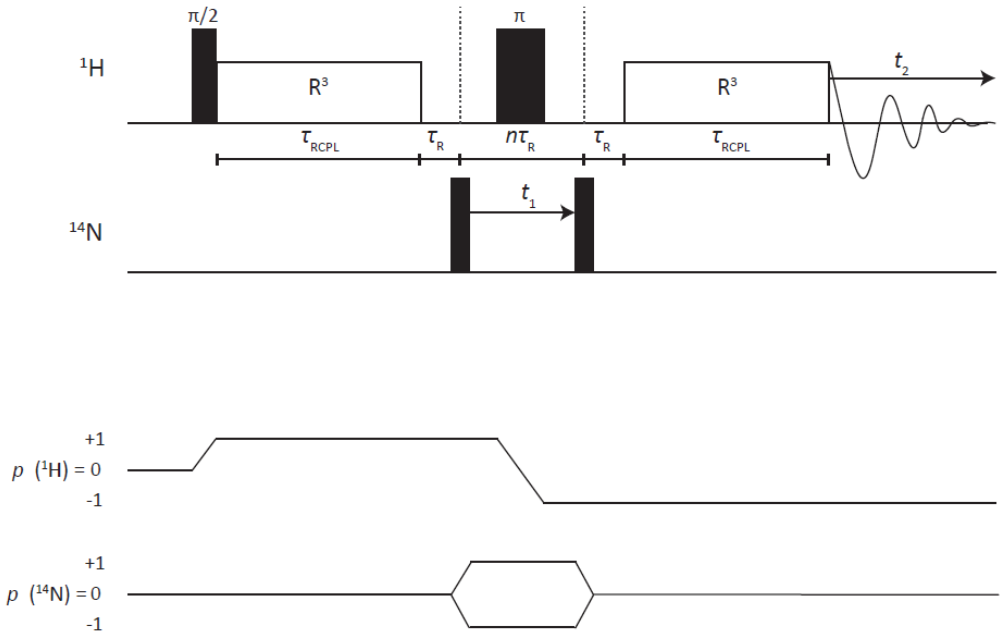


Figure 2.19: A ^{14}N-^{1}H HMQC pulse sequence, with R^3 employed as the recoupling scheme, alongside the coherence transfer pathway

2.2.7. Experimental Details

Solid-state NMR experiments were performed on:

1. A Bruker Avance III spectrometer, operating at ^{1}H and ^{13}C Larmor frequencies of 500.0 MHz and 125.8 MHz, respectively.
2. A Bruker Avance II+ spectrometer, operating at ^{1}H, ^{13}C and ^{14}N Larmor frequencies of 600.0 MHz, 150.7 MHz and 43.4 MHz, respectively
3. A Bruker Avance III HD spectrometer, operating at ^{1}H, ^{13}C and ^{14}N Larmor frequencies of 700.0 MHz, 176.0 MHz and 50.6 MHz, respectively

1.3 mm HXY probes in double resonance mode were used for ^{1}H one-pulse MAS, 2D ^{1}H SQ spin-diffusion (NOESY-type), 2D ^{1}H DQ (with one rotor period of BaBa recoupling[165, 166]) and 2D ^{14}N-^{1}H HMQC[169-172] (with R^3 recoupling[173, 174]) experiments. Experimental details for each set of experiments run for the systems studied are recorded in Table 4. The spectrometer (S), number of coadded transients (ns), recycle delay (d1), number of t_1 FIDs and spin-diffusion mixing time (d8) are reported as appropriate. In all cases (except during CP), a ^{1}H nutation frequency of 100 kHz was used corresponding to a ^{1}H 90° pulse duration of 2.5 μs.

Each 2D fast MAS experiment employed a rotor synchronized t_1 increment of 16.67 μs and 133 μs of R^3 recoupling was employed for all 2D ^{14}N-^{1}H HMQC.

Table 2.4: Experimental parameters for fast MAS NMR experiments for each of the systems.

	^{1}H one-pulse			**2D ^{1}H DQ**				**2D ^{1}H SQ**					**^{14}N-^{1}H HMQC**			
	S	ns	d1 (s)	S	ns	FIDs	d1 (s)	S	ns	FIDs	d1 (s)	d8 (s)	S	ns	FIDs	d1 (s)
26L:HF	**1**	16	100	**2**	32	224	10	**2**	4	82	100	3	**2**	16	240	60
25L:$F_{0.5}$:FA	**2**	8	10	**2**	32	256	10	**2**	16	128	10	0.5	**2**	16	192	15
25AMP:$F_{0.5}$:$FA_{0.5}$	**3**	8	5	**3**	16	256	20	**1**	8	128	20	0.5	**2**	32	128	50
52AMP:HF	**2**	8	80	**2**	8	128	80	**2**	8	128	80	1	**2**	8	128	80
26AMP:$F_{0.5}$:$(H_2O)_2$	**2**	4	75	**2**	16	128	75	**2**	8	136	75	0.5	**3**	8	136	80
26AMP_2:F:H_2O	**2**	4	80	**2**	8	128	80	**2**	8	128	80	0.5	**2**	8	216	80
26AMP_2:F:$FA_{0.5}$	**2**	8	60	**2**	8	192	60	**2**	8	152	60	0.3	**2**	32	82	65

Details of the ^{1}H-^{13}C experiments, CP MAS and HETCOR, are listed in Table 5, including the CP mixing time (CP), t_1 increment and scaling factor (SF). In the HETCOR experiments, eDUMBO-1_{22}[161, 162] homonuclear decoupling was used with a 32 μs cycle, with 320 divisions of 100 ns each. In the HETCOR pulse sequence, the following phase cycling was employed: ^{1}H 90° pulse (90º 270°), ^{13}C CP contact pulse (2{0°} 2{180°} 2{90°} 2{270°}), receiver (0° 180° 180° 0° 90° 270° 270° 90°). For both CP MAS and HETCOR ^{1}H-^{13}C experiments, SPINAL64 ^{1}H heteronuclear decoupling was applied during the acquisition of the ^{13}C FID, with a pulse duration of 5.9 μs at a nutation frequency of 100 kHz, and a 70 to 100% ramp[164] on the ^{1}H channel was used for the CP contact time with nutation frequencies of 47.5 and 60 kHz for ^{13}C and ^{1}H, respectively. In 2D experiments, the States-TPPI method was employed for sign discrimination.[159]

Table 2.5: Experimental details for ^{1}H-^{13}C experiments for each of the systems.

			^{1}H-^{13}C CP			^{1}H-^{13}C HETCOR				
	S	Probe	ns	CP (μs)	d1 (s)	ns	FIDs	t1	CP (μs)	SF
26L:HF	**1**	4 mm HX	32	750	78	16	72	80	200	1.8
25L:$F_{0.5}$:FA	**1**	4 mm HX	32	1500	60	16	128		500	1.4
25AMP:$F_{0.5}$:$FA_{0.5}$	**2**	2.5 mm HX	64	2000	70	40	60		500	1.8
52AMP:HF	**3**	3.2 mm HX	16	1500	80	24	120	36	500	1.6
26AMP:$F_{0.5}$:$(H_2O)_2$	**2**	3.2 mm HX	32	1500	80	24	72	80	500	1.8
26AMP_2:F:$FA_{0.5}$	**2**	3.2 mm HX	16	1500	30	64	128	73	500	1.6
						64	52	40	200	1.6

^{13}C and ^{1}H chemical shifts are referenced with respect to TMS via L-alanine at natural abundance as a secondary reference (1.1 ppm for the CH_3 ^{1}H resonance and 177.8 ppm for the CO ^{13}C resonance) corresponding to adamantane at 1.85 ppm (^{1}H)[175] and 38.5 ppm (^{13}C)[176]. ^{14}N shifts are referenced with respect to a saturated NH_4Cl aqueous solution via spectra of L-β-aspartyl-L-alanine at natural abundance (−284 ppm for the lower NH resonance at a Larmor frequency of 43.4 MHz) corresponding to liquid CH_3NO_2 at 0 ppm.[171, 177] ^{1}H, ^{13}C and ^{14}N shifts can be experimentally determined to an accuracy of ± 0.2, ± 0.1 and ± 5 ppm, respectively. Solution-state NMR was carried out on a Bruker Avance III spectrometer operating at a ^{1}H Larmor frequency of 400.0 MHz. Samples were dissolved in DMSO-d_6 at a concentration of approximately 10 mg/mL. 1D ^{1}H one-pulse experiments were referenced to TMS via the residual solvent peak.

2.3. DFT

2.3.1. Schrödinger Equation

The Schrödinger equation (SE) is the starting point of quantum chemistry. The wavefunction is a function of coordinates of all electrons and ions (and depends parametrically on time). It cannot be solved exactly for any atom larger than hydrogen but most quantum chemical approaches use approximations to come to a valid solution. The time-independent, non-relativistic SE can be written as:

$$\hat{H}\psi_i(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, \vec{R}_1, \vec{R}_2, \dots, \vec{R}_M) = E_i\psi_i(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, \vec{R}_1, \vec{R}_2, \dots, \vec{R}_M) \quad (143)$$

where $\vec{r}_i$ and $\vec{R}_i$ are the positions of the i^{th} electron and nucleus, respectively, $\hat{H}$ is the Hamiltonian operator for the total energy of a system containing *N* electrons and *M* nuclei in the absence of external fields and E_i is the energy of the state ψ_i. $\hat{H}$ is of the form:

$$\hat{H} = -\frac{1}{2}\sum_{A=1}^{M}\frac{1}{M_A}\nabla_A^2 - \frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} \quad (144)$$

using atomic mass units and where:

$M_A \longrightarrow$ mass of nucleus A

$Z_A \longrightarrow$ charge of nucleus A

$r_{iA} \longrightarrow$ interaction distance between electron i and nucleus A

$r_{ij} \longrightarrow$ interaction distance between electron i and electron j

$R_{AB} \longrightarrow$ the internuclear distance between nucleus A and nucleus B

2.3.2. Born-Oppenheimer approximation

In the Born-Oppenheimer approximation, the nuclei are considered fixed as they move so slowly as compared to the electrons. This means that Eqn. 144 can easily be separated into two parts, the electronic Hamiltonian, $\widehat{H}_{elec}$, which, applied to the electronic wavefunction, gives the electronic energy, E_{elec}, and a constant nuclear repulsion energy, E_{nuc}:

$$\widehat{H}_{elec} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} = \hat{T} + \hat{V}_{ee} + \hat{V}_{Ne} \tag{145}$$

$$E_{nuc} = \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} \tag{146}$$

In Eqn. 145, $\hat{T}$ is the electronic kinetic energy, $\hat{V}_{ee}$ the electronic potential energy and $\hat{V}_{Ne}$ the mixed potential energy. The total energy of the system under the Born-Oppenheimer approximation is simply:

$$E_{tot} = E_{elec} + E_{nuc} \tag{147}$$

2.3.3. Hohenberg-Kohn

Hohenberg and Kohn developed DFT as an exact theory describing many-body systems.[178]

Theorem 1 For an electronic system in any external potential $V_{ext}(\boldsymbol{r})$, the said potential is determined, uniquely, save for a constant, by the ground state electron density, $\rho_0(\boldsymbol{r})$. $V_{ext}(\boldsymbol{r})$ in turn fixes $\widehat{H}$, fully determining it.

Theorem 2 A universal functional of the energy, $E[\rho(\boldsymbol{r})]$, can be defined in terms of the density $\rho(\boldsymbol{r})$, which is valid for any external potential $V_{ext}(\boldsymbol{r})$. For any $V_{ext}(\boldsymbol{r})$, the exact ground state is determined by the global minimum value of $E[\rho(\boldsymbol{r})]$.

The consequence of Theorem 1 is that the many body wavefunctions are determined. Therefore, all properties of the system are determined from the ground state density. By Theorem 2, $E[\rho(\boldsymbol{r})]$ is therefore all that is required to determine the ground state density and thus the ground state energy. Note that DFT is a ground state theory as the above theorems rely on application of the variational principle which is limited to the ground state – therefore,

excited states have to be determined by other means. Separating $E[\rho(\boldsymbol{r})]$ out into the system dependent (depends on $\rho(\boldsymbol{r})$) and independent (no dependence on *N*, *etc.*) parts, it can be written as:

$$E[\rho(\boldsymbol{r})] = F[\rho(\boldsymbol{r})] + \int V_{ext}(\boldsymbol{r})\rho(\boldsymbol{r})d\boldsymbol{r} \tag{148}$$

Although the Hohenberg-Kohn theorems reformulate the many-body problem in terms of the electron density, they do not actually solve it as the energy functional is *a priori* unknown.

2.3.4. Kohn-Sham

Within the Kohn-Sham scheme,[179] the many-body problem is replaced with a set of single particle problems that can be summed over. This is achieved by replacing the fully interacting electrons with a system of imaginary non-interacting particles in an effective potential. It is based on the assumption that the ground state density of the fully interacting system is equal to some non-interacting system that is exactly soluble with all the complex parts (exchange and correlation) included in some approximate density functional, $E_{xc}(\boldsymbol{r})$. Under this, the independent part of Eqn. 148 becomes:

$$F[\rho(\boldsymbol{r})] = T_0[\rho(\boldsymbol{r})] + \frac{1}{2}\int\int \frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r}d\boldsymbol{r}' + E_{xc}(\boldsymbol{r}) \tag{149}$$

In Eqn. 149, $T_0[\rho(\boldsymbol{r})]$ corresponds to the non-interacting electron kinetic energy and the remainder of the energy is made up of the classical Coulomb energy, $E_H[\rho(\boldsymbol{r})]$, and $E_{xc}(\boldsymbol{r})$, which accounts for the non-classical self-interaction correction.

Alongside the system dependent component of the energy functional, for which $V_{ext}(\boldsymbol{r})$ is merely the potential energy from the nuclei, the total energy of the system can be written as:

$$E[\rho(\boldsymbol{r})] = T_0[\rho(\boldsymbol{r})] + \int d\boldsymbol{r}\rho(\boldsymbol{r})V_{ext}(\boldsymbol{r}) + E_H[\rho(\boldsymbol{r})] + E_{xc}(\boldsymbol{r}) \tag{150}$$

The only unknown is now $E_{xc}(\boldsymbol{r})$, which accounts for around 10% of the total energy. In practice, DFT is implemented by approximating $E_{xc}(\boldsymbol{r})$. The simplest approach is to use a Local Density Approximation, where, with the imaginary electrons modelled as a homogeneous electron gas, the contribution to $E_{xc}(\boldsymbol{r})$ from $\boldsymbol{r}$ is considered to only depend on $\rho(\boldsymbol{r})$. Many approximations also utilise the gradient of ρ and are thus termed Generalised Gradient Approximations (GGAs) – the functional developed by Perdew, Burke and Ernzerhof (PBE),[180] that is used in this thesis, is of this type and has two non-empirical derivations. Meta-GGAs utilise the kinetic energy density to improve the approximation but are often semi-empirical (*e.g.* PKZB[181]), so liable to fail outside their fitting sets, or provide only moderate improvement over standard GGAs (*e.g.* TPSS[182]) while reducing computational efficiency. Global hybrid GGAs (*e.g.* B3LYP[183] and PBE0[184]), on the other hand, replace a fixed proportion of the local exchange with Hartree-Fock exchange (this proportion is not fixed but dependent on density/density gradient etc. in non-global hybrids) but are not as easily evaluated for solids.[182, 185]

2.3.5. Dispersion Correction

One common flaw in many of the exchange correlation functionals discussed above is in the lack of description of long range, dispersive forces like van der Waals. Standard functionals only consider local properties, within the region where there is an overlap between electron densities, which decay exponentially with $\boldsymbol{r}$. Dispersive forces, however, are attractive interactions caused by the response of one region's electrons to the charge density fluctuations of another and follow a dipole-dipole relationship, so their intensity drops off as $-1/\boldsymbol{r}^{6}$.[186] This can cause significant problems within DFT calculations, leading to the development of numerous correction schemes. In molecular crystals, when unit cell parameters are allowed to vary, an expansion is seen as molecules move away from each other if dispersive forces are absent, whereas slight contraction is generally seen (as would be expected for calculations that are performed at nominal 0 K) when a correction scheme is included. The most basic of these

(DFT-D methods) simply add an additional energy term based on dispersion coefficients C_6, summing over elemental pairs but this excludes many-body effects and faster decaying terms as well as raising questions over the best generation of C_6 coefficients. Methods such as DFT-D3[187] and TS[188] (utilised in this work) follow the same approach but with variable C_6 coefficients depending on the atomic environment (rather than fixed for all instances of an atom regardless of, *e.g.*, oxidation state).[186] TS achieves this by simply combining reference atomic polarizabilities and C_6 coefficients, resulting in high accuracy although it is unclear how it will cope with either increasing volumes or ionic systems.[188]

Schemes incorporating the dispersive forces within the exchange correlation functional, determining the interactions directly form the electron density, have also been developed but, although they are less likely to fail for systems outside a fitting space (as they are non-empirical) and the computational cost has been successfully reduced,[189] they are both complex and generally lower in precision. Calculation of the correlation energy by utilising orbitals (ACFDT[190] and DFT/post-HF methods)[191] produce very accurate results but at a high computational cost. Incorporation of many-body dispersion (MBD) forces is also possible by treating each atomic position as a quantum harmonic oscillator with a characteristic frequency and determining the shifts in frequency as each is switched on, however, it is challenging to get accurate relations between oscillator models and atoms.[192, 193]

2.3.6. Computational Details

DFT calculations were performed using CASTEP[194] Academic Release version 16.1. All calculations used the PBE exchange correlation functional,[180] a plane-wave basis set with ultrasoft pseudopotentials and a plane-wave cut-off energy of 700 eV. Integrals over the Brillouin zone were taken using a Monkhorst-Pack grid of minimum sample spacing $0.1 \times 2\pi$ $Å^{-1}$ (unless otherwise stated). For all structures for which the NMR parameters were calculated, geometry optimisation was performed with the unit cell parameters fixed. Geometry optimisations were also run for numerous structures allowing the unit cell

parameters to vary (26L:HF, discussed in Chapter 3, and the 26AMP:F hydrate structures following removal of the water molecules, discussed in Chapter 4), with TS dispersion correction.[188] MOGUL[195] searches were performed both before and after geometry optimisation to ensure that the bond lengths, torsion and angles were consistent with the CCDC database.

NMR parameters were calculated using the GIPAW[107] method and were performed for both the geometry optimised crystal structures as a whole and for each of the isolated molecules in the asymmetric unit. For the isolated molecule calculations, each molecule in the asymmetric unit was extracted from the geometry optimised unit cell and placed in a sufficiently large box such that it could not interact with repeated molecules across periodic boundary conditions[118] (unit cell dimensions increased by 10 Å in each direction). As each individual molecule carried a charge, this was specified in the .param file.[196]

The calculated isotropic chemical shifts (δ_{iso}^{calc}) were determined from the calculated chemical shieldings (σ_{calc}) by $\delta_{iso}^{calc} = \sigma_{ref} - \sigma_{calc}$, with calculated σ_{ref} values as recorded in Table 6. σ_{ref} was determined for ^{1}H and ^{13}C by taking the sum of the experimental chemical shift and the GIPAW calculated absolute isotropic chemical shieldings. The resulting y-intercept was taken as σ_{ref}.[117, 197] A literature value of −153 ppm was used for ^{14}N.[198] It is noted that it is common practice to calculate a specific reference shielding for each system[199] (see, *e.g.*, Supplementary Table 8 in reference), though average values over a range of compounds are also available.[200]

Table 2.6: Reference shielding, σ_{ref}, for each system.

	σ_{ref} (ppm)	
	^{1}H	^{13}C
26L:HF	30.5	169.7
25L:$F_{0.5}$:FA	29.9	169.5
25AMP:$F_{0.5}$:$FA_{0.5}$	29.9	169.7
52AMP:HF	30.0	170.8
26AMP:$F_{0.5}$:$(H_2O)_2$	29.9	170.7
26AMP_2:F:H_2O	29.9	170.4
26AMP_2:F:$FA_{0.5}$	30.0	170.4

Chapter 3. Investigation of the Structure of 2,6-Lutidinium Hydrogen Fumarate

3.1. Introduction

The crystal structure of 26L with HF (Fig. 3.1a and b) was first published in 2001 by Pan *et al.*, CCDC ref.: MIBYEB, no.: 181445.[139] It is a 1 : 1 salt, acid : base, which crystallises in the P21/c space group. Hydrogen fumarate molecules, which take the s-*cis*/s-*trans* conformation, form into chain structures running through the system, each with a 2,6-lutidinium (26L) molecule hydrogen bonded to the carboxylate group and lying at 76.5° to the acid chain (Fig. 3.1c). Pairs of chains align to form stacks of lutidine molecules which pack together to form the 3D structure (Fig. 3.1d). The interaction between the acid and base has been determined from SXRD to be ionic and so, as 26L is liquid at room temperature, the system can be described as a salt solvate.[6] The melting point of 26L is –6°C.

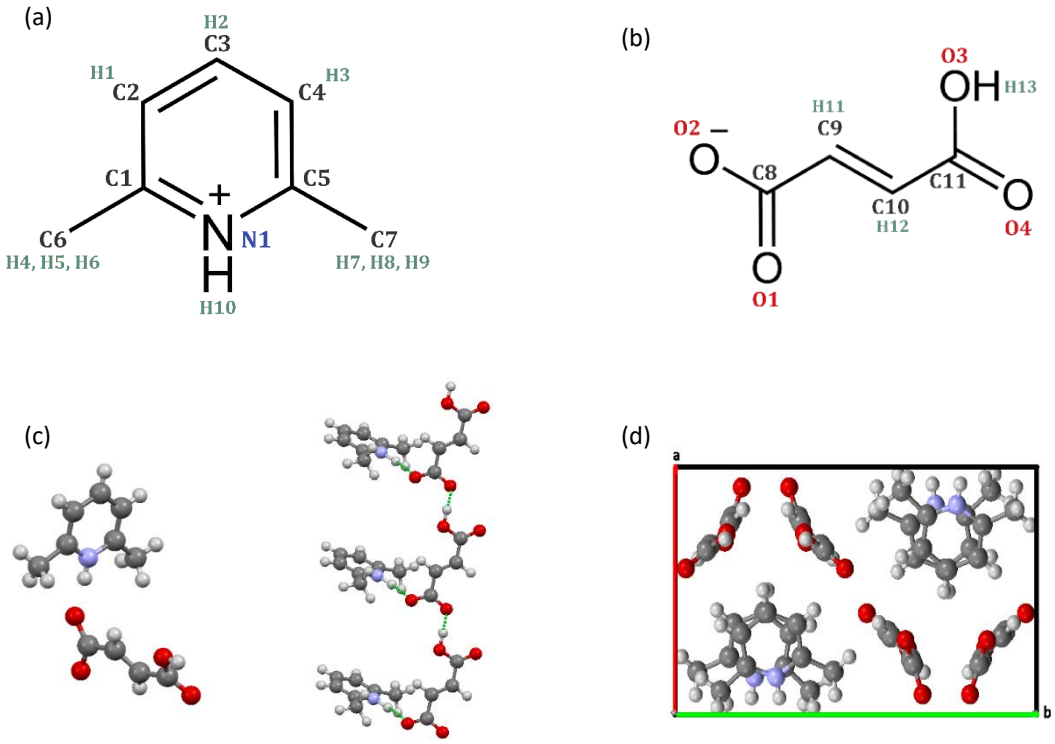


Figure 3.1: Molecular structures of (a) 2,6-lutidinium and (b) hydrogen fumarate molecules with the atomic labels used in this work; (c) asymmetric unit of the DFT-optimised crystal structure of 26L:HF (left) and stacking of the asymmetric units to form hydrogen-bonded chain structures (right); (d) packing of the unit cell viewed along the c axis. Carbon, oxygen, nitrogen and hydrogen atoms are shown as grey, red, purple/blue and white circles.

In this chapter, the intermolecular interactions and stability of 26L:HF are investigated using a combination of NMR, XRD and DFT, using the NMR crystallography approach discussed in Chapter 1. As published in 2019,[201] alongside the structural investigation, additional thermal analysis techniques are utilised to aid elucidation of 26L:HF degradation.

3.2. Sample Preparation

All chemicals were obtained from Sigma Aldrich (UK) at purities of 98% or higher and used without further purification. Purity was verified by ^{1}H solution-state NMR. FA was dissolved in isopropanol and 26L was added in a 1:1 molar ratio. Samples were made on a ~50 mg scale with 24 mg FA and 26 μL 26L (with density 0.92 mg / L) in minimum amount of solvent required to dissolve all the solute (~10 mL). Crystallisation was achieved by slow solvent evaporation under ambient conditions over approximately 4 days. Crystal growth in subsequent crystallisation attempts was found to be improved by the addition of seed crystals to the crystallisation media. These seed crystals were selected from the small crystals present in the initial samples and were added when the saturated solution was placed under ambient conditions.

3.3. NMR Crystallography Approach

3.3.1. XRD

Single crystals of 26L:HF were successfully obtained by slow evaporation. Crystals exhibited a thin plate morphology and, upon closer inspection with an optical microscope, thicker regions of the specimen could be seen as stacks of multiple thin plates, with some evidence of intergrowths.

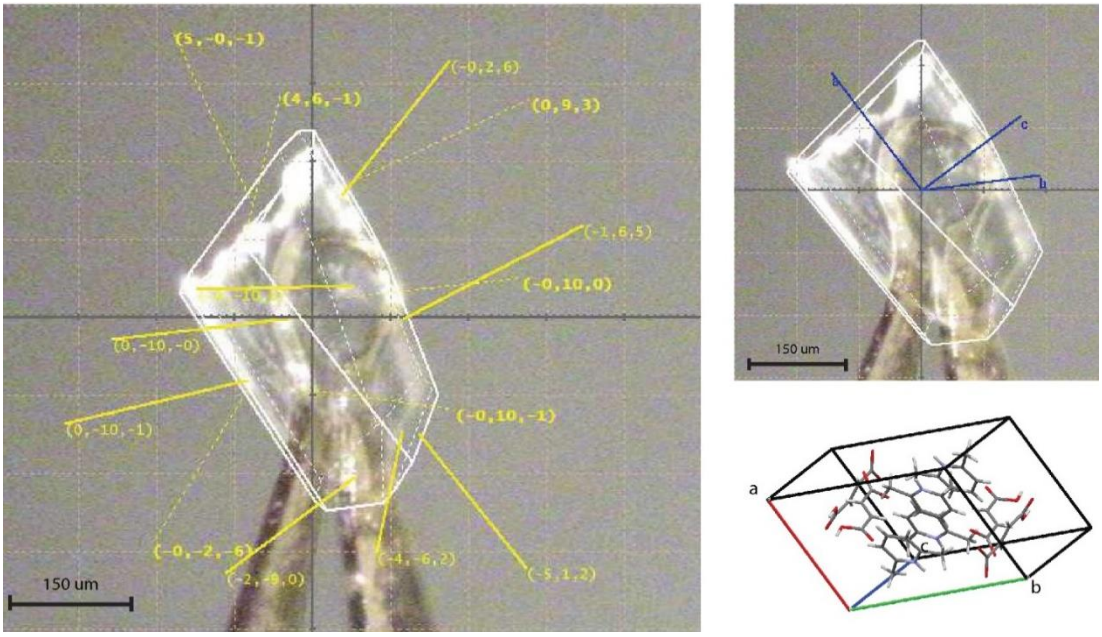


Figure 3.2: A 26L:HF plate crystal of sufficient quality for SXRD with (left) labelling of the (hkl) planes for each face and (top right) the corresponding alignment of the unit cell axes. The literature structure (181445) unit cell, oriented the same way, is also shown (bottom right).

To ensure that each sample was homogeneous, room temperature SXRD at 293 K was utilised to screen each crystallisation for unit cell variations and polymorphism. Most of the crystals found to be suitable for diffraction refined to the structure, published by Pan *et al.*[139] (CCDC ref.: MIBYEB, no.: 181445, see Fig. 3.1). The single exception to this is discussed below in section 3.4. During the crystal screening, the unit cell of 181445 was also seen in many other crystals judged unsuitable for full data collection due to poor diffraction. For all the plate crystals analysed by SXRD, the (0 1 0) plane constituted the largest face, exhibiting least growth (Fig. 3.2). This corresponds to a low drive for aligning the hydrogen-bonded fumarate chains (Fig. 3.1c) along the *b* axis. Conversely, the smallest crystal face, corresponding to the most favourable growth, relates to alignment of the chain structures along the *a* axis forming the (1 0 0) crystal plane. The growth of the chains themselves, that results in the stacks of lutidine molecules (Fig. 3.1c), is moderate.

Following grinding to a powder, room temperature PXRD patterns were run to ensure no changes had been induced during the grinding process. Fig. 3.3 shows the result of a Rietveld refinement against a high-resolution synchrotron scan of 26L:HF carried out under ambient conditions with a R_{bragg} of 5.59%. As can be seen in Table 3.1, there is good agreement

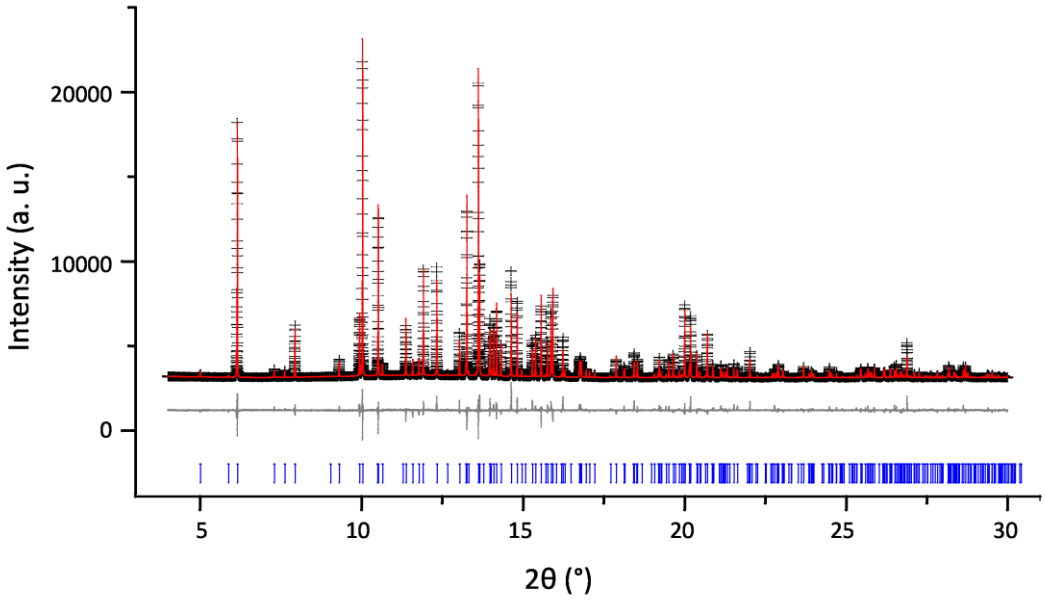


Figure 3.3: Final Rietveld fit for 26L:HF at 300 K, λ = 0.8249 Å, showing the experimental (black crosses), calculated (red upper line) and difference (grey lower line) PXRD profiles. Tick marks (bottom) indicate allowed peak positions. For refinement parameters, see Table 3.1 and Tables A1.1 and A1.2.

between the refined unit cell (MAC 300 K) and that of structure 181445 published by Pan *et al.*,[139] with only small differences in the unit cell parameters consistent with the small difference in temperature between data collections. A more detailed comparison of atomic coordinates is presented in Table A1.1 and the final refinement details are shown in Table A1.2 in Appendix 1.

In the Rietveld refinement all the experimental reflections (as shown by the tick marks) are replicated in the calculated pattern and their 2θ positions are in excellent agreement. Although there was some evidence of differences in peak height and shape between experiment and calculation, as seen in the difference plot in Fig. 3.3, this can be explained by residual preferred orientation effects. The natural plate morphology of the crystals results in strong preferred orientation effects due to the alignment of crystallites. Minimisation of this by grinding the powder more finely, to allow the collection of better PXRD data, was hindered by breakdown of the crystal structure if the sample was ground for too long (discussed in section 3.5, below). Most powders analysed therefore still contained larger micro-crystallites, exacerbating the preferred orientation effects. Taking these effects into account, the Rietveld refinement of the PXRD data was therefore considered sufficient to

confirm that no structural change had occurred on grinding and the published structure of 26L:HF is a suitable model.

Table 3.1: Comparison of unit cell parameters for 26L:HF structure, determined at low and room temperature by both XRD and subsequent DFT geometry optimisation allowing the unit cell to vary.

	***a* (Å)**	***b* (Å)**	***c* (Å)**	***β* (°)**	**Volume (Å^3)**	**Energy[f] (kJ/mol)**
181445[a] (292 K)	9.898(2)	15.347(2)	7.4970(10)	107.810(10)	1084.3(3)	756.3
MAC[b,c] (300 K)	9.906752(18)	15.35923(3)	7. 501959(12)	107.78973(12)	1086.917(4)	-
MAC[d] (100 K)	9.83096(4)	15.16926(8)	7.48196(3)	108.8946(3)	1055.651(9)	-
DFT 181445	9.8382	15.3210	7.4735	108.477	1068.42	0.9
1876100[e] (293 K)	9.8451(3)	15.1918(4)	7.4842(2)	108.573(3)	1061.07(5)	740.8
DFT 1876100	9.8364	15.2651	7.4771	108.936	1061.96	0.0

[a]Structure determined by Pan *et al.*[131] (CCDC ref.: MIBYEB, 181445), *R*1 [$I > 2\sigma(I)$] = 4.05 %
[b]Multi-Analysing Crystals, high resolution synchrotron scan
[c]R_{bragg} = 5.59 %
[d]R_{bragg} = 4.74 %
[e]Structure determined in this study, *R*1 [$I > 2\sigma(I)$] = 3.73 %
[f]Relative to the structure of lowest energy (DFT 1876100)

3.3.2. NMR

Fig. 3.3a and 3.3b presents ^{1}H one-pulse MAS and ^{1}H-^{13}C CP MAS NMR spectra of 26L:HF together with stick spectra that represent the ^{1}H and ^{13}C chemical shifts calculated using the GIPAW method for the geometry optimised crystal structure (see Table A1.6 for full listing). Generally, the level of agreement between experiment and calculation was within the established discrepancy of such calculations of 1% of the chemical shift range.[93, 120, 202-204] The broad peak at 5 – 9 ppm in the ^{1}H MAS NMR spectrum (Fig. 3.3a) is in agreement with GIPAW calculation. It contains the CH protons from both the aromatic ring and the fumarate carbon-carbon double bond backbone. There is a distinct shoulder at 6.3 ppm, as expected for H1 (δ_{iso}^{calc}= 6.3 ppm), the CH proton with the lowest calculated chemical shift, and a second shoulder at 7.9 ppm, corresponding to H3 and H11 ((δ_{iso}^{calc}= 7.9 ppm and δ_{iso}^{calc}= 8.0 ppm, respectively). The single methyl peak in both experiment and calculation shows the similarity in the local environments of the two methyl groups due to the symmetry of the 2,6-lutidinium

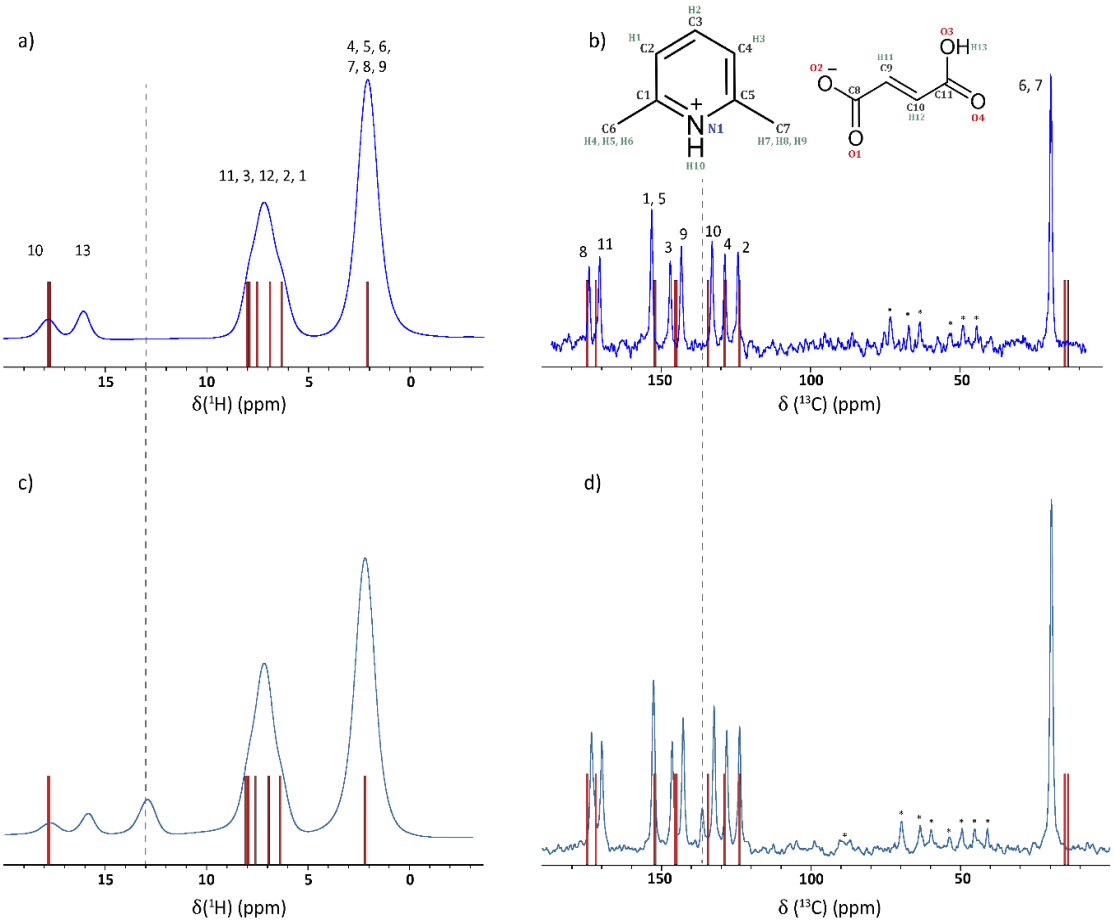


Figure 3.4: (a and c) ^{1}H (600 MHz) one-pulse MAS (60 kHz) NMR spectra and (b and d) ^{1}H (500 MHz)-^{13}C CP MAS (12.5 kHz) NMR spectra of 26L:HF, with stick spectra corresponding to GIPAW calculated chemical shifts shown as vertical bars (red). Spectra obtained initially for the pure crystalline powder are shown at the top (a and b), while spectra obtained later after the formation of FA (see discussion in text) are shown below (c and d). The positions of the anomalous peaks that developed over time are indicated by vertical dashed lines. Asterisks denote spinning sidebands.

molecule and its interaction with the fumarate molecule. The only discrepancy between experiment and calculation is the occurrence of two distinct peaks at high chemical shift rather than one, as discussed below.

In the ^{1}H-^{13}C CP MAS NMR spectrum (Fig. 3.3b), there is also good agreement between experimental and GIPAW calculated chemical shifts, with only small discrepancies. One of the most noticeable differences is the clear separation of the C3 and C9 resonances, which are calculated as being only 0.5 ppm apart (145.8 and 145.3 ppm, respectively). By comparison, two peaks are evident in the experimental spectrum, at 142.7 and 146.4 ppm. Owing to their closeness in the calculated values, it was unclear how they should be assigned on the basis of a 1D spectrum alone.

A 2D ^{1}H-^{13}C HETCOR MAS NMR spectrum of 26L:HF, whereby CP was used to transfer magnetisation via through space dipolar couplings, is shown in Fig. 3.5. This was recorded using a CP contact time of 200 μs, such that cross peaks for longer-range C···H proximities are apparent as well as direct one-bond C-H connectivities. The ^{1}H-^{13}C HETCOR spectrum is shown together with crosses that represent the GIPAW calculated chemical shifts for the C-H dipolar correlations up to 2.8 Å (see Table A1.3 in Appendix 1). This allows an upper distance to be assigned to the observed experimental correlations by comparison to the interatomic distances extracted from the 26L:HF crystal structure. For the zoomed region in the inset of Fig. 3.5, crosses for GIPAW calculated chemical shifts correspond to the directly bonded CH moieties (~1.1 Å). The 2D correlation peaks show good agreement between experiment and GIPAW calculation, with only very few of the expected peaks missing, all of which are at a separation of more than 2.2 Å in the 26L:HF crystal structure. For example, no

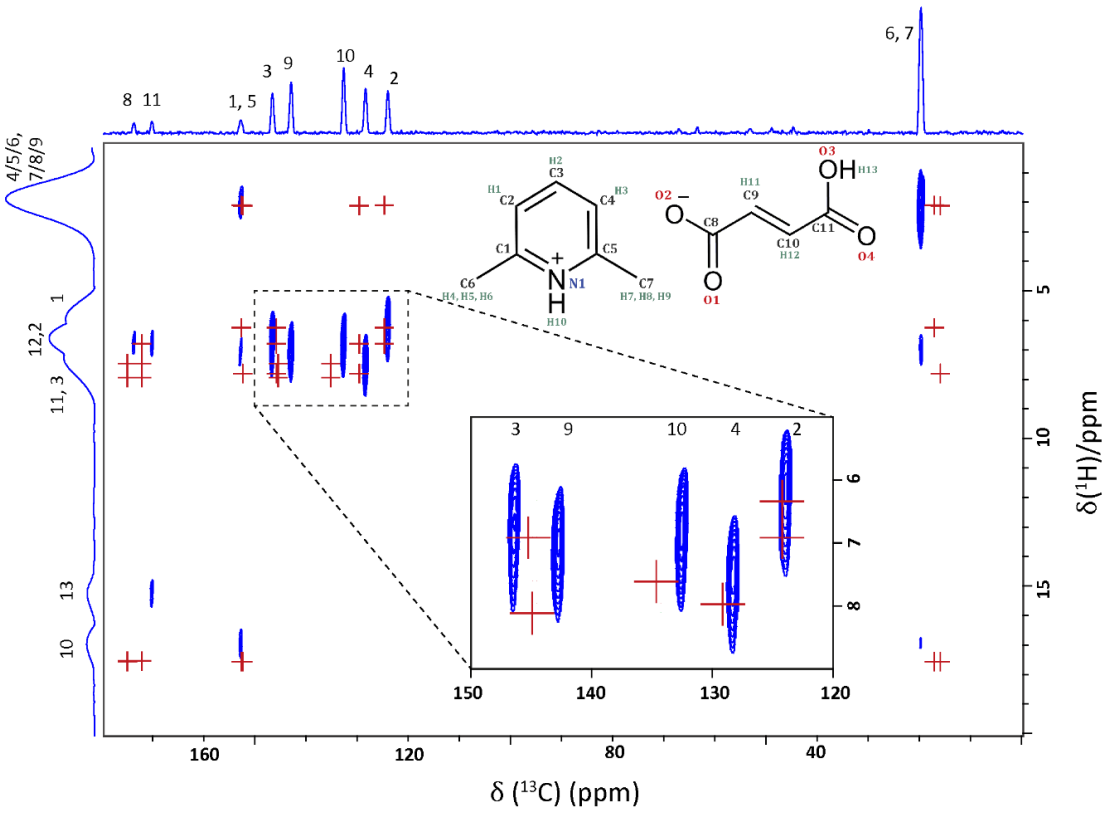


Figure 3.5: A ^{1}H (500 MHz)-^{13}C CP (200 μs) HETCOR MAS (12.5 kHz) NMR spectrum of 26L:HF recorded using FSLG ^{1}H homonuclear decoupling in t_1 with calculated (GIPAW) chemical shifts shown as red crosses, for CH proximities out to 2.8 Å and for directly bonded CH connectivities (~1.1 Å) in the zoomed-in region shown as an inset. Proximities for the quaternary carbons are listed in Table A1.3. The base contour level is at 7.5% of the maximum peak height.

experimental cross peak is seen for the correlations of C8 with either H10 or H13, which correspond to distances of 2.38 and 2.53 Å, respectively.

The small difference in experimental correlations apparent in the zoomed-in region in Fig. 3.5 supports assigning C3 to the ^{13}C resonance at 146.4 ppm, the higher chemical shift as per the GIPAW calculation, as its cross peak corresponds to a slightly lower proton chemical shift value than that of the resonance at 142.7 ppm. As can be seen from the GIPAW calculated chemical shifts for the directly bonded C-H moieties, this is expected as C3 (δ_{iso}^{exp} = 146.4 ppm) is directly bonded to H2 (δ_{iso}^{calc} = 6.9 ppm) and C9 (δ_{iso}^{exp} = 142.7 ppm) is directly bonded to H11 (δ_{iso}^{calc} = 8.0 ppm). C3 also has a proximity of 2.16 Å to H1 (the lowest of the CH proton resonances, as stated above) as well as its directly bonded H2. C9 does not share this correlation to H1 so its cross peak with the CH region is expected to be at a higher ^{1}H chemical shift, as observed. The discrepancy between experiment and GIPAW calculation for these ^{13}C chemical shifts is –2.6 and 0.6 ppm for C9 and C3, respectively. For C3, this is well within the established discrepancy compared to experiment for such calculations of 1% of the chemical shift range (~2 ppm for ^{13}C chemical shifts) although, with a difference of –2.6 ppm, C9 is slightly further from calculated than would be expected. The only peaks that fall well outside this usual discrepancy range are the ^{13}C peaks for the methyl groups, which fall ~5 ppm higher than calculated. This is a consequence of the known deviation from negative one in the gradient of a plot of experimental chemical shift against calculated chemical shielding (see, *e.g.*, Fig. 4b in Ref. 47);[205] an alternative approach would be to use different reference shieldings for different parts of the spectrum.[204]

As noted above, two peaks are observed above 10 ppm in the ^{1}H MAS NMR spectrum (Fig. 3.4a) whereas, as can be seen from the stick spectra, the GIPAW calculated chemical shifts predict that both H13 and H10 (the OH and NH protons, respectively) are at the same value of 17.7 ppm. In the 2D ^{1}H-^{13}C HETCOR spectrum, the absence of cross peaks between the 17.7 ppm proton environment and C11 (δ_{iso}^{exp} = 169.9 ppm) suggests that H13 resonates at a lower ppm value compared with calculation and can be assigned to the peak at 15.8 ppm,

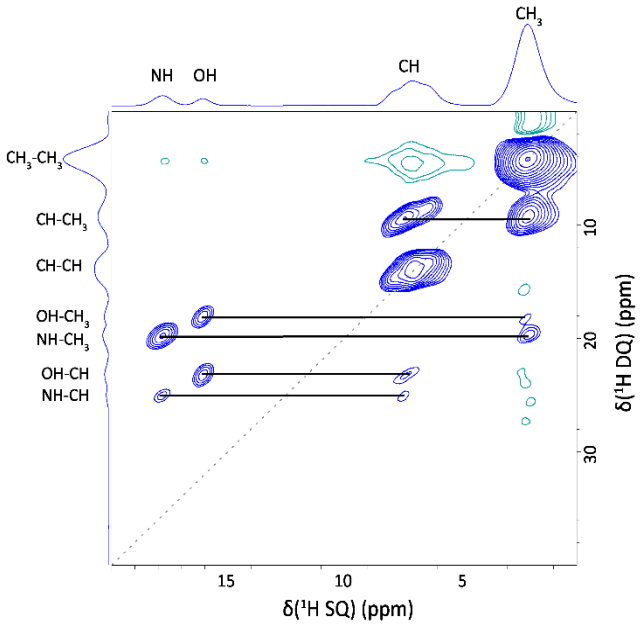


Figure 3.6: A ^{1}H (600 MHz) DQ MAS NMR spectrum of 26L:HF recorded with one rotor period of BaBa recoupling. The base contour level is at 3.3% of the maximum peak height. Blue and green contours correspond to positive and negative intensity, respectively. DQ/SQ correlations for the NH (H10) and OH (H13) resonances are listed in Table 3.2. The dashed diagonal line indicates the $\delta_{DQ} = 2\delta_{SQ}$ diagonal, while horizontal lines indicate a DQ peak at the sum of the two SQ peaks for dipolar coupled unlike protons.

which correlates with C11. A 2D ^{14}N-^{1}H HMQC spectrum, as shown and discussed later, is also in good agreement with calculation. A single N environment correlates with the highest proton peak (δ_{iso}^{exp} = 17.7 ppm), assigned as H10, whereby the low ^{14}N shift, –102 ppm, also indicates that proton transfer to the N has occurred and the structure is sustained by an ionic salt interaction due to its excellent agreement with the calculated ^{14}N shift (δ^{calc} = –102 ppm).

In the 2D DQ-SQ ^{1}H-^{1}H MAS NMR spectrum presented in Fig. 3.6, cross peaks occur in the DQ dimension at the sum of the two SQ resonances for protons close in space (generally accepted as up to a distance of 3.5 Å).[106] Table 3.2 lists the H···H proximities under 3.5 Å for the NH (H10, δ_{iso}^{exp} = 17.7 ppm) and OH (H13, δ_{iso}^{exp} = 15.8 ppm) protons of 26L:HF in the DFT geometry optimised structure. The closest proximities of the OH and NH protons are to the CH_3 groups, giving peaks at δ_{DQ} = 2.1 + 15.8 = 17.9 ppm and δ_{DQ} = 2.1 + 17.7 = 19.8 ppm, respectively. The next closest proximity to the NH, H10, is to H11, which also falls within 3.5 Å of the OH, H13 (see Table 3.2), corresponding to cross peaks at δ_{DQ} = 15.8 + 7.9 = 23.7 ppm (H13-H11) and δ_{DQ} = 17.7 + 7.9 = 25.4 ppm (H10-H11). Unlike H10, H13 has a

sufficiently close proximity to both H2 and H3 such that cross peaks are also observed at δ_{DQ} = 15.8 + 7.0 = 22.8 ppm (H13-H2) as well as at δ_{DQ} = 15.8 + 7.9 = 23.7 ppm (H13-H3).

Table 3.2: H-H proximities (<3.5 Å) and corresponding ^{1}H DQ chemical shifts for the NH and OH protons in 26L:HF.

Proton 1	δ_{iso}^{exp} SQ1 (ppm)	Proton 2	δ_{iso}^{exp} SQ2 (ppm)	δ_{iso}^{exp} DQ (ppm)	Separation[a] (Å)
H13 (OH)	15.8	***H9 (CH$_3$)***	*2.1*	*17.9*	*2.52*
		H4 (CH$_3$)	*2.1*	*17.9*	*2.87*
		H11 (CH)	*7.9*	*23.7*	*2.88*
		H3 (CH)	*7.9*	*23.7*	*2.97*
		H2 (CH)	*7.0*	*22.8*	*3.04*
		H2 (CH)	*7.0*	*22.8*	*3.26*
		H5 (CH$_3$)	*2.1*	*17.9*	*3.41*
		H11 *(CH)*	7.9	23.7	3.47
H10 (NH)	17.7	**H5 *(CH$_3$)***	2.1	19.8	2.37
		H8 *(CH$_3$)*	2.1	19.8	2.51
		H11 (CH)	*7.9*	*25.4*	*2.57*
		H7 *(CH$_3$)*	2.1	19.8	2.71
		H4 *(CH$_3$)*	2.1	19.8	2.92
		H4 (CH$_3$)	*2.1*	*19.8*	*3.04*
		H11 (CH)	*7.9*	*25.4*	*3.27*
		H5 (CH$_3$)	*2.1*	*19.8*	*3.28*

[a] H-H distances are taken from the DFT (CASTEP) optimised structure. Intermolecular proximities are denoted using italic font.

3.3.3. Isolated Molecule GIPAW Chemical Shift Calculations

Isolated molecule GIPAW calculations were performed on CCDC 181445 following geometry optimisation and the changes with respect to the full crystal structure are listed in Table 3.3. In these calculations, δ_{iso}^{calc} of the OH and NH (H13 and H10, respectively) are significantly higher for the crystal structure than for the isolated molecules, with both being calculated at 17.7 ppm in the crystal rather than 6.4 and 10.5 ppm, respectively, for the isolated molecules. The hydroxyl H13 shows the largest change in chemical shift between the crystal structure and isolated molecule calculations, corresponding to the strong hydrogen bonding between adjacent hydrogen fumarate molecules to form the chains of asymmetric units seen in the geometry optimised 181445 structure (see Fig. 3.1; the O···O and O···H distances are

Table 3.3: A comparison of GIPAW calculated ^{1}H shifts (in ppm) for the full crystal structure of 26L:HF and for isolated molecules extracted from the crystal structure.

Atom	δ_{Expt}	$\delta_{Crystal}$	$\delta_{Molecule}$	$\Delta\delta_{Crystal\ -\ Molecule}$
H13	15.8	17.7	6.4	11.3
H10	17.7	17.7	10.5	7.1
H12	7.0	7.5	6.1	1.4
H11	7.9	8.0	6.8	1.2
H3	7.9	7.9	7.3	0.6
H4/H5/H6	2.1	2.1	2.4	–0.3
H7/H8/H9	2.1	2.1	2.4	–0.3
H1	6.3	6.3	7.3	–1.0
H2	7.0	6.9	7.9	–1.0

2.54 and 1.48 Å, respectively, with a 175° OHO angle). The NH (H10) also shows evidence in isolated molecule calculations of a strong H bond to O2 as expected (the N···O and O···H distances are 2.64 and 1.55 Å, respectively, with a 169° NHO angle), with the smaller $\Delta\delta_{Crystal\ -\ Molecule}$ values corresponding to the slightly larger distance and an angle further from the ideal 180° compared to the case for H13. A further difference is that H10 corresponds to a charged NH^{+} moiety while H13 is in a neutral COOH group. Two CH protons, H11 and H12, also show some indication of a weak hydrogen bonding interaction[118, 125] as they exhibit a change in chemical shift of 1.2 and 1.4 ppm, respectively. H11 is weakly bonded to O2 in the stacked chain (C···O distance of 3.43 Å, CH···O distance of 2.38 Å and CHO angle 161°) and H12 is weakly bonded to O4 in the adjacent fumarate chain (C···O distance of 3.50 Å, CH···O distance of 2.45 Å and CHO angle 161°).

H1 and H2 exhibit weak ring current effects (–1.0 ppm), possibly due to the stacking of the pyridine rings within the crystal structure, although it is interesting to note that this effect seems to be offset for H3 which instead shows a very slight positive change (0.6 ppm). This is probably due to weaker effects from ring currents, as the stacking is slightly misaligned (Fig. 3.1d, bottom right), placing it further out from the π-π stack. This π-π stacking is thought to be the driving force behind the alignment of the pairs of acid chains.[140]

3.4. Skewed Cell Contraction

Fig. 3.7 shows a Rietveld refinement against a PXRD high resolution synchrotron scan that was collected at 100 K. As with the room temperature high resolution scan shown in Fig. 3.3, residual preferred orientation effects prevent accurate refinement of atomic positions, but the tick marks, corresponding to 2θ values of the reflections expected for the refined structure, are in excellent agreement with those recorded experimentally. As can be seen in Table 3.1, this Rietveld refinement showed a contraction accompanied by an increase in the β angle. This effect was also observed in SXRD crystal screens conducted at a range of temperatures between 100 K and 300 K (Fig. A1.2).

Compared to 181445, the Rietveld refinement for the 300 K PXRD pattern presented above (Fig. 3.3) shows a small, but similarly skewed expansion, with a marginal decrease in β angle, alongside increases in cell lengths, consistent with being recorded at a slightly higher temperature than the literature structure (300 K compared to 292 K). The skew in the cell, when going to low temperature during contraction, is also evident in the DFT calculations (also shown in Table 3.1). Geometry optimisations, performed allowing the unit cell

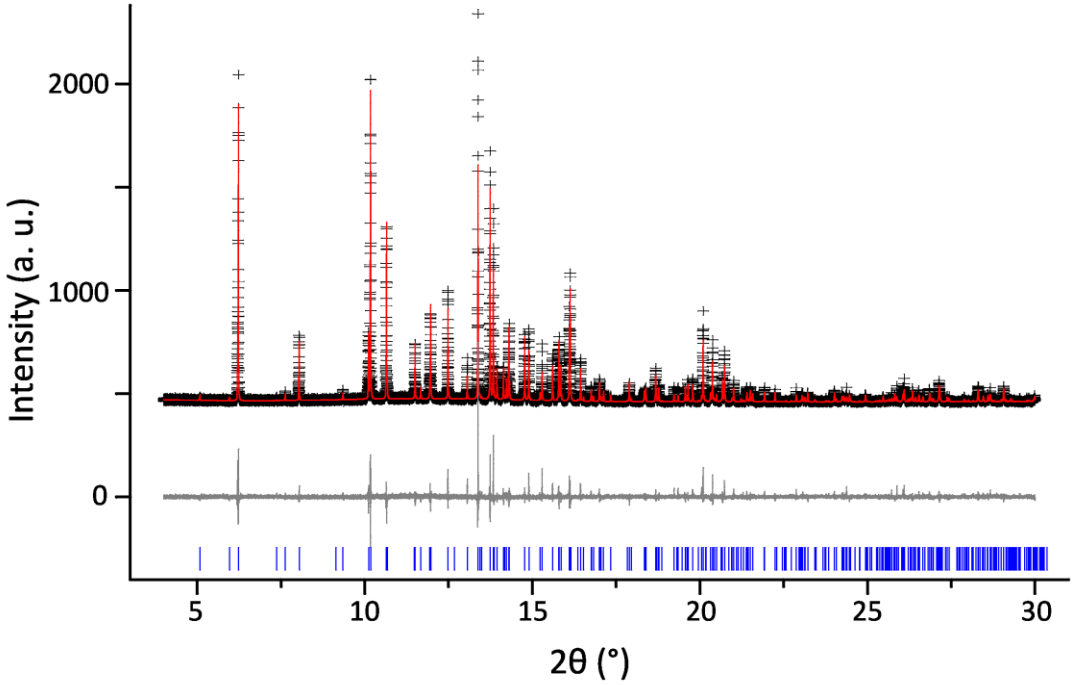


Figure 3.7: Final Rietveld fit for 26L:HF at 100 K, showing the experimental (black crosses), calculated (red upper line) and difference (grey lower line) PXRD profiles. Tick marks indicate peak positions. For refinement parameters, see Table 3.1 and Tables A1.4 and A1.5.

Table 3.4: Selected crystal data for 1876100

Chemical formula	$C_{11}H_{13}O_4N$	Temperature (K)	293(2)
Formula weight	223.22	μ (mm^{-1})	0.900
Crystal system	Monoclinic	Independent reflections	2076
Space group	$P2_1/n$	R_{int}	0.0232
Z	4	$R1$ [$I > 2\sigma(I)$]	0.0317
a (Å)	9.8451(3)	α (°)	90
b (Å)	15.1918(4)	β (°)	108.573(3)
c (Å)	7.4842(2)	γ (°)	90

parameters to vary, showed convergence towards the low temperature structure. As no external temperature is included in the calculations, they are effectively performed at 0 K, so it is unsurprising that they exhibit such a tendency.

During the room temperature crystal SXRD screening to check each crystallisation for variations or new polymorphs, one crystal was identified that differed slightly from the previously identified structure. This newly identified structure has been deposited with the CCDC (no. 1876100) and selected crystal data is given in Table 3.4. Although the molecular packing of the crystal remained unchanged, with only small shifts in relative atomic positions (Fig. 3.8), the unit cell parameters presented in Table 3.1 show both a slight contraction in the unit cell lengths (the most significant being a reduction of 0.15 Å on the b axis) and a 0.76° widening of the β angle compared with 181445, which was recorded at 292 K. Intramolecular one-bond atomic distances actually increased by an average of 0.006 Å, with the mean C-C/N

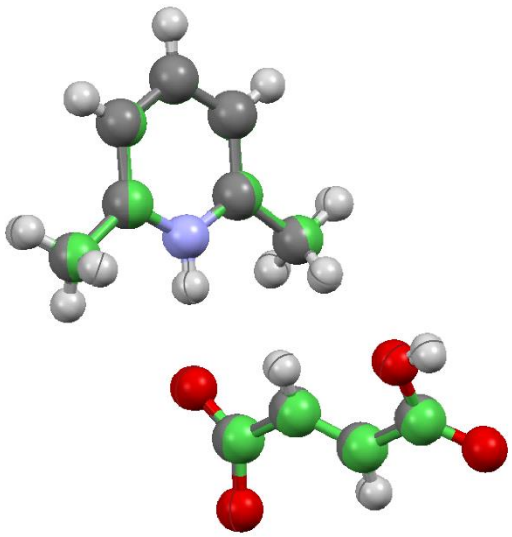


Figure 3.8: Overlay of the asymmetric unit of newly identified structure deposited to the CCDC as 1876100 (green) and of the previously published structure (ref.: 181445).

one-bond separation in the pyridine ring changing from 1.367 Å in the original structure to 1.373 Å in 1876100. However, intermolecular distances between the atoms of neighbouring molecules decreased by an average of 0.03 Å (where intermolecular distances up to 6 Å were considered), with this value dominated by the distances in the *b* dimension, with the O···O hydrogen bond (aligned with the *a* dimension) decreasing by 0.01 Å compared to a decrease of 0.06 Å for the O3 to C8 distance between neighbouring hydrogen fumarate molecules (aligned with the *b* dimension). The form of the contraction and its corresponding skew in the β angle are similar to that exhibited by the contraction at low temperature, although to a lesser extent.

GIPAW calculations for 1876100, following geometry optimisation, showed that the minimal shifts in atomic positions produced only small changes to the calculated chemical shifts (the largest being 0.2 ppm for ^{1}H and 0.6 ppm for ^{13}C, with mean differences of 0.01 ppm and 0.1 ppm, respectively, see Table A1.6). These mean differences are below the experimental error and therefore make the two structures effectively indistinguishable by NMR. This in turn suggests that the skewed contraction produces no significant changes to the chemical environments for the two molecules. However, the unit cell changes, particularly the increase in β angle, are sufficient to visibly change the 2θ positions of reflections in the simulated PXRD pattern with respect to the literature structure (Fig. A1.3).

A second example of 1876100 was not identified within any of the single crystal screens, for more than 20 crystals, with all other crystals converging to the 181445 structure of 26L:HF at room temperature. This observation that no more crystals showing this contraction were identified, may be due to the minimal presence of 1876100 or a bias in which single crystals were analysed. A difference in crystal quality, exhibited either by their appearance or quality of diffraction, could create such a bias.

Extensive PXRD analysis failed to provide conclusive evidence for even a minor second structural phase within the powdered compound, suggesting that 1876100 may have been an isolated case. However, in the case of only partially ground powders (so as to as minimise break down of the co-crystal, see below discussion), some of the reflections

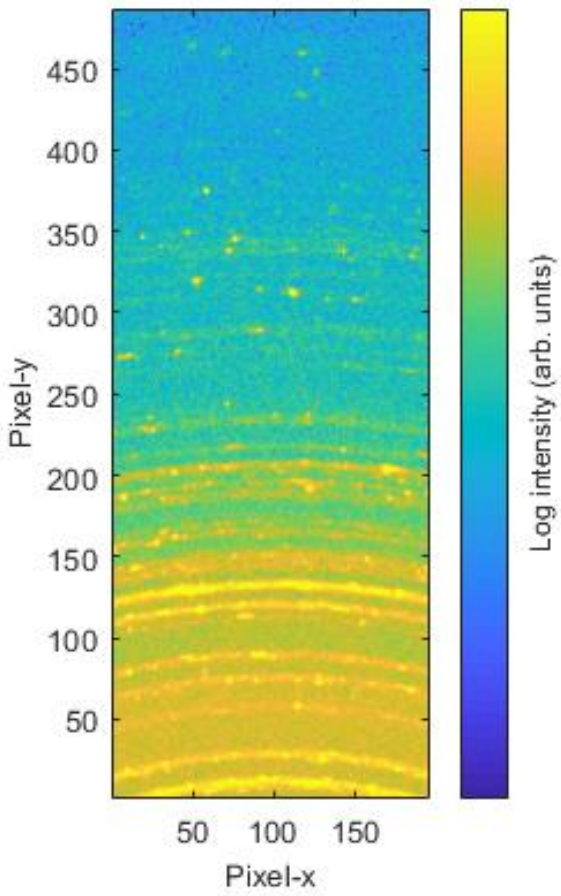


Figure 3.9: Static transmission PXRD of 26L:HF with a 2D detector.

appeared to show a splitting. Several larger crystallites were observed in static transmission PXRD of these samples, using a 2D detector, that may explain this. They lie slightly off the main powder ring, as predicted due to a small change in β angle (Fig. 3.9), and might therefore be linked to 1876100. As 1876100 seems to be related to the low temperature contraction, it is unclear how it existed within a room temperature SXRD screen. The energies of 181445 and 1876100, determined by DFT (see Table 3.1), differed by 16.5 kJ/mol prior to geometry optimisation of the crystal structure, with 1876100 lower in energy. A significantly smaller difference of 4.1 kJ/mol was seen following DFT geometry optimisation with fixed unit cell parameters, with 181445 now lower in energy, showing that they are of comparable stability. Only a 0.9 kJ/mol change was seen following DFT geometry optimisation with the unit cell allowed to vary.

Considering Table 3.1, it is notable that, for the DFT calculations, variable cell geometry optimisations that were started from both the previously published structure and 1876100 converged towards the 100 K structure but neither reached it. Each satisfied the convergence conditions of the DFT calculation whilst the volume of their unit cells was still larger than that of the 100 K structure. Although the difference between the two unit cells was smaller following variable cell geometry optimisation, the output of the calculation based on 1876100 was still slightly more contracted than that based on the original CCDC structure, 181445. This discrepancy is probably due to the limitations of the dispersion correction scheme, as TS does not account for MBD forces, which can be integral to molecular crystals, and is also based on reference data, therefore struggling for instances that fall outside its reference set. Application of an MBD correction scheme[192, 193] or a method that determines the dispersive forces from the electron density may provide a more accurate model.[189]

It can also clearly be seen (Table 3.1) that despite effectively being at a lower temperature (0 K), the outputs of neither calculation exhibit as significant a contraction as is evident for the 100 K synchrotron case, although the optimisation of 1876100 did produce a similar skew of the β angle.

3.5. Formation of Fumaric Acid

Samples that had been stored for more than a week as powders rather than single crystals showed additional peaks in the ^{1}H MAS and ^{1}H-^{13}C CP MAS NMR spectra, as shown in Fig. 3.3c and 3.3d. The high chemical shift of the new ^{1}H peak, 12.9 ppm, is indicative of strong hydrogen bonding. A 2D ^{14}N-^{1}H HMQC MAS NMR spectrum showed only the peak identified during the initial analysis (Fig. 3.10b), which confirmed the presence of a single N environment and therefore indicated that the new ^{1}H resonance likely corresponds to an OH group. A 2D ^{1}H-^{13}C HETCOR MAS NMR spectrum (Fig. A1.4) did not show the additional ^{1}H peak, further supporting the assignment as an OH, and also implied no short-range

proximity (~<2.8 Å) to a carbon atom. The extra ^{13}C peak that had been identified in the ^{1}H-^{13}C CP MAS spectrum did correlate with protons in the CH region, however, suggesting a second new ^{1}H resonance lies under the CH region. The closeness in chemical shift of this

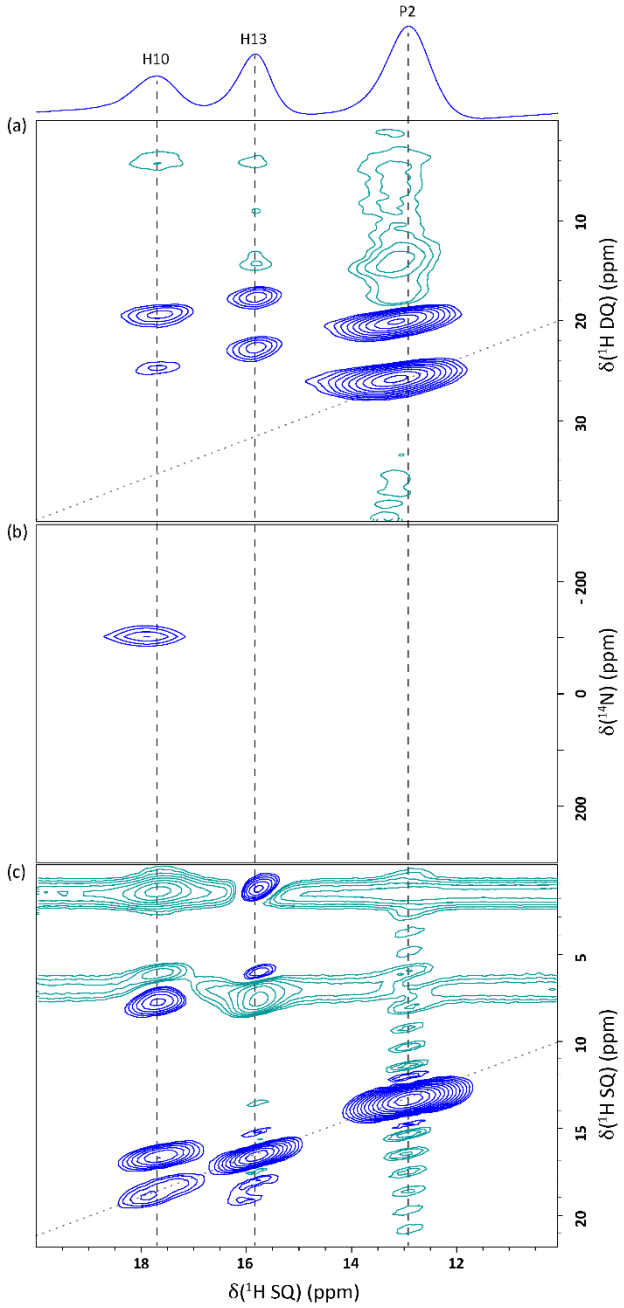


Figure 3.10: 2D MAS (60 kHz) NMR spectra of 26L:HF sample (not freshly ground into a powder) showing the 10 – 20 ppm ^{1}H region: (a) a ^{1}H DQ spectrum recorded with one rotor period of BaBa recoupling; (b) a ^{14}N-^{1}H HMQC spectrum with 8 rotor periods of R^3 recoupling; and (c) a ^{1}H SQ NOESY spectrum with t_{mix} = 300 ms. All spectra were recorded at a ^{1}H Larmor frequency of 600 MHz. Base contour levels are at 5.3 %, 36.2 % and 2.3 % of the maximum peak height, respectively. Blue and green contours correspond to positive and negative intensity, respectively. The negative intensities seen at the CH_3 and CH F_1 (vertical axis) SQ frequencies in (c) are due to the much greater intensity of their auto-correlation peaks. The dashed diagonal lines in (a) and (c) indicate the (a) $\delta_{DQ} = 2\delta_{SQ}$ and (c) $\delta_{SQ} = \delta_{SQ}$ diagonals.

new ^{13}C resonance to that of C10 hinted that it may correspond to a carbon in the fumarate backbone.

A 2D ^{1}H-^{1}H DQ-SQ MAS NMR spectrum (Fig. 3.10a) shows that two cross peaks are observed for the newly appearing proton resonance. These correspond to an auto-correlation (δ_{DQ} = 13.0 + 13.0 = 26.0 ppm) and a correlation with the CH region (δ_{DQ} = 13.0 + 7.0 = 20.0 ppm), corresponding to the same ^{1}H chemical shift as H11 or H12. A 2D SQ ^{1}H-^{1}H NOESY spectrum (Fig. 3.10c) clearly shows the presence of two distinct phases. A mixing time of 300 ms was used to allow spin diffusion throughout the entire system which should result in cross peaks between all protons within the same phase.[206, 207] These are present for H10 and H13, but no correlations are seen for the peak at 12.9 ppm. The absence of cross peaks for the 12.9 ppm resonance with the methyl and NH protons in either of the ^{1}H-^{1}H 2D correlation experiments, coupled with the proximity to the CH region shown in the DQ experiment, suggests that the secondary phase does not contain 26L and may correspond to crystalline FA. This is supported by the correlation between the anomalous C peak with the CH region in the ^{1}H-^{13}C HETCOR spectrum (Fig. A1.4), and the OH ^{1}H chemical shift of 13 ppm[208] and the

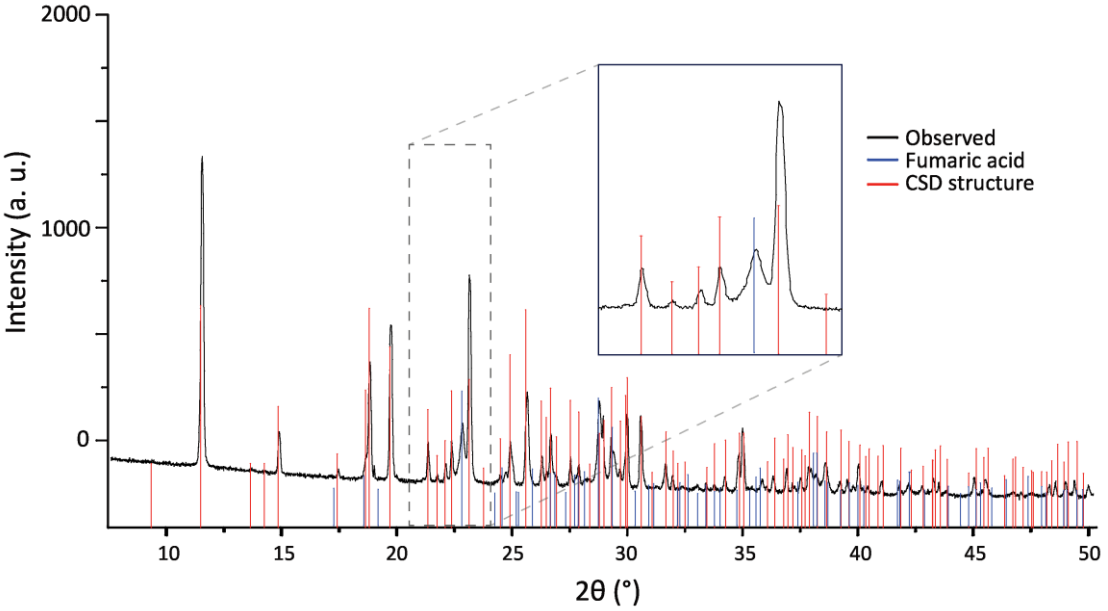


Figure 3.11: PXRD pattern of 26L:HF recorded more than a week after first being ground to powder, with tick marks representing the reflection positions simulated for CCDC structure 181445 (red) and crystalline FA (blue). The zoomed-in region (inset) shows the agreement between the additional experimental reflections and those of FA.

two ^{13}C chemical shifts of 136.3 ppm and 172.3 ppm[209] reported in the literature for FA, with the higher resonance lying beneath the existing C8 peak.

This conclusion is supported by the position of reflections seen in a subsequent PXRD pattern recorded with a Panalytical X'Pert Pro MPD diffractometer (Fig. 3.11). Samples that had been stored as powders contained reflections that were not present in patterns recorded on freshly ground crystals. Although these new reflections matched well those expected for crystalline FA,[210] a multiphase refinement was not possible. This was due to both the presence of splittings (possibly caused by the presence of larger crystallites of 26L:HF with different β angles, as shown in Fig. 3.9), preferred orientation effects and, crucially, significant overlap of many of the FA reflections with those produced by 26L:HF.

This formation of FA can be explained by evaporation of 26L from the structure, suggesting it is appropriate to think of this structure as behaving more like a solvate than a salt. Solution-state ^{1}H NMR of samples that had been stored as powders under ambient conditions showed a small decrease in the ratio of 26L to FA compared to either a freshly ground powder or a dissolved single crystal (Fig. A1.5). Storage as single crystals did not result in measurable degradation of the salt or formation of FA by XRD or MAS NMR, even

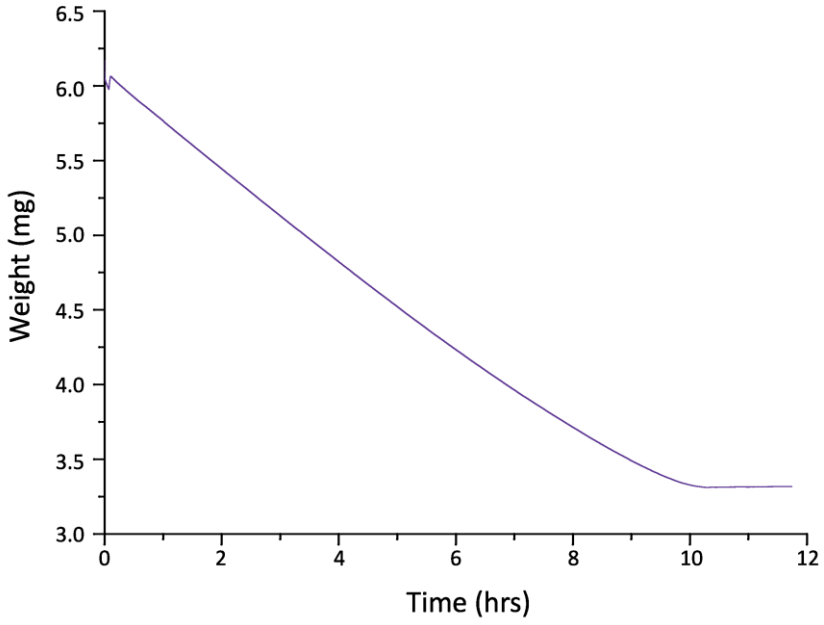


Figure 3.12: TGA of 26L:HF powder recorded on a Mettler Toledo Star^{e} instrument with a ramp of 10°C/min from 25-70°C. The sample was then held at 70°C for 12 hours.

over the course of several months. The apparent relative stability of the single crystals compared to powdered samples implies gradual evaporation of the 26L from the crystal surface, negligible at the relatively smaller surface area to volume ratio of single crystals compared to a crystallite within the powder. As 2,6-lutidinium ions sit in stacks through the structure, we speculate that the loss of 26L through evaporation would leave wide pores in the remaining structure that only contains FA, resulting in collapse of the structure to the more stable crystalline FA form. It is unclear at what stage transfer of a proton from 2,6-lutidinium to hydrogen fumarate or conversion to the two s-*cis* conformation occurs, as both are required for the packing within the crystalline FA structure of neutral molecules rather than hydrogen fumarate ions.

DSC data for 26L:HF showed no evaporation with only a single melting point at 112 °C (Fig. A1.6). This falls in between the melting points of lutidine and FA (−6 and 298 °C, respectively) as expected for a multicomponent structure. As the boiling point of 26L is 143 °C, any evaporation would be expected to be slow and so unlikely to be evident relative to a speed of heating of 10 °C/min. TGA of powdered 26L:HF held at 70°C for 12 hrs showed a gradual loss in mass over the first 10 hrs, which then plateaued (Fig. 3.12). The loss in mass corresponds to 96.3% of the 26L that was present in the complete crystal structure originally. If the mass loss is due to evaporation of 26L, as proposed, the plateau prior to complete loss suggests that the remaining 3.7% of 26L molecules are trapped in the centre of the crystallites, with insufficient energy at 70 °C to escape. This could be due to collapse of the majority of the crystallites structure to FA preventing evaporation of this residual 3.7% of 26L molecules.

3.6. Summary

A combined NMR and XRD crystallographic investigation of 26L:HF has been presented together with a computational study based on DFT geometry optimisation and GIPAW calculation of NMR chemical shifts.

The use of this combined approach enabled the identification of FA formation within powder samples over time. Based on a corresponding reduction in the ratio of 26L to FA (by solution-state NMR) and significant weight loss of a sample at 70 °C (as observed by solution-state NMR and TGA analysis, respectively), it is proposed that this is due to slow evaporation of 26L, which is liquid at room temperature, from the crystal structure. Monitoring the process under different TGA and DSC parameters as well as by Raman spectroscopy may give more insight, however.

It was not possible to determine the formation of FA solely by PXRD as both reflection overlap and splittings present in the in-house data prevented a multiphase refinement, making the addition of solid-state NMR analysis crucial to understanding the structures stability. This highlights both the complementarity of XRD and NMR methods and also the benefits of following a multi-technique crystallographic approach.

26L:HF was found to form plate crystals and to undergo a skewed unit cell contraction at low temperatures. A second room temperature structure, CCDC 1876100, was also identified by SXRD. This shows a similar variation in unit cell parameters to the 100 K structure, both in the contraction along a and b axes and in the skew of the β angle. Although it may exist as a minor component within the crystalline powder, it is unlikely to have any significant effect on the physical properties as the changes in the relative atomic positions compared to the previously published structure are so small. This results in only very small changes in the GIPAW calculated chemical shifts, suggesting it is invisible to solid-state NMR.

Chapter 4. Exploring the Polymorphism of 2-Amino-6-Methylpyridinium Salts

4.1. Introduction

26AMP differs from 26L (discussed in Chapter 3) in the simple replacement of the methyl group at position 2 with an amino group (Fig. 4.1). This single change serves to raise the melting point of the molecule by around 50 °C (from −6 °C for 26L[211] to 40-44 °C for 26AMP[131]). Although the melting point is still not high, 26AMP is solid at room temperature so it therefore cannot be classified as a 'solvent' molecule, implying the formation of more stable multi-component forms may be possible as compared to that seen for 26L.

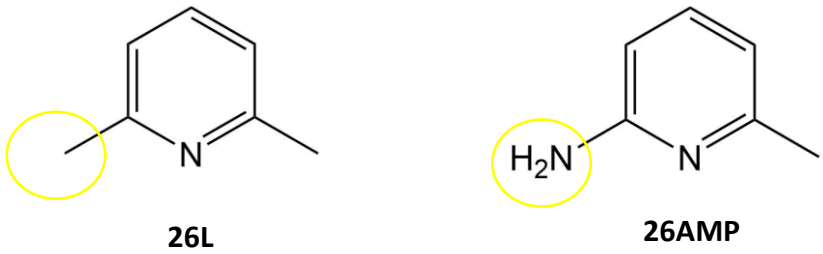


Figure 4.1: The difference between 26L (left) and 26AMP (right), consisting of replacement of a methyl group with an amino

Despite the ubiquitous nature of salt formation as a technique for improving biophysical properties, targeting a stable multicomponent system that positively alters a compound's characteristics relies largely on manual screening.[212] This is key not only to discover a desirable stable form but also to explore form diversity, identifying polymorphs and solvates that may need to be taken into account when developing a manufacturing approach. The continued need for large manual screens is due both to the difficulty in predicting potential forms ab initio, particularly for multicomponent forms, and the lack of information such predictions provide about the experimental accessibility of the structures produced. Many factors affect nucleation and growth of a crystal form, including the counter ion, pH, component ratio, solvent and numerous process parameters (heating/slurry/evaporation/*etc.*).[213] In some cases, multiple forms can occur under the same

crystallisation conditions although this is rare and such forms generally differ in stoichiometry and/or ionisation.[140, 214] If one of these co-occurring forms is potentially desirable, this may cause additional problems when ensuring product purity, as a selective crystallisation method will need to be developed.

It is estimated that around a third of organic molecules can form solvates.[215] It is crucial within the pharmaceutical industry to understand solvate formation of a product to allow the design of an appropriate crystallisation route.[216] Generally, this is to ensure that solvates are avoided as many of the solvents used are toxic and not acceptable within a drug form. Hydrates are a potentially desirable form but are difficult to utilize as, like other solvates, they are inherently unstable, with a phase change generally occurring near to room temperature.[6] It is therefore generally considered most desirable to obtain an anhydrous, non-solvated form instead. Pyridine derivatives are of interest to the pharmaceutical industry as the known structures, reported in the CCDC, show less hydrate formation than other base

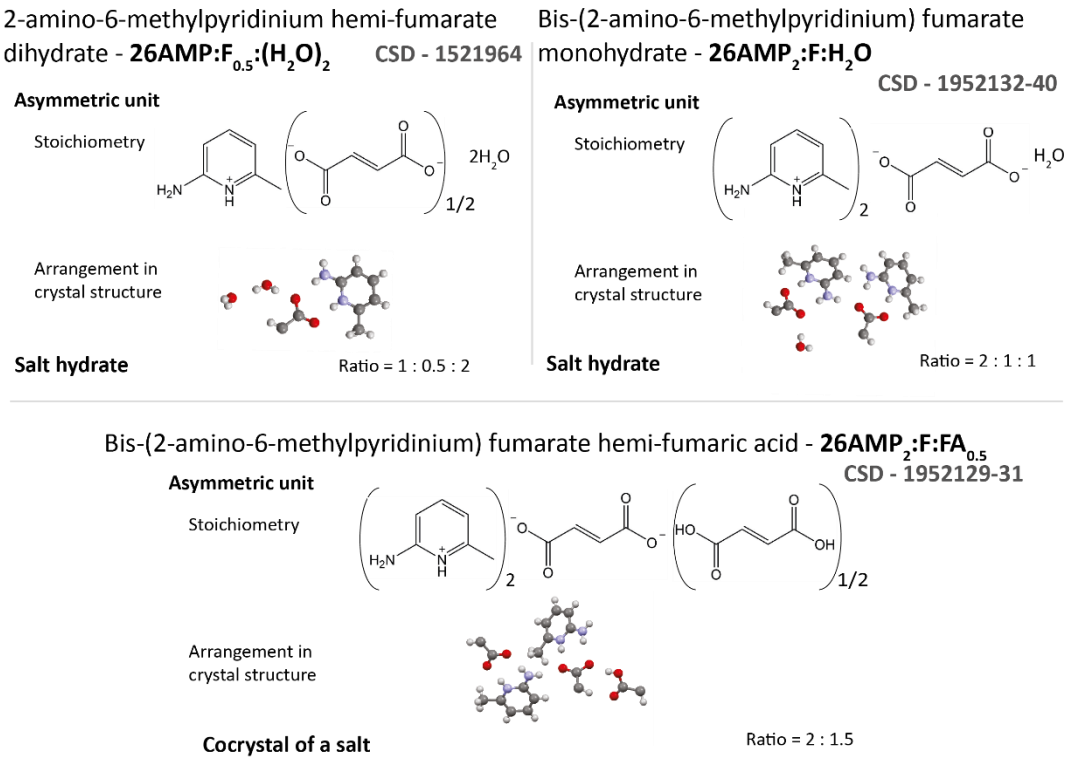


Figure 4.2: Three multicomponent forms obtained from crystallisations of 26AMP and FA. The constituent components of each are shown, alongside the asymmetric units for each crystal structure and the ratio of components (base : acid : water). Based on the number of distinct molecules in the asymmetric units, 26AMP$_2$:F:H$_2$O and 26AMP$_2$:F:FA$_{0.5}$ can be named as bis-(2-amino-methylpyridinium) fumarate mono-hydrate and bis-(2-amino-methylpyridinium) fumarate hemi-FA, respectively, but here the naming convention used is for one molecule of 26AMP, consistent with 26AMP:F$_{0.5}$:(H$_2$O)$_2$.

molecules when forming multicomponent carboxylate salts (9.1 % of stuctures are hydrates, while, for other cations, 16.7 to 24.6 % of structures are).[217]

In this chapter, the effect of altering the crystallisation solvent and hydration on the resulting form of the co-crystallisation of 26AMP and FA is investigated. The resultant structures, two salt hydrates and one cocrystal of a salt (Fig. 4.2), are characterised and their stability and relation to each other investigated.

4.2. Sample Preparation

All chemicals were obtained from Sigma Aldrich (UK) at purities of 98% or higher and used without further purification. Salts were prepared by mixing a 1:1 molar ratio of 26AMP (81 mg) and FA (87 mg) in the minimum amount of hot solvent (warmed over a hot plate at 80 °C and stirred) required to dissolve all the solute (~ 10 mL). The resulting solutions were allowed to cool slowly at room temperature under and crystals appeared after a few days. Methanol, ethanol and isopropanol were each tested as solvents and a slight dependence was seen in the resultant crystal form, with the hydrate system favoured in methanol and the anhydrous form the major form produced in either ethanol or isopropanol. Addition of ~ 0.5 mLwater to the crystallisation media was found to force hydrate crystallisation regardless of solvent choice. Growth of good quality crystals in subsequent crystallisation attempts was found to be improved by co-grinding the base and acid with < 1 mL of methanol before dissolution in hot methanol and/or the addition of seed crystals to the crystallisation media. These seed crystals were selected from the small crystals present in the initial samples (chosen depending on which form was being targeted) and were added when the hot solvent solution was placed at room temperature to cool. Use of seed crystals of a particular form was not found to prevent growth of the other forms.

Table 4.1: Selected crystal data for 26AMP:$F_{0.5}$:$(H_2O)_2$, 26AMP$_2$:F:H_2O and 26AMP$_2$:F:$FA_{0.5}$.

	26AMP:$F_{0.5}$:$(H_2O)_2$	26AMP$_2$:F:H_2O	26AMP$_2$:F:$FA_{0.5}$
Stoichiometry[a] (base : acid : water)	1 : $^1/_2$: 2	2 : 1 : 1	2 : 1 $^1/_2$: 0
Chemical formula[a]	$C_8H_{14}N_2O_4$	$C_{16}H_{22}N_4O_5$	$C_{18}H_{22}N_4O_6$
Formula weight[a] (g mol^{-1})	202.21	350.37	366.36
Crystal system	Monoclinic	Triclinic	Triclinic
Space group	$P2_1/c$	$P\bar{1}$	$P\bar{1}$
a (Å)	9.6866(2)	7.6213(3)	8.4426(4)
b (Å)	14.4243(3)	10.0308(4)	10.5523(4)
c (Å)	7.41090(10)	12.7136(5)	12.3026(4)
α (°)	90	93.561(3)	85.716(3)
β (°)	96.494(2)	107.316(4)	73.190(4)
γ (°)	90	103.160(4)	67.717(4)
Z	4	2	3
Temperature (K)	250	250	200
*R*1 [$I > 2\sigma(I)$]	0.0313	0.0404	0.0397

[a] For the asymmetric unit

4.3. Structures

Three crystal structures of 2-amino-6-methylpyridinium salts have been determined (with assistance of Dr. Guy Clarkson, University of Warwick), described below, with selected crystal data for their low temperature structures given in Table 4.1. The stated stoichiometry, chemical formula and formula weight in the Table correspond to the contents of the asymmetric unit.

In all cases, non-CH hydrogen atoms were found in the electron density map and CH hydrogens were added at calculated positions and refined using a riding model. Initial

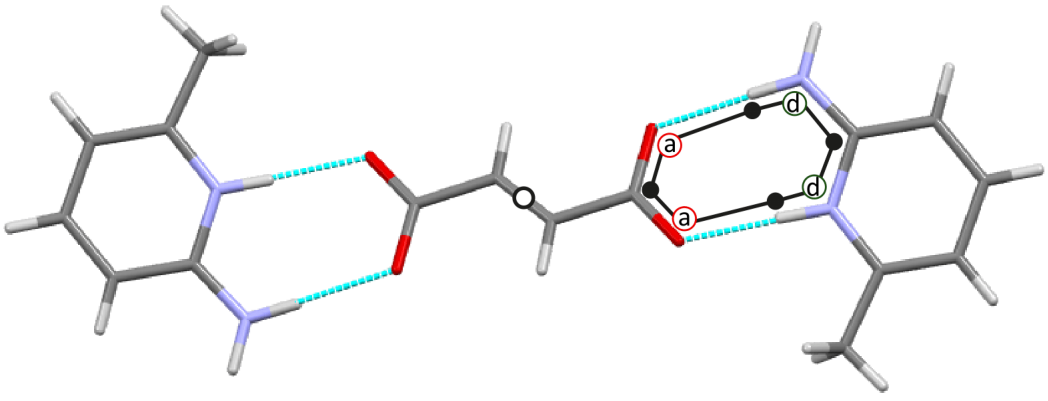


Figure 4.3: The base-acid-base (b-a-b) unit seen in all three 2-amino-6-methylpyridinium systems with the inversion centre in the middle of the fumarate shown and one of the $R_2^2(8)$ motifs outlined. The two acceptor and two donor atoms are labelled a and d, respectively.

verification of proton transfer was completed by comparison of the carboxylate/carboxylic acid C-O and C=O bond lengths which are closer to parity following ionisation while the single bond to the OH group is longer than the double bond in the neutral molecule. Transfer was then confirmed by ^{14}N-^{1}H HMQC NMR experiments (see Section 4.5.2).

In all three systems, two 2-amino-6-methylpyridinium ions and a fumarate ion form centrosymmetric base-acid-base (b-a-b) units (Fig. 4.3). These three ions are H-bonded via two pairs of intermolecular H-bonds, with each pair consisting of two donors, the pyridinium nitrogen and the 2-amino group nitrogen, and two acceptors, the carboxylate oxygen atoms on the fumarate, forming a ring motif, $R_2^2(8)$.[70] The two pairs of H-bonds are identical as the midpoint of the C=C double bond of the fumarate is an inversion centre. The H-bond parameters for the b-a-b units are given in Table 4.2, along with the parameters for the other significant H-bonding motifs identified. In the monohydrate and anhydrous cases, the fumarate ion is in the two s-*cis* conformation. Identification of conformation for the fumarate ions is more complicated than for the nuetral molecule due to the sharing of charge that often

Table 4.2: H-bonding parameters for the three 2-amino-6-methylpyridinium crystal structures.

	N···O distance (Å)	**NH$^+$···$^-$O angle (°)**	**NH···O angle (°)**	**O···O distance (Å)**	**OH···O angle (°)**	**OH···$^-$O angle (°)**
26AMP:F$_{0.5}$:(H$_2$O)$_2$	*2.69*	*172.2*	-			
	2.85	-	176.2			
	2.94	-	170.8			
				2.79	173.3	-
				2.76	171.0	-
				2.86	167.9	-
				2.85	179.2	-
26AMP$_2$:F:H$_2$O	*2.72*	*174.3*	-			
	2.79	-	170.9			
	2.80	-	175.8			
	2.71	*177.2*	-			
	2.81	-	169.7			
	2.86		163.9			
				2.82	176.3	
				2.85	170.8	
26AMP$_2$:F:FA$_{0.5}$	*2.69*	*174.1*	-			
	2.76	-	175.9			
	2.88	-	156.1			
	2.76	*172.5*	-			
	2.82	-	156.5			
	2.78	-	159.9			
				2.55	-	169.2

*Parameters in italics are those relating to b-a-b units

occurs between the two carboxylate oxygens, which contributes to the aforementioned (see section 1.2) change between carboxylate/carboxylic acid C-O and C=O bond lengths. Although the two C-O bond lengths are far more similar following ionisation, whichever remained slightly longer was taken to correspond to the C-O^-. Applying this to the C-O bonds of each carboxylate group suggests that the two s-*trans* conformation, the highest in energy, is present in the dihydrate system.

4.3.1. 26AMP:$F_{0.5}$:$(H_2O)_2$

The dihydrate form has a 1 : $^1/_2$: 2 stoichiometry, base : acid : water, and crystallises in the monoclinic space group $P2_1/c$. It has previously been reported by Selyani *et al.* who recorded it at 296 K (CCDC ref.: COGCIN, no.: 1521964)[142] but all parameters and calculations

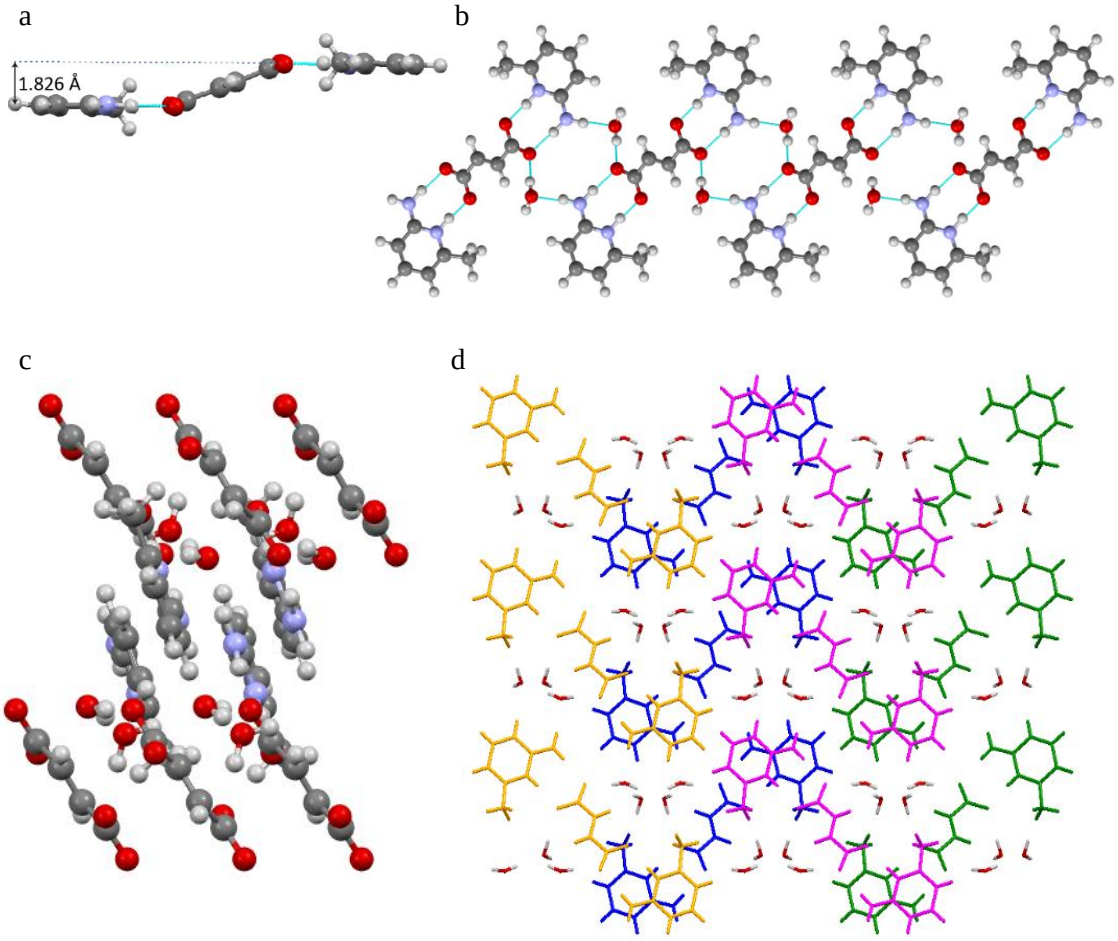


Figure 4.4: Packing in 26AMP:$F_{0.5}$:$(H_2O)_2$: (a) b-a-b unit with base molecules in parallel planes; (b) a band of b-a-b units stabilised by water molecules; packing of the bands into alternately angled layers (c) viewed along the b axis and (d) viewed along the c axis (yellow and pink bands form the layer above the green and blue bands). H-bonds are shown as bright blue lines in (a) and (b).

referred to in this work correspond to those of the structures determined in this work. In the b-a-b units, the two cations of 2-amino-6-methylpyridinium lie in planes parallel to each other that are offset by 1.826 Å (Fig. 4.4a). These b-a-b units align with each other to form chains, stabilised by H-bonding from the amine group via a water molecule to the carboxylate. This forms a larger H-bonded 12 sided ring motif between the $R_2^2(8)$[70] motifs of subsequent b-a-b units (Fig. 4.4b), forming a band of linked b-a-b units through the structure.

The bands then pack alongside one another to form layers across the $[10\bar{1}]$ plane (viewed from above in Fig. 4.4d). A second water molecule provides stabilisation between the bands in subsequent layers (Fig. 4.4c) which are offset and angled in the other direction (Fig. 4.4d). The resulting 3D structure has clear channels of water running through it, as shown in Fig. 4.4d, but does not function as a channel hydrate due to the strong H-bonds fixing them into distinct crystallographic locations.

4.3.2. 26AMP$_2$:F:H$_2$O

This salt hydrate has a stoichiometry in the asymmetric unit of 2 : 1 : 1, base : acid : water, and crystallises in the triclinic space group $P\bar{1}$. Two distinct b-a-b units are formed between 2-amino-6-methylpyridinium and fumarate ions (Fig. 4.5a). As there is an inversion point at the midpoint of each fumarate C=C double bond, the asymmetric unit contains two hemi fumarate molecules (see Fig. 4.2). The two distinct b-a-b units are linked by H-bonds between the second NH_2 proton of each base (*i.e.* the NH_2 proton that is furthest from the pyridine nitrogen and not involved in maintaining the b-a-b unit) and one of the carboxylate oxygens of the neighbouring b-a-b unit as well as water mediated H-bonds with one of the carboxylate oxygens (Fig. 4.5b). The two b-a-b units lie 65.5° out of plane with each other (Fig. 4.5c) whereas subsequent repeats of each unit are in parallel planes to themselves. As so few crystals were obtained from the crystallisation, bulk confirmation by PXRD was not possible.

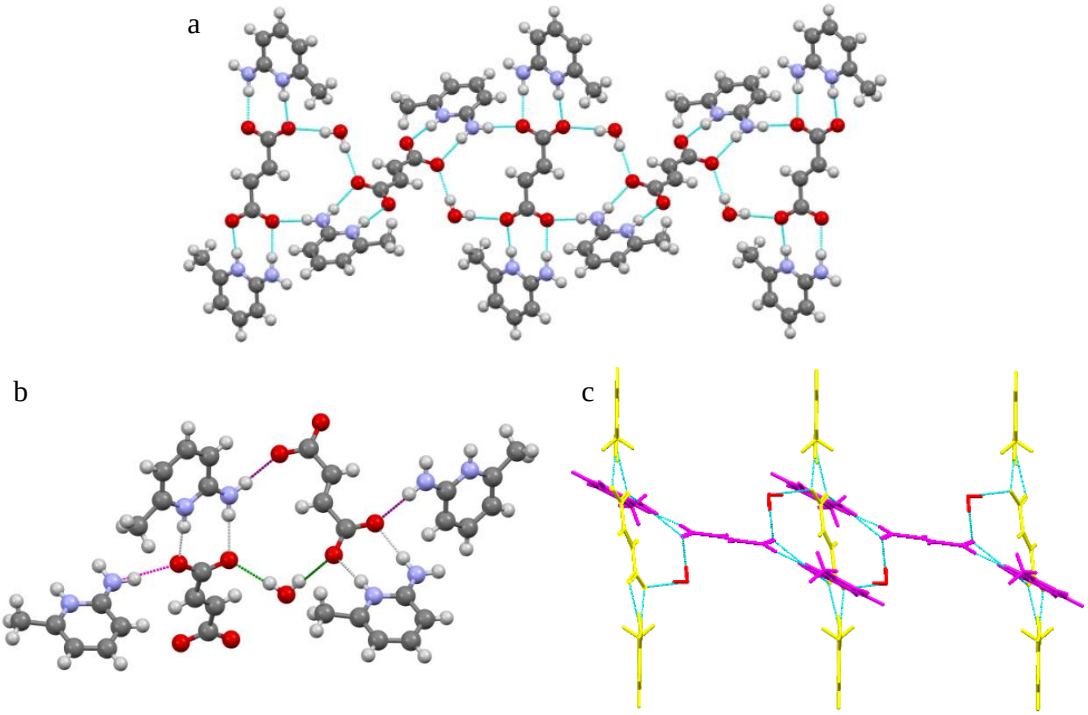


Figure 4.5: Packing in 26AMP$_2$:F:H$_2$O: (a) chain of b-a-b units viewed from above the (120) plane, (b) network of H-bonds linking b-a-b units with the symmetry related H-bonds coloured the same (only light greys and purple occur twice in figure) and (b) chain viewed along the c axis. H-bonds are shown as bright blue lines in (a) and (c).

4.3.3. 26AMP$_2$:F:FA$_{0.5}$

This anhydrous cocrystal of a salt has stoichiometry 2 : 1 $^1/_2$: 0, base : acid, and crystallises in the triclinic space group $P\bar{1}$. Two distinct b-a-b units are formed by the 2-amino-6-methylpyridinium and fumarate ions. One b-a-b unit is joined into a chain by molecules of FA (Fig. 4.6a) and the other sits in the spaces created by the FA at 60.3° to the linked units (Fig. 4.6b). As seen for the fuamrate ions, the FA molecule also has the lowest energy two s-*cis* conformation. The independent/unlinked b-a-b unit is close to being planar, with a difference of only 10.2° between the pyridinium and fumarate, whereas the fumarate of the linked b-a-b unit is twisted with respect to its pyridinium due to its interaction with the FA, making the NH···O H-bond between the amine and the carboxylate less ideal, with an NHO angle of 156.5°. Like the b-a-b units, the FA molecule is also centrosymmetric with the midpoint of the C=C double bond of the fumarate at an inversion centre.

The unlinked units bridge the spaces between chains of FA linked units, packing them into a layer (Fig. 4.6c). This layer is stabilised by amino-carboxylate H-bonds between FA linked and bridging b-a-b units (Fig. 4.6d). The layers are then stacked along the *a* dimension (Fig. 4.6e).

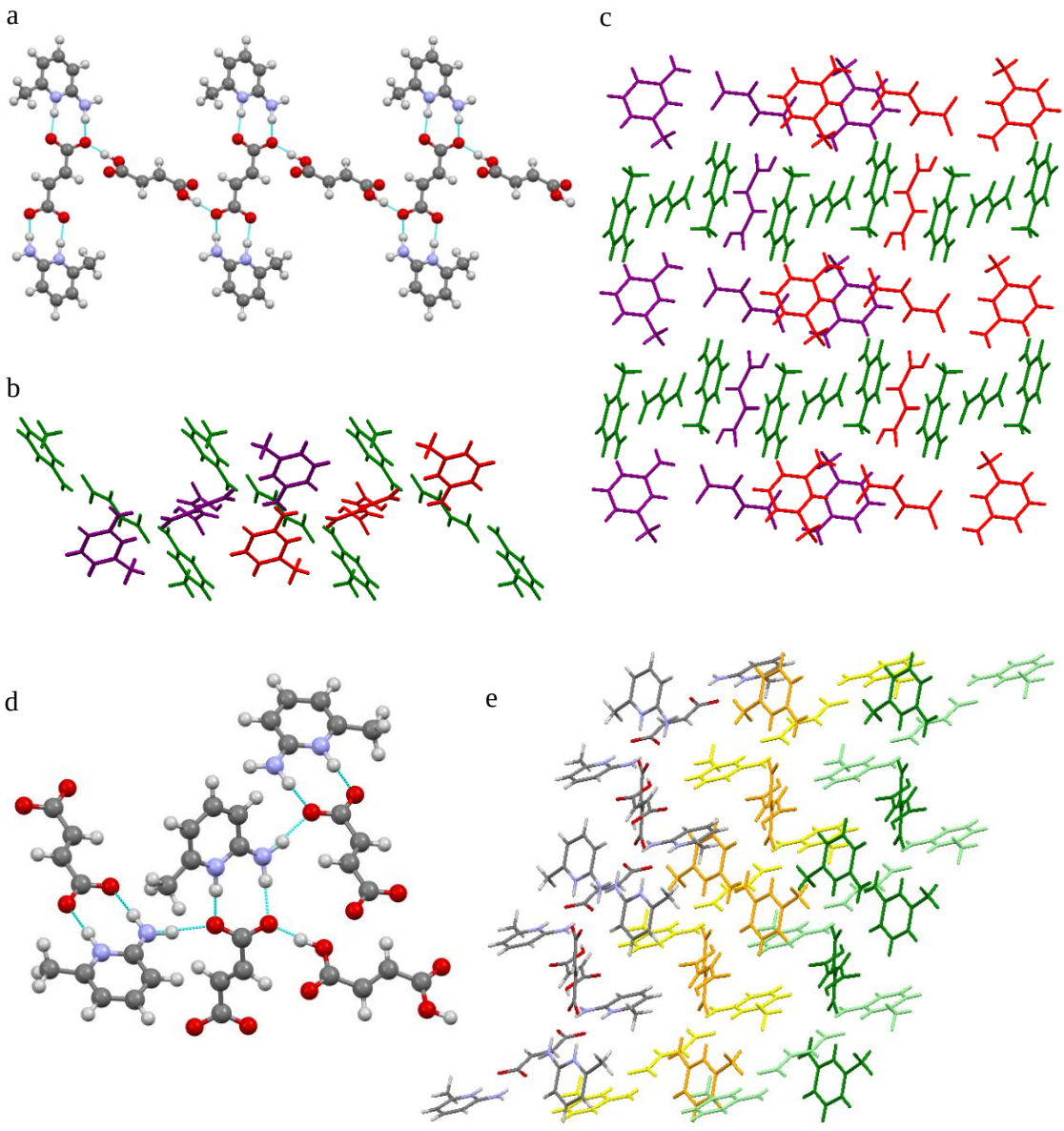


Figure 4.6: Packing in 26AMP$_2$:F:FA$_{0.5}$: (a) planar chain of FA linked b-a-b units; (b) bridging of the space between planar chains by a second b-a-b unit to form a layer, viewed along the a axis and (c) same, viewed along the c axis (red and purple correspond to molecules within two neighbouring chains, with the bridging b-a-b units shown in green – the central bridging b-a-b units correspond to those directly linking the displayed chains); (d) H-bonds between NH$_2$ and carboxylate groups linking the distinct b-a-b units; and (e) packing of the H-bonded layers (changes in colour correspond to subsequent layers, with the lighter colour in each representing the bridging b-a-b unit). H-bonds are shown as bright blue or light grey lines in (a) and (d).

4.4. Dehydration

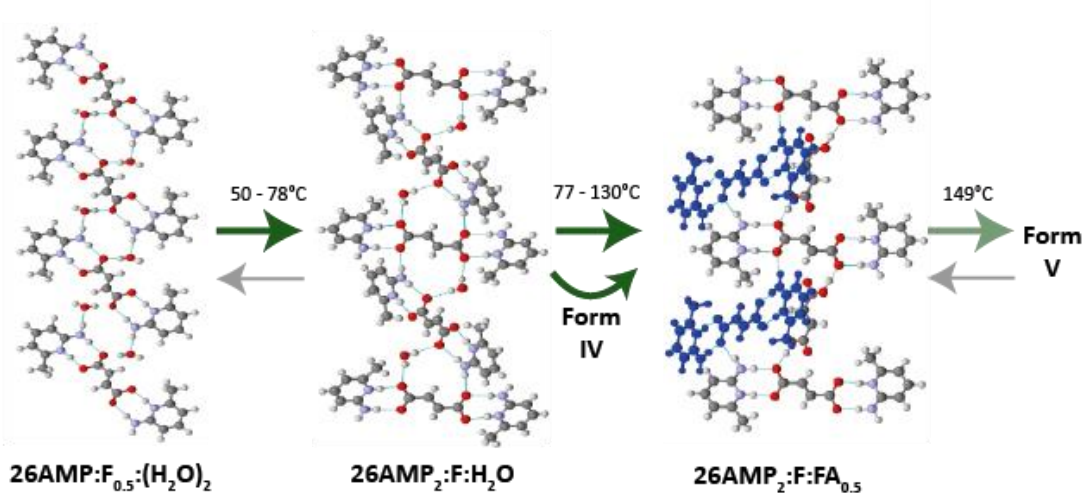


Figure 4.7: Transformations between crystal forms (transparent arrows indicate an unconfirmed transition). Left to right: the crystal structures show corresponding chain structures from 26AMP:$F_{0.5}$:$(H_2O)_2$, 26AMP$_2$:F:H_2O and 26AMP$_2$:F:$FA_{0.5}$.

This work will show that these three forms are directly related, with 26AMP:$F_{0.5}$:$(H_2O)_2$ converting first to 26AMP$_2$:F:H_2O and then a further transformation to 26AMP$_2$:F:$FA_{0.5}$ occurring on heating (Fig. 4.7). These transformations have been investigated by a range of techniques to better understand the structures, transitions and their relative stabilities. SXRD was conducted over a wide range of temperatures for each of the structures to track the variation in unit cell parameters and provide starting points for DFT calculations closer to the temperatures that were reached during VT-NMR experiments. This also provided an opportunity to see the conversion of 26AMP$_2$:F:H_2O to 26AMP$_2$:F:$FA_{0.5}$ at 77 °C, before the crystal quality had reduced too far for usable data collection (resulting in only faint, low angle, powder rings remaining).

The clearest evidence for these transitions is seen in a VT-PXRD experiment conducted on a sample of 26AMP:$F_{0.5}$:$(H_2O)_2$. As can be seen in Fig. 4.8, the experimental powder pattern originally showed good agreement with the pattern simulated for 26AMP:$F_{0.5}$:$(H_2O)_2$ at room temperature, but significant change in both structure and crystallinity is seen as the temperature was increased to 150 °C. A comparison of the PXRD scans recorded at 50 °C and 60 °C to the simulated pattern for 26AMP$_2$:F:H_2O (Fig. 4.9, top) shows no discrepancies although the experimental reflections have broadened compared to

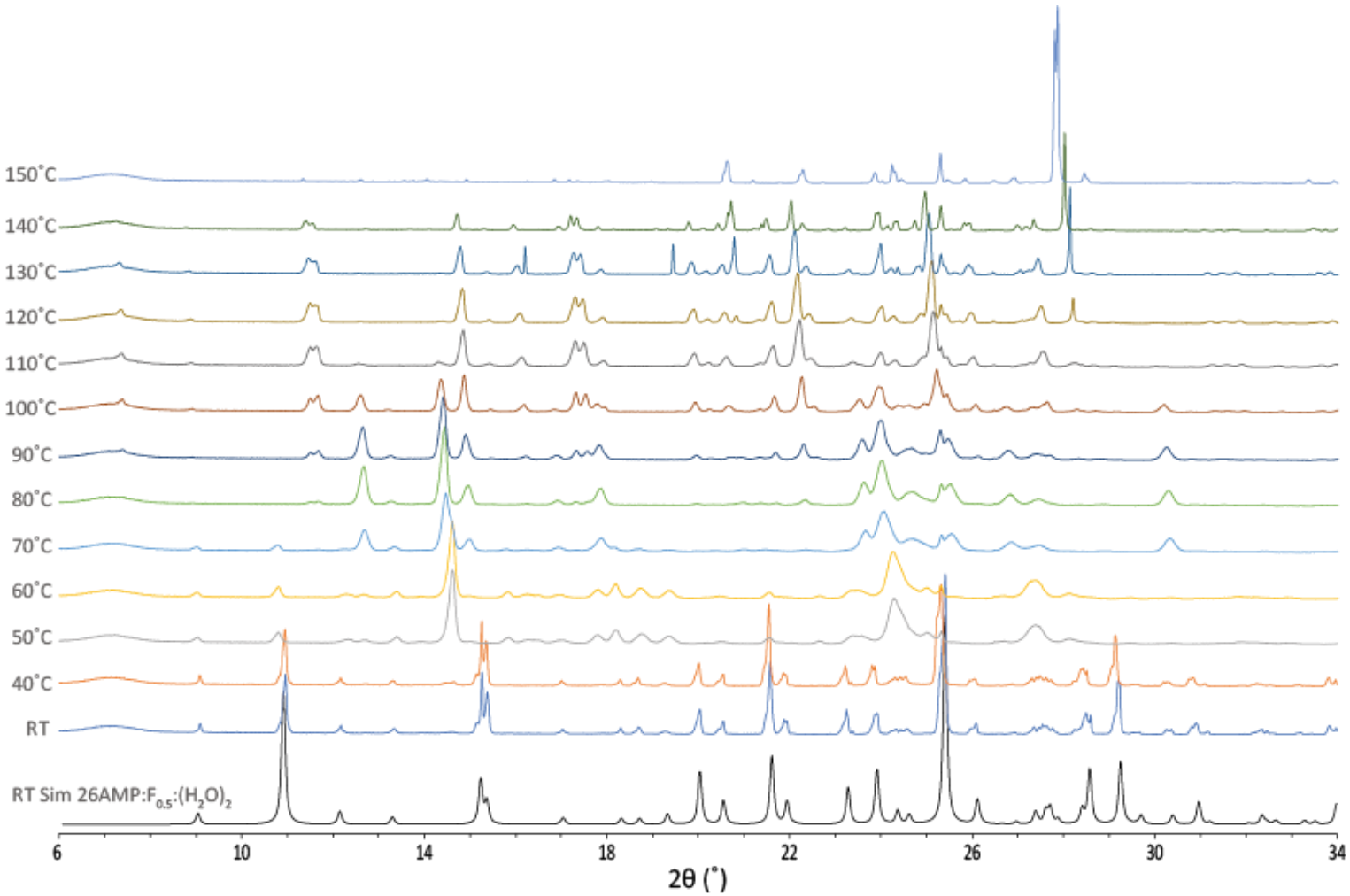


Figure 4.8: VT-Experimental PXRD patterns of a sample that was initially 26AMP:$F_{0.5}$:$(H_2O)_2$, recorded from room temperature to 150 °C, shown with the simulated powder pattern for 26AMP:$F_{0.5}$:$(H_2O)_2$ (bottom).

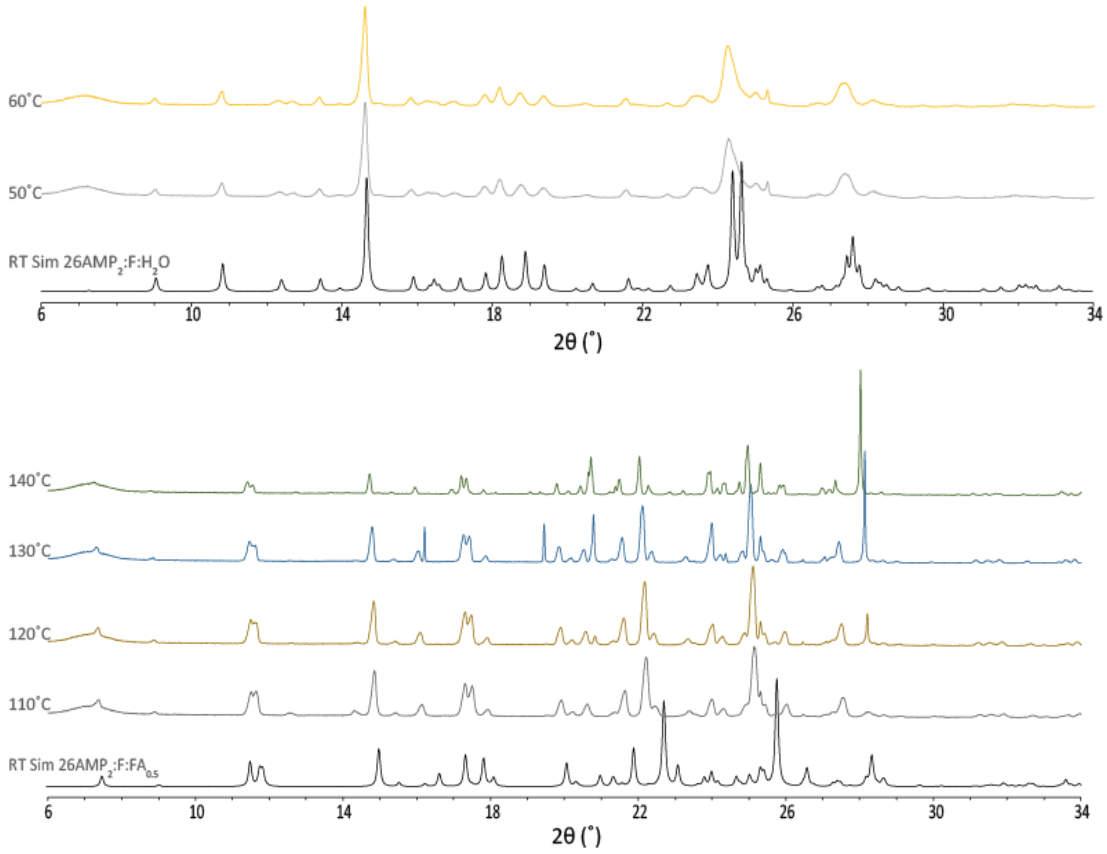


Figure 4.9: Selected experimental PXRD patterns from Fig. 4.8, shown with the simulated powder patterns for 26AMP$_2$:F:H$_2$O (top) and 26AMP$_2$:F:FA$_{0.5}$ (bottom).

those at room temperature, suggesting that the conversion to the less hydrated form is accompanied by a reduction in sample crystallinity. It is interesting to note that, based on the SXRD structures, the conformation of fumaric acid may also change from the two s-*trans* conformation to the slightly lower energy two s-*cis* conformation during this transition. As discussed below, sharper, more crystalline peaks begin to emerge again above 100 °C. The reduction in the intensity of most reflections at 150 °C is attributed to the loss of most of the sample at this point due to sample contraction and the flat plate set up employed. A scan of just the sample holder allows the reflection at 25.3°, which increases in intensity due to sample contraction and subsequent loss, to be ruled out as being due to the Al_2O_3 sample holder (although not shown in Fig. 4.8, additional reflections are also seen at 34.8°, 37.4° and 43.0° which show the same behaviour and confirm the assignment to the sample holder). The broad amorphous peak around 7° is thought to be due to the furnace windows. The higher temperature transitions do not seem to be as straightforward as that seen at 50 °C (Fig. 4.9,

top). This is not unexpected as both a loss of the remaining water and also a change in the base : acid ratio is required to allow the formation of $26AMP_2$:F:$FA_{0.5}$. Although peaks associated with $26AMP_2$:F:$FA_{0.5}$ start to appear at 70 °C, unrelated reflections also form simultaneously at 13.4° and 30.5° (Fig. 4.8). These reflections, thought to correspond to an unknown polymorph, Form IV, decrease again from 90 °C onwards and have vanished by 120 °C leaving only reflections corresponding to $26AMP_2$:F:$FA_{0.5}$ (Fig. 4.9, bottom). It is therefore assumed that $26AMP_2$:F:$FA_{0.5}$ can form both directly or via Form IV. As the melting point of 26AMP in isolation is ~50 °C, the change in base : acid ratio required could be due to slow evaporation of base molecules, as seen in 2,6-lutidinium hydrogen fumarate, discussed in Chapter 3.[201] If the sample is single phase by 120 °C, this indicates that a proportion of the 26AMP molecules have been lost as well as the remaining water, accounting for the significant size reduction of the crystal during the SXRD monitored transition (discussed below, Fig. 4.11). However, there are sharp reflections present in the PXRD pattern at 16.20° and 19.45° in the scan recorded at 130 °C that are not present at either 120 °C or 140 °C. These may be associated with either another intermediate or a secondary phase that is immediately lost again at higher temperatures (or changes in preferred orientation within the sample upon heating that cause them to be masked at 140 °C). The remaining reflections are attributed to $26AMP_2$:F:$FA_{0.5}$, although several show a significant shift in angle as would be expected due to thermal expansion (as the simulated pattern is for a room temperature structure). This shift is clearly evident for the simulated reflections at 15.08°, 22.82°, 25.86° and 28.44° which can still be seen to shift within the relatively smaller temperature range from 110 °C to 140 °C (14.87°, 22.23°, 25.17° and 28.30° and 14.74°, 22.02°, 24.98° and 28.02° for 110 °C and 140 °C, respectively).

Another transformation is observed to occur beyond $26AMP_2$:F:$FA_{0.5}$ to a fifth crystalline structure (Form V) as the reflections remaining in the 150 °C pattern (Fig. 4.8) clearly differ to both those at 140 °C and the $26AMP_2$:F:$FA_{0.5}$ simulated pattern. Most noticeably, the reflections are at 27.81° and 27.88° at 150 °C, whereas only a single reflection is seen at 28.02° at 140 °C. The structure of this second additional form, Form V, is unknown.

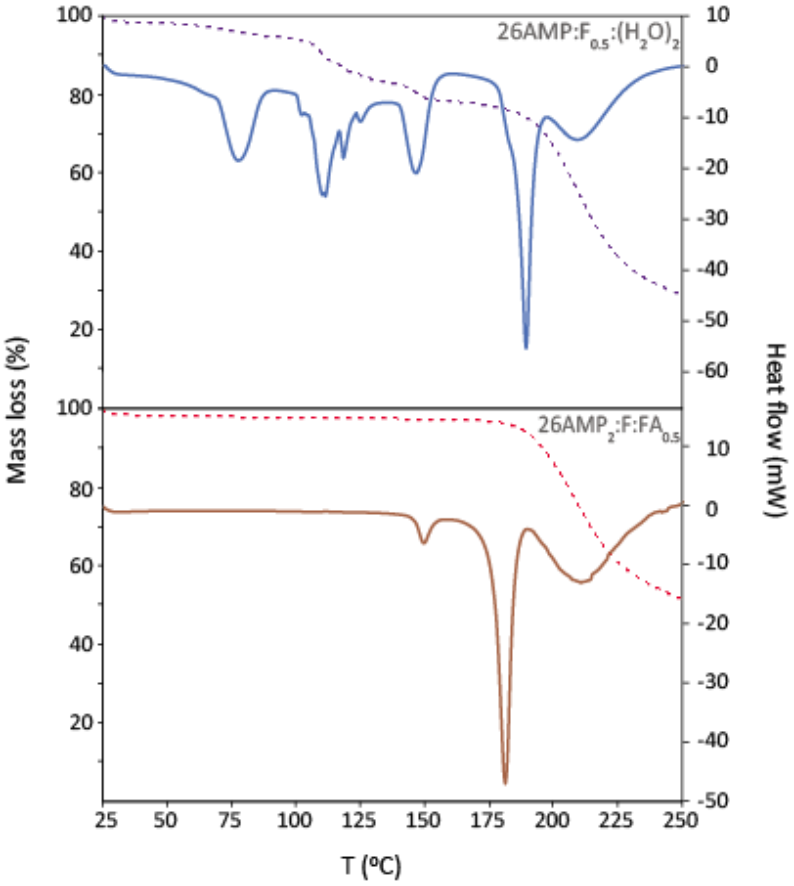


Figure 4.10: DSC (solid lines) and TGA (dashed lines) of 26AMP:$F_{0.5}$:$(H_2O)_2$ and $26AMP_2$:F:$FA_{0.5}$ recorded on a Mettler Toledo Star^e instrument with a ramp of 10 °C/min from 25-250 °C.

This further transformation is supported by the DSC data for $26AMP_2$:F:$FA_{0.5}$, which shows a solid-solid transition prior to melting that occurs at 149.3 °C (Fig. 4.10). The difference in rate of change of temperature between the VT-PXRD and DSC, as well as flat plate vs. lidded DSC pan set up, is an explanation as to why all the transitions are observed at slightly lower temperatures in the VT-PXRD. DSC of 26AMP:$F_{0.5}$:$(H_2O)_2$ shows three solid-solid transitions prior to melting, with minima in the heat flow at 78.0 °C, 112.5 °C and 148.7 °C. These are thought to correspond to the transitions to $26AMP_2$:F:H_2O, $26AMP_2$:F:$FA_{0.5}$ and Form V, with the second transition considerably more complex and presenting multiple minima, due to the additional trasition pathway via Form IV, and the third transition matching that seen within DSC of a $26AMP_2$:F:$FA_{0.5}$ sample. However, the mass losses corresponding to the first two transitions for 26AMP:$F_{0.5}$:$(H_2O)_2$ account for only about half the losses expected (17.57 % of initial mass observed in total for both transitions, compared to the 35.64 % expected). This discrepancy, as well as the absence of recrystallisation peaks in the DSC and higher transition

temperatures, may in part be due to the rate of change in temperature in the DSC compared to that utilised for the VT-PXRD but also suggests the sample is not phase pure, most likely with some of the subsequent forms or FA already present. (Note that loss of $26AMP_2$:F:H_2O sample through recording VT-SXRD and packing a 1.3 mm MAS NMR rotor meant there was not enough single-phase sample remaining to run TGA/DSC.)

Heating of a single crystal of 26AMP:$F_{0.5}$:$(H_2O)_2$ above room temperature resulted in both clouding and fracturing of the crystal (Fig. 4.11), with this reduction in crystal quality corresponding to a gradual loss in diffraction peaks, leaving only faint, low angle, powder rings remaining in each frame. Although a structure was successfully recorded at 67 °C, the crystal degraded during data collection. A similar clouding was seen for a crystal of $26AMP_2$:F:H_2O but the loss in diffraction peaks was far more gradual, allowing the same crystal to be analysed over a range of temperatures between −23 °C and 77 °C (see Table 4.3 and discussion, below). At 77 °C, the structure recorded for this crystal matched that of $26AMP_2$:F:$FA_{0.5}$, not $26AMP_2$:F:H_2O. The significant clouding and slight contraction seen in Fig. 4.11 had occurred by the beginning of the data collection at 67 °C and may correspond to water loss from more clouded outer regions of the crystal. The more extreme size reduction was in evidence at the start of the following data collection at 77 °C. As a change in base : acid ratio is needed to transform between $26AMP_2$:F:H_2O and $26AMP_2$:F:$FA_{0.5}$, this is thought to be due to not only the loss of the remaining water but also some of the base molecules to allow the transition.

As the potential loss of 26AMP is assumed to be possible only through evaporation of the base molecules, as seen in 26L:HF,[201] it was initially thought that it must occur slowly and may therefore have been in progress during the previous experiment at 67 °C. However, no evidence of this was seen in the 67 °C data, with R_{int} remaining low and relatively stable

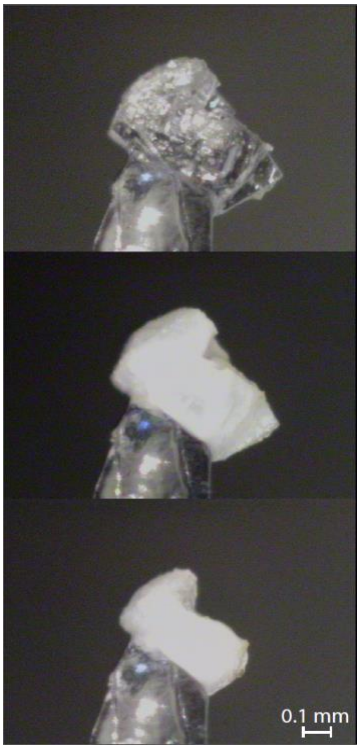


Figure 4.11: A single crystal, initially 26AMP:$F_{0.5}$:$(H_2O)_2$, when first mounted (top), at 67 °C (middle) and at 77 °C (bottom).

between frames throughout collection (bar a single frame early on, see Fig. A2.1 in Appendix 2). This was also true for the data collected at 77 °C, suggesting that the transition took place during the temperature change. This phase change in a single crystal is the most conclusive evidence that the 26AMP_2:F:H_2O transition is possible and occurs but does not elucidate the mechanism by which 1 in 3 base molecules are lost. It is significantly lower in temperature than that observed in the VT-PXRD above (Fig. 4.8 and 4.9).

4.5. NMR Crystallography Study

4.5.1. Geometry Optimisation

The relative energies of each structure following either fixed or variable unit cell geometry optimisation are listed in Table 4.3, alongside their unit cell parameters. Fixing the unit cell maintains an association to the temperature at which the crystal structure was recorded. As expected, all three systems show thermal expansion with increasing temperature, although this

is very slight for 26AMP$_2$:F:FA$_{0.5}$ and falls within the estimated standard deviation of the measurements (Fig. 4.12). Although there is a direct correlation between temperature and unit cell volume, none of the systems expands uniformly, with all exhibiting changes to the unit cell angles and therefore a reduction in one or more of the unit cell lengths. It would generally be expected that lower temperature structures with smaller unit cell volumes would be lower in energy but that does not appear to be the case. In fact, the 150 K structures of both 26AMP:F$_{0.5}$:(H$_2$O)$_2$ and 26AMP$_2$:F:H$_2$O have the highest energies of their sets, although the differences are relatively small (< 5 kJ mol^{-1}). The room temperature (290 K) structure has the lowest energy for 26AMP$_2$:F:H$_2$O and surprisingly, given both its high temperature and larger unit cell volume, the 340 K structure has the lowest energy for 26AMP:F$_{0.5}$:(H$_2$O)$_2$. The published structure of 26AMP:F$_{0.5}$:(H$_2$O)$_2$, COGCIN,[142] was recorded at RT and, as expected, is similar to the 300 K 26AMP:F$_{0.5}$:(H$_2$O)$_2$ structure.

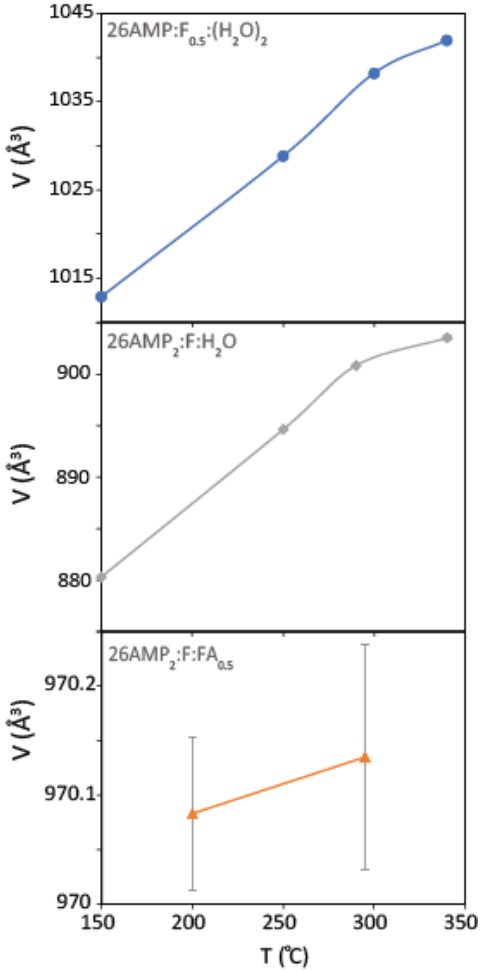


Figure 4.12: Variation in volume of the unit cell with temperature for 26AMP:F$_{0.5}$:(H$_2$O)$_2$ (top), 26AMP$_2$:F:H$_2$O (middle) and 26AMP$_2$:F:FA$_{0.5}$ (bottom). Error bars show the estimated standard deviation in the unit cell volume but are too small to see for 26AMP:F$_{0.5}$:(H$_2$O)$_2$ and 26AMP$_2$:F:H$_2$O.

Table 4.3: Unit cell parameters of 26AMP:$F_{0.5}$:$(H_2O)_2$, AF16-H0.5 and 26AMP_2:F:$FA_{0.5}$ following a geometry optimisation with the unit cell fixed (i.e., as solved by SXRD) and allowed to vary, alongside the respective energy of each optimised structure.

	Structure	*a* (Å)	*b* (Å)	*c* (Å)	*α* (°)	*β* (°)	*γ* (°)	Volume (Å^3)	Energy (kJ mol^{-1})*
26AMP:$F_{0.5}$:$(H_2O)_2$	150K (fixed)	9.6598(3)	14.4288(6)	7.3093(3)	90	96.158(3)	90	1012.89	4.14
	150K	9.5378	14.4175	7.19	90	95.22	90	984.603	0.50
	250K (fixed)	9.6866(2)	14.4243(3)	7.41090(10)	90	96.494(2)	90	1028.83	1.98
	250K	9.5511	14.4245	7.21322	90	95.005	90	989.973	0.74
	COGCIN (fixed)	9.7260(4)	14.4355(6)	7.4679(4)	90	96.691(4)	90	1041.35	0.22
	COGCIN	9.5369	14.4089	7.2091	90	94.978	90	986.91	0.49
	300K (fixed)	9.7012(4)	14.4259(6)	7.4682(3)	90	96.607(4)	90	1038.22	0.69
	300K	9.551	14.4317	7.197	90	94.965	90	988.297	0.56
	340K (fixed)	9.6906(10)	14.4135(16)	7.5135(10)	90	96.856(11)	90	1041.95	0
	340K	9.5344	14.4339	7.1801	90	95.075	90	984.236	0
26AMP_2:F:H_2O	150K (fixed)	7.5217(2)	10.0604(3)	12.6922(4)	93.195(2)	107.236(3)	104.283(3)	880.295	0.16
	150K	7.4769	10.1273	12.6615	93.126	107.711	105.162	872.075	0.02
	250K (fixed)	7.6213(3)	10.0308(4)	12.7136(5)	93.561(3)	107.316(4)	103.160(4)	894.647	0.06
	250K	7.4498	10.1473	12.6579	93.223	107.942	105.183	868.742	0.02
	290K (fixed)	7.6640(3)	10.0215(4)	12.7140(4)	93.693(3)	107.314(3)	102.633(3)	900.848	0.02
	290K	7.4598	10.1433	12.6514	93.326	107.933	104.961	870.01	0
	340K (fixed)	7.7084(13)	9.9834(17)	12.689(2)	93.802(14)	107.271(15)	101.985(14)	903.541	0
	340K	7.4431	10.1803	12.67	93.309	108.196	105.284	869.429	0.01
26AMP_2:F:$FA_{0.5}$	200K (fixed)	8.4426(4)	10.5523(4)	12.3026(4)	85.716(3)	73.190(4)	67.717(4)	970.083	0
	200K	8.3511	10.4192	12.1944	86.931	74.204	69.293	953.957	0
	RT (fixed)	8.44545(4)	10.5589(7)	12.3000(6)	85.679(4)	73.080(6)	67.701(5)	970.135	0.60
	RT	8.36347	10.4544	12.1978	86.787	74.1539	69.1314	957.656	0.25

*Energies are given relative to the lowest energy structure for each set. The difference between the lowest energy structures in the fixed and variable unit cell sets was 150.98 kJ mol^{-1}, 5.32 kJ mol^{-1} and 0.99 kJ mol^{-1} for 26AMP:$F_{0.5}$:$(H_2O)_2$, 26AMP_2:F:H_2O and 26AMP_2:F:$FA_{0.5}$, respectively.

It is also interesting to note that although the variable unit cell optimisation moves all the structures to a lower energy than any of their respective 'temperature' structures (as expected given that the calculations have no temperature dependence included so are nominally at 0 K) they do not converge to exactly the same point. This is likely due to how fine the tolerances are. The optimisation enters a new cycle until the forces and energy changes between subsequent configurations are within a tolerance limit, at which point the calculation is judged to be at the local minimum and no more changes are made. The smaller these tolerances are, the closer to the absolute minimum the calculation must be at the end point. The default tolerances are fine enough that the small variations remaining do not affect the NMR parameters calculated subsequently but do mean that the calculations from different starting points do not perfectly converge. It can clearly be seen that the variation between the fixed unit cell 'temperature' structures (mentioned above) is significantly larger than that seen for the variable '0 K' structures.

Table 4.4: GIPAW calculated and experimental ^{1}H chemical shifts for 26AMP:$F_{0.5}$:$(H_2O)_2$, 26AMP$_2$:F:H_2O and 26AMP$_2$:F:$FA_{0.5}$.

	26AMP:$F_{0.5}$:$(H_2O)_2$		**26AMP$_2$:F:H_2O**		**26AMP$_2$:F:$FA_{0.5}$**	
	$\delta^{iso}_{calc}(^1H)$ **(ppm)**	$\delta^{iso}_{exp}(^1H)$ **(ppm)**	$\delta^{iso}_{calc}(^1H)$ **(ppm)**	$\delta^{iso}_{exp}(^1H)$ **(ppm)**	$\delta^{iso}_{calc}(^1H)$ **(ppm)**	$\delta^{iso}_{exp}(^1H)$ **(ppm)**
Me	1.5	1.5	1.4	1.5	1.0	1.8
	-		1.4	1.5	1.6	1.8
H_2O	4.7	4.2	4.3	4.1	-	
	5.3	5.3	4.5	4.1	-	
	5.3	6.0	-		-	
	6.1	6.0	-		-	
CH	5.8	6.0	5.3	5.7	4.9	4.9
	6.3	6.0	5.4	5.7	4.9	5.1
	6.5	6.8	5.5	5.7	5.7	5.8
	6.7	6.8	5.9	5.7	6.2	6.7
	-		6.3	6.3	6.6	6.7
	-		6.4	6.3	6.7	6.7
	-		7.0	7.4	6.7	6.7
	-		7.4	7.4	7.1	7.7
	-		-		7.5	7.7
NH_2	7.9	8.2	8.7	9.2	8.1	7.7
	8.3	8.2	8.7	9.2	9.0	9.1
	-		9.5	9.8	9.3	9.1
	-		10.5	9.8	12.1	11.4
NH	14.8	14.9	15.0	14.8	14.7	14.4
	-		15.0	14.8	16.0	16.4
OH	-		-		16.3	14.9

The difference in final energy between the fixed and variable unit cell optimisations for each system appears to be related to its level of hydration. The anhydrous system, $26AMP_2:F:FA_{0.5}$, differs by only ~1 kJ/mol, whereas $26AMP_2:F:H_2O$ changes by ~ 5 kJ/mol

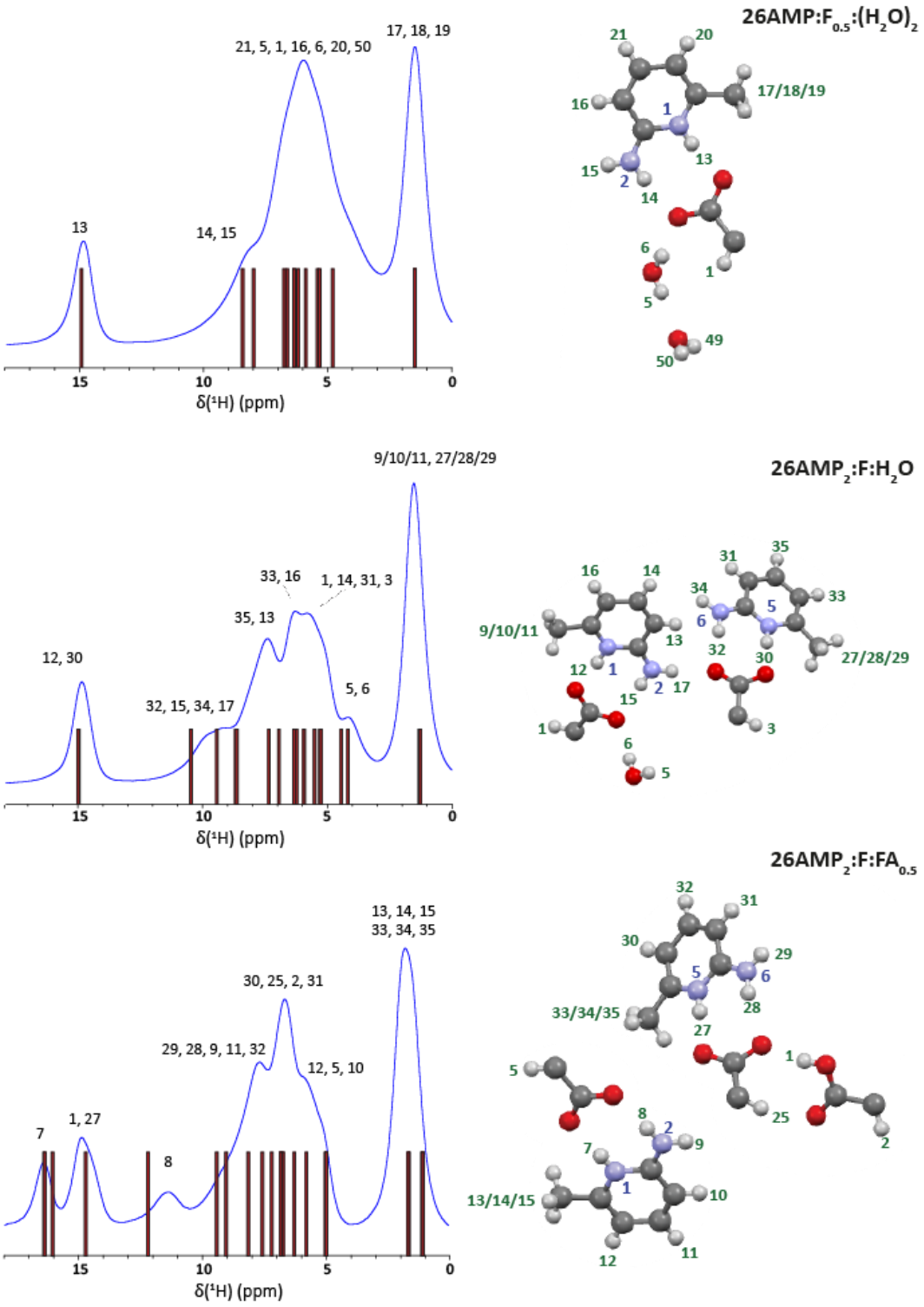


Figure 4.13: ^1H (600 MHz) one-pulse MAS (60 kHz) spectra of $26AMP:F_{0.5}:(H_2O)_2$ (top), $26AMP_2:F:H_2O$ (middle) and $26AMP_2:F:FA_{0.5}$ (bottom) with stick spectra corresponding to GIPAW calculated chemical shifts for the geometry optimised crystal structures. The assignments to each proton, labelled in the structures on the right, are given.

and 26AMP:$F_{0.5}$:$(H_2O)_2$ changes by ~150 kJ/mol. This may be due to the additional space and flexibility within the system due to the inclusion of water molecules.

4.5.2. ^{1}H NMR

Experimental solid-state MAS NMR analysis of all three systems showed good agreement between the GIPAW calculated and experimental 1H chemical shifts as well as verifying, as expected from calculation, that the forms can be easily distinguished from each other due to the change in 1H chemical shifts of the NH2 protons (Table 4.4 and Fig. 4.13). The anhydrous form is also easily identified due to the separation between the NH protons (H7 and H27) and the presence of the OH proton (H1). The OH proton of 26AMP$_2$:F:$FA_{0.5}$ also showed the largest discrepancy between experiment and calculation of the chemical shift with the experimental value 1.4 ppm lower than calculated. A similar difference was also observed for 26L (Chapter 3). This discrepancy is either due to a genuine difference in the actual proton position compared to the optimised crystal structure or it is simply that the discrepancy compared to experiment for the calculations for such a high chemical shift proton in an OH···O H-bond is greater than those of other environments (a similar difference was also seen for the OH···O proton in 26L:HF – see further discussion in section 6.3.2).

The ^{1}H-^{1}H DQ MAS spectra are also very distinct (Fig. 4.14). The number of proton correlations due to the proximities between the water molecules and the CH protons in the dihydrate, 26AMP:$F_{0.5}$:$(H_2O)_2$, leads to a very broad, many shouldered resonance between 3 ppm and 10 ppm (Fig. 4.14a). There are very clear NH^+-H_2O correlations (δ_{DQ} = 14.8 + 5.3 = 20.2 ppm) as well as an apparent NH^+-NH^+ correlation (δ_{DQ} = 14.9 + 14.9 = 29.8 ppm). The NH^+ protons in 26AMP:$F_{0.5}$:$(H_2O)_2$ are brought into proximity due to the stacking of 2-amino-6-methylpyridinium ions between the layers of b-a-b chains and the distance between subsequent NH^+s is 3.80 Å, so the visible correlation is unexpected, as the BABA recoupling used is generally considered to only show cross-peaks for proximities up to ~3.5 Å (see Table 4.5), although it is still possible. As discussed in Chapter 2, the relative intensities of DQ peaks

at a particular SQ frequency is an indicator of the relative ^{1}H-^{1}H distances, by the relationship $I_{AB}/I_{AC} = r_{AC}^6/r_{AB}^6$.[167] As expected, the NH^+-NH^+ cross-peak has the lowest intensity of the

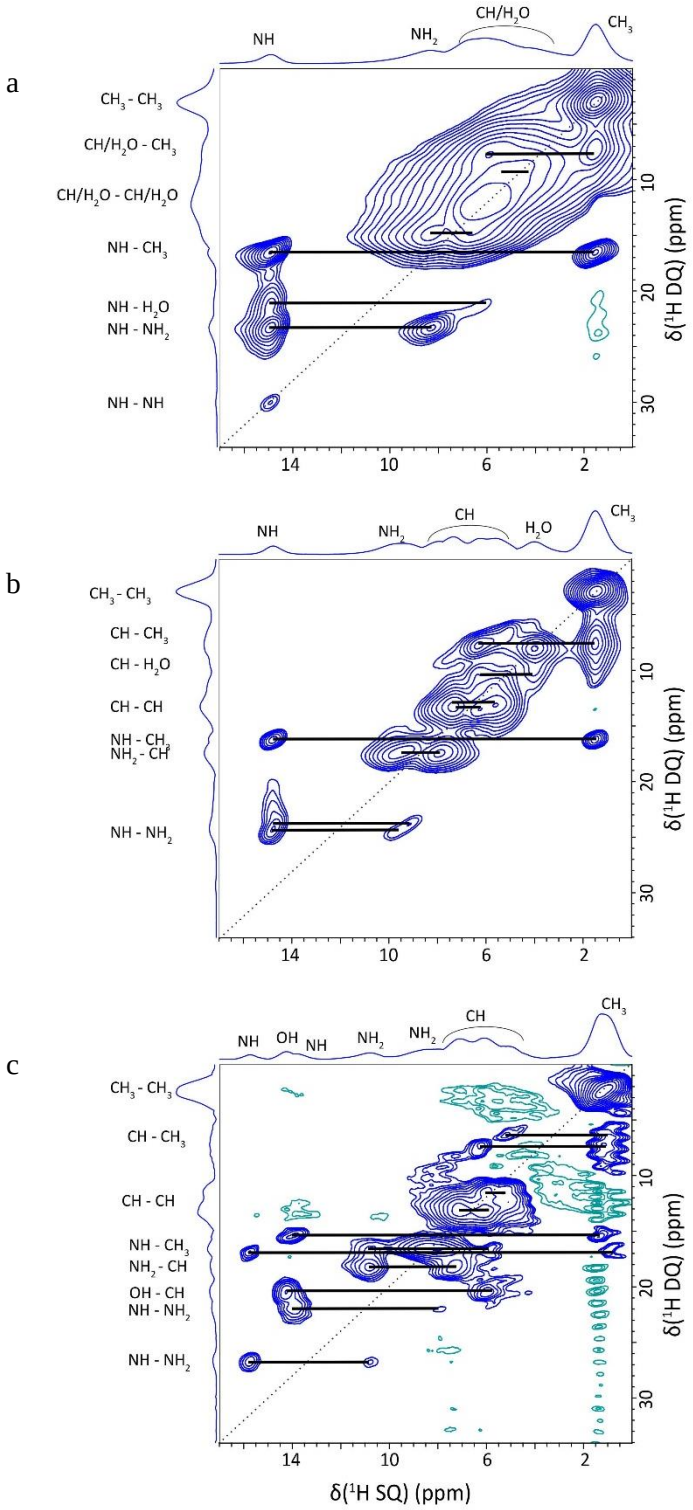


Figure 4.14: ^{1}H (600 MHz) DQ MAS (60 kHz) NMR spectra of 26AMP:$F_{0.5}$:$(H_2O)_2$ (top), 26AMP$_2$:F:H_2O (middle) and 26AMP$_2$:F:$FA_{0.5}$ (bottom) recorded with one rotor period of BaBa recoupling. The base contour level is at 4.5%, 6.2% and 28.6% of the maximum peak height, respectively. Blue and green contours correspond to positive and negative intensity respectively. The dashed diagonal line indicates the $\delta_{DQ} = 2\delta_{SQ}$ diagonal, while horizontal lines indicate a DQ peak at the sum of the two SQ peaks for dipolar coupled unlike protons.

DQ peaks at the SQ frequency for H13, the NH^+ proton. Its relative intensity compared to the NH_2 cross-peak is complicated due to multiple resonances associated with different distances contributing to the intensity of as the NH_2 cross-peak (both H14 and H15 at 2.30 Å and 3.56 Å, respectively). Similarly, proximity with the three methyl protons is averaged between the listed, nearest distance in Table 4.5 and further away. As the correlation with a water proton, H49, covers only a single resonance, this should be a reliable measure. However, it is twice as intense as it should be relative to the cross-peak to H49, $I_{H13-H13}/I_{H13-H49} = 0.40$ while $r^6_{H13-H49}/r^6_{H13-H13} = 0.20$ for the geometry optimised crystal structure. As there is only one inequivalent 2-amino-6-methylpyridinium ion in the asymmetric unit, there is therefore a neighbouring NH at an equal distance in the layer above and below which leads to a doubling

Table 4.5: H-H proximities (<4 Å) and corresponding ^{1}H DQ chemical shifts for the pyridinium NH's of each crystal structure

Structure	NH^+	δ_{iso}^{exp} SQ1 (ppm)	Proton 2	δ_{iso}^{exp} SQ2 (ppm)	δ_{iso}^{exp} DQ (ppm)	Separation[a] (Å)
26AMP:$F_{0.5}$:$(H_2O)_2$	H13	14.9	H14	8.2	23.1	2.30
			H17/H18/H19	1.5	16.4	2.59
			H49	*5.3*	*20.2*	*2.91*
			H15	8.2	23.1	3.56
			H13	*14.9*	*29.8*	*3.80*
26AMP_2:F:H_2O	H12	14.8	H15	9.8	24.6	2.26
			H9/H10/H11	1.5	16.3	2.44
			H34	*9.2*	*24.0*	*2.92*
			H17	9.2	24.0	3.54
			H32	*9.8*	*24.6*	*3.99*
	H30	14.8	H32	9.8	24.6	2.24
			H27/H28/H29	1.5	16.3	2.50
			H5	*4.1*	*18.9*	*2.89*
			H35	*7.4*	*22.2*	*3.48*
			H34	9.2	24.0	3.53
			H9/H10/H11	*1.5*	*16.3*	*3.77*
			H33	*6.3*	*21.1*	*3.98*
26AMP_2:F:$FA_{0.5}$	H27	14.4	H28	9.1	23.5	2.27
			H9	*7.7*	*22.1*	*2.79*
			H33/H34/H35	1.8	16.2	2.80
			H29	9.1	23.5	3.53
			H8	*11.4*	*25.8*	*3.93*
	H7	16.4	H8	11.4	27.8	2.26
			H13/H14/H15	1.8	18.4	2.55
			H2	*6.7*	*23.1*	*3.17*
			H33/H34/H35	*1.8*	*18.4*	*3.47*
			H9	7.7	24.1	3.54

[a] H-H distances are taken from the DFT (CASTEP) optimised structure. Intermolecular proximities are denoted using italic font.

of the resonance intensity, resulting in both the doubling relative to H49 and explaining how such a distant proximity has a visible correlation peak.

Conversely, the expected NH-H_2O correlation peak for 26AMP$_2$:F:H_2O, between H30 and H5, is not seen despite the optimised crystal structure proximity being 2.89 Å, well below the anticipated boundary of ~3.5 Å (Fig. 4.14b). This suggests that the water molecule does not sit in that position. The strong agreement of the SQ chemical shifts would seem to make this unlikely but does not rule out loss of the correlation due to motion of the H_2O molecule. The shoulder in the DQ dimension of the NH^+ cross peak of the NH^+-NH_2 pairings (δ_{DQ} = 14.8 + 9.8 = 24.6 ppm to δ_{DQ} = 14.8 + 9.2 = 24.0 ppm for both H12 and H30) extends to a lower DQ shift (δ_{DQ} = 14.8 + 4.1 = 18.9 ppm) suggesting that the water correlation may be present at a very low intensity. This NH^+-H_2O correlation between H30 and H5 is the only environment correlation expected for one b-a-b unit and not the other. Despite being crystallographically distinct, the similarity between b-a-b units is such that H12 and H30, the NH^+ protons of 26AMP$_2$:F:H_2O, lie at the same chemical shift and their NH_2 and CH_3 correlations overlap (H12/H30-methyl: δ_{DQ} = 14.8 + 1.5 = 16.3 ppm). The pair of cross peaks for the CH-H_2O correlations are visible (δ_{DQ} = 6.7 + 4.1 = 10.8 ppm), as well as the water self-correlation confirming that it is present in the expected local environment (δ_{DQ} = 4.1 + 4.1 = 8.2 ppm). The overlap of the two NH resonances means that analysis of the relative intensities for the NH DQ cross-peaks was not conducted for 26AMP$_2$:F:H_2O.

Comparing Fig. 4.14b and c, the distinction between inequivalent b-a-b units is far more apparent in 26AMP$_2$:F:FA$_{0.5}$, where the NH^+ protons, H7 and H27, are separated by 2 ppm and the nearest NH_2 proton for each are also significantly different, falling at 11.4 ppm and 9.0 ppm for H8 and H28, respectively (H28 is assigned to the shoulder centred at 9.0 ppm in the 1D spectra with several other protons but appears to lie closer to the main peak centre at δ_{SQ} = 8.1 ppm from the DQ spectrum). The OH proton of the FA at 14.7 ppm, unique to 26AMP$_2$:F:FA$_{0.5}$, is also clearly distinguishable from the NH^+ protons at 16.4 ppm and 14.4 ppm. This distinction between the three highest chemical shift protons of 26AMP$_2$:F:FA$_{0.5}$ can be easily made as there are two sets of peaks for each of the NH^+ protons, to NH_2 and CH_3

(H27: δ_{DQ} = 14.4 + 9.1 = 23.5 ppm and δ_{DQ} = 14.4 + 1.8 = 16.2 ppm; H7: δ_{DQ} = 16.4+ 11.4 = 25.8 ppm and δ_{DQ} = 16.4 + 1.8 = 18.2 ppm), and then a single correlation (δ_{DQ} = 14.7 + 6.0 = 23.7 ppm) corresponding to the OH proximity to H25. This helped to confirm the assignment of the OH peak, alongside the proton shifts observed for the ^{14}N correlations in the ^{14}N-^{1}H HMQC spectrum for 26AMP$_2$:F:FA$_{0.5}$ (presented below).

For 26AMP$_2$:F:FA$_{0.5}$, the overlap of SQ resonances of H27, one of the NH$^+$ protons, with H1, OH, means that the intensities of its double quantum correlations are unreliable measures due to the contribution from H1 proximities. H7, the higher ^{1}H chemical shift NH$^+$, is distinct, however. Taking a slice from the spectrum corresponding to this SQ resonance shows both the clearly visible cross-peaks to the methyl protons and H8, the closest NH$_2$ proton but also a low intensity peak for the expected intermolecular proximity to H2, the FA CH proton, at 3.17 Å. As for 26AMP:F$_{0.5}$:(H$_2$O)$_2$, the intensity of the methyl peak is not reliable due to the averaging of couplings to the three protons, particularly for 26AMP$_2$:F:FA$_{0.5}$ as two methyl groups fall under the same resonance. In fact this results in the correlation to the methyl groups having the highest intensity despite the H7-H8 proximity being nearer. Likely due to how close the H2 correlation peak is to the noise level, the relative intensity compared to H7-H8 is about half that expected, $I_{H7-H2}/I_{H7-H8} = 0.13$ while $r^6_{H7-H8}/r^6_{H7-H2} = 0.33$ for the geometry optimised crystal structure.

^{14}N-^{1}H HMQC MAS NMR spectra (Fig. 4.15) were recorded at 14.1 T for 26AMP$_2$:F:H$_2$O and 26AMP$_2$:F:FA$_{0.5}$ and at 16.44 T for 26AMP:F$_{0.5}$:(H$_2$O)$_2$. Calculated and experimental shifts for the nitrogen environments are recorded in Table 4.6. All of the ^{14}N-^{1}H HMQC spectra confirmed the transfer of protons to the pyridine nitrogens as the corresponding ^{14}N shifts for 26AMP$_2$:F:H$_2$O (Fig. 4.15b) were both −109 ppm (calculated at −73.7 ppm and −74.7 ppm for N1 and N5, respectively, at 14.1 T) and for 26AMP$_2$:F:FA$_{0.5}$ (Fig. 4.15c) were −87 ppm and −91 ppm (calculated at −74.3 ppm and −73.4 ppm for N1 and N5, respectively). For 26AMP:F$_{0.5}$:(H$_2$O)$_2$ (Fig. 4.15a), the experimental ^{14}N shift (recorded

at 16.44 T) was −158 ppm (δ^{calc} = –123 ppm), also confirming proton transfer to the pyridine nitrogen (see extended discussion in section 6.3.1). As the shift range for ^{14}N is so large, the expected error in the calculations is correspondingly greater.

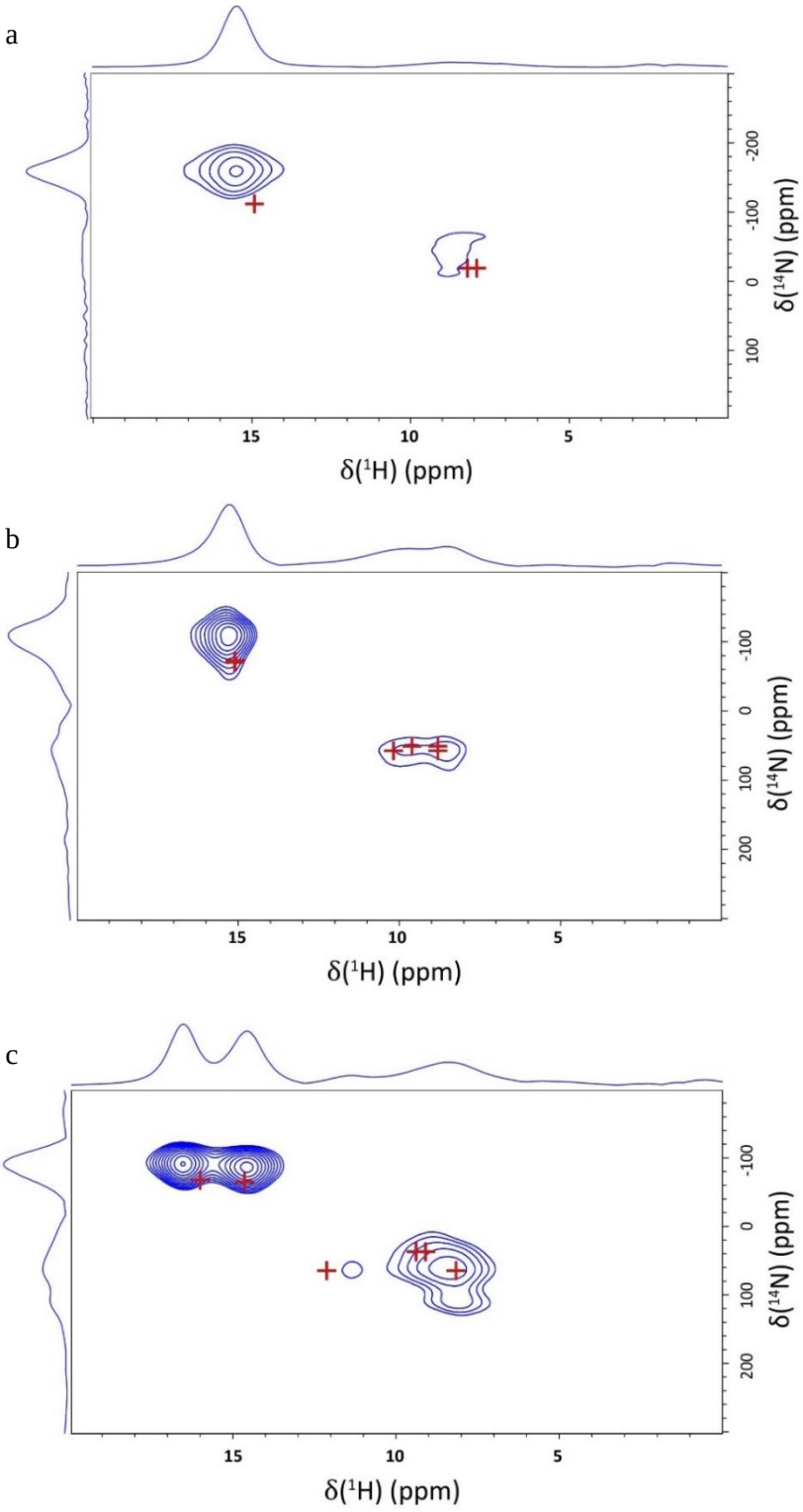


Figure 4.15: ^{14}N-^{1}H HMQC MAS (60 kHz) spectra of (a) 26AMP:$F_{0.5}$:$(H_2O)_2$, (b) $26AMP_2$:F:H_2O and (c) $26AMP_2$:F:$FA_{0.5}$ recorded with 8 rotor periods of R^3 recoupling (τ_{RCPL} = 133.6 μs). Spectra were recorded at ν_0 (^{1}H) = 700 MHz (a) and 600 MHz (b and c).

The NH_2 correlation peaks for each system also show similar agreement and confirm the assignments of each of the NH_2 protons. Their intensities are significantly weaker than those of the NH^+ environments. This is likely due to the difference in coupling between the moieties, as the NH_2 group is significantly affected by H-H couplings. The recoupling within the HMQC experiment was also optimised for the NH^+ interaction and the optimum for the NH_2 groups is shorter, with the magnetisation already decaying by the time it is transferred back to ^{1}H for detection.

Table 4.6: GIPAW calculated ^{15}N chemical shifts, quadrupolar parameters and calculated ^{14}N shifts for each geometry optimised crystal structure, alongside the experimental ^{14}N shifts.

Structure	NH^+	$^{15}N\ \delta_{iso}^{calc}$ (ppm)	η_Q	C_Q (MHz)*	P_Q (MHz)	δ_{iso}^{Q} (ppm)	$^{14}N\ \delta^{calc}$ (ppm)	$^{14}N\ \delta^{expt}$ (ppm)
26AMP:$F_{0.5}$:$(H_2O)_2$	N1	−213.0	0.8	1.7	1.9	105.4	−107.5	−158
	N2	−293.4	0.5	−3.0	−3.1	278.3	−15.2	−40
26AMP_2:F:H_2O	N1	-210.6	0.8	1.7	1.8	136.9	-73.7	−109
	N2	-291.2	0.6	-2.8	-2.9	343.8	52.6	57
	N5	-210.0	0.8	1.7	1.8	135.3	-74.7	−109
	N6	-289.1	0.6	-2.8	-2.9	347.1	58.1	57
26AMP_2:F:$FA_{0.5}$	N1	−207.7	0.7	1.7	1.8	133.4	−74.3	−87
	N2	−286.9	0.6	−2.8	−3.0	353.4	66.5	108
	N5	−211.8	0.9	1.7	1.9	138.4	−73.4	−91
	N6	−291.9	0.7	−2.7	−2.9	329.7	37.8	64

*A C_Q scaling factor of 0.95 is employed.

4.5.3. ^{13}C CP MAS NMR and Multiphase Identification

^{1}H-^{13}C CP MAS NMR spectra also allow one to distinguish between forms, as with the ^{1}H MAS NMR spectra, despite there being less impact relative to the chemical shift range from packing variations and changes in H-bonding (see Table 4.7 and Fig. 4.16). There are also no additional ^{13}C environments – whereas the hydrates contain additional proton environments from the water molecules (compared to the anhydrous form) and the anhydrous form has an OH due to the FA. The difference is due instead to the number of molecules contained in the asymmetric unit. 26AMP:$F_{0.5}$:$(H_2O)_2$ has only a single base and a hemi-fumarate molecule in the asymmetric unit (not counting water molecules) whereas 26AMP_2:F:H_2O and 26AMP_2:F:$FA_{0.5}$ both have two independent base molecules and two independent hemi-fumarate molecules, as well as 26AMP_2:F:$FA_{0.5}$ having a hemi-FA. The narrower lineshapes

Table 4.7: GIPAW calculated and experimental ^{13}C chemical shifts

	26AMP:$F_{0.5}$:$(H_2O)_2$		**$26AMP_2$:F:H_2O**	**$26AMP_2$:F:$FA_{0.5}$**	
	$\delta^{iso}_{calc}(^{13}C)$ (ppm)	**$\delta^{iso}_{exp}(^{13}C)$ (ppm)**	**$\delta^{iso}_{calc}(^{13}C)$ (ppm)**	**$\delta^{iso}_{calc}(^{13}C)$ (ppm)**	**$\delta^{iso}_{exp}(^{13}C)$ (ppm)**
Me	18.3	20.9	14.1	14.5	18.9
	-		15.2	16.9	19.5
CH/C	110.6	110.4	110.4	108.7	109.5
	111.4	110.4	111.0	112.2	112.0
	140.1	137.6	112.4	112.4	112.5
	144.7	145.4	112.7	115.6	114.8
	150.6	149.2	138.6	137.7	135.0
	151.4	155.5	140.8	137.7	135.6
	-		142.6	140.4	142.2
	-		144.3	140.7	137.9
	-		150.2	142.7	143.9
	-		150.8	149.0	147.4
	-		151.5	151.1	148.8
	-		152.8	152.6	156.9
	-		-	153.2	156.9
COOH/$^-$	176.8	174.4	175.3	174.8	170.9
	-		175.4	176.1	174.4
	-		-	177.0	175.3

for ^{13}C mean that, where repeated environments in the forms are present, the small differences between the multiple resonances can often be resolved. As can be seen in Fig. 4.16a, 26AMP:$F_{0.5}$:$(H_2O)_2$ therefore has the simplest spectrum, while there is an evident doubling of most carbon sites in Fig. 14.16b for $26AMP_2$:F:$FA_{0.5}$. Due to the small number of crystals obtained for $26AMP_2$:F:H_2O, there was not enough sample to pack a large enough rotor for ^{13}C experiments, with the sample's long T_1 relaxation time making running them in a 1.3 mm rotor unfeasible. The 1D ^1H-^{13}C CP MAS NMR spectra were assigned as shown in Fig. 4.16 using both the GIPAW calculated shifts and the correlations evident within HETCOR experiments described below.

In a ^1H-^{13}C CP HETCOR MAS NMR spectrum of 26AMP:$F_{0.5}$:$(H_2O)_2$ (Fig. 4.17), the presence of a cross peak between 149.2 ppm and the methyl protons confirms the assignment

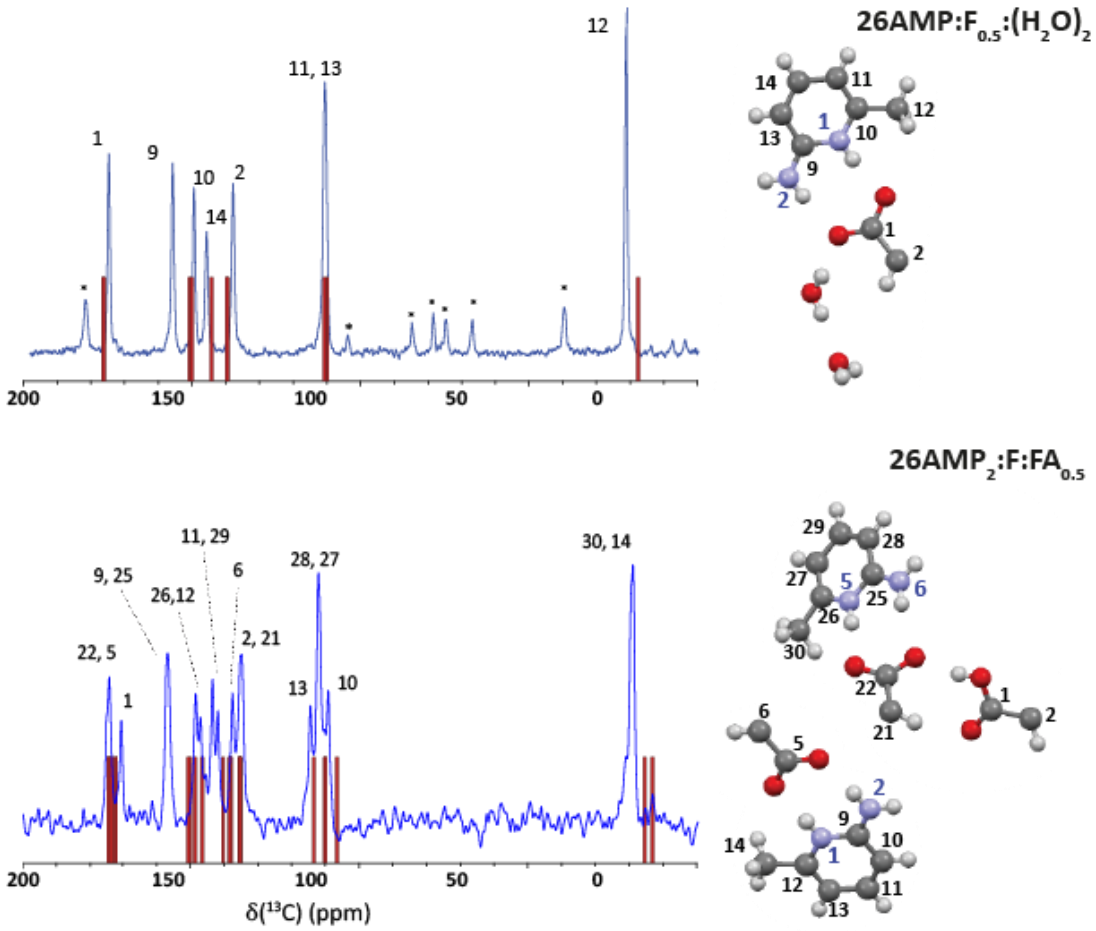


Figure 4.16: ^{1}H-^{13}C CP-MAS (12.5 kHz) spectra of 26AMP:$F_{0.5}$:$(H_2O)_2$ (top), ν_0 (^{1}H) = 600 MHz, and 26AMP$_2$:F:$FA_{0.5}$ (bottom), ν_0 (^{1}H) = 500 MHz, with stick spectra corresponding to GIPAW calculated chemical shifts. The assignments to each proton, labelled in the structures on the right, are given.

of that resonance to C10 (δ_{iso}^{calc} = 150.6 ppm). The peak at 155.5 ppm, as it does not have a close proximity to the methyl, is therefore assigned as C9 (δ_{iso}^{calc} = 151.4 ppm) despite the larger discrepancy between calculation and experiment than is normally expected (~2 ppm for ^{13}C).[202-204] This may be due to the position of C9 within the 2-amino-6-methylpyridinium as it is directly bound to both the pyridine and amino nitrogen's. There appears to be cross peaks between both C12 and C2 and H50, one of the water protons. Although a CP contact time of 500 μs was employed, this seems unusual, particularly for C2 for which the smallest separation within the crystal structure is 3.95 Å. The fact that no cross peak is evident for the closer C11 to H13 proximity (3.35 Å) highlights the range in efficiency of the CP transfer for different environments in spite of the use of a ramped contact time. The correlation between C1 and the methyl protons (shorter still at 2.96 Å) is also barely visible above the noise.

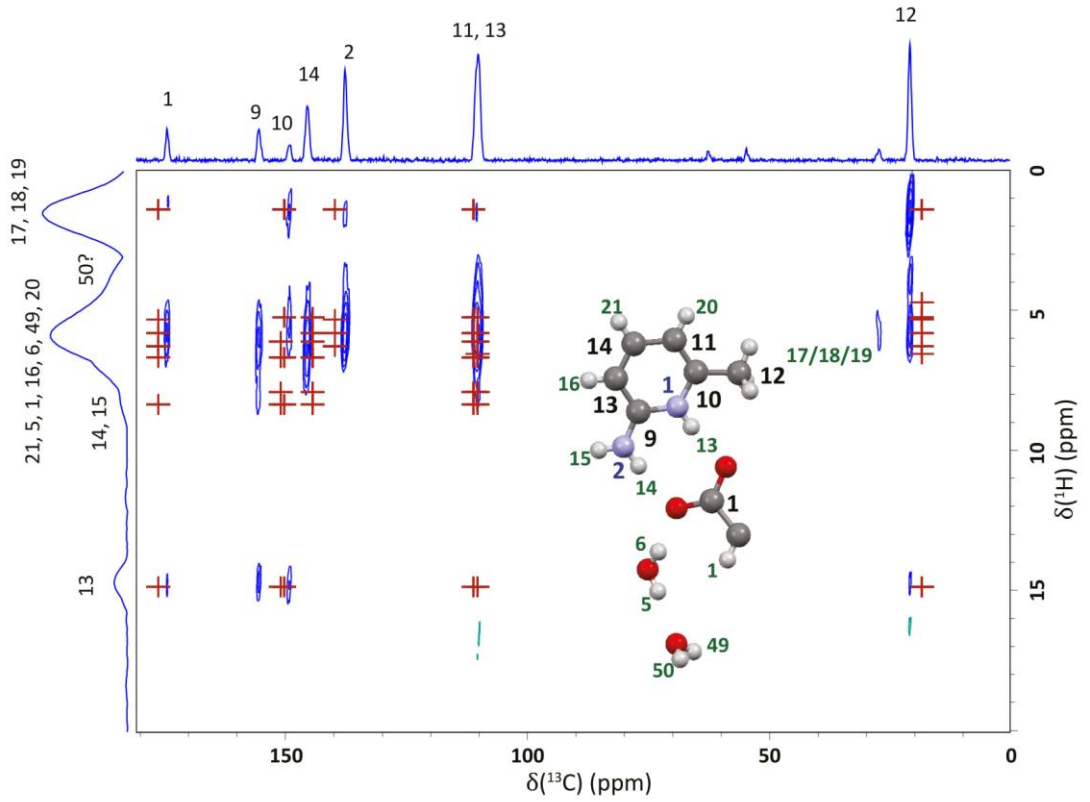


Figure 4.17: A ^{1}H (600 MHz)-^{13}C CP (500 μs) HETCOR MAS (12.5 kHz) NMR spectrum of 26AMP:$F_{0.5}$:$(H_2O)_2$ recorded using FSLG ^{1}H homonuclear decoupling in t_1 with calculated (GIPAW) chemical shifts shown as red crosses out to a maximum C···H distance of 3.5 Å. The base contour level is at 7.8% of the maximum peak height.

As all the ^{13}C assignments could be confidently confirmed, only one HETCOR experiment was conducted for 26AMP:$F_{0.5}$:$(H_2O)_2$. To allow a clearer understanding of the proximities within $26AMP_2$:F:$FA_{0.5}$, however, an experiment with a shorter contact time of 200 μs, was also recorded so as to remove cross peaks for the more distant correlations (Fig. 4.18a). The only resonance that does not seem to relate to a one bond correlation that is visible in Fig. 4.18a experiment is for the close C1-H1 proximity (carboxyl carbon to OH).

The longer range ^{1}H-^{13}C CP HETCOR MAS NMR spectrum (Fig. 4.18b), recorded with a 500 μs CP contact time, suggests that the ^{13}C resonance at 137.9 ppm should be assigned to C6 rather than C29, which was calculated at a slightly lower value, due both to its correlation with a proton at a lower CH chemical shift of δ_{iso}^{exp} = 4.9 ppm as expected for the C6 correlation to H5 (δ_{iso}^{calc} = 4.9 ppm and seen in the shorter range HETCOR spectrum) and also a cross peak with a methyl proton. C6's closest proximity to a methyl proton is 3.81 Å, whereas for C29 it is more than 4 Å. C29 has therefore been assigned to the resonance at 142.2

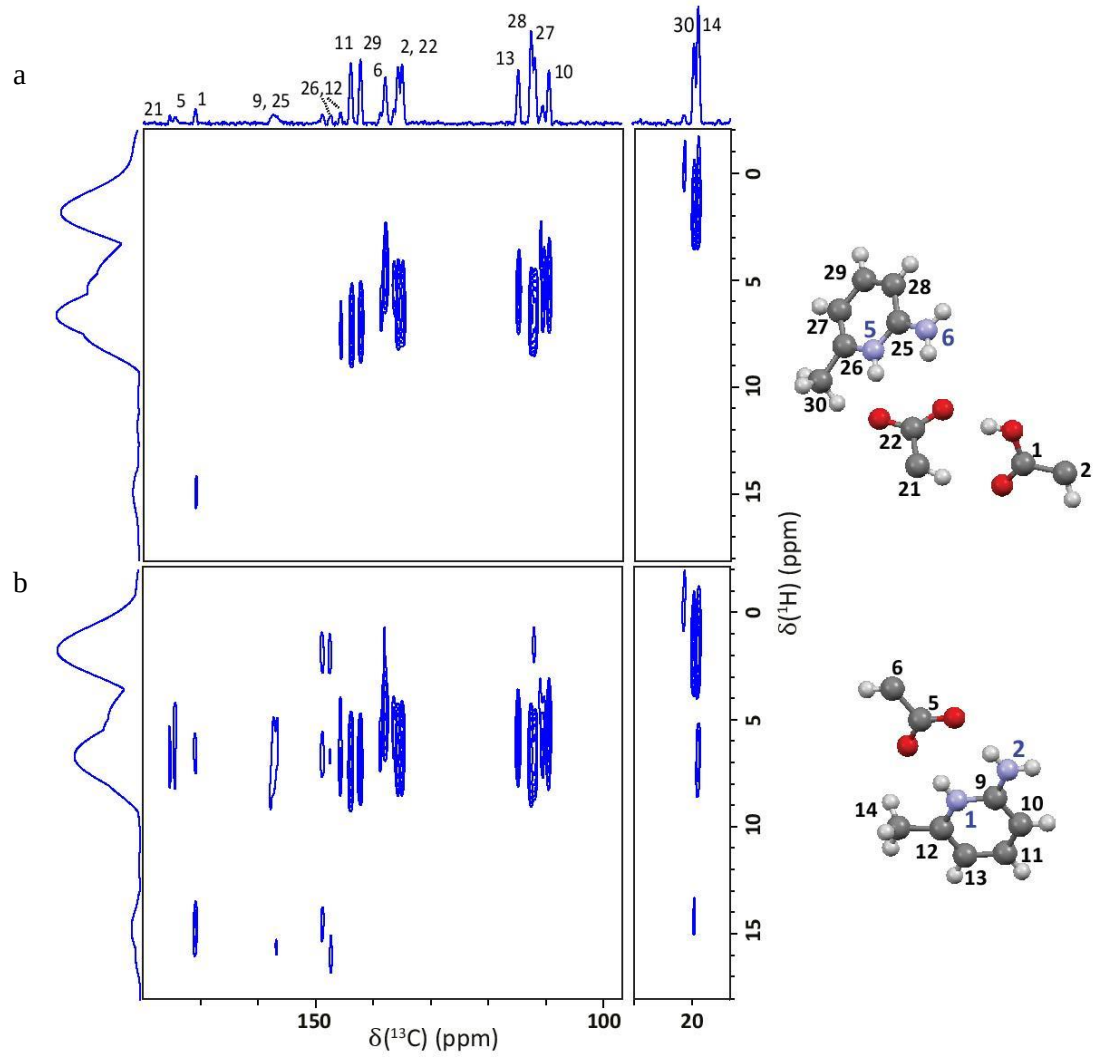


Figure 4.18: ^{1}H (600 MHz)-^{13}C HETCOR MAS (12.5 kHz) NMR spectra of 26AMP$_2$:F:FA$_{0.5}$ recorded using FSLG ^{1}H homonuclear decoupling in t_1 and a CP transfer duration of (a) 200 μs and (b) 500 μs. The base contour level is at 6.4% of the maximum peak height.

ppm and C11, its equivalent carbon on the other base molecule, to 143.9 ppm. These two carbons are assigned in this order on the basis of the GIPAW calculations as the similarity in their environments and correlations means that neither has a distinct distinguishing CH correlation. This is also the case for C22 and C2, the fumarate CH carbons, which appear as a single peak in the 1D spectrum and are assigned to resonances at 135.0 ppm and 135.7 ppm in the HETCOR spectrum, with both calculated at ^{13}C δ_{iso}^{calc} = 137.7 ppm and only 0.1 ppm difference between the directly bound ^{1}H calculated chemical shifts.

C12 and C26 are confidently assigned to 147.3 ppm and 148.8 ppm, respectively, due to the presence of cross peaks in the long range HETCOR spectrum, with the CH and methyl ^{1}H regions and to their respective nearby NH, and the absence of any cross peaks for a 200 μs

CP contact time, as they are both quaternary carbons. Their relative assignment was confirmed by the difference in the ^{1}H chemical shift of their respective NH protons, with C12 (δ_{iso}^{exp} = 147.3 ppm) correlating with H7 at 16.0 ppm and C26 (δ_{iso}^{exp} = 148.8 ppm) correlating with H27 at 14.6 ppm. C9 and C25, the other quaternary carbons which sit between the pyridine and amide nitrogens, have been assigned on the same basis, although the cross peak between C25 and H27 does not appear. As they are also low intensity, it may be that it cannot be seen above the noise level, with the expected correlations for both C9 and C25 to NH_2 protons missing for the same reason (although the CH cross peak for C9 does extend to higher chemical shift towards the region where the NH_2 protons fall). As seen for the corresponding carbon in 26AMP:$F_{0.5}$:$(H_2O)_2$, C9 and C25 also show the largest discrepancy between experiment and calculation, with the calculation determining them to be at a lower chemical shift (δ_{iso}^{calc} = 152.6 ppm and 153.2 ppm compared to δ_{iso}^{exp} = 156.7 ppm and 157.4 ppm, respectively). The similarity in discrepancy between the two systems (~ 4 ppm) for carbons of the same chemical environment suggests that this is a systematic error in the calculation (see further discussion in section 6.3.2).

Although both methyl carbons, C14 and C30, are expected to have long-range correlations to the nearest aromatic CH proton and to their respective NH protons, however, only one out of the two is observed in each case. C14, δ_{iso}^{exp} = 18.9 ppm, has a cross peak with the CH region at δ_{iso}^{exp} = 6.4 ppm and C30, δ_{iso}^{exp} = 19.6 ppm, has a cross peak with H27 at δ_{iso}^{exp} = 14.2 ppm. As the proximities are very similar, as they are all intra-molecular, there is no reason they should not have cross peaks to both. As those present are low intensity, the other cross peaks may not differ greatly and be just below the noise level and therefore not seen, as the result of subtle differences in local structure and consequent small variations in separation and/or CP efficiency.

Although a Rietveld refinement of PXRD data of the sample was initially taken to show acceptable agreement to $26AMP_2$:F:$FA_{0.5}$, with an R_{Bragg} of 5.7, the HETCOR MAS NMR spectra suggested that the bulk of the $26AMP_2$:F:$FA_{0.5}$ sample may be a mixture also containing $26AMP_2$:F:H_2O and 26AMP:$F_{0.5}$:$(H_2O)_2$. There are four low intensity peaks at

110.6 ppm, 136.4 ppm, 138.7 ppm and 145.7 ppm (the sample packed in a 1.3 mm rotor from a few large crystals appears to be single phase, with additional peaks only present in the larger volume sample packed in a 3.2 mm rotor for ^{13}C experiments). As they are present in both the

Table 4.8:Assignment of CH correlation peaks in the ^{1}H-^{13}C HETCOR spectra of 26AMP$_2$:F:FA$_{0.5}$ (see Fig. 4.15).

C	δ_{iso}^{exp} (ppm)	H*	δ_{iso}^{exp} (ppm)
C14	18.9	*H13/H14/H15 (CH_3)*	*1.7*
		H12 (CH)	6.4
C30	19.6	*H33/H34/H35 (CH_3)*	*1.9*
		H27 (NH)	14.2
C10	109.5	*H10 (CH)*	*5.6*
?	110.6	*?*	*5.8*
C27	111.9	H33/H34/H35 (CH_3)	1.6
		H30 (CH)	*6.7*
C28	112.5	*H31 (CH)*	*6.7*
C13	114.8	*H12 (CH)*	*5.8*
C2	135.0	*H2 (CH)*	*6.4*
C22	135.7	*H25 (CH)*	*6.4*
?	136.4	*?*	*5.4*
C6	137.9	H13/H14/H15 (CH_3)	1.2
		H5 (CH)	*4.9*
?	138.7	*?*	*6.2*
C29	142.2	*H32 (CH)*	*7.0*
C11	143.9	*H11 (CH)*	*7.2*
?	145.7	*?*	*7.3*
C12	147.3	H13/H14/H15 (CH_3)	2.0
		H12 (CH)	6.7
		H7 (NH)	16.0
C26	148.8	H33/H34/H35 (CH_3)	2.0
		H30 (CH)	6.6
		H27 (NH)	14.6
C9	156.7	H10 (CH)	6.1
		H7 (NH)	15.7
C25	157.4	H31 (CH)	7.3
C1	170.9	H2 (CH)	6.6
		H1 (OH)	14.8
C5	174.4	H5 (CH)	5.4
C21	175.4	H25 (CH)	6.7

*^{1}H correlations written in italics correspond to those seen in both the short and long range experiments

Table 4.9: Candidate correlations from 26AMP$_2$:F:H_2O and 26AMP:$F_{0.5}$:$(H_2O)_2$ for the unassigned cross peaks in the ^{1}H-^{13}C HETCOR experiments of 26AMP$_2$:F:FA$_{0.5}$

Structure	C	δ_{iso}^{calc} (ppm)	H	δ_{iso}^{calc} (ppm)
26AMP$_2$:F:H_2O	C23	110.4	H31	5.4
	C6	138.6	H3	5.3
	C26	144.3	H35	7.4
	C11/C13	110.6/111.4 (110.1/110.5)*	H16/H20	6.1/5.3
26AMP:$F_{0.5}$:$(H_2O)_2$	C2	140.1 (137.6)	H1	6.3
	C14	144. 7 (145.4)	H21	6.7

*^{13}C chemical shifts stated in brackets correspond to the experimentally determined values. Due to the peak overlap within the ^{1}H dimension, only the calculated chemical shifts are given.

short and long-range HETCOR experiments and only correlate with the CH region, they are assumed to be CH environments. None are discernible in the 1D ^{1}H-^{13}C CP MAS NMR spectrum due to their proximity to resonances from $26AMP_2$:F:$FA_{0.5}$ and the resultant peak overlap. Where they fall close to assigned cross peaks which also only correlate with the CH proton region, the judgement on which to assign to $26AMP_2$:F:$FA_{0.5}$ carbons was made based on the relative intensities, with the lower intensity peaks judged less likely to be a directly bound carbon within the majority phase. Final assignments are shown in Table 4.8.

By comparing the positions of the unassigned cross peaks to those expected and/or seen for the two hydrate structures, a set of possible matches was constructed (Table 4.9). The resonance at ^{13}C 110.6 ppm may correspond to either 26AMP:$F_{0.5}$:$(H_2O)_2$ C13 - as although it falls at the calculated value of C11 (δ_{iso}^{calc} = 110.6 ppm), the experimental values for C13 are closer (δ_{iso}^{exp} = 110.5 ppm and δ_{iso}^{calc} = 5.3 ppm, for C13 and its CH proton, H20, respectively). Alternatively, it may be $26AMP_2$:F:H_2O C23 which shows equally good agreement although there is no experimentally determined confirmation of the chemical shift. As the remaining three peaks cannot be assigned to only one of the hydrate structures but only to a combination, resonances for both likely lie underneath. This would also explain why the δ_{iso}^{exp} (^{13}C) = 110.5 ppm/δ_{iso}^{calc} (^{1}H) = 5.3 ppm peak is the highest in intensity of the four non-$26AMP_2$:F:$FA_{0.5}$ peaks. The resonances at ^{13}C 136.4 ppm and 138.7 ppm are thought to correspond to 26AMP:$F_{0.5}$:$(H_2O)_2$ C2 (δ_{iso}^{exp} = 137.6 ppm and δ_{iso}^{calc} = 6.3 ppm) and $26AMP_2$:F:H_2O C6 (δ_{iso}^{calc} = 138.6 ppm and δ_{iso}^{calc} = 5.3 ppm), respectively. On the basis of the ^{1}H chemical shift, the remaining cross peak, at ^{13}C 145.7 ppm and ^{1}H 7.3 ppm, is judged more likely to be $26AMP_2$:F:H_2O C26 (δ_{iso}^{calc} = 144.3 ppm and δ_{iso}^{calc} = 7.4 ppm) than 26AMP:$F_{0.5}$:$(H_2O)_2$ (δ_{iso}^{exp} = 145.4 ppm and δ_{iso}^{calc} = 6.7 ppm) due to smaller absolute expected discrepancies between calculation and experiment for ^{1}H, due to having a smaller chemical shift range. A

multiphase Rietveld refinement (Fig. 4.19) against all three structures improved the R_{wp} to 9.3% and the R_{Bragg} values to 1.5%, 1.8% and 3.7% for $26AMP_2$:F:$FA_{0.5}$, $26AMP_2$:F:H_2O and 26AMP:$F_{0.5}$:$(H_2O)_2$ (with a phase content of 74:17.5:8.5), respectively. A Pawley refinement against the three phases had produced a R_{wp} of 7.3%.

A single-phase large-scale crystallisation of $26AMP_2$:F:$FA_{0.5}$ has not yet been achieved and, as the crystals cannot be phase separated by eye, it has not been possible to obtain a single phase HETCOR spectrum.

4.5.4. Isolated Molecule GIPAW Chemical Shift Calculations

In the isolated molecule calculations, it can clearly be seen in Table 4.10 that the largest $\Delta\delta_{Crystal\ -\ Molecule}$, and therefore potentially the strongest H-bond, corresponds to the OH···O^- interaction in $26AMP_2$:F:$FA_{0.5}$. This is the only non-water OH···O interaction exhibited. Its apparent strength must be treated with some caution as the experimental shift was 1.4 ppm lower than calculated, however, as $\Delta\delta_{Crystal\ -\ Molecule}$ = 9.1 ppm, even if it is reduced by the experimental discrepancy this will still be the largest change for any system. The next biggest

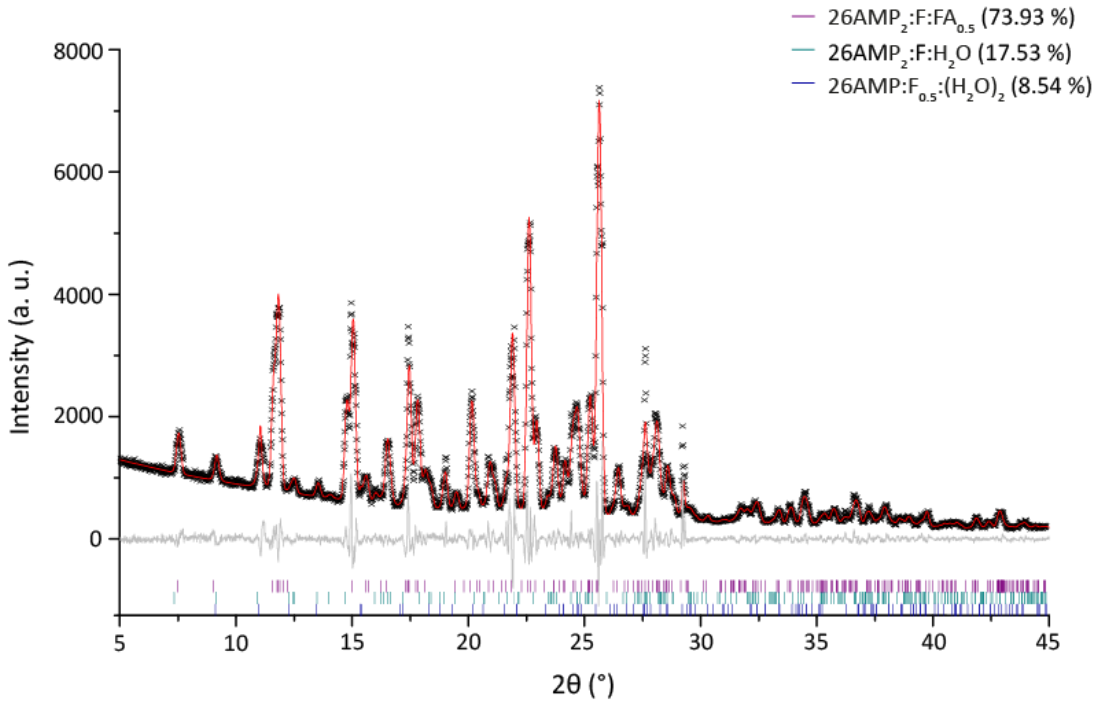


Figure 4.19: Final multiphase Rietveld fit for $26AMP_2$:F:$FA_{0.5}$, showing the experimental (black crosses), calculated (red upper line) and difference (grey lower line) PXRD profiles. Tick marks indicate peak positions. For refinement parameters, see Table A2.2 in Appendix 2.

Table 4.10: A comparison of GIPAW calculated ^{1}H shifts (in ppm) for the full geometry optimised crystal structures of 26AMP:$F_{0.5}$:$(H_2O)_2$, 26AMP$_2$:F:H_2O and 26AMP$_2$:F:$FA_{0.5}$ and for isolated molecules extracted from these crystal structures.

Structure	Atom	δ_{Expt}	$\delta_{Crystal}$	$\delta_{Molecule}$	$\Delta\delta_{Crystal\ -\ Molecule}$
26AMP:$F_{0.5}$:$(H_2O)_2$	**H20**	5.3	5.3	5.9	–0.6
	H21	6.8	6.7	6.9	–0.2
	H17/H18/H19	1.5	1.5	1.5	0.0
	H16	6.0	6.1	5.7	0.4
	H1	6.0	6.3	5.8	0.5
	H15	8.2	7.9	5.0	2.9
	H14	8.2	8.3	4.7	3.6
	H50	4.2	4.7	–0.5	5.2
	H49	6.0	5.3	–0.4	5.8
	H6	6.0	5.8	–0.2	6.0
	H13	14.9	14.8	8.3	6.5
	H5	6.8	6.5	–0.2	6.7
26AMP$_2$:F:H_2O	**H9/H10/H11**	1.5	1.4	1.5	–0.1
	H3	5.7	5.3	5.8	–0.5
	H31	5.7	5.4	5.8	–0.4
	H14	5.7	5.5	5.7	–0.2
	H27/H28/H29	1.5	1.4	1.6	–0.2
	H13	7.4	7.0	6.9	0.1
	H1	5.7	5.9	5.8	0.1
	H16	6.3	6.3	5.9	0.4
	H35	7.4	7.4	6.9	0.4
	H33	6.3	6.4	5.9	0.5
	H34	9.2	8.7	5.0	3.7
	H17	9.2	8.7	4.9	3.7
	H5	4.1	4.2	–0.6	4.8
	H15	9.8	9.5	4.6	4.8
	H6	4.1	4.5	–0.6	5.1
	H32	9.8	10.5	4.5	6.0
	H12	14.8	15.0	8.3	6.6
	H30	14.8	15.0	8.3	6.6
26AMP$_2$:F:$FA_{0.5}$	**H5**	5.1	4.9	5.8	–0.9
	H10	5.1	4.9	5.8	–0.9
	H13/H14/H15	1.8	1.0	1.6	–0.6
	H12	5.8	5.7	5.9	–0.2
	H33/H34/H35	1.8	1.6	1.6	–0.1
	H32	7.7	7.1	7.0	0.1
	H31	6.7	6.2	5.8	0.4
	H11	7.7	7.5	7.1	0.4
	H2	6.7	6.6	6.0	0.6
	H30	6.7	6.7	6.0	0.7
	H25	6.7	6.7	5.8	0.9
	H9	7.7	8.1	5.0	3.0
	H29	9.1	9.3	5.1	4.3
	H28	9.1	9.0	4.6	4.4
	H27	14.4	14.7	8.3	6.4
	H8	11.4	12.1	4.9	7.2
	H7	16.4	16.0	8.5	7.5
	H1	14.7	16.3	7.2	9.1

$\Delta\delta_{Crystal\ -\ Molecule}$ are for the NH^+ and NH_2 protons involved in the H-bonds that form the independent b-a-b unit within the same system. The b-a-b unit linked into chains by FA has a less planar geometry and less ideal H-bond distances and angles (as shown in Table 4.2, above). This is reflected in a slightly lower $\Delta\delta_{Crystal\ -\ Molecule}$ (6.4 ppm rather than 7.5 ppm for the NH^+). The corresponding interactions in the b-a-b units of 26AMP:$F_{0.5}$:$(H_2O)_2$ and 26AMP_2:F:H_2O are comparable to those for the linked 26AMP_2:F:$FA_{0.5}$ unit.

One of the water molecules in 26AMP:$F_{0.5}$:$(H_2O)_2$ exhibits similarly strong H-bonds to the NH^+ protons, while the other appears to be slightly less tightly bound ($\Delta\delta_{Crystal\ -\ Molecule}$ = 6.7/6.0 ppm and $\Delta\delta_{Crystal\ -\ Molecule}$ = 5.8/5.2 ppm for H5/6 and H49/50, respectively). Although all the water H-bond angles are close to ideal (all > 165°), the O···O distances are shorter for H5 and H6 than for H49 and H50, also suggesting a corresponding difference in the interaction strength. It is suggested that this difference in how tightly bound the water molecules are is in part responsible for the ability of the system to transform into 26AMP_2:F:H_2O on initial dehydration rather than either collapsing or going straight to 26AMP_2:F:$FA_{0.5}$. The water molecule of AF26-0.5 is comparable to the weaker 26AMP:$F_{0.5}$:$(H_2O)_2$ water interaction with $\Delta\delta_{Crystal\ -\ Molecule}$ = 5.1 and 4.8 ppm. The less firmly bound water may account for the relative instability of the system however, making it more a transitional state, explaining why both the dihydrate and anhydrous forms are so much more readily crystallised (only ~5 small crystals were ever successfully isolated).

It is interesting to note that all the interactions identified are H-bonds between classical donors and acceptors. The only possible non-classical H-bond ($\Delta\delta_{Crystal\ -\ Molecule}$ = 0.9 ppm and the cut-off for significance is generally considered ±1 ppm) is in 26AMP_2:F:$FA_{0.5}$ between the fumarate CH and the carboxylic acid of the linking FA molecule. The only values that may correspond to π interactions are $\Delta\delta_{Crystal\ -\ Molecule}$ = −0.9 ppm for H5 and H10 in 26AMP_2:F:$FA_{0.5}$ but this may be chance as they are changes of less than 1 ppm and there is no clear indication in the structure as to what would cause an interaction.

4.5.5. VT NMR

As hydrates generally have relatively low stability, all MAS NMR spectra of $26AMP_2$:F:H_2O and 26AMP:$F_{0.5}$:$(H_2O)_2$ were run with cooling applied to prevent the sample heating above 5 °C. When this heating is removed, the sample warms to ~70 °C under fast MAS (60 kHz). As can be seen in Fig. 4.20, the ^{1}H spectrum of 26AMP:$F_{0.5}$:$(H_2O)_2$ changes, with new resonances appearing as shoulders at 15.9 ppm, 9.6 ppm and 7.1 ppm. The new peak at 15.9 ppm is of particular interest as the corresponding hydrate resonance, at 14.9 ppm, is for H13 which is integral to the b-a-b unit and has no direct interaction with the water molecules. H14 is also shifted to a higher ppm which suggests strengthening of the two H-bonds forming the b-a-b unit whereas the new resonance at 7.1 ppm may represent a change to a lower ppm value for the other NH_2 proton, H15, due to weakening of its H-bond to a water molecule.

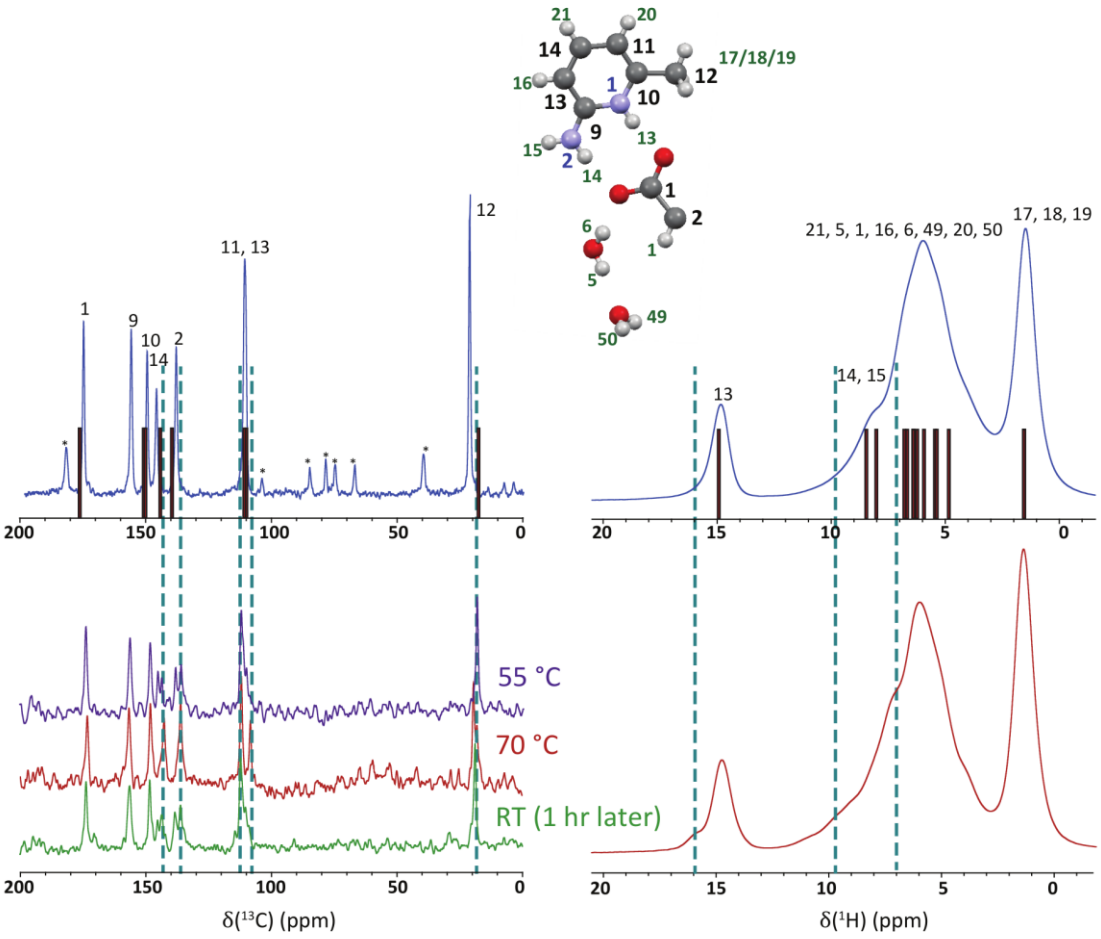


Figure 4.20: ^{1}H (600 MHz)-^{13}C CP-MAS (12.5 kHz) NMR spectra (left) and ^{1}H (600 MHz) one-pulse MAS (60 kHz) NMR spectra (right) of 26AMP:$F_{0.5}$:$(H_2O)_2$ before (top) and after (bottom) heating to 70 °C, with stick spectra corresponding to GIPAW calculated chemical shifts.

Heating of 26AMP:$F_{0.5}$:$(H_2O)_2$ when under 12.5 kHz MAS allowed the same temperature range to be analysed for ^{13}C resonances. The ^{1}H-^{13}C CP-MAS spectrum, recorded at 55 °C, shows the reduction of resonances corresponding to the fully hydrated structure as new resonances grow. By 70 °C, the initial resonances have disappeared completely. There also seems to be evidence of an amorphous component identified from both broadening of the resonances at the bases of the peaks and the appearance of a broad, low intensity peak between 50 and 100 ppm, clearly visible at 70 °C (Fig. 4.20). Conversion appears to be far more complete than in the ^{1}H MAS spectrum at 60 kHz, likely due to the 1.3 mm rotor used for fast MAS being more tightly packed and sealed than the 3.2 mm rotor used for the ^{13}C experiments, which was only half full. The initial state was partially recovered after the sample had sat at room temperature for one hour. The reversibility of this change suggests a change in arrangement/position of the water molecules within the channels but is unlikely to be complete dissociation of the water, which is tightly bound and integral to the structure (there is no direct join between adjacent b-a-b units). It may be the beginning of transition to $26AMP_2$:F:H_2O as there is no evidence that this is not reversible provided water is still available. The new peaks in the ^{1}H spectrum definitely do not correspond to a completed transition to $26AMP_2$:F:H_2O as the NH peaks would not be resolvable from the 26AMP:$F_{0.5}$:$(H_2O)_2$ NH peak, with their δ_{iso}^{exp} = 14.8 ppm compared to the new resonance seen at δ_{iso}^{exp} = 15.8 ppm.

DFT calculations have been run with water removed and for a high temperature (340 K) SXRD structure. Following variable unit cell geometry optimisations, the ‘dehydrated’ structure showed no increase in the atomic forces relative to the standard optimisations and the general structure and packing of b-a-b units was maintained despite the lack of strong interactions to stabilise their packing (as the calculations are at a nominal 0 K, this is not particularly surprising). GIPAW calculated chemical shifts of the ‘dehydrated’ structure showed significant changes for the NH^+ and NH_2 protons (Table 4.11). H14 and H13, integral to the b-a-b unit, both increased as the loss of the fumarate interaction with water allowed more ideal alignment between the fumarate and pyridine group resulting in stronger H-bonding. Conversely, the H15 ^{1}H chemical shift decreased by 5.4 ppm as its H-bond to a water

molecule was completely removed. The calculated changes are far more significant than those observed experimentally but the direction in which the chemical shifts change is the same. For the calculations conducted on the original SXRD crystal structures (following fixed unit cell geometry optimisation), there was no significant difference seen for any of protons between 150 K and 340 K. In combination, these results support the idea that the H-bonds to the water molecules have weakened due to a shift in their position relative to the b-a-b units, allowing the b-a-b interactions to strengthen.

Although no changes to the ^{1}H NMR changes were observed for 26AMP$_2$:F:H$_2$O on removal of cooling, a complementary set of calculations were performed as for 26AMP:F$_{0.5}$:(H$_2$O)$_2$ (Table A2.2 in Appendix 2). Again, there was no change due to thermal expansion between 150 K and 340 K but the 'dehydrated' structure did show some. These were, however, less significant than seen for 26AMP:F$_{0.5}$:(H$_2$O)$_2$, with only one NH$^+$, H30, showing any change and only increasing by half as much as 26AMP:F$_{0.5}$:(H$_2$O)$_2$ H13 (1.6 ppm compared to 3.0 ppm). This may explain why no effect was observed experimentally under the breadth of the ^{1}H MAS NMR linewidth.

Table 4.11: A comparison of GIPAW calculated ^{1}H shifts (in ppm) for the geometry optimised (with fixed unit cells) crystal structures of 26AMP:F$_{0.5}$:(H$_2$O)$_2$ that were recorded at a range of temperatures, alongside a calculation with the water molecules removed.

Atom	$\delta_{\text{LT Crystal}}$ (150 K)	δ_{Crystal} (250 K)	$\delta_{\text{RT Crystal}}$ (290 K)	$\delta_{\text{HT Crystal}}$ (340 K)	$\delta_{\text{Dehydrated}}$ (250 K)	$\Delta\delta_{\text{Crystal – Dehydrated}}$
H17/H18/H19	1.5	1.5	1.4	1.4	1.4	0.1
H50	5.1	4.7	4.8	5.1	-	-
H20	5.3	5.3	5.2	5.2	4.9	0.4
H49	5.5	5.3	5.3	5.4	-	-
H6	6.0	5.8	5.9	5.6	-	-
H16	6.2	6.1	6.1	6.1	5.0	1.1
H1	6.3	6.3	6.3	6.2	5.9	0.3
H5	6.7	6.5	6.5	6.6	-	-
H21	6.9	6.7	6.6	7.0	7.1	–0.4
H15	8.0	7.9	7.9	7.8	2.5	5.4
H14	8.1	8.3	8.0	8.2	11.6	–3.3
H13	14.8	14.8	14.8	14.8	17.8	–3.0

4.6. Summary

These systems show the sensitivity of crystal structure packing to very small variations in the crystallisation conditions. Although the b-a-b unit forms in all three, the subsequent H-bonding of these units into chains alters the final packing of the structures significantly. The b-a-b units are analogous to those reported by Haynes et al. for lutidine succinic/FA cocrystals as well as several lutidinium succinate/fumarate salts.[140] It is also exhibited by 2-amino-5-methylpyridinium fumarate FA,[141] a cocrystal of a salt system showing a similar chain formation of b-a-b units linked by FA molecules (discussed further in Chapter 5) as seen in 2-amino-6-methylpyridinium fumarate FA.

That the change of crystallisation solvent results in a difference in the packing of the b-a-b units, but not in their formation, suggests that this unit is very strongly favoured and forms early in the crystallisation process. The relative size of the solvents used may affect the ease with which these units can align to form higher structures. Of the three solvents used here (Fig. 4.21), only methanol showed no formation of $26AMP_2{:}F{:}FA_{0.5}$ in the absence of additional water. Methanol is smaller than both ethanol and isopropanol and may therefore allow subsequent units to come closer together before being displaced. Stabilisation of the units is facilitated by the residual water, forming close packed chains. Correspondingly, the slightly larger solvent molecules prevent units coming too close before a FA molecule can bridge the gap, reducing the need for water stabilisation of the chain. None of the solvents were completely selective to a single crystal structure although increasing the hydration level did seem to force the formation of $26AMP{:}F_{0.5}{:}(H_2O)_2$ with several of the prepared samples believed to be phase pure. However, one of these raised hydration set-ups produced only microcrystals of $26AMP{:}F_{0.5}{:}(H_2O)_2$ and several large crystals of $26AMP_2{:}F{:}H_2O$. The crystallisation parameter responsible for this variation has yet to be identified and $26AMP_2{:}F{:}H_2O$ remains difficult to produce.

All three of the structures discussed are related, lying on the same transformation pathway (Figure 4.21). Within the hydrate region, these transformations may be reversible

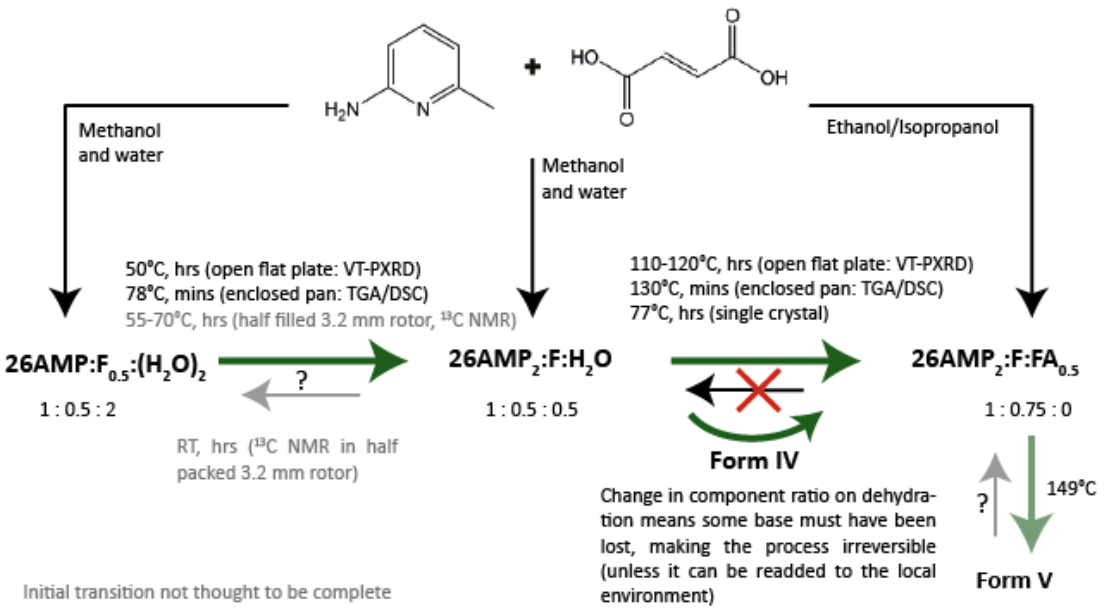


Figure 4.21: Transformation diagram for the three multicomponent forms determined for 26AMP and FA.

but, as $26AMP_2$:F:$FA_{0.5}$ contains a different acid : base ratio, the transition to anhydrous forms is believed to be irreversible. This $26AMP_2$:F:H_2O to $26AMP_2$:F:$FA_{0.5}$ transformation is assumed to be possible due to loss of 26AMP, similar to that seen for 26L in chapter 3,[201] but only occurring at higher temperatures due to the higher relative stability of the base molecule. VT-PXRD and DSC results suggest that this second transition can occur both directly and also via formation of an unknown structure, Form IV, as well as providing evidence for a third transition region to another crystalline solid form, accessed from $26AMP_2$:F:$FA_{0.5}$ at ~150 K, resulting in Form V, the structure of which is similarly unknown.

VT-NMR of 26AMP:$F_{0.5}$:$(H_2O)_2$ saw the appearance of new ^1H and ^{13}C resonances on heating up to 70 °C. For the ^{13}C CP MAS experiments, performed in a partially filled 3.2 mm rotor, a corresponding reduction in the original resonances was seen with their complete disappearance by 70 °C (Fig. 4.20). The reversal of this conversion following the return to room temperature suggests a change in the H-bonding of the water molecules, possibly an intermediate state in the transition to $26AMP_2$:F:H_2O, with partial but incomplete disassociation of at least one water molecule. Complete disassociation is not expected both due to the importance of the water mediated H-bonds to maintain the overall structure and as the changes in ^1H chemical shift are not as extreme as expected from GIPAW calculations for

the same structure following complete removal of the water molecules (section 4.5.5, Fig. 4.20, Table 4.11). The use of boron nitrate rotor caps would allow higher temperatures to be accessed in NMR experiments, possibly allowing one or more of the transitions to be monitored.

Good agreement was seen between the GIPAW calculated shifts of all three structures and the experimental NMR parameters. Two seemingly systematic discrepancies were, however, noted. The first corresponds to the ^{13}C GIPAW calculated chemical shift of the carbon at the 2 position on the pyridine ring, which sits between two nitrogen's, and varied from the experimental resonances by ~4 ppm for both of the 26AMP cations in $26AMP_2{:}F{:}FA_{0.5}$ and in the single example in $26AMP{:}F_{0.5}{:}(H_2O)_2$ (the lack of ^{13}C NMR experiments for $26AMP_2{:}F{:}H_2O$ means that it cannot be confirmed in this case). The second discrepancy seen was for the OH proton of the hemi-fumarate in $26AMP_2{:}F{:}FA_{0.5}$, which exhibited a similar variation to that seen for the OH of the hydrogen fumarate anion in 2,6-lutidinium fumarate, 1.4 ppm.

Even taking the aforementioned error into account, the isolated molecule GIPAW calculations (Table 4.10) suggest that the $OH\cdots O^-$ interaction between the fumarate and FA in $26AMP_2{:}F{:}FA_{0.5}$ is the strongest interaction, followed by the amide/pyridinium-carboxylate interaction that forms the independent b-a-b unit in the same system. This is unsurprising as this b-a-b unit exhibits the most ideal H-bond geometry. The larger $\Delta\delta_{Crystal\ -\ Molecule}$ values seen for both protons of one of $26AMP{:}F_{0.5}{:}(H_2O)_2$'s water molecules compared to those of the second water molecule suggest that the former is more tightly bound. This difference in H-bond strength may be why $26AMP{:}F_{0.5}{:}(H_2O)_2$ dehydrates first to a second stable hydrate, $26AMP_2{:}F{:}H_2O$, before dehydrating to the anhydrous form, $26AMP_2{:}F{:}FA_{0.5}$. The higher scattering power of protons within neutron diffraction, compared to XRD, means that tracking the transitions by neutron diffraction may provide additional insight, although deuterated samples would likely be required for powder analysis as the large inelastic scattering of the hydrogen atoms can produce a high background.[113]

Chapter 5. 5-Amino-2-Methylpyridine and Related Systems

5.1. Introduction

Replacement of a single functional group or changing its position of substitution can have a significant impact on crystallisation. 25L and 25AMP are the 2,5 substitutional equivalents of the base molecules discussed in Chapters 3 and 4, 26L and 26AMP. 25L and 25AMP differ from the two previously discussed bases by a move of the 6 position methyl group to the 5 position. As between 26L and 26AMP, these two 2,5 bases differ from each other by the replacement of the 2 position methyl with an amine (Fig. 5.1). Both 25L and 25AMP have multicomponent structures with FA that have already been published (Fig. 5.2, Table 5.1).[140, 141]

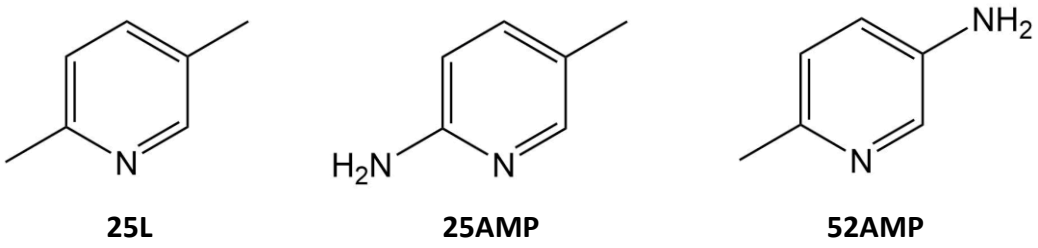


Figure 5.1: The difference between 25L (left), 25AMP (centre) and 52AMP (right), consisting of replacement of a methyl group with an amino and reversal of the substitution.

2-Amino-5-methylpyridinium hemi-fumarate hemi-fumaric acid (25AMP:$F_{0.5}$:$FA_{0.5}$) and 2,5-lutidinium hemi-fumarate fumaric acid (25L:$F_{0.5}$:FA) are both cocrystals of salts (Fig. 5.2a and b, respectively), like $26AMP_2$:F:$FA_{0.5}$, and they both exhibit the base-acid-base (b-a-b) motif seen in all the structures discussed in Chapter 4, with all fumarate ions and FA molecules taking the low energy two s-*cis* conformation. In 25AMP:$F_{0.5}$:$FA_{0.5}$ the b-a-b units are formed into chains through a H-bonded ring, with graph seet notation $R_2^4(8)$, incorporating both the amino proton involved in the b-a-b unit and a H-bond from the other (non-b-a-b unit) amino proton to the carboxyl of the fumarate ion in the neighbouring b-a-b unit. The FA molecules within this structure form crosslinks between b-a-b chains, sitting between fumarate molecules. The replacement of the amino group with a methyl means that the b-a-b unit of

25L:F0.5:FA lacks stabilisation from a second H-bond. As a consequence, two independent FA chains run between the fumarate anions of the b-a-b units, at 49.6° to each other, resulting in a diamond shaped lattice (Fig. 5.2c) with b-a-b units at the nodes. In both 25AMP:$F_{0.5}$:$FA_{0.5}$ and 25L:F0.5:FA, a two dimensional plane of molecules is connected by intermolecular H-

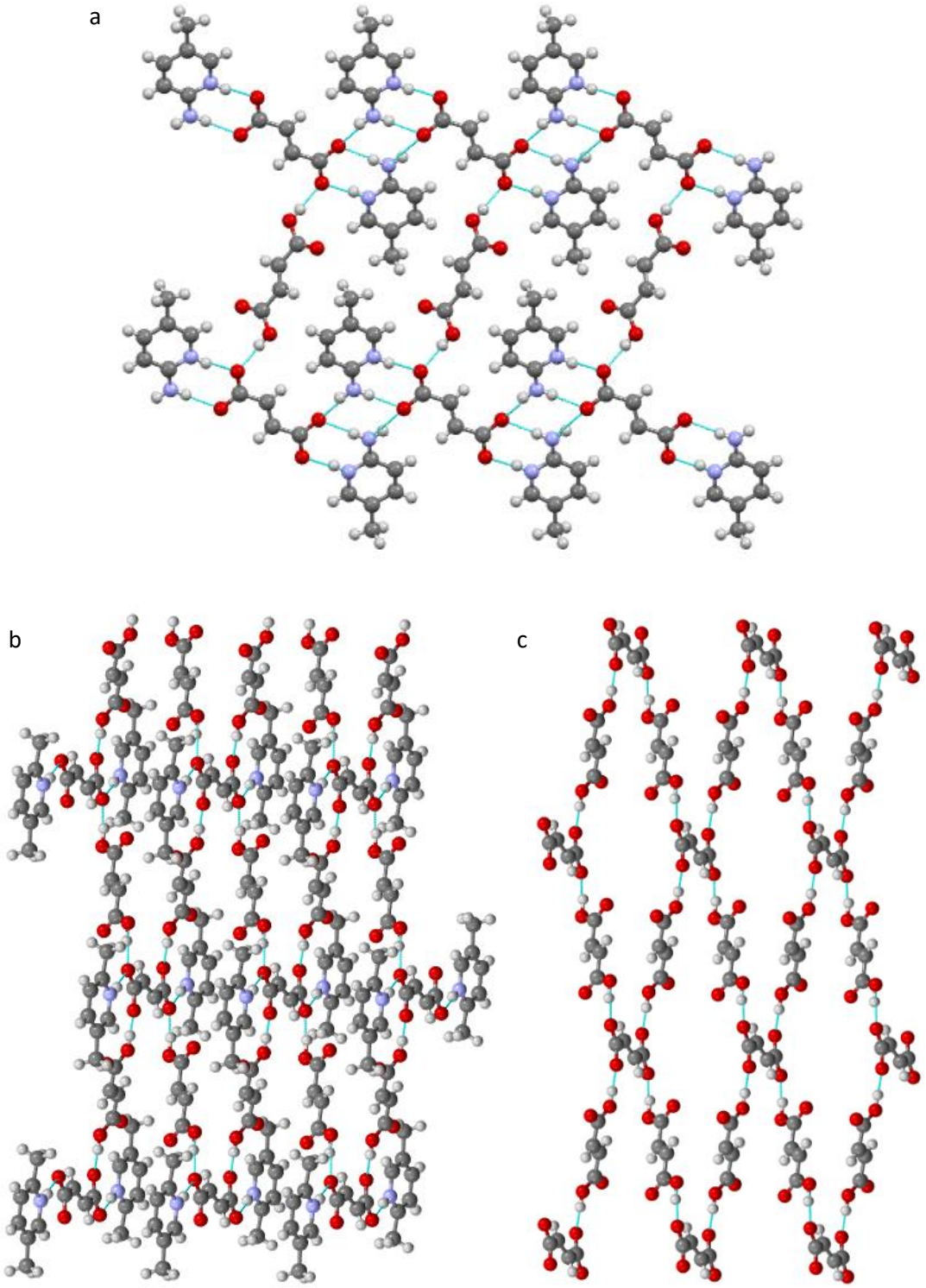


Figure 5.2: A single plane in the packing of (a) 25AMP:$F_{0.5}$:$FA_{0.5}$ and (b) 25L:$F_{0.5}$:FA. (c) The two FA-fumarate chains running through the structure of 25L:$F_{0.5}$:FA to create a diamond shaped lattice. H-bonds are shown as bright blue lines.

bonding and these layers then stack to give the full three-dimensional crystal structure. Neither a stability study nor an NMR characterisation of 25AMP:$F_{0.5}$:$FA_{0.5}$ or 25L:F0.5:FA have previously been reported.

Swapping the amino/methyl substitution positions compared to 25AMP results in 5-amino-2-methylpyridine (52AMP, Fig. 5.1) for which no multicomponent structure has been previously published with FA. In this chapter, a new salt form is reported from SXRD and thoroughly characterised using an NMR crystallography approach. Alongside this a comparison of the structure and stability of 25AMP:$F_{0.5}$:$FA_{0.5}$ and 25L:$F_{0.5}$:FA is also presented, to which NMR crystallography has also been applied to provide a fuller understanding of their forms.

5.2. Sample Preparation

All chemicals were obtained from Sigma Aldrich (UK) at purities of 98% or higher and used without further purification. Crystallisations were performed at a range of acid : base ratios, from 3 : 1 to 1 : 2, but all systems produced crystals with a 1 : 1 molar ratio. Although ethanol, methanol and isopropanol were all trialled for each system, as described below, isopropanol was found to be most successful.

25L:$F_{0.5}$:FA was crystallised by slow evaporation under ambient conditions from isopropanol on ~100 mg scale with 48 mg FA and 52 μL 25L (with density 0.92 mg / L). Crystals began to grow after 4 days. Crystal growth in subsequent crystallisation attempts was found to be improved by the addition of seed crystals to the crystallisation media. These seed crystals were selected from the small crystals present in the initial samples and were added when the saturated solution was set down for evaporation. Crystallisation was also attempted by co-grinding the components with < 1 mL of solvent.

52AMP:HF and 25AMP:$F_{0.5}$:$FA_{0.5}$ were crystallised by mixing each base (81 mg) and FA (87 mg) in the minimum amount of hot methanol required to dissolve all the solute (~ 15

mL) and then allowing the resulting solution to cool slowly at room temperature. Crystals began to form after a couple of days for both systems. Crystal growth in subsequent crystallisation attempts was found to be improved by co-grinding the base and acid with < 1 mL before dissolution in hot methanol and/or the addition of seed crystals to the crystallisation media. These seed crystals were selected from the small crystals present in the initial samples and were added when the hot methanol solution was placed at room temperature to cool.

5.3. XRD

5.3.1. SXRD of 52AMP:HF

The crystal structure of 52AMP:HF has been determined, described below, with selected crystal data given in Table 5.1 (alongside corresponding details for 25L:$F_{0.5}$:FA and 25AMP:$F_{0.5}$:$FA_{0.5}$). In all cases, hydrogen atoms were found in the electron density map. Initial verification of proton transfer was completed by comparison of the carboxylate/carboxylic acid C-O and C=O bond lengths and was then confirmed by ^{14}N-^{1}H HMQC NMR experiments (see sections 5.4.2.1 and 6.3.1).

Table 5.1: Selected crystal data for 52AMP:HF, 25L:$F_{0.5}$:FA and 25AMP:$F_{0.5}$:$FA_{0.5}$.

	52AMP:HF	25L:$F_{0.5}$:FA	25AMP:$F_{0.5}$:$FA_{0.5}$
Stoichiometry (base : acid)	1 : 1	1 : 1 $^{1}/_{2}$	1 : 1
Chemical formula	$C_{10}H_{12}N_2O_4$	$C_{13}H_{15}NO_6$	$C_{10}H_{12}N_2O_4$
Formula weight (g mol^{-1})	224.22	281.26	224.22
Crystal system	Triclinic	Triclinic	Triclinic
Space group	$P\bar{1}$	$P\bar{1}$	$P\bar{1}$
a (Å)	8.0181(3)	7.2696(2)	4.0366(4)
b (Å)	8.0467(3)	9.1353(3)	9.3145(10)
c (Å)	9.3998(3)	11.6311(3)	14.0077(14)
α (°)	93.715(3)	84.115(2)	94.030(3)
β (°)	109.263(3)	72.513(2)	95.060(3)
γ (°)	110.974(3)	88.763(2)	90.903(3)
Z	2	2	2
Temperature (K)	150	180	100
*R*1 [$I > 2\sigma(I)$]	0.0354	0.0426	0.0359

52AMP:HF has a stoichiometry of 1 : 1, base : acid, and crystallises in the triclinic space group $P\bar{1}$. Hydrogen fumarate molecules, in the two s-*cis* conformation, form acid chains along the a-axis with graph set notation $C_1^1(7)$ (Fig. 5.3a).[70] Pairs of these chains, running in opposite directions, are linked through the 52AMP ions associated with each hydrogen fumarate (Fig. 5.3b). Through pyridinium-carboxylate and amine-carboxylate H-bonds, a base-acid-base-acid ring is formed, $R_4^4(18)$ which supports this pairing (Fig. 5.3c).[70] A H-bond via the other amino proton, to the carboxylic acid O=C, allows crosslinking between paired chains, forming a H-bonded layer on the (010) crystal plane (Fig. 5.3d). These layers then stack to form the 3D structure.

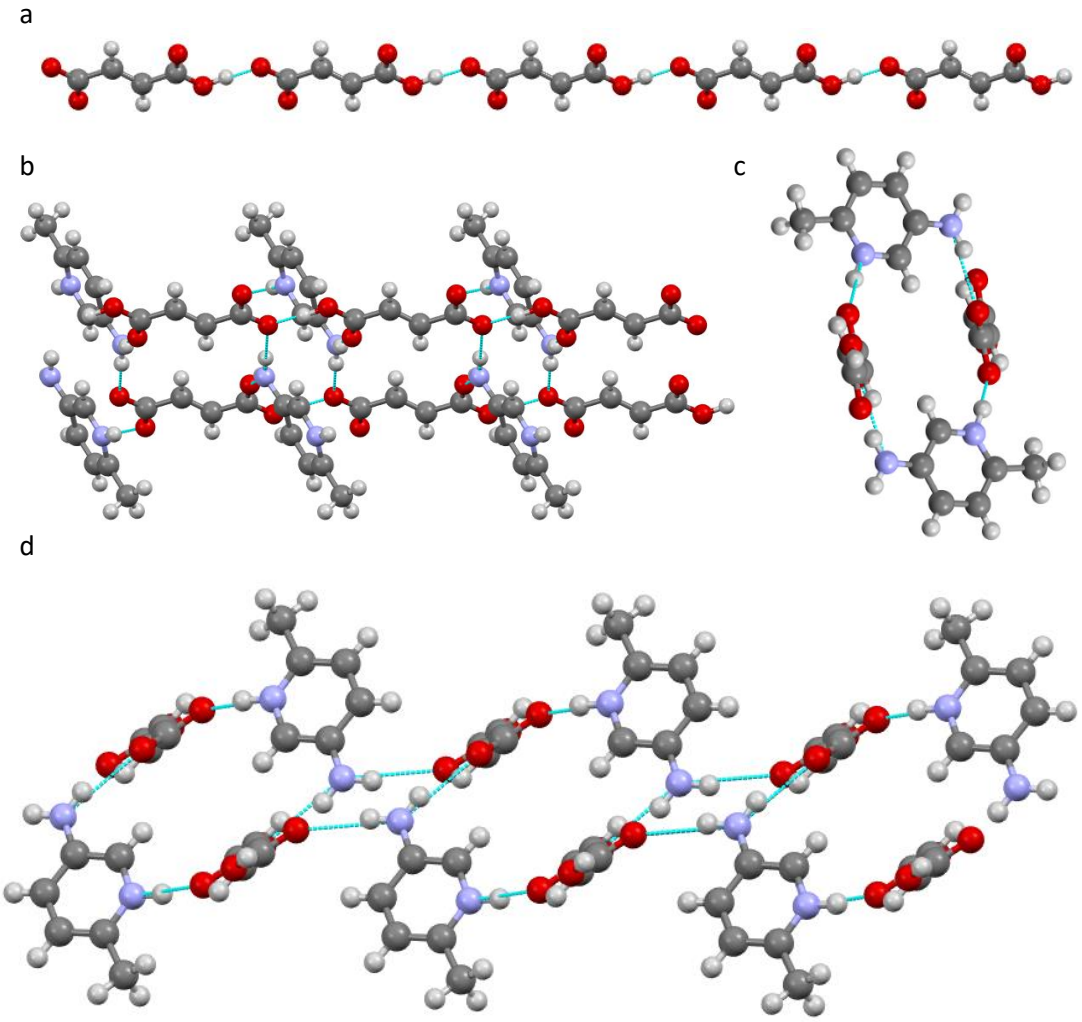


Figure 5.3: Packing of 52AMP:HF showing (a) the acid chain, the paired acid chains linked through the 52AMP molecules viewed (b) along the c axis and (c) along the a axis, and (d) a set of paired chains joined by the smaller ring motif.

Table 5.2: Selected structural parameters for 52AMP:HF, 25L:$F_{0.5}$:FA and 25AMP:$F_{0.5}$:$FA_{0.5}$.

	N···O distance (Å)	$NH^+\cdots^-O$ angle (°)	NH···O angle (°)	O···O distance (Å)	OH···O angle (°)	$OH\cdots^-O$ angle (°)
52AMP:HF	2.65	160.4	-			
	2.92	-	160.1			
	2.98	-	163.3			
				2.50	178.1	-
25L:$F_{0.5}$:FA	2.78	173.1	-			
				2.59	176.0	-
				2.60	-	167.3
25AMP:$F_{0.5}$:$FA_{0.5}$	2.73	167.2	-			
	2.83	-	174.8			
	2.83	-	153.6			
				2.56	-	169.5

The packing structure of 52AMP:HF differs significantly from those seen in 25AMP:$F_{0.5}$:$FA_{0.5}$ and 25L:$F_{0.5}$:FA. Neither FA nor doubly ionised fumarate are present, with the occurrence of hydrogen fumarate instead preventing b-a-b unit formation. This difference may be due to the combined loss of b-a-b unit stabilisation/strengthening from having the amino group at the 2 position on the pyridinium ring, as in the case of $26AMP_2$:F:$FA_{0.5}$ and 25AMP:$F_{0.5}$:$FA_{0.5}$, and the addition of the functional group to the other side of molecule. Acid chains are seen in all three structures but are of different forms as both 25AMP:$F_{0.5}$:$FA_{0.5}$ and 25L:$F_{0.5}$:FA alternate FA and fumarate within their chains (Fig. 5.2) whereas the 52AMP:HF packing structure is formed of repeats of crystallographically equivalent hydrogen fumarates (Fig. 5.3a). The H-bond parameters for the significant H-bonding motifs identified are given in Table 5.2.

5.3.2. PXRD

PXRD of 52AMP:HF showed no change in structure on grinding. A Rietveld refinement of the experimental powder pattern against the SXRD structure gave R_{wp} = 8.83% and R_{Bragg} = 1.39% (Appendix 3, Table A3.1). The other two systems also showed no grinding induced change although, as discussed below, neither refined as well against the single crystal structures as 52AMP:HF (Fig. 5.4).

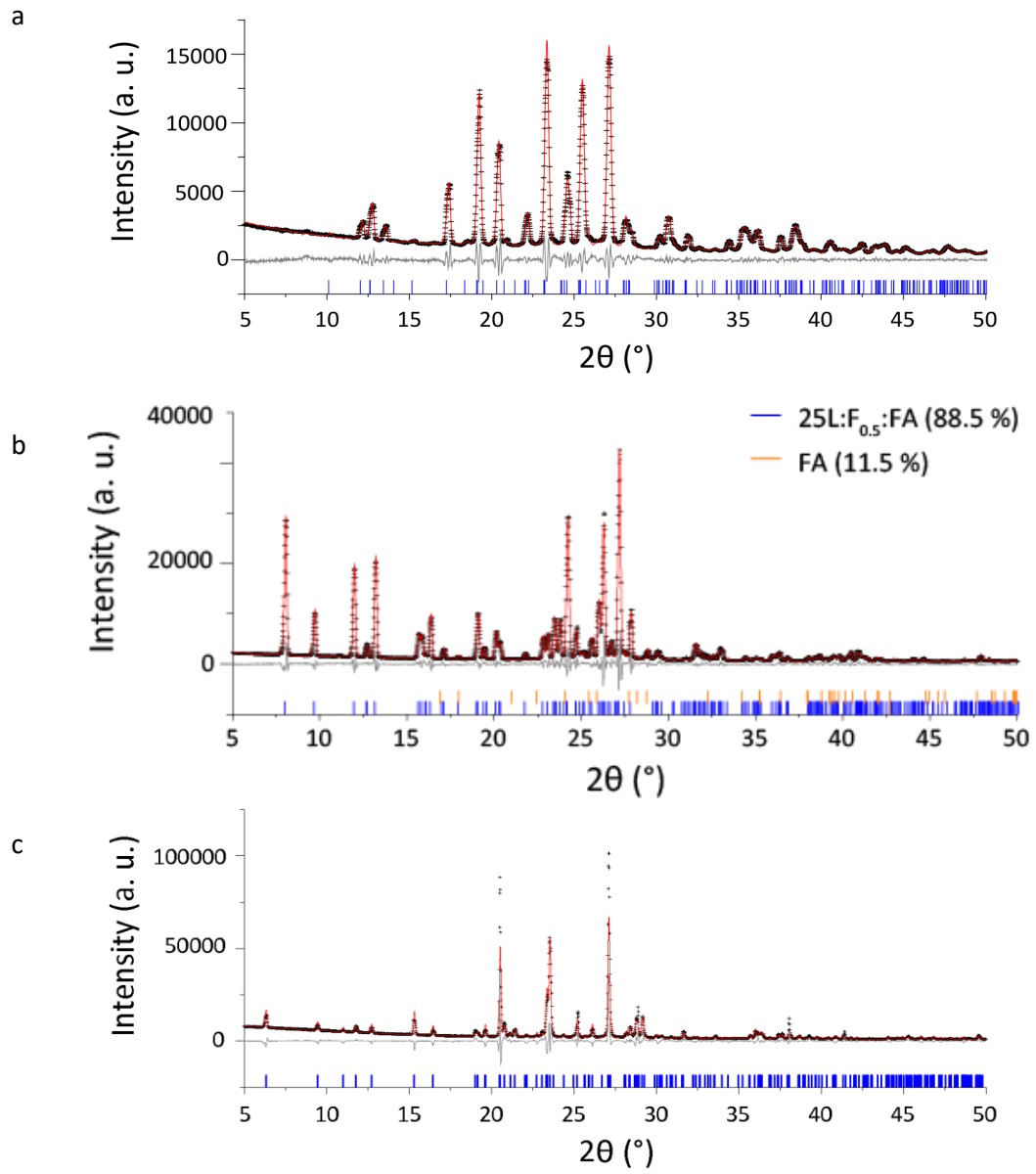


Figure 5.4: Final Rietveld fit for (a) 52AMP:HF, (b) 25L:$F_{0.5}$:FA and (c) 25AMP:$F_{0.5}$:$FA_{0.5}$ at room temperature, showing the experimental (black crosses), calculated (red upper line) and difference (grey lower line) PXRD profiles. Tick marks (bottom of each) indicate allowed peak positions. For refinement parameters, see Tables A3.1. A3.2 and A3.3, respectively.

Examination of the PXRD pattern recorded for 25L:$F_{0.5}$:FA produced by slow evaporation suggested that free crystalline FA remained within the bulk crystallised sample. The small crystals of FA could not be successfully isolated and removed prior to grinding, so a two-phase refinement was necessary and showed that FA accounted for 11.5% of the sample (Appendix 3, Table A3.2). Selection of large crystals to pack a 1.3 mm rotor for fast MAS experiments ensured that the FA phase was absent from these NMR experiments. PXRD for 25L:$F_{0.5}$:FA produced by grinding will be discussed in Section 5.4.3, below.

The PXRD pattern of 25AMP:$F_{0.5}$:$FA_{0.5}$ shows some evidence of splittings, as seen for 26L:HF (chapter 3), although they are not as distinct, suggesting that there are small variations in unit cell parameters within the crystallised sample and that the powder was insufficiently well ground to prevent preferred orientation effects from highlighting this. As a result, the refinement showed larger discrepancies in peak shape and intensity although all experimental reflections (bar their splitting) corresponded to the simulated reflections, with R_{wp} = 13.28% and R_{Bragg} = 5.00% (Appendix 3, Table A3.3). These splittings were not investigated further for 25AMP:$F_{0.5}$:$FA_{0.5}$. In 26L:HF, for which a second structure was determined (see section 3.4), they had no significant effect on the GIPAW calculated shifts. Therefore, the small changes in individual atomic positions associated with these splittings are thought to impart a negligible impact on the resulting local interactions.

5.4. NMR Crystallography

5.4.1. 1D ^{1}H one-pulse and ^{13}C CP MAS NMR

Agreement between the experimental solid-state MAS NMR and the GIPAW calculated chemical shifts of the geometry optimised structures varied between the three systems. This was in part due to the phase purity of each sample but several other differences were also apparent. Fig. 5.5 shows 1D ^{1}H MAS and ^{1}H-^{13}C CP MAS spectra for each of the systems, with the labelling scheme for each structure outlined in Fig. 5.6. Assignments are based on both the GIPAW calculated chemical shifts (Appendix 3, Tables A3.4 and A3.5) and 2D spectra, presented in the following section.

All three systems show good agreement for ^{1}H between experimental NMR chemical shifts and calculated GIPAW chemical shifts. Surprisingly, as no indication was seen in the PXRD pattern in Fig. 5.4a, the ^{1}H MAS spectrum for 52AMP:HF has a resonance thought to correspond to crystalline FA at 12.8 ppm (Fig. 5.5a) as seen with 26L:HF in Chapter 3. Other than this minor secondary phase, the only discrepancy for 52AMP:HF corresponds to the

apparent overcalculation of H3, the OH proton involved in an OH···O H-bond, which experimentally is 1.6 ppm lower than the calculated value (δ_{iso}^{calc} = 18.4 ppm compared to δ_{iso}^{exp} = 16.8 ppm). This is the same environment for which a discrepancy was previously

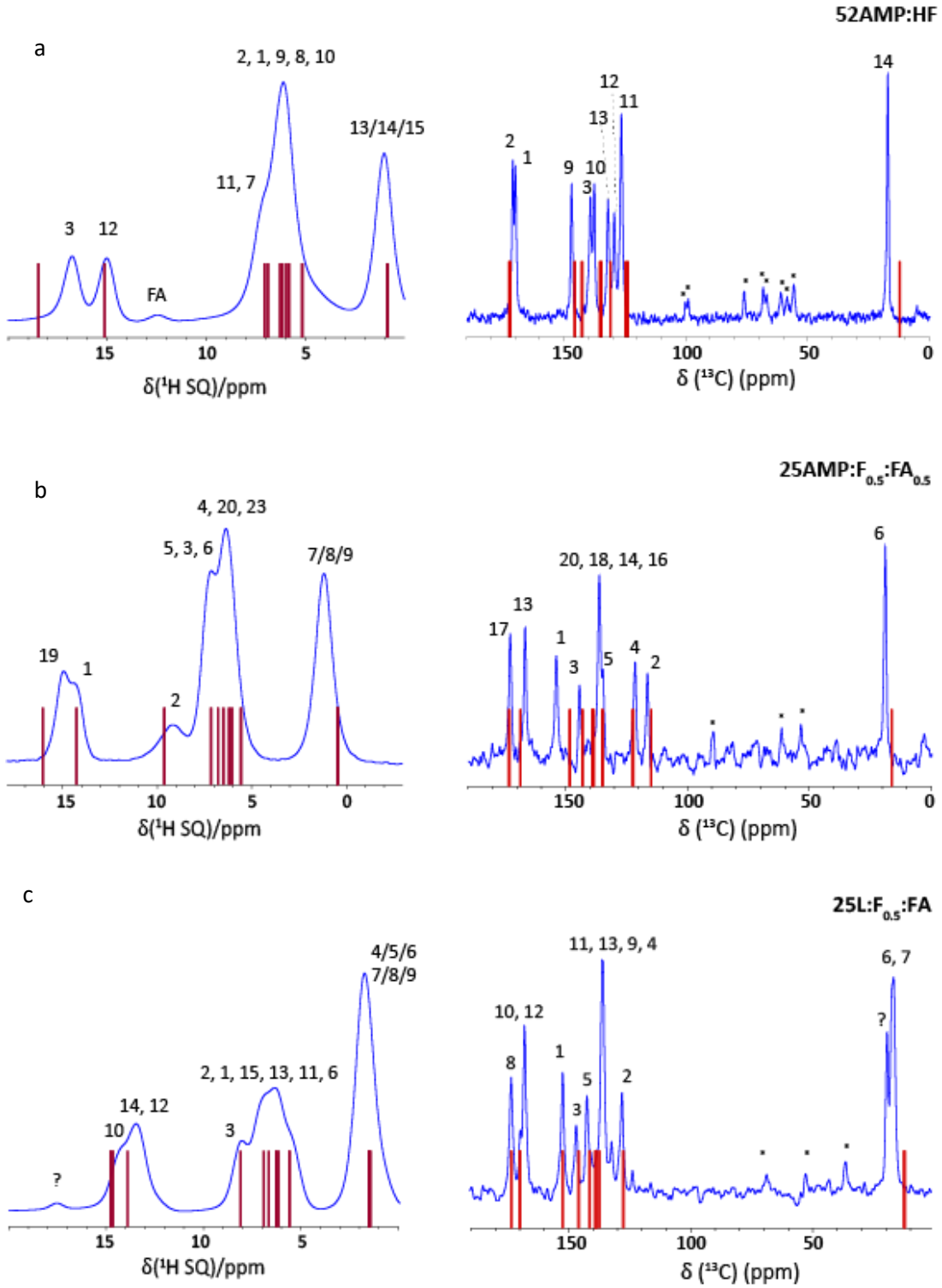


Figure 5.5: ^{1}H (600 MHz) one-pulse MAS (60 kHz) and ^{1}H-^{13}C CP-MAS (12.5 kHz) spectra of (a) 52AMP:HF, (b) 25AMP:$F_{0.5}$:$FA_{0.5}$ and (c) 25L:$F_{0.5}$:FA with stick spectra corresponding to GIPAW calculated chemical shifts. The assignments to each proton, labelled in the structures in Fig. 5.5, are given.

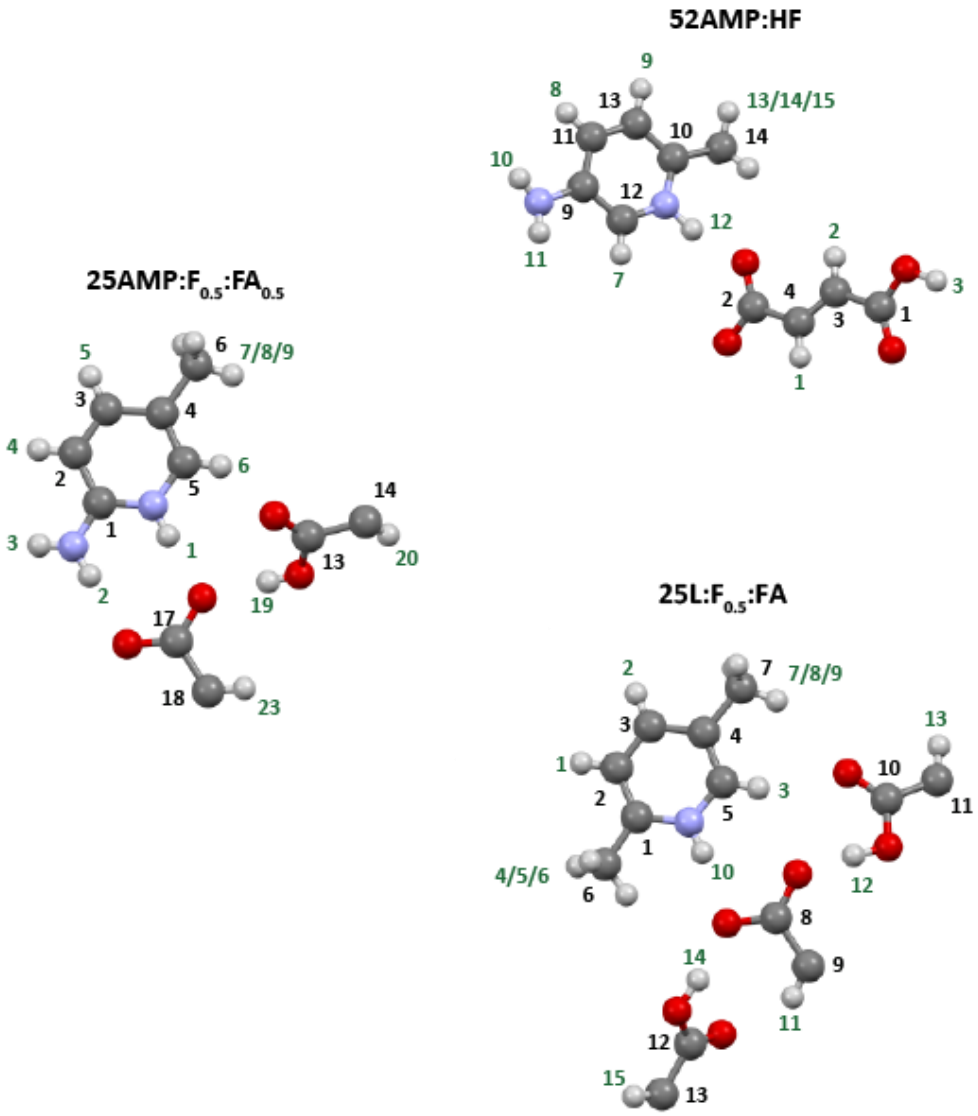


Figure 5.6: Labelling of the atoms within the asymmetric unit of 52AMP:HF (top), 25AMP:$F_{0.5}$:$FA_{0.5}$ (middle) and 25L:$F_{0.5}$:FA (bottom).

described for both 26L:HF and 26AMP_2:F:$FA_{0.5}$ and it is also seen for 25AMP:$F_{0.5}$:$FA_{0.5}$ and 25L:$F_{0.5}$:FA (Fig. 5.5b and c). In the latter case, the experimental chemical shifts for the two OH protons lie at δ_{iso}^{exp} = 13.4 ppm, below the NH proton, despite both being calculated at a higher chemical shift than this environment.

The level of variation seen in each structure is consistent in both size and direction of change with $\Delta\delta^{exp\text{-}calc}$ = –1.9, –1.4, –1.3 and –1.3/–1.4 for the OH···O protons of 26L:HF, 26AMP_2:F:$FA_{0.5}$, 25AMP:$F_{0.5}$:$FA_{0.5}$ and 25L:$F_{0.5}$:FA, respectively. In all cases the assignment has been confirmed by the correlations present in 2D ^{1}H-^{1}H DQ and ^{1}H-^{13}C HETCOR spectra as well as by confirming the assignment of the NH proton with a ^{14}N-^{1}H HMQC spectrum.

This occurrence of a similar discrepancy for the same chemical moiety across multiple structures supports the conclusion that this is a systematic error in the calculation. Similarly, C1 of 25AMP:$F_{0.5}$:$FA_{0.5}$, the carbon at the 2 position on the pyridine ring, sits between two nitrogens and is calculated 5.9 ppm lower than it is seen experimentally, in the same way as for the 2 position carbon in the 26AMP systems (see sections 4.5.3 and 6.3.2).

While the OH···O proton chemical shift is the sole ^{1}H divergence between calculation and experiment for 25AMP:$F_{0.5}$:$FA_{0.5}$, this is not the case for 25L:$F_{0.5}$:FA. As can be seen in Fig. 5.5c, there is a proton resonance at 17.5 ppm, higher than the NH assigned peak at δ^{exp} = 14.3 ppm. The high chemical shift is indicative of a proton involved in a strong hydrogen bond and is downfield of the expected OH ^{1}H chemical shift for crystalline FA (δ^{exp} = 12.8 ppm). Along with the analysis and assignments from the 2D spectra, below, the relative intensity of this resonance suggest that it corresponds to an additional unknown phase.

Fig. 5.5 shows that all the methyl carbon ^{13}C chemical shifts are calculated at a lower shift than seen experimentally, as expected due to the deviation from linearity of the referencing noted in section 3.3.2, but otherwise 25AMP:$F_{0.5}$:$FA_{0.5}$ and 25L:$F_{0.5}$:FA show good agreement between experiment and GIPAW calculation, although the experimental spectrum for 25L:$F_{0.5}$:FA has several additional resonances attributed to the unknown secondary phase (Fig. 5.5c). Despite the excellent agreement seen for the ^{1}H chemical shifts for 52AMP:HF, there is relatively poor agreement for the ^{13}C chemical shifts. As previously discussed, the established discrepancy between GIPAW calculations and experimental values is ~2 ppm for ^{13}C. It can be clearly seen in the 1D ^{1}H-^{13}C CP MAS spectrum in Fig. 5.5a that this is exceeded for numerous carbon environments in 52AMP:HF. Not only is C14, the methyl carbon, calculated at a lower chemical shift due to temperature effects, but C10 is also calculated to lie 2.8 ppm lower than it is seen experimentally. Conversely, both C3 and C4 are calculated to lie more than 3 ppm higher than their experimental chemical shifts. These latter two are the CH carbons in the hydrogen fumarate anion. These variations cannot be attributed to an error in the referencing due to the mix of positive and negative variation. Unlike 26AMP and 25AMP, 52AMP does not have an amino-carbon-pyridinium environment so this

apparently systematic error is not expected. It is, however interesting to note that C10, although not differing by as much, is also a 2 position, quaternary carbon.

5.4.2. 2D NMR

5.4.2.1. 52AMP:HF

Fig. 5.7a shows that there are clear correlations in a ^{1}H-^{1}H DQ MAS spectrum of 52AMP:HF for H12 (NH) with H3 (OH), H7 (aromatic CH) and the methyl group at δ_{DQ} = 15.0 + 16.8/7.1/1.0 = 31.8/22.1/16.0 ppm. The low intensity resonances that appear as shoulders on the pair of H12-H7 cross-peaks are thought to correspond to the H12-H9 proximity (δ_{DQ} = 15.0 + 6.1 = 21.1 ppm). The H12-H9 correlation is expected to be weaker as it is significantly

Table 5.3: H···H proximities (<3.5 Å) and corresponding ^{1}H DQ chemical shifts for the pyridine NH and the OH peaks of each crystal structure

Structure	NH	δ_{iso}^{exp} SQ1 (ppm)	Proton 2	δ_{iso}^{exp} SQ2 (ppm)	δ_{iso}^{exp} DQ (ppm)	Separation[a] (Å)
52AMP:HF	H12 (NH)	15.0	H7	7.1	22.1	2.31
			H13/H14/H15	1.0	16.0	2.35
			H3	*16.8*	*31.8*	*2.71*
			H9	*6.1*	*21.1*	*3.38*
	H3 (OH)	16.8	H2	*6.1*	*22.9*	*2.78*
			H8	*6.1*	*22.9*	*2.93*
			H7	*7.1*	*23.9*	*2.93*
			H11	*7.1*	*23.9*	*2.94*
25L:F$_{0.5}$:FA	H10 (NH)	14.3	H3	8.1	22.4	2.32
			H4/H5/H6	1.7	16.0	2.66
			H14	*13.7*	*28.0*	*3.02*
	H14 (OH)	13.7	*H11*	*5.6*	*19.3*	*2.39*
			H7/H8/H9	*1.7*	*15.4*	*2.84*
			H2	*7.1*	*20.8*	*3.18*
			H13	*6.4*	*19.8*	*3.39*
	H12 (OH)	13.4	*H11*	*5.6*	*19.0*	*2.28*
			H5	*1.7*	*15.1*	*2.66*
			H3	*8.1*	*21.5*	*2.82*
			H1	*6.4*	*19.8*	*3.32*
			H13	6.4	19.8	3.49
25AMP:F$_{0.5}$:FA$_{0.5}$	H1 (NH)	14.0	H2	8.8	22.8	2.26
			H6	6.8	20.8	2.34
			H19	*14.7*	*28.7*	*2.61*
	H19 (OH)	14.7	*H23*	*6.1*	*22.8*	*2.58*
			H1	*14.7*	*28.7*	*2.61*
			H4	*6.1*	*22.8*	*2.98*
			H5	*6.8*	*21.5*	*3.29*
			H6	*6.8*	*21.5*	*3.29*

[a] H-H distances are taken from the DFT (CASTEP) optimised structure. Intermolecular proximities are denoted using italic font.

further than the next nearest, 3.38 Å compared to 2.71 Å for H3 (see Table 5.3). No correlation with the amino protons is expected due to the position of the amino group on the opposite side

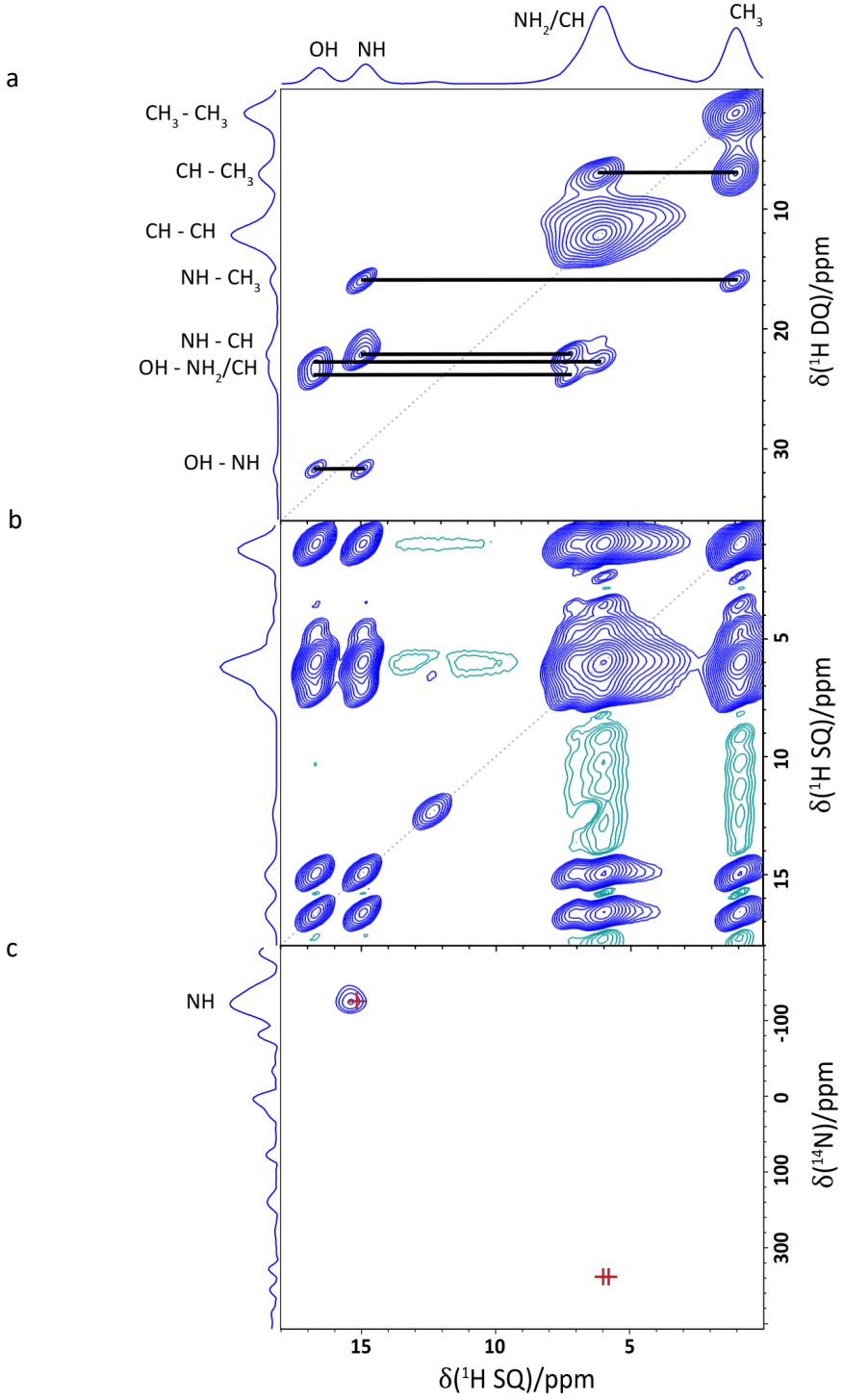


Figure 5.7: 2D MAS (60 kHz) NMR spectra of 52AMP:HF: (a) a ^{1}H-^{1}H DQ spectrum recorded with one rotor period of BaBa recoupling; (b) a ^{1}H-^{1}H SQ NOESY spectrum with t_{mix} = 500 ms; and (c) a ^{14}N-^{1}H HMQC spectrum with 8 rotor periods of R^3 recoupling with GIPAW calculated chemical shifts for N···H proximities < 2 Å shown as red crosses. All spectra were recorded at a ^{1}H Larmor frequency of 600 MHz. Base contour levels are at 8.4%, 1.5% and 40.1% of the maximum peak height, respectively. Blue and green contours correspond to positive and negative intensity, respectively. The negative intensities seen at the CH_3 and CH F_1 (vertical axis) SQ frequencies in (b) are due to the much greater intensity of their auto-correlation peaks. The dashed diagonal lines in (a) and (b) indicate the (a) $\delta_{DQ} = 2\delta_{SQ}$ and (b) $\delta_{SQ} = \delta_{SQ}$ diagonals.

of the base molecule, unlike in 25AMP:$F_{0.5}$:$FA_{0.5}$ where it is immediately adjacent. The absence of a H12-H11 correlation cannot be confirmed, however, as it is expected to fall at the same place as the H12-H7 pair (δ_{DQ} = 15.0 + 7.1 = 22.1 ppm) within the SQ ^{1}H resolution. A 2D SQ ^{1}H-^{1}H NOESY spectrum (Fig. 5.7b) was used to confirm the presence of FA as it shows the existence of two distinct phases. In the same method employed for 26L:HF (Chapter 3), a mixing time of 500 ms was used to allow spin diffusion throughout the entirety of each phase,[206, 207] with the clear separation of cross-peaks indicating the occurrence of more than one phase. Only a single correlation was seen for the resonance at δ^{exp} = 12.8 ppm, corresponding to a proximity to a proton with a chemical shift in the CH region, as expected for the OH proton of crystalline FA. The relatively low ratio of FA within the sample is evidenced by the lack of DQ correlations for the FA peak in the ^{1}H-^{1}H DQ MAS spectrum, unlike in 26L:HF, where they were clearly apparent.

A ^{14}N-^{1}H HMQC spectrum of 52AMP:HF (Fig. 5.7c) does not show any cross-peaks for the NH_2 protons but confirms proton transfer between the O and pyridine N (δ^{exp} = −126.1 ppm, δ^{calc} = −126.6 ppm - Appendix 3, Table A3.6) as well as the assignment of H12.

A 2D ^{1}H-^{13}C HETCOR MAS NMR spectrum of 52AMP:HF is shown in Fig. 5.8, recorded using a CP contact time of 500 μs such that cross-peaks for longer-range C···H proximities are apparent as well as direct one-bond C-H connectivities. The ^{1}H-^{13}C HETCOR spectrum is shown together with crosses that represent the GIPAW calculated chemical shifts for the C-H dipolar correlations up to 3.3 Å (see Table A3.6 in Appendix 3). More correlations with the methyl protons are present experimentally than expected for this cut off distance, with small cross-peaks apparent for C3 (δ^{exp} = 132.5 ppm, δ^{calc} = 135.9 ppm), C9 (δ^{exp} = 147.5 ppm, δ^{calc} = 146.3 ppm) and C1 (δ^{exp} = 170.6 ppm, δ^{calc} = 172.7) which have methyl proton proximities of 3.51 Å (C3 and C9) and 3.40 Å (C1). Crosses for the GIPAW calculated chemical shifts for longer range connectivities were not included as they are missing for the other proton environments. The relative intensity of the methyl peak compared to other environments, due to its containing three protons, means that CP transfer is more efficient and correlations are visible out to a greater distance.

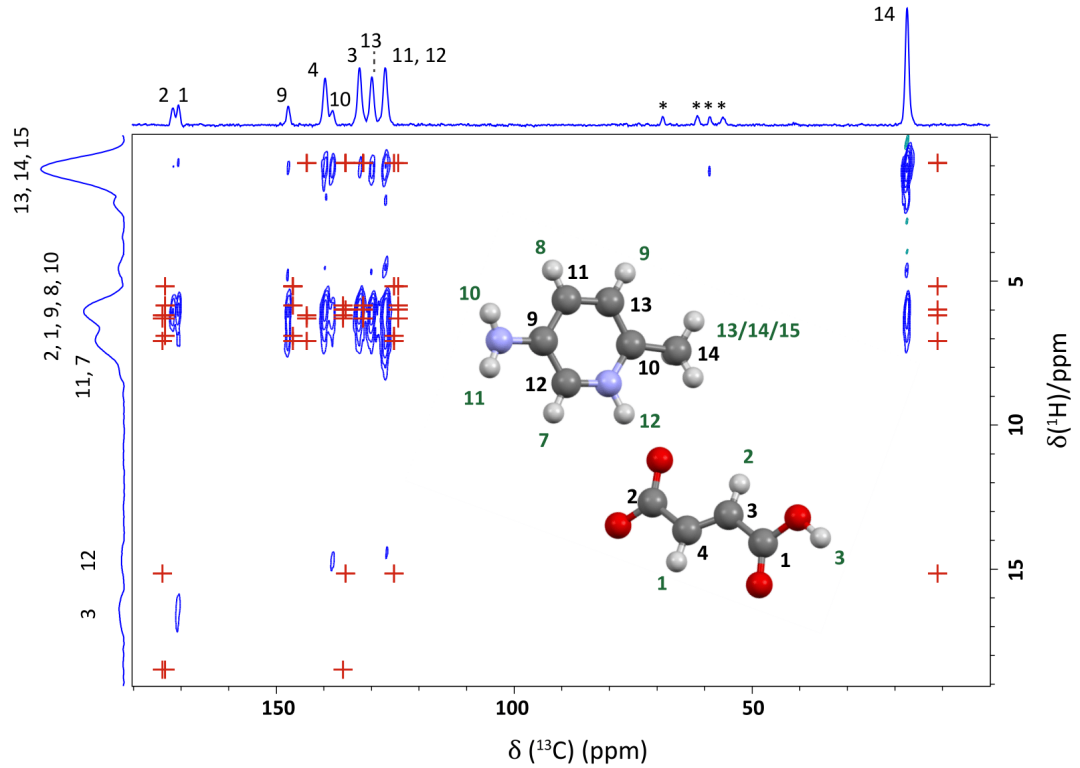


Figure 5.8: A ^{1}H (700 MHz)-^{13}C CP (500 μs) HETCOR MAS (12.5 kHz) NMR spectrum of 52AMP:HF recorded using FSLG ^{1}H homonuclear decoupling in t_1 with GIPAW calculated chemical shifts shown as red crosses for C···H proximities < 3.3 Å. The base contour level is at 4.9% of the maximum peak height.

The aforementioned discrepancies between experiment and GIPAW calculation for the ^{13}C chemical shifts mean that C10 and C3 are seen experimentally in the opposite order to which they are calculated. Their calculated ^{13}C chemical shifts lie only 0.4 ppm apart with C10 at the lower chemical shift. The assignment of C10 to the resonance at δ^{exp} = 138.2 ppm is, however, confirmed by its cross-peak with H12 at δ^{exp} = 15.0 ppm, with a C10-H12 proximity of 2.08 Å compared to a distance of 3.68 Å for C3-H12. The H12 cross-peak for the ^{13}C resonance at δ^{exp} = 127.0 ppm also confirms that C12 lies under this peak along with C11. Correlations are missing for C2-H12, C2-H3, C3-H3 and C14-H12. The first two carbons are quarternary so this is not surprising, despite the relevant distances for C2 being significantly shorter than for other correlations present at 2.30 Å and 2.68 Å for H12 and H3, respectively. As C2 has the longest C-H distance to its nearest proton, a lower CP efficiency is expected. The absence of the C3-H3 cross-peak is expected as not only is C3 quaternary, but this corresponding distance is 3.29 Å, on the limit of what is seen for any other carbon

environment. The absence of the C14-H12 cross-peak is rather inexplicable as this is a methyl carbon and the C-H separation is 2.60 Å.

5.4.2.2. 25AMP:$F_{0.5}$:$FA_{0.5}$

Fig. 5.9a shows that there are clear correlations in a ^{1}H-^{1}H DQ MAS spectrum of 25AMP:$F_{0.5}$:$FA_{0.5}$ between the NH and NH_2 protons. Unlike in 52AMP:HF, the amino group is immediately adjacent to the pyridinium NH so a correlation is expected (see Table 5.3, above). The strong H-bond to the carboxylate results in a higher ^{1}H chemical shift of H2, the amine proton involved on the b-a-b unit, meaning that its cross-peaks with the pyridinium proton (δ_{DQ} = 14.0 + 8.8 = 22.8 ppm) are clearly distinct from the CH correlations (δ_{DQ} = 14.0 + 6.8 = 20.8 ppm). The closeness in ^{1}H chemical shift of H19 (OH) and H1 (NH) mean that their cross-peaks to the CH protons (and for H1 to H3, the non-b-a-b amino proton) are overlapping (δ_{DQ} = 14.0/14.7 + 6.8 = 28.7 ppm). The NH_2- NH_2 (δ_{DQ} = 8.8 + 6.8 = 15.6 ppm) correlation is clearly visible alongside the broad on-diagonal peak for the correlations between the CH/NH_2 environments which are close in ^{1}H chemical shift (δ_{SQ} = ~6.1-6.8 ppm). 25AMP:$F_{0.5}$:$FA_{0.5}$ also clearly lacks an NH-CH_3 correlation, as expected with the methyl group at the 5 position of the pyridinium ring. Fig 5.8b shows that proton transfer is also confirmed for 25AMP:$F_{0.5}$:$FA_{0.5}$ with the NH cross-peak at $^{-}\delta^{exp}$ (^{14}N) = −103.8 ppm (δ^{calc} = −88.5) in a ^{14}N-^{1}H HMQC spectrum. The NH_2 cross-peaks are also visible for 25AMP:$F_{0.5}$:$FA_{0.5}$, although at a much lower intensity, and are also in good agreement with the GIPAW calculated ^{14}N shifts as seen by the alignment (Appendix 3, Table A3.6).

A 2D ^{1}H-^{13}C HETCOR MAS NMR spectrum of 25AMP:$F_{0.5}$:$FA_{0.5}$ is shown in Fig. 5.10, recorded using a CP contact time of 500 μs. Crosses represent the GIPAW calculated chemical shifts for the C-H dipolar correlations up to 2.8 Å (see Table A3.7 in Appendix 3). As could be seen in the 1D ^{1}H one-pulse MAS NMR spectrum (Fig. 5.5b), the methyl protons, H7/H8/H9, are observed at a slightly higher chemical shift than calculated, δ^{exp} = 0.8 ppm rather than δ^{calc} = 0.4 ppm. The correlations with C13, C1 and C5 also confirm the assignments of the highest chemical shift protons with the C13-H19 cross-peak confirming that H19 (OH)

is the highest chemical shift proton, δ^{exp} = 14.7 ppm, but 1.3 ppm lower than calculated, δ^{calc} = 16.0 ppm, which is believed to be a systematic error as discussed above. The C5 and C1 cross-peaks to H1 confirm the good agreement between calculation and experiment for the

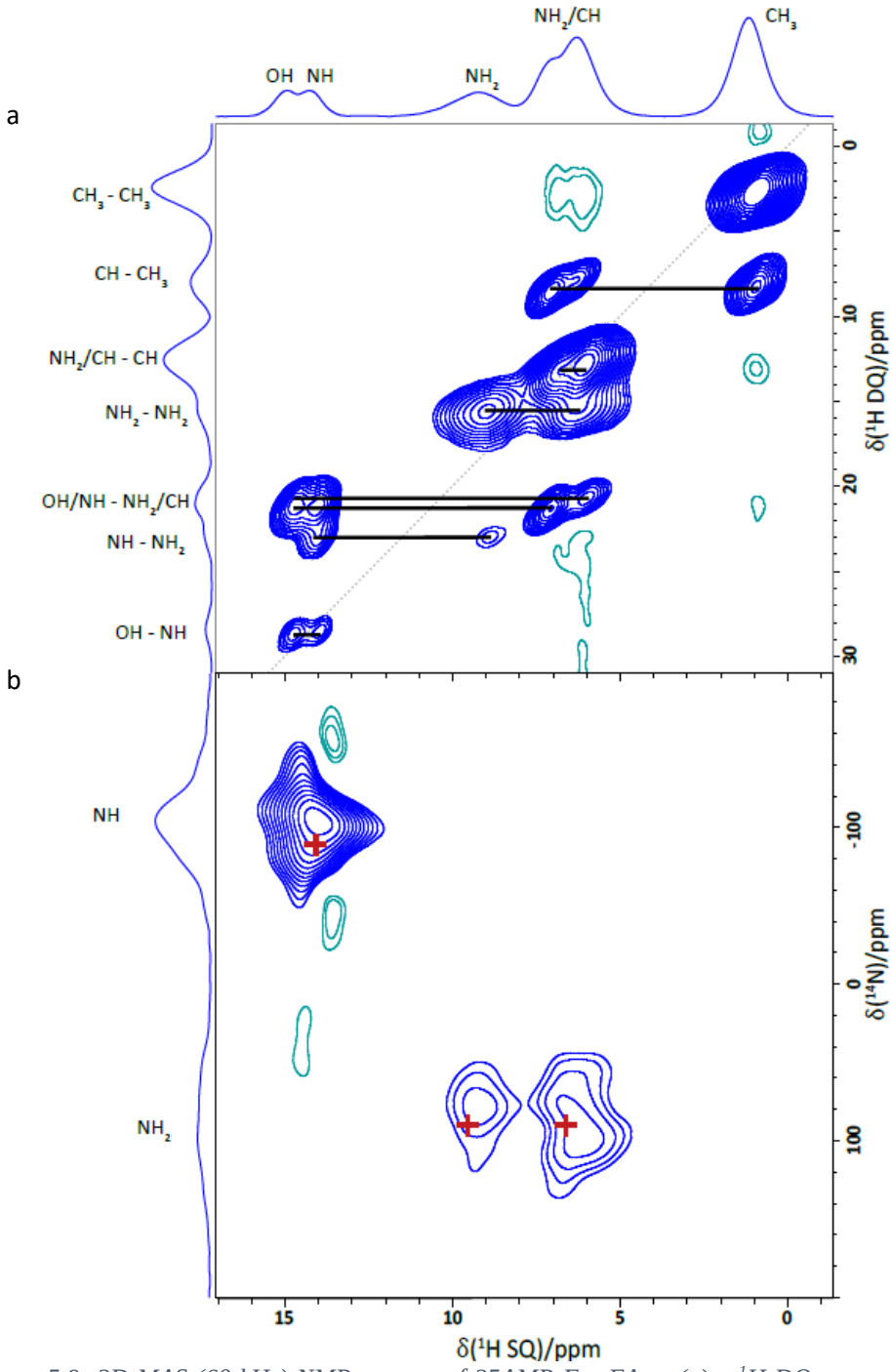


Figure 5.9: 2D MAS (60 kHz) NMR spectra of 25AMP:$F_{0.5}$:$FA_{0.5}$: (a) a ^{1}H DQ spectrum recorded with one rotor period of BaBa recoupling at a ^{1}H Larmor frequency of 700 MHz; and (b) a ^{14}N-^{1}H HMQC spectrum recorded with 8 rotor periods of R^3 recoupling (with GIPAW calculated chemical shifts for N···H proximities < 2 Å shown as red crosses) recorded at a ^{1}H Larmor frequency of 600 MHz. Base contour levels are at 5.3% and 12.0% of the maximum peak height, respectively. Blue and green contours correspond to positive and negative intensity, respectively. The dashed diagonal line in (a)) indicates the $\delta_{DQ} = 2\delta_{SQ}$ diagonal.

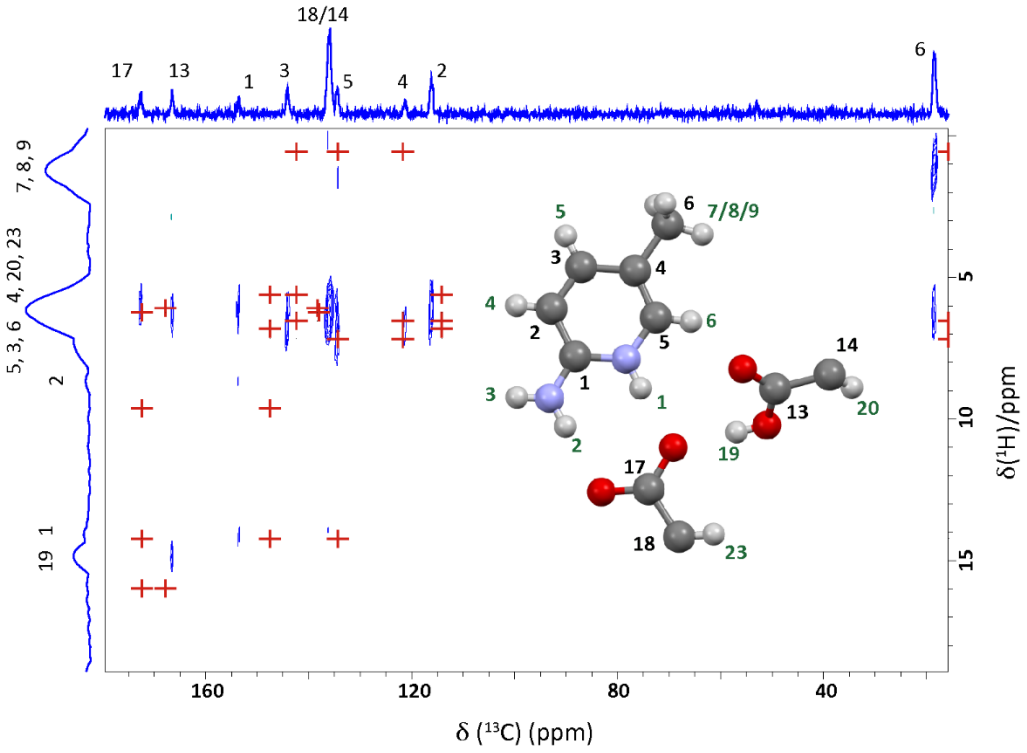


Figure 5.10: A ^{1}H (600 MHz)-^{13}C CP (500 μs) HETCOR MAS (12.5 kHz) NMR spectrum of 25AMP:$F_{0.5}$:$FA_{0.5}$ recorded using FSLG ^{1}H homonuclear decoupling in t_1 with GIPAW calculated chemical shifts shown as red crosses for C···H proximities < 2.8 Å. The base contour level is at 4.9% of the maximum peak height.

NH while also confirming the assignment of C1 to a higher than calculated chemical shift, this being the aforementioned ^{13}C environment that appears to systematically be calculated at a lower chemical shift than observed experimentally.

As seen for 52AMP:HF, the carboxylate carbon, C17, is missing many of the expected cross-peaks. The one present, C17-H23, corresponds to the shortest distance at 2.22 Å while the proximities to H2, H1 and H19 are all over 2.5 Å. The other missing cross-peaks are those for the correlations of C3 and C4 with H7/H8/H9. C3's nearest methyl proximity is on the limit of the distances observed at 2.79 Å so its absence is not surprising whereas the absence of the cross-peak for C4 is thought to be due to it being a quaternary carbon but is still somewhat unexpected due to the closer proximity of 2.14 Å.

5.4.2.3. 25L:$F_{0.5}$:FA

The apparent presence of an unknown secondary phase within 25L:$F_{0.5}$:FA is most readily analysed in the 2D MAS NMR spectra. The ^{1}H resonance at δ^{exp} = 17.5 ppm in the 1D one-

pulse MAS NMR spectrum is at a very low intensity relative to the other peaks, implying that this phase accounts for a small proportion of the sample. Its presence is not observed in either

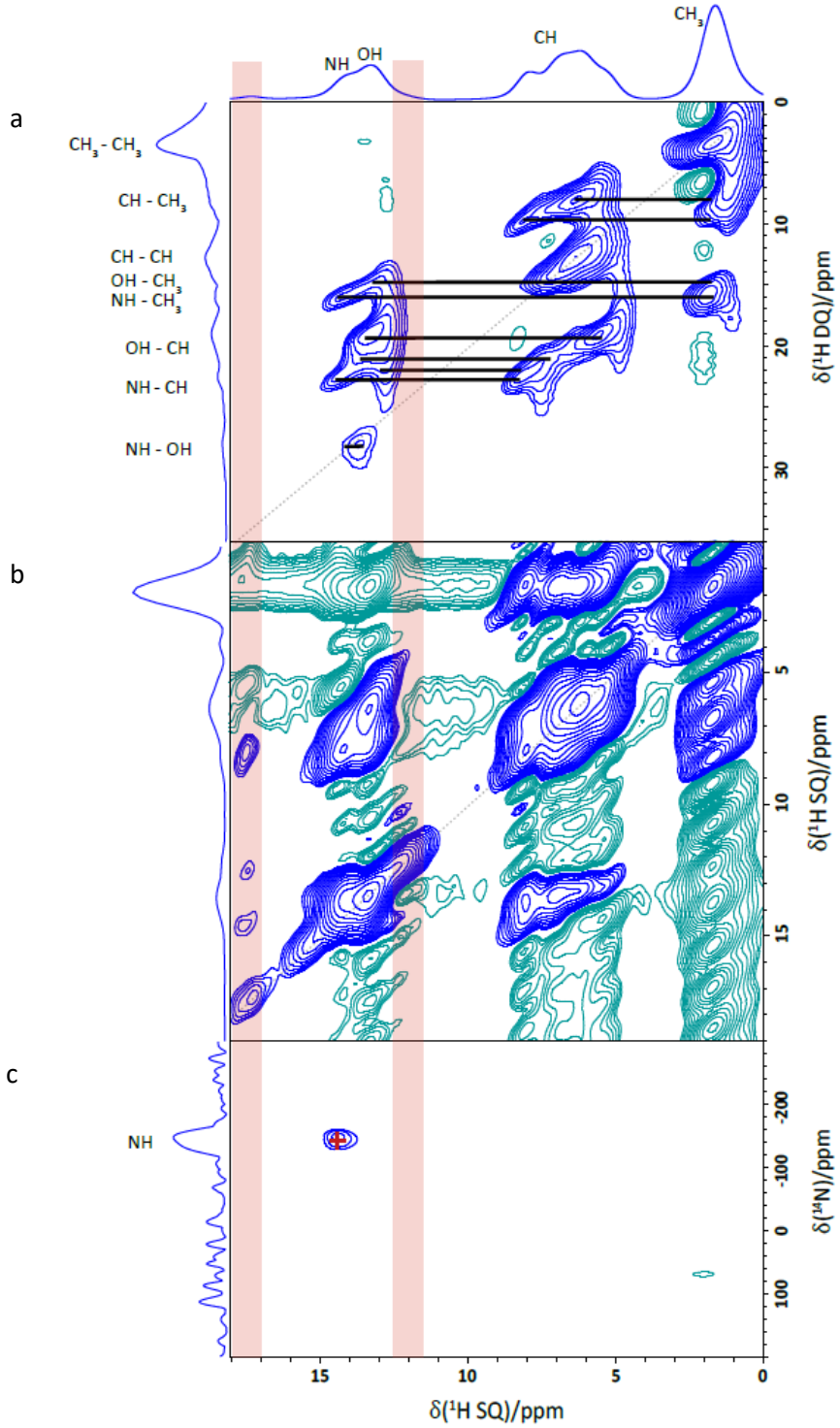


Figure 5.11: 2D MAS (60 kHz) NMR spectra of 25L:$F_{0.5}$:FA: (a) a ^{1}H DQ spectrum recorded with one rotor period of BaBa recoupling; (b) a ^{1}H SQ NOESY spectrum with t_{mix} = 500 ms; and (c) a ^{14}N-^{1}H HMQC spectrum recorded with 8 rotor periods of R^3 recoupling with GIPAW calculated chemical shifts for N···H proximities < 2 Å shown as red crosses. The SQ projection in the F_2 dimension is the ^{1}H MAS one-pulse spectrum. All spectra were recorded at a ^{1}H Larmor frequency of 600 MHz. Base contour levels are at 3.7%, 0.2% and 46.7% of the maximum peak height, respectively. Blue and green contours correspond to positive and negative intensity, respectively. The dashed diagonal lines in (a) and (b) indicate the (a) $\delta_{DQ} = 2\delta_{SQ}$ and (b) $\delta_{SQ} = \delta_{SQ}$ diagonals.

the ^{1}H-^{1}H DQ MAS or ^{14}N-^{1}H HMQC spectra (Fig. 5.11a and c), possibly as a result of this low intensity.

A ^{1}H-^{1}H DQ MAS spectrum (Fig. 5.11a) shows correlation of both H10 (HN) and H14 (one of the OH protons) to methyl protons (δ_{DQ} = 14.3 + 1.7 = 16.0 ppm and δ_{DQ} = 13.7 + 1.7 = 15.4 ppm, respectively), as not only does the substitution of the methyl group at the 2 position on the pyridinium ring result in the NH proximity (like in 52AMP:HF) but the packing of layers within the structure also brings the 5 position methyl group into close proximity to one of the FA molecules (see Table 5.3, above). Some separation is seen between the FA OH ^{1}H chemical shifts in the ^{1}H-^{1}H DQ MAS spectrum, with their correlations to the CH protons in particular suggesting that H14 lies at slightly higher chemical shift, δ^{exp} = 13.7 ppm rather than δ^{exp} = 13.4 ppm, although, due to the overlap with both H12 and H10, it is difficult to place with precision (therefore, although H14 is recorded at 13.7 ppm in Table 5.3, which lists DQ correlations, it has been left assigned to 13.4 ppm elsewhere). The OH-CH cross-peak at the SQ CH region forms an extended diagonal resonance covering the correlations from H12-H1/H13 (δ_{DQ} = 13.4 + 6.4 = 19.8 ppm) through H14-H13, H12/H14-H11 and H14-H2 to H12-H3 (δ_{DQ} = 13.4 + 8.1= 21.5 ppm). This also overlaps with the H10-H3 cross-peak ((δ_{DQ} = 14.3 + 8.1= 22.4 ppm). The corresponding many shouldered cross-peak in the SQ OH region narrows at DQ 21.5 ppm to centre around SQ 13.4 ppm and exclude 13.7 ppm. The similarity in ^{1}H chemical shift of the H10 and H14 results in the cross-peaks for their spatial correlation merging into a single peak on the diagonal (δ_{DQ} = 14.3 + 13.7 = 28.0 ppm).

Fig 5.10c shows that proton transfer is also confirmed for 25L:$F_{0.5}$:FA with the ^{14}N shift δ^{exp} = −150 ppm (δ^{calc} = −141.7 ppm) in a ^{14}N-^{1}H HMQC spectrum (Appendix 3, Table A3.6). This also serves as additional confirmation of the assignment of H10 to the higher shift shoulder rather than the OH protons as the ^{1}H chemical shift of the cross-peak is at δ^{exp} = 14.2 ppm.

A 2D SQ ^{1}H-^{1}H NOESY spectrum (Fig. 5.11b) was used to confirm the presence of the secondary phase. With a mixing time of 500 ms, the resonance at 17.5 ppm has cross-

peaks with both the regions corresponding to the CH (δ^{exp} = 8.0 ppm) and NH group (δ^{exp} = 14.5 ppm) in 25L:$F_{0.5}$:FA as well as a low intensity cross-peak with 12.5 ppm, which also has a corresponding diagonal resonance. This 12.5 ppm resonance is thought to appear as the tail of the OH peak in the ^{1}H MAS one-pulse due to its low intensity. The extremely high ^{1}H chemical shift of three of these resonances (12.5 ppm, 14.5 ppm and 17.5 ppm) suggest that strong H-bonds from NH or OH protons are present. The relatively high chemical shift of the CH region proton may indicate a CH involved in a non-classical H-bond but, the exact composition of the secondary phase being unknown, could also correspond to a solvate-molecule interaction. At 8.0 ppm, it is above the chemical shift region of the H-bonded protons of the water molecules in 26AMP:$F_{0.5}$:$(H_2O)_2$ and 26AMP_2:F:H_2O (discussed in Chapter 4) so a hydrate is unlikely but, although CH moieties are almost certainly present within the phase, they cannot be confirmed as the cause of this cross-peak with confidence. No correlation is seen to the methyl region but this is not in evidence for the NH and OH protons of the main 25L:$F_{0.5}$:FA phase either.

A 2D ^{1}H-^{13}C HETCOR MAS NMR spectrum of 25AMP:$F_{0.5}$:$FA_{0.5}$ is shown in Fig. 5.12, recorded using a CP contact time of 500 μs, including crosses representing the GIPAW calculated chemical shifts for the C-H dipolar correlations up to 2.9 Å. It is important to note the assignment, above, of H10 to a higher ^{1}H chemical shift than H12 and H14, contrary to the order in which they are calculated, due to the calculation of the OH···O protons at a higher chemical shift than seen experimentally, as the crosses in Fig. 5.12 are therefore in the wrong order in the ^{1}H dimension. This assignment is confirmed by the proximity of 2.07 Å between C5 and H10 with a cross-peak at δ^{exp} (^{13}C) = 142.5 ppm and δ^{exp} (^{1}H) = 14.3 ppm. The separation between the two OH protons that was noted in the ^{1}H-^{1}H DQ MAS spectrum is supported by the weak correlation seen between C9 (δ^{exp} = 136.0 ppm) and H14 (δ^{exp} = 13.7 ppm), corresponding to an intermolecular C-H distance of 2.83 Å. The breadth and intensity of the cross peak between the ^{13}C resonance at δ^{exp} = 169.8 ppm and the OH protons confirm that C10 and C12 lie under the same peak, with correlations to H12 and H14, respectively.

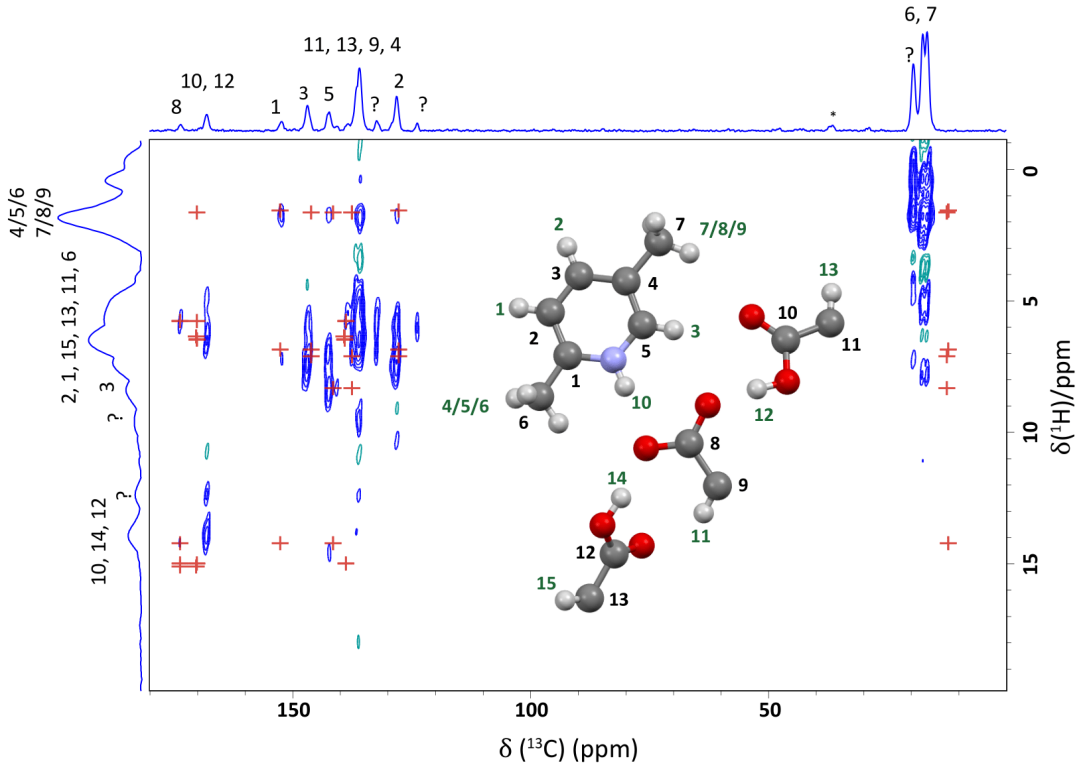


Figure 5.12: A ^{1}H (600 MHz)-^{13}C CP (500 μs) HETCOR MAS (12.5 kHz) NMR spectrum of 25L:$F_{0.5}$:FA recorded using FSLG ^{1}H homonuclear decoupling in t_1 with GIPAW calculated chemical shifts shown as red crosses for C···H proximities < 2.9 Å. The base contour level is at 5.4% of the maximum peak height.

Unsurprisingly, the cross peak from the quarternary C8 (δ^{exp} = 173.5 ppm), the carboxylate carbon, to the carboxylic acid OH protons is missing, corresponding to C-H distances of 2.50 Å and 2.51 Å, with only the closer CH and NH proximities seen: H11 at 2.20 Å and H10 at 2.36 Å. Similarly, the C12 correlation with the methyl protons is also not present experimentally, at 2.84 Å.

The additional correlation peaks, not assigned to 25L:$F_{0.5}$:FA, are listed in Table 5.4. Unlike for the fast MAS spectra, packed in a 1.3 mm rotor, the larger sample volume required to pack a 3.2 mm rotor for the ^{13}C experiments mean that crystal selection to avoid FA was more difficult and it is likely that some was present, as seen in the PXRD presented in Section 5.3.2, above. Environments that could correspond to carboxylate/carboxylic acid, pyridinium/acid CH and methyl groups are present with similar C-H correlations to those in the main phase (as expected for the intramolecular proximities and conserved intermolecular proximities) as well as several distinct regions corresponding to either significant changes in

Table 5.4: CH correlation peaks in the ^{1}H-^{13}C HETCOR spectra of 25L:$F_{0.5}$:FA thought to correspond to a secondary phase (see Fig. 5.11).

$^{13}C\ \delta_{iso}^{exp}$ (ppm)	Possible ^{13}C environments	$^{1}H\ \delta_{iso}^{exp}$ (ppm)	Possible ^{1}H environments
19.7	CH_3	1.6	CH_3
		3.9	CH
123.8	CH	6.7	CH
		5.6	CH
128.1	CH	5.6	CH
		9.6	CH/OH
132.3	CH	5.6	CH
		6.4	CH
136.0	CH	8.8	CH/OH
		12.5	OH/NH
142.6	CH	6.6	CH
146.8	CH	5.6	CH
168.0	$COOH/^{-}$	4.5	CH
		12.5	OH/NH^{+}

packing and intermolecular interactions or the inclusion of an additional component, such as solvent.

25L:$F_{0.5}$:FA was crystallised by slow evaporation from isopropanol but its formation was also attempted by co-grinding of the components in a pestle and mortar, also with a small amount (< 1 mL) of isopropanol. PXRD of the resulting powder sample showed multiple reflections that corresponded to the structures of neither 25L:$F_{0.5}$:FA or FA, most notably at 5.3° (Fig. 5.13). The same effect was seen in PXRD of a sample that was produced at small scale in a 1.3 mm rotor under fast MAS, 60 kHz. In this case, 25L:$F_{0.5}$:FA crystals were selected and packed directly from the crystallisation media without allowing sufficient time for the residual solvent on them to evaporate off. This dampness was immediately evident from the ^{1}H MAS spectrum which showed sharp resonances in the CH and methyl regions (Fig. 5.14, S2 - initial). The 17.5 ppm resonance seen in Fig. 511c was not present. Under fast MAS with no cooling, the spectrum changed over the course of 2 days and resonances at 17.5 ppm (as before), as well as 18.5 ppm, 11.8 ppm and 3.9 ppm formed, in addition to there being distinct changes to the shape/shoulders of the broad NH/OH and CH peaks and the appearance of shoulders to the methyl peak (Fig. 5.14, S2 - final). The sharp resonances vanished, which was taken to mean the residual solvent was no longer present, having evaporated off or been incorporated into the crystal structure of the second polymorph. This suggests that the two

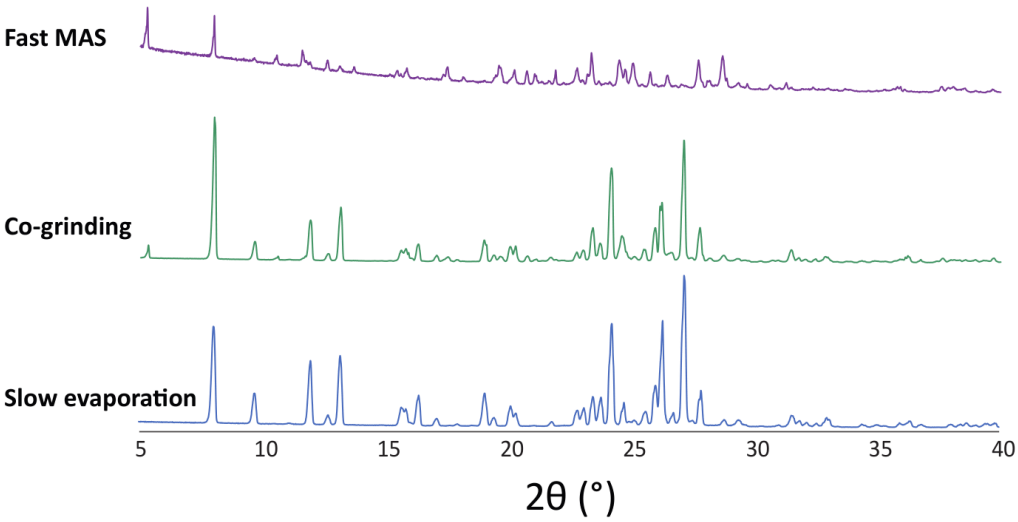


Figure 5.13:PXRD patterns of samples of 25L:$F_{0.5}$:FA: following fast MAS in the presence of isopropanol (top), produced by co-grinding 25L and FA in <1 mL isopropanol (middle) and produced by slow-evaporation (bottom).

forms are related and a transition occurred from 25L:$F_{0.5}$:FA to the second phase. This would explain why PXRD of the bulk sample produced by slow evaporation did not show any indication of the second phase but the ^{13}C MAS NMR, conducted on the same sample at a later time, did.

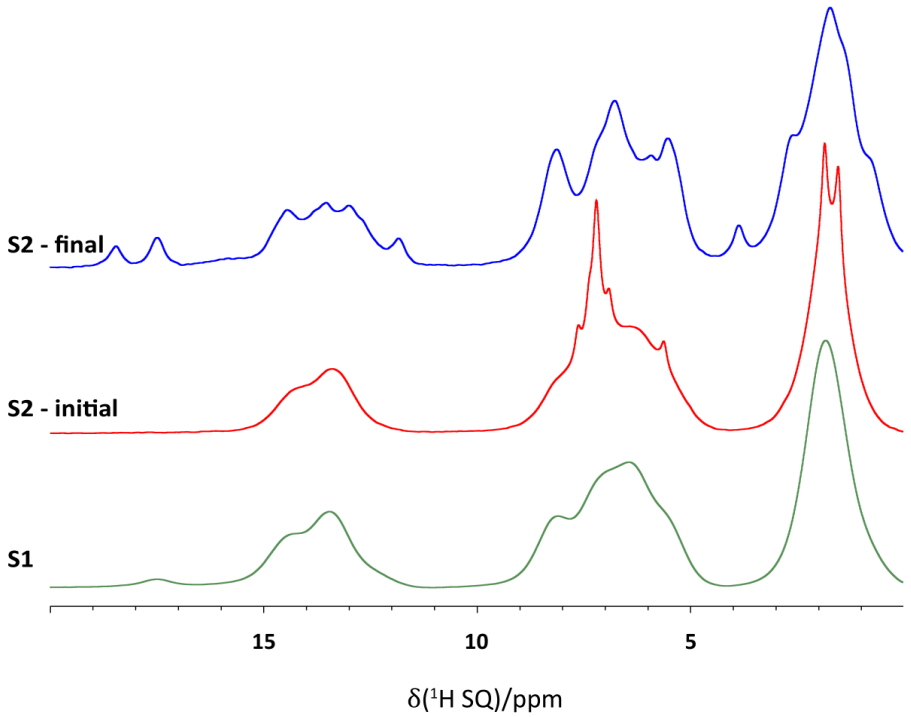


Figure 5.14: One-pulse ^1H (600 MHz) MAS (60 kHz) NMR spectra of 25L:$F_{0.5}$:FA for a dry sample (S1, bottom) and a sample that initially contained residual solvent (S2, middle) and changed under fast MAS (60 kHz) over the course of 2 days (S2, top).

As in the original sample, discussed above, no change/growth was apparent in the ^{1}H resonances despite many days under fast MAS. One possible implication of this is that some solvent is required to facilitate the transition.

5.5. Intermolecular Interactions and Stability

An isolated molecule GIPAW calculation for 52AMP:HF identifies, by means of $\Delta\delta_{Cryst\text{-}Mol}$ exceeding 1 ppm, only the four classical H-bonds (Table 5.5) that were assumed from proximities and angles within the crystal structure (Table 5.2). The OH···O interaction involving H3 is by far the strongest with $\Delta\delta_{Cryst\text{-}Mol}$ = 11.9 ppm and corresponding to both the shortest distance, at 2.50 Å, and the most linear bond, with an OHO angle of 178.1°. This is followed in strength by the pyridinium NH···O at $\Delta\delta_{Cryst\text{-}Mol}$ = 5.5 ppm, which is the next shortest. Even allowing for the over estimation of OH···O ^{1}H chemical shifts (discussed for 26L:HF and $26AMP_2$:F:$FA_{0.5}$ in Chapters 3 and 4 as well as in section 5.4.1, above) it is still nearly twice as strong as the pyridinium interaction. The other two significant interactions correspond to the amino NH···O H-bonds which are both weaker than the pyridinum H-bond with one (H10, which forms cross-links between paired acid chains) substantially so, only forming a weak H-bond with $\Delta\delta_{Cryst\text{-}Mol}$ = 1.3 ppm. Even the more tightly bound amino NH···O H-bond only has $\Delta\delta_{Cryst\text{-}Mol}$ = 3.0 ppm.

Similarly to 52AMP:HF, there are four significant interactions identified by comparison of the full crystal structure with isolated molecule GIPAW calculations for 25AMP:$F_{0.5}$:$FA_{0.5}$ which are all H-bonds corresponding to the same groups, OH···O (which is the strongest at $\Delta\delta_{Cryst\text{-}Mol}$ = 9.0 ppm), pyridinium NH···O which is similar if slightly stronger in strength to that seen in 52AMP:HF (at $\Delta\delta_{Cryst\text{-}Mol}$ = 5.9 ppm, a slightly longer N···O distance for 25AMP:$F_{0.5}$:$FA_{0.5}$ is compensated for by a H-bond angle closer to linearity), and two amino NH···O interactions. The H2 amino NH···O proton has a comparable $\Delta\delta_{Cryst\text{-}Mol}$ to the pyridinium as it is involved in the strong H-bonded $R_2^2(8)$ synthon that stabilises the b-a-b unit

Table 5.5: A comparison of GIPAW calculated ^{1}H shifts (in ppm) for the full geometry optimised crystal structures of 52AMP:HF, 25L:$F_{0.5}$:FA and 25AMP:$F_{0.5}$:$FA_{0.5}$ and for isolated molecules extracted from the geometry optimised crystal structures.

Structure	Atom	δ_{Expt}	$\delta_{Crystal}$	$\delta_{Molecule}$	$\Delta\delta_{Crystal\ -\ Molecule}$
52AMP:HF	**H3**	16.8	18.4	6.5	11.9
	H12	15.0	15.1	9.6	5.5
	H11	7.1	7.0	4.0	3.0
	H10	6.1	5.2	3.9	1.3
	H7	7.1	6.9	6.4	0.5
	H2	6.1	6.3	6.0	0.3
	H1	6.1	6.2	6.1	0.1
	H9	6.1	6.0	6.5	–0.5
	H8	6.1	5.8	6.5	–0.7
	H13/H14/H15	1.0	0.9	1.63	–0.7
25L:$F_{0.5}$:FA	**H12**	13.4	14.8	8.6	6.2
	H14	13.4	14.7	9.0	5.7
	H10	14.3	13.9	9.9	4.0
	H13	6.4	6.2	4.7	1.5
	H11	5.6	5.6	4.4	1.3
	H15	6.3	6.3	5.1	1.3
	H3	8.1	8.2	7.1	1.1
	H1	6.4	6.7	6.7	0.0
	H7/H8/H9	1.7	1.6	1.6	0.0
	H2	7.1	7.0	7.2	-0.3
	H4/H5/H6	1.7	1.6	1.9	-0.4
25AMP:$F_{0.5}$:$FA_{0.5}$	**H19**	14.7	16.0	7.0	9.0
	H1	14.0	14.2	8.3	5.9
	H2	8.8	9.6	4.5	5.0
	H3	6.8	6.7	4.7	2.0
	H6	6.8	7.1	6.5	0.6
	H23	6.1	6.1	5.8	0.3
	H4	6.1	5.5	5.8	-0.3
	H20	6.1	6.0	6.4	-0.4
	H5	6.8	6.4	6.9	-0.4
	H7/H8/H9	0.8/	0.4	1.3	-0.9

whereas in 52AMP:HF, with the amino group on the other side of the molecule, the amino and pyridinium interactions compete with each other rather than strengthening the same structural unit. The other 25AMP:$F_{0.5}$:$FA_{0.5}$ amino NH···O, H3 which is not involved in the b-a-b unit, is likewise considerably weaker ($\Delta\delta_{Cryst\text{-}Mol}$ = 2.0 ppm) indicating that it is of lower importance to the primary structure formation. There is evidence of a potential weak π-interaction ($\Delta\delta_{Cryst\text{-}Mol}$ = −0.9 ppm) involving the protons of the 25AMP:$F_{0.5}$:$FA_{0.5}$ methyl group, as they point towards the centre of the pyridine ring of the layer below – for fixed

atoms, rather than the dynamic methyl group, this would likely have been a stronger interaction.

Crystal to isolated molecule GIPAW calculations for 25L:$F_{0.5}$:FA indicate a higher number of H-bonds (for $\Delta\delta_{Cryst\text{-}Mol} > 1$) but none of the same strength as those identified in the other systems. Its two OH···O interactions are more comparable to the pyridinium NH···O H-bonds seen in 52AMP:HF and 25AMP:$F_{0.5}$:$FA_{0.5}$, with $\Delta\delta_{Cryst\text{-}Mol}$ of 6.2 ppm and 5.7 ppm for H12 and H14, respectively, despite their physical parameters appearing similar in both distance and angle to the OH···O interactions in 25AMP:$F_{0.5}$:$FA_{0.5}$ (taking the aforementioned overcalculation of OH···O protons in to account, they are actually slightly weaker). The pyridinium NH···O interaction is also weaker at $\Delta\delta_{Cryst\text{-}Mol}$ = 4.4 ppm. Unlike 52AMP:HF and 25AMP:$F_{0.5}$:$FA_{0.5}$, however, 25L:$F_{0.5}$:FA also has four weak H-bonds with non-classical CH donors. The three strongest of these CH···O weak H-bonds (one at $\Delta\delta_{Cryst\text{-}Mol}$ = 1.5 and two at $\Delta\delta_{Cryst\text{-}Mol}$ = 1.3) correspond to the fumarate and FA CHs and provide stabilisation between the two inequivalent acid chains. The weakest case here is from an aromatic CH, H3.

Given that 52AMP has a higher melting point than 25AMP (95 °C compared to 76 °C for 25AMP) and 52AMP:HF has a notably stronger OH···O H-bond than 25AMP:$F_{0.5}$:$FA_{0.5}$, one might expect 52AMP:HF to have a higher melting point. However, when the relative strength of the other three H-bonds within the two systems are considered, it is less surprising that 25AMP:$F_{0.5}$:$FA_{0.5}$ is more thermally stable, melting at 182 °C, 20 °C higher than 52AMP:HF (Fig. 5.15). Similarly, despite having the most independent H-bond motifs, the interactions in 25L:$F_{0.5}$:FA are all relatively weak compared to those within both 52AMP:HF and 25AMP:$F_{0.5}$:$FA_{0.5}$, so it is understandable that it melts at a much lower temperature (128 °C). As 25L has a drastically lower melting point than the other base molecules (−15 °C), it would have been expected to have the lowest thermal stability without analysis of the interaction strengths. It does, however, show the largest increase in stability when co-crystallised with FA compared to its base in isolation out of the three bases.

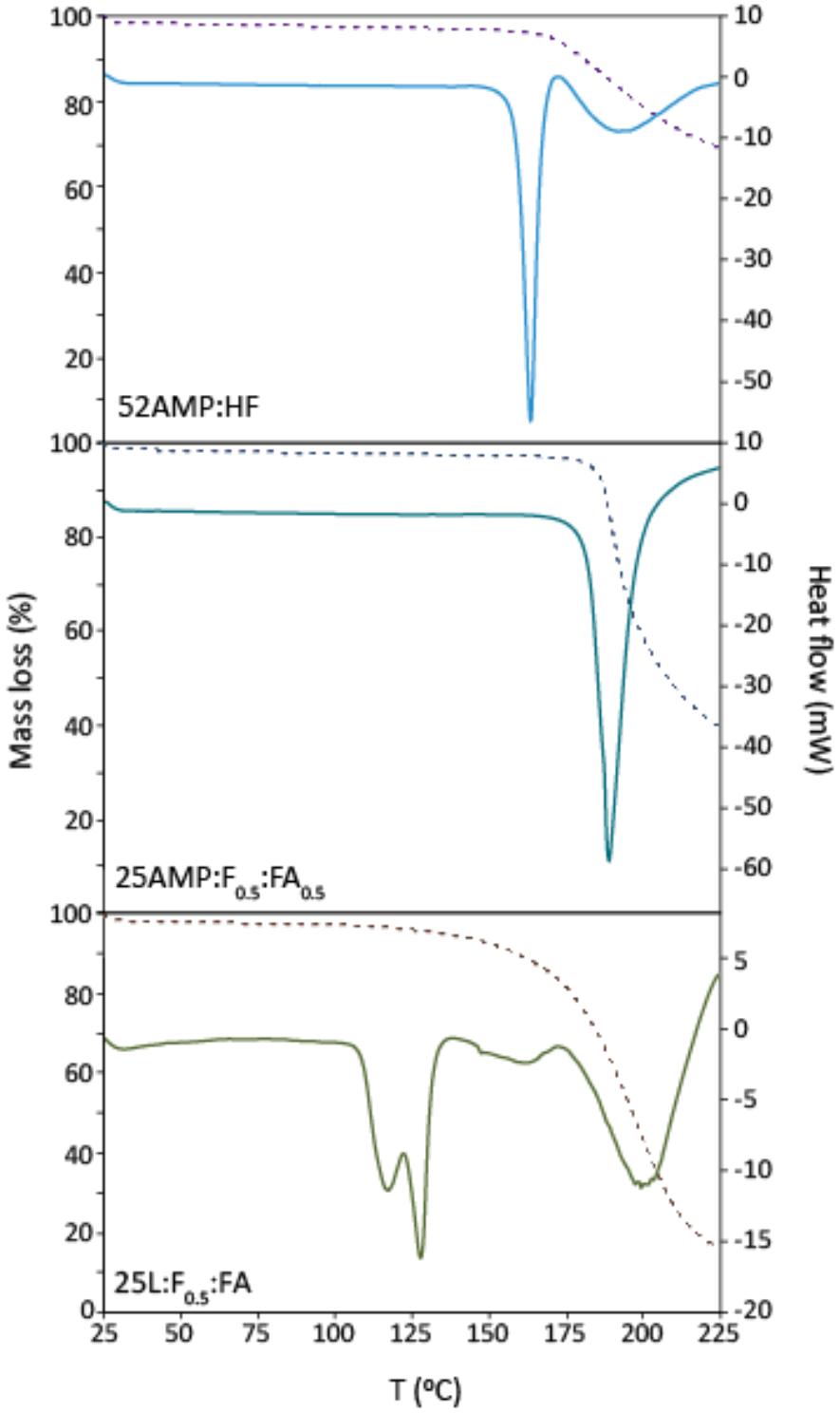


Figure 5.15: DSC (solid lines) and TGA (dashed lines) of 52AMP:HF, 25AMP:$F_{0.5}$:$FA_{0.5}$ and 25L:$F_{0.5}$:FA recorded on a Mettler Toledo Star^e instrument with a ramp of 10 °C/min from 25-225 °C.

It is also interesting to note that, although it occurs below the boiling points of the bases, both 52AMP:HF and 25AMP:$F_{0.5}$:$FA_{0.5}$ begin to lose mass once they melt, 25AMP:$F_{0.5}$:$FA_{0.5}$ more rapidly than 52AMP:HF, which also shows evidence of a further liquid-liquid transition post-melting. This is not unexpected given the apparent evaporation

below the base boiling point exhibited by both of the related bases 26L and 26AMP (see Chapter 3 and 4). What is interesting is that this is not seen for 25L:$F_{0.5}$:FA, with mass loss due to base evaporation not occurring significantly until a liquid-liquid transition beginning at 175 °C, well above the system's melting point. This is above the boiling point for 25L (157 °C, although a small dip indicating an initial liquid-liquid transition is also centred on this temperature). This difference may be due to the fact that 25L:$F_{0.5}$:FA is the only crystal structure in which there is a lower proportion of base than acid, with a ratio of 1 : 1.5 where the next lowest is 1 : 1 (base : acid). As FA is significantly more stable, with a melting point in its single component form of 287 °C, it is possible that when in excess it traps the base molecules within the melt more effectively until the higher transition is reached.

It is also noted that the pre-melting minima/shape of the melting dip may be due to the range of crystallite sizes present within the sample, with the substantial split in 25L:$F_{0.5}$:FA due to two well distinct particle sizes – alternatively this may be due to impurities.

5.6. Summary

The crystallisation of 52AMP with FA results in a 1:1 salt, 52AMP:HF, based upon a hydrogen fumarate acid chain akin to that seen for 26L:HF,[139] although with the hydrogen fumarates in the two s-*cis* conformation rather than the slightly higher energy s-*cis*/s-*trans* conformation. This is distinctly different to the form of the related 25AMP and 25L systems, which both form cocrystals of salts with strong b-a-b units, as seen for all of the 26AMP systems, linked by fumarate/FA acid chains. This difference is thought to be due to the position of the amino group on the opposite side of the pyridinium ring to the pyridinium N, introducing a set of H-bond interactions competing with the acid-base H-bond rather than supporting it, as in 25AMP:$F_{0.5}$:$FA_{0.5}$.

As discussed in chapters 3 and 4, protons involved in OH···O interactions were once again found to be calculated 1.3-1.6 ppm higher than observed experimentally for 52AMP:HF,

25AMP:$F_{0.5}$:$FA_{0.5}$ and 25L:$F_{0.5}$:FA, while similarly high chemical shift NH protons had good agreement between GIPAW calculations and experimental peaks. This further evidence of a consistent discrepancy supports the view that there is a systematic error in the calculation of this moiety. Analysis of the 25AMP:$F_{0.5}$:$FA_{0.5}$ ^{13}C chemical shifts also lends weight to the possibility of a systematic error in the calculation of the certain carbon environments, namely a quaternary carbon bonded to two nitrogen atoms (discussed in Chapter 4). Other than these particular ^{1}H and ^{13}C environments, the agreement between experiment and calculation for 25AMP:$F_{0.5}$:$FA_{0.5}$ was very good, providing good support for the crystal structure.

For 52AMP:HF, the error between the GIPAW calculated ^{13}C chemical shifts and the experimentally observed resonances was higher than expected despite excellent agreement for the ^{1}H chemical shifts. Although the average absolute error is 2.5 ppm, not much higher than the anticipated error of ~2 ppm, the largest difference (excluding the methyl carbon which differs due to temperature effects) is −3.7 ppm. They are therefore all significantly smaller than the systematic ^{13}C discrepancy in 25AMP:$F_{0.5}$:$FA_{0.5}$ (5.9 ppm) but have no obvious cause.

Analysis of 25L:$F_{0.5}$:FA provided evidence not only for the common presence of recrystallised FA within samples but also for the existence of a second polymorph. A study of evolution of the ^{1}H chemical shifts over two days under fast MAS (60 kHz) suggests that these forms are related as a transition from 25L:$F_{0.5}$:FA to the second form is apparent. However, conversion does not seem to occur in completely dry samples but only in the presence of a solvent. This may explain why conversion was not complete in the sample S2 (Fig. 5.14) as it stopped when the residual solvent had evaporated off. The second polymorph may require an additional energy input to form preferentially as it is only seen in low quantities from a slow evaporation crystallisation but a higher proportion formed from co-grinding and under fast MAS.

Chapter 6. Identifying Crystal Forms and Structural Patterns

6.1. Introduction

In this chapter, the 7 systems that have been thoroughly characterised in Chapters 3-5 (Fig. 1.5, reproduced at the end of this thesis) are discussed terms of their relative stability, crystal form and structural components. They form the basis of a wider investigation into the crystal form and structure of related multicomponent crystals within the CSD.

As new forms of an API or AI, whether of different crystallinity, polymorph or components, constitute an IP opportunity,[13] their correct characterisation is of importance, to ensure correct regulation and to allow both filing and successful protection of patents. The potential for similarities in molecular geometry between crystal forms can make this particularly challenging,[108, 109] especially as the hydrogen atoms, the association of which are often the defining factor, are almost invisible to XRD due to their low electron density. Even in cases where structural features or molecular geometries may aid in the determination, the difficulty of producing high quality single crystals, particularly from greener, solvent free chemistry methods such as cogrinding,[85-87] mean that a SXRD may not be an option, and single crystal neutron diffraction impossible, especially during early API development phases.

As discussed in Chapter 1, structural features (motifs/synthons/*etc.*) are often conserved within families of related molecules such that they can be a useful indicator of preferred packing and stability. Variations in packing within or between features can act as an indirect confirmation of crystal form (salt/cocrystal). For example, the difference in C-O and C=O bond lengths is commonly used to distinguish between carboxylate and carboxylic acid groups with similar bond lengths taken as confirmation of proton transfer.[38, 139, 218] Trends within the CSD can also be used for crystal engineering, to try to target desirable features by inclusion of certain molecular moieties associated with key biophysical properties. These can be both architectural aims, *i.e.* formation of a synthon as seen with Swapna *et al.* designing

cocrystals of Isoniazid (a tuberculosis drug),[65] and form related, for instance reducing the chance of hydrate formation by incorporating pyridine based compounds.[140, 217]

As one of the key biophysical properties targeted for improvement by salt and cocrystal formation is the thermal stability,[9] this chapter begins by comparing the stabilities of the systems already presented in this work, both to highlight which of the three forms, salt, salt hydrate or cocrystal of salt, is more thermally stable in absolute terms and also which systems exhibit the largest gains in stability relative to their parent base compound. The commonality and variation in the H-bond networks and structural patterns that support each system is examined. The features identified are then used as the starting point for an investigation into the structural trends in the CSD for both fumarate/FA and succinate/succinic acid based multicomponent crystals. Also presented is a potential solid-state NMR method using the ^{14}N shift for confidently determining the ionicity or neutral nature of a pyridine containing crystal form which could be either in the absence of/prior to of a full crystal structure or as an additional verification (as it has been used throughout this work).

All structures downloaded from the CCDC are referred to throughout by their CSD codes.

6.2. Crystal Form Stability

As was discussed in Chapter 1, the effect of forming multicomponent systems on a product's melting point can vary significantly, but most commonly the resultant system has an intermediate value, between the melting points of the individual components.[76] Looking at Table 1.1 and 6.1, this is seen to be true of all of the core systems, with the high melting point of FA dramatically increasing the thermal stability of all the bases. It is interesting that those systems that undergo solid-solid (S-S) transitions prior to melting (see Table 6.1) also do this at a higher temperature than they would have melted in isolation. It is important to note that the TGA/DSC analysis does not necessarily capture the slower acting instabilities (for

example, the slow loss of 26L from 26L:HF, which occurs even at room temperature for microcrystalline samples).

Despite 52AMP having the highest melting point of the bases, 52AMP:HF has only the second lowest melting point. It is notable in fact that the three systems that exhibit the most stability with increasing temperature are the cocrystal of a salt forms, with the two hydrogen fumarate salts melting at lower temperatures. One of the key structural differences between the two forms is the formation of the b-a-b unit and the form of the acid chain (which alternates between doubly ionised and neutral molecules in the cocrystals of salts with the doubly ionised molecules central to the b-a-b unit, as discussed below in section 6.4). It is possible that the combined b-a-b/acid chain provides more stability than the hydrogen fumarate chains seen in the salt systems. The chosen interpretation of the 25L:$F_{0.5}$:FA DSC (see Chapter 5) indicates it melts at a higher temperature than 26L, despite having the lowest melting point out of all the bases. 25L:$F_{0.5}$:FA does, however, undergo a solid-solid transition around the same point that 26L:HF melts. An alternative interpretaton of the DSC is that this is followed by a second solid-solid transition at 125°C during which it becomes an amorphous rather than crystalline solid, explaining the broadness of its 'melting point' at ~200°C (see Chapter 5, noting that VT-PXRD was not performed for this system to verify this).[20] Its increased stability may be due in part to the component ratio as 25L:$F_{0.5}$:FA has the highest proportion of FA of all the local systems. As both 25L:$F_{0.5}$:FA and 26AMP$_2$:F:$FA_{0.5}$ convert to unknown structures before the point of melting, it is difficult to infer what impact the known structural patterns (discussed below in section 6.4) have on the thermal stability.

Table 6.1: Thermal properties of each multicomponent system.[a] The number of runs for each is given by n.

Structure	Base melting point (°C)	S-S 1 (°C)	S-S 2 (°C)	S-S 3 (°C)	Melting point (°C)	*n*
25AMP:$F_{0.5}$:$FA_{0.5}$	76-77[133]	-	-	-	185±5	2
25L:$F_{0.5}$:FA	-15[136]	117±12	-	-	125±8	3
52AMP:HF	95-99[135]	-	-	-	158±7	3
26L:HF	-6[211]	-	-	-	114±4	2
26AMP:$F_{0.5}$:$(H_2O)_2$ / 26AMP$_2$:F:$FA_{0.5}$	40-44[131]	73±5	106±10	150±8	186±3	7[b]

[a]The solid-solid (S-S) transitions were identified by TGA/DSC analysis (see Figures A1.6, 4.10 and 5.14)
[b]n = 5 for the first two solid-solid transitions

Both salt hydrates, precursors to $26AMP_2$:F:$FA_{0.5}$, are less thermally stable than the hydrogen fumarate salts. These both possess b-a-b units, but they have no acid chain providing stabilisation between them, with water molecules fulfilling that role instead. The transitions believed to lead to conversion to $26AMP_2$:F:$FA_{0.5}$ both occur below the melting point of 26L:HF (114°C) and the first solid-solid transition of 25L:$F_{0.5}$:FA (117°C). Unsurprisingly, as these both involve dehydrations, the second occurs at 106°C, just above the boiling point of water. This second transition also involves the loss of some 26AMP.

6.3. GIPAW Calculations of NMR Parameters

6.3.1. ^{14}N Shifts

As has been previously reported, there is a significant change in ^{15}N chemical shift depending on whether or not the nitrogen is protonated to become positively charged.[219-224] The value and range of chemical shifts observed depends on what type of N moiety is involved (*e.g.* amine or amide). A significant difference is also expected for ^{14}N shifts, although the exact separations are not so commonly reported, nor the ranges within which each form falls.[225] This is in part due to the fact that ^{14}N is quadrupolar, meaning that the shift observed in the solid-state differs depending on the strength of the magnetic field (see Eqn. 120 in section 2.2.5.3), which also serves to change the observed separation and shift. However, ^{14}N has an isotopic abundance of 99.6 % meaning it can be interrogated without the need for isotopic labelling.

Each of the systems within the core set possesses a pyridinium nitrogen, the protonation of which is integral to determining their crystal form. In addition to the core set, 46 more multicomponent crystal structures containing a pyridine derivative were downloaded from the CCDC and DFT calculations performed. The same cut-off energy, exchange correlation functional and k-point spacing was used for all calculations. Following geometry optimisation, GIPAW calculations were performed and the ^{14}N calculated shifts determined

for both 14.1 T and 16.4 T. Referencing here is performed with $\sigma_{ref} = -153$ ppm, corresponding to liquid CH_3NO_2 at 0 ppm.[171, 177] These were then compared to the N···H distance, normalised to the N···X distance (where X is the H-bond donor, mostly O but also N and C).

A further 10 structures containing non-pyridine cyclic/tertiary amines were also included to increase the number of such environments, allowing investigation of how generally the pyridine ranges can be applied (giving a total of 58 additional structures, listed in Table A4.1 in Appendix 4). As many structures contained multiple N environments, this resulted in 68 pyridine/pyridinium nitrogen atoms and 26 alternatives, including nitrogens in:

- a tertiary amine incorporated into 6-membered ring that includes a second N
- a tertiary amine incorporated into 6-membered ring that includes an O
- a secondary amine incorporated into 6-membered ring that includes an O
- a tertiary amine
- a tertiary amine incorporated into 5-membered ring containing multiple N

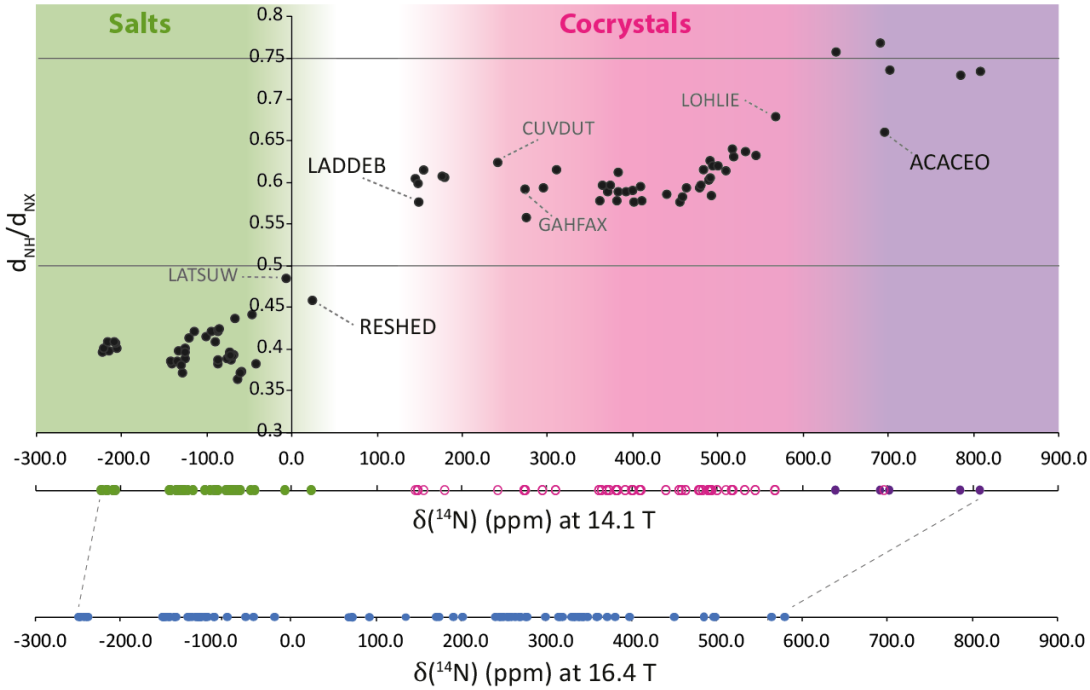


Figure 6.1: A plot of d(N···H)/d(N···X) against the ^{14}N GIPAW calculated shifts at 14.1 T for 86 chemically distinct nitrogen environments in the 58 structures from the CSD (see Table A4.1) and the systems studied in this work.

For the nitrogen of a pyridine/pyridinium, there is a marked difference between the calculated ^{14}N shift when considering unprotonated cocrystals (high shift) or protonated salts (low shift) (Fig. 6.1). At 14.1 T, a ^{14}N Larmor frequency of 43.4 MHz, the smallest difference between an ionic and neutral interaction is 126 ppm, including both salt and cocrystal outliers. All of the analysed N environments are shown on Fig. 6.1, making it clear that similar ^{14}N shift ranges are also appropriate for the moieties discussed. However, the cocrystal systems with the lowest ^{14}N shifts consist of 5 of the 5-membered ring nitrogens, between δ_{iso}^{calc} = 144-178 ppm, and LADDEB, at δ_{iso}^{calc} = 148 ppm, these may be considered special cases.

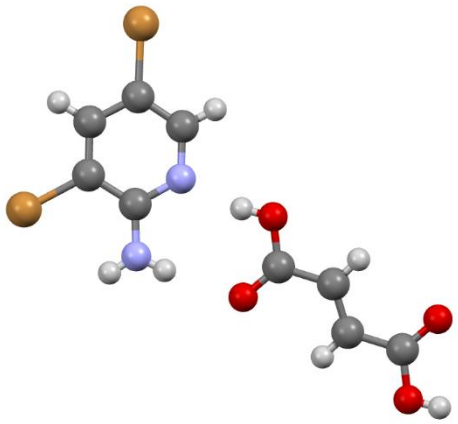


Figure 6.2: Asymmetric unit of LADDEB

LADDEB has two Br atoms substituted at the 3 and 5 positions which are likely the cause in the reduction of the ^{14}N shift (Fig. 6.2). Whether this is due to a genuine chemical impact due to the electronegativity of Br or the result of inaccuracies in the DFT calculations (as the pseudo-potentials for Br are far less well developed) is unclear. Other than LADDEB, the low ^{14}N shift cocrystals correspond to 3 stuctures with 5-membered rings that each contain more than one N atom. Although overall this N moiety, for which there are 8 calculated shifts in total, does follow largely the same trend as the other N environments, the range of ^{14}N shifts for cocrystal interactions show more variation in ^{14}N shift, for the same centrality of the proton (d_{NH}/d_{NX} = ~0.6), than the other N environments, extending down to the same region as LADDEB (< 200 ppm) and up towards the highest H-bonded neutral N (> 500 ppm).

Discounting both LADDEB, as an exception, and the 5-membered ring nitrogens, as falling into a slightly different range, leads to an increased separation of 220 ppm between RESHED and CUVDUT. This is very significant and suggests these ranges may provide a quick way to distinguish between salt and cocrystal interactions for both pyridines and several other similar environments with confidence.

It can also be seen from Fig. 6.1 that it may also be possible to determine whether the N is involved in H-bonding. There is overlap between non-interacting N sites and those with H-bonding since the highest H-bonded N (ACACEO at δ_{iso}^{calc} = 695 ppm) has a ^{14}N shift 128 ppm higher than the rest. This N is H-bonded to an amino N, as well as being the second most distant acceptor, and is the only example of this interaction within the set. The next highest H-bonded N is the LOHLIE N at d_{NH}/d_{NX} = 0.68, which also corresponds to a non-classical H-bond to a CH, followed by systems that have either 2 N within a six-membered ring or also have an N···HN interaction, just not to an amino group. Other than the interacting N at 695 ppm, all the H-bonded neutral N fall below 600 ppm and the non-interacting ones are all higher than it.

The highest salt N belongs to RESHED, δ_{iso}^{calc} = 22 ppm, the published structure for which is a cocrystal.[140] The proton within the key H-bond in the literature structure is almost central between the N and O atoms, d_{NH}/d_{NO} = 0.56, although no closer than seen for RESFOL. During geometry optimisation of RESHED, the proton migrated over the centre point to sit closer to the N atom, d_{NH}/d_{NO} = 0.46 (Fig. 6.3). This is somewhat surprising as this proton within the crystal structure was found in the SXRD analysis rather than being placed. The calculated shift is also still unusually high for the proton's final position given the absence of any uncommon functional groups within the system. The crystallisation of RESHED, published as a 3,5-lutidine FA cocrystal, was attempted but single crystals were not obtained. The resulting powder, 35L:F, was determined to be most likely multiple phases by PXRD, possibly containing both crystalline FA and a phase resembling RESHED but additional reflections were also present (see Appendix 4, Fig. A4.1). A 2D ^{14}N-^{1}H (600 MHz) HMQC

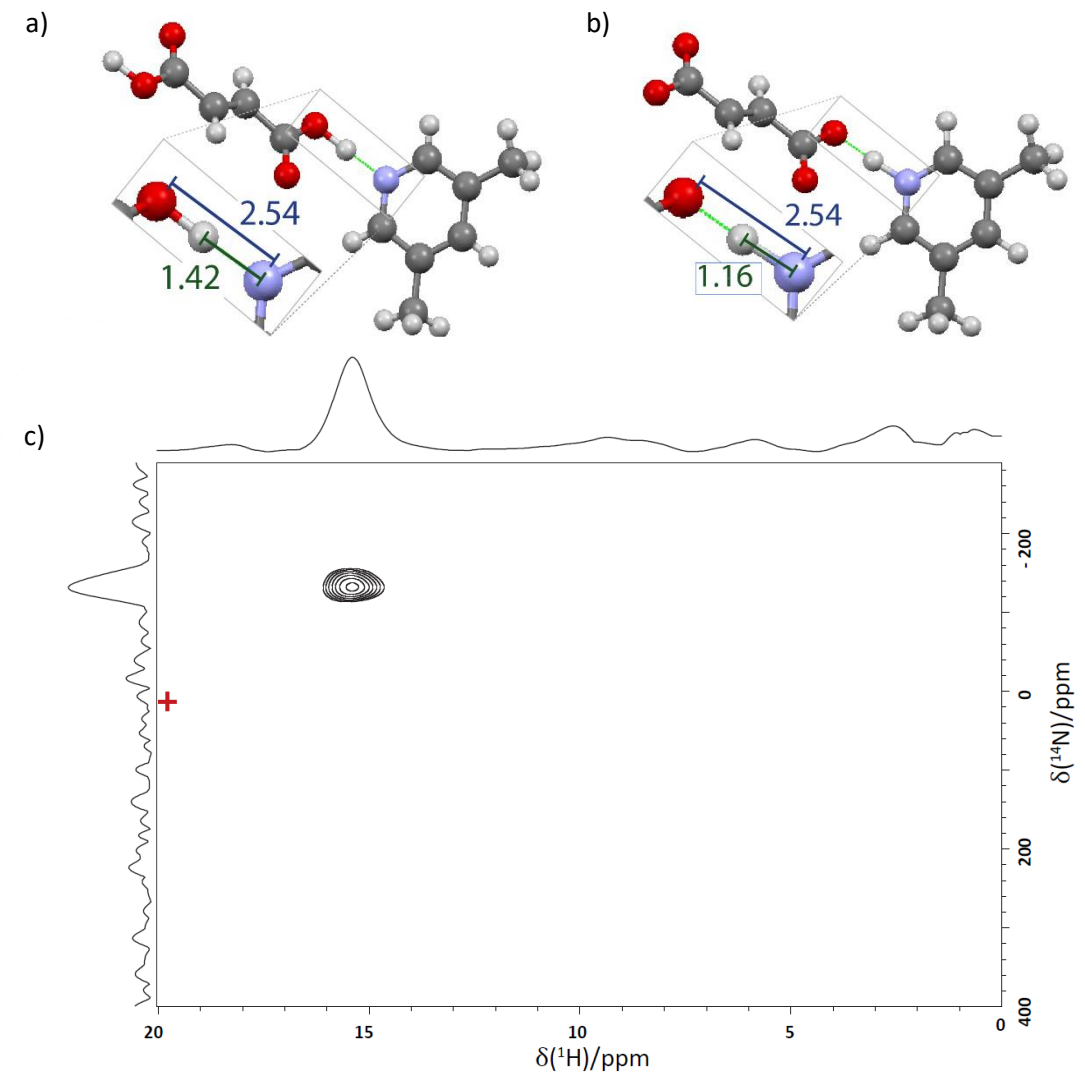


Figure 6.3: Asymmetric unit of RESHED (a) as published by Haynes et al. and (b) following geometry optimisation alongside (c) a ^{14}N-^{1}H (600 MHz) HMQC MAS (60 kHz) spectrum of P35L:F, recorded with 8 rotor periods of R^3 recoupling (τ_{RCPL} = 133.6 μs).

spectrum contained only a single correlation which lay at δ_{iso}^{exp} = −131 ppm, suggesting that at least one of the phases present contained 3,5-lutidinium in a salt form (Fig. 6.3c). It is unlikely that a non-H-bonded pyridine N would be visible within this experiment so the presence of additional 3,5-lutidine containing phases cannot be entirely ruled out. Other than RESHED, the highest salt ^{14}N shift is LATSUW at −8 ppm with a nearly central H-bond proton (d_{NH}/d_{NO} = 0.48). If RESHED is also discounted, the salt-cocrystal separation rises to 250 ppm. A further slight increase is apparent if, rather than the more general case, only the pyridine/pyridinium N are considered with 260 ppm between LATSUW and GAHFAX.

As ^{14}N is quadrupolar and the shift therefore field dependent, it was also determined for 16.4 T, the other field strength of magnet on which several ^{14}N experiments within this work were conducted. It can be seen (Fig. 6.1) that all the calculated shifts move to a lower ppm but with a corresponding narrowing in range. Discounting LADDEB and 5-membered rings containing more than one N, the separation between salt and cocrystal narrows from 250 ppm to 192 ppm. This is still distinct but suggests that for clarity it may be preferable to run experiments at 14.1 T instead.

Based on these calculations, the general rule can be determined that, at 14.1 T, pyridinium N atoms (or equivalent) will have an ^{14}N shift below 0 ppm, pyridine N atoms involved in H-bonding will fall above 250 ppm and those not H-bonded above 600 ppm, with similar N moieties following similar trends. Exceptions may occur for dynamic H-bonds or in close proximity to very electronegative atoms. This potentially represents a simple method to determine the form of a multicomponent crystal in the absence of single crystals of high enough quality to locate the protons and to validate the form of XRD crystal structures where they were found.

6.3.2. Discrepancies: N=C-N ^{13}C and OH···O ^{1}H chemical shifts

Throughout this work, discrepancies have been identified for two specific chemical environments that show significantly larger differences between their GIPAW calculated and experimental chemical shifts than the expected 1% of the chemical shift range. One discrepancy is for ^{1}H in an OH···O hydrogen bond (which occurs in crystalline FA as well as in most of the multicomponent systems) and the other discrepancy is for a quaternary ^{13}C which is covalently bound to both a pyridinium nitrogen and an amino nitrogen (Fig. 6.4).

For the proton discrepancy, the level of variation seen in each structure is relatively consistent, with $\Delta\delta^{exp\text{-}calc}$ ranging from −1.9 to −1.2 in 26L:HF and FA, respectively. In

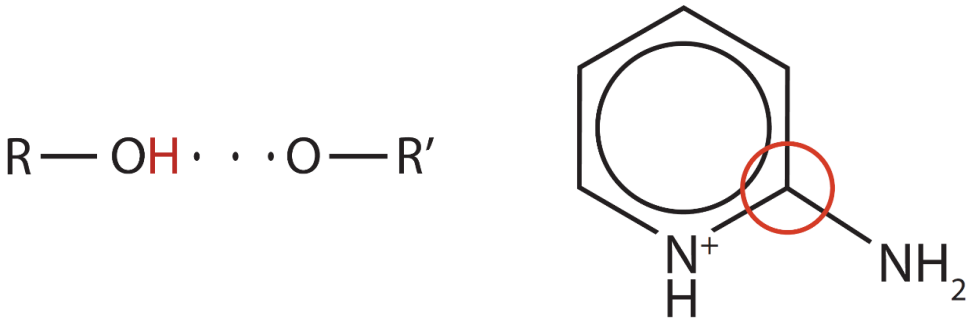


Figure 6.4: Chemical structures of the two environments which show large discrepancies between GIPAW calculated and experimental chemical shift: a ^{1}H in a strong OH···O hydrogen bond (left) and a quaternary ^{13}C between a pyridinium nitrogen and an amino nitrogen (right).

comparison to ^{1}H in an OH···O H-bond, the $NH^{+}\cdots{}^{-}O$ environment, found in all the multicomponent systems, is at a similarly high chemical shift ^{1}H but in each case shows good agreement with the GIPAW calculated chemical shift (Table 6.2). This suggests that the OH···O discrepancy is not simply explained by the known temperature dependence of H-bonded chemical shifts, due to the change in shielding between the excited vibrational states of H-bonds, as this would also effect the $NH^{+}\cdots{}^{-}O$ proton.[226-233]

Table 6.2: GIPAW calculated and experimental ^{1}H chemical shifts (in ppm) for the OH and NH^{+} moieties in each of the systems discussed in chapters 3-5 and FUMAAC.

System	OH			NH		
	δ^{exp}	δ^{calc}	$\Delta\delta^{exp-calc}$	δ^{exp}	δ^{calc}	$\Delta\delta^{exp-calc}$
25L:$F_{0.5}$:FA	13.4/13.4	14.7/14.8	–1.3/–1.4	14.3	13.9	0.4
25AMP:$F_{0.5}$:$FA_{0.5}$	14.7	16.0	–1.3	14.0	14.2	0.2
52AMP:HF	16.8	18.4	–1.7	15.0	15.1	–0.1
26L:HF	15.8	17.7	–1.9	17.7	17.7	0.0
26AMP:$F_{0.5}$:$(H_2O)_2$	-	-	-	14.9	14.8	0.1
26AMP_2:F:H_2O	-	-	-	14.8/14.8	15.0/15.0	–0.2
26AMP_2:F:$FA_{0.5}$	14.9	16.3	–1.4	14.4/16.4	14.7/16.0	–0.3/0.4
FA[a]	12.9[208]	14.1	–1.2	-	-	-

[a]The experimental value for FA was taken from literature[208]

It is interesting to note that FA, which contains an entirely neutral carboxylic acid/carboxylic acid hydrogen bond, exhibits the smallest discrepancy. By contrast, the salt systems all have carboxylic acid/carboxylate interactions.

For the carbon discrepancy, $\Delta\delta^{exp-calc}$ shows a positive difference between experiment and GIPAW calculation when the quaternary ^{13}C is covalently bound to both a pyridinium nitrogen and an amino nitrogen (Table 6.3). The quaternary carbons that sits at the analogous

position in both 25L:$F_{0.5}$:FA and 26L:HF, directly bound to the pyridinium nitrogen, show excellent agreement between experiment and calculation with the largest discrepancy only 0.2 ppm. As these are substituted with a methyl group rather than an amino group, the combination of amino and pyridinium interactions is considered to be the origin of the discrepancy. 52AMP:HF did exhibit a higher difference for the 2-methyl but not as large as seen for the 2-amino systems (its own 5-amino shows a smaller difference of $\Delta\delta^{exp\text{-}calc}$ = 1.2 ppm) and, as all the ^{13}C chemical shifts showed poor agreement for this system, it is therefore not considered relevant to the discussed discrepancy.

Table 6.3: GIPAW calculated and experimental ^{13}C chemical shifts (in ppm) for the quaternary carbons at the 2-position on the pyridinium ring in each of the systems discussed in chapters 3-5.

System	^{13}C δ^{exp}	^{13}C δ^{calc}	$\Delta\delta^{exp\text{-}calc}$
26L:HF	152.6/152.6	152.4/152.6	0.2/0.0
25L:$F_{0.5}$:FA	152.4	152.5	0.1
52AMP:HF	138.2	135.4	2.8
25AMP:$F_{0.5}$:$FA_{0.5}$	153.8	147.9	5.9
26AMP:$F_{0.5}$:$(H_2O)_2$	155.5	151.3	4.2
26AMP_2:F:$FA_{0.5}$	156.7/157.4	152.6/153.2	4.1/4.2

No change in the ^{13}C chemical shift was observed when recorded at different fields (see Appendix 4, Fig. A4.3), ruling out a shift of the ^{13}C chemical shift due to enhanced second-order quadrupolar effects from the two adjacent ^{14}N atoms. There are very few examples in the literature of such large discrepancies for ^{13}C: one example is that in 2006, Harris reported for the quaternary C5 site (fused between two 6-membered ring with one C=C and two C-C bonds) in testosterone GIPAW calculation at 182.6 and 182.7 compared to 170.6 and 172.0 ppm experimentally for the two distinct molecules in the asymmetric unit cell.[234]

6.4. Structural Motifs for the Systems Studied in this Thesis

The acid-base interaction, discussed in section 6.2.1 for either a neutral or ionic N-H···X H-bonding interaction, is central in an ionic form to all seven of the systems but its relative strength differs greatly between them due to the variation in H-bond geometry. This is

Table 6.4: Structural parameters for the H-bonds identified in the isolated molecule calculations performed for each system.

	$\Delta\delta_{Crys-Mol}$ (ppm)	N···O distance (Å)	$NH^+\cdots{}^-O$ angle (°)	NH···O angle (°)	O···O distance (Å)	OH···O angle (°)	OH···$^-$O angle (°)
52AMP:HF	**5.5**	2.65	160.4	-			
	3.0	2.92	-	160.1			
	1.3	2.98	-	163.3			
	11.9				2.50	178.1	-
25L:$F_{0.5}$:FA	**4.0**	2.78	173.1	-			
	6.2				2.59	176.0	-
	5.7				2.60	-	167.3
25AMP:$F_{0.5}$:$FA_{0.5}$	**5.9**	2.73	167.2	-			
	5.0	2.83	-	174.8			
	2.0	2.83	-	153.6			
	9.0				2.56	-	169.5
26L:HF	**7.1**	2.64	169.1	-			
	11.3				2.54		175.0
26AMP:$F_{0.5}$:$(H_2O)_2$	**6.5**	2.69	172.2	-			
	3.6	2.85	-	176.2			
	2.9	2.94	-	170.8			
	6.0				2.79	173.3	-
	6.7				2.76	171.0	-
	5.2				2.86	167.9	-
	5.8				2.85	179.2	-
$26AMP_2$:F:H_2O	**6.6**	2.72	174.3	-			
	4.8	2.79	-	170.9			
	3.7	2.80	-	175.8			
	6.6	2.71	177.2	-			
	6.0	2.81	-	169.7			
	3.7	2.86		163.9			
	5.1				2.82	176.3	
	4.8				2.85	170.8	
$26AMP_2$:F:$FA_{0.5}$	**7.5**	2.69	174.1	-			
	7.2	2.76	-	175.9			
	3.0	2.88	-	156.1			
	6.4	2.76	172.5	-			
	4.4	2.82	-	156.5			
	4.3	2.78	-	159.9			
	9.1				2.55	-	169.2

influenced by the other intermolecular interactions present, with the optimal geometry of each interaction balanced against the others to give the overall configuration. As the position of substitution has a significant impact on which intermolecular interactions are present it therefore shows a direct link with the ^{1}H $\Delta\delta_{Crys-Mol}$ value determined from the isolated molecule GIPAW calculations (discussed in detail for each system in Chapter 3, 4 and 5) associated with pyridinium-fumarate H-bonds (Table 6.4). The 2,6 substitutions, both amino-methyl and dimethyl (lutidine), show the largest pyridinium $\Delta\delta_{Crys-Mol}$, ranging between 7.5 and 6.4 ppm. For the 2,5 substitutions, the highest pyridinium $\Delta\delta_{Crys-Mol}$ is 5.9 ppm and the lowest 4.0 ppm, with the strength decreasing from 25AMP:$F_{0.5}$:$FA_{0.5}$, through 52AMP:HF to

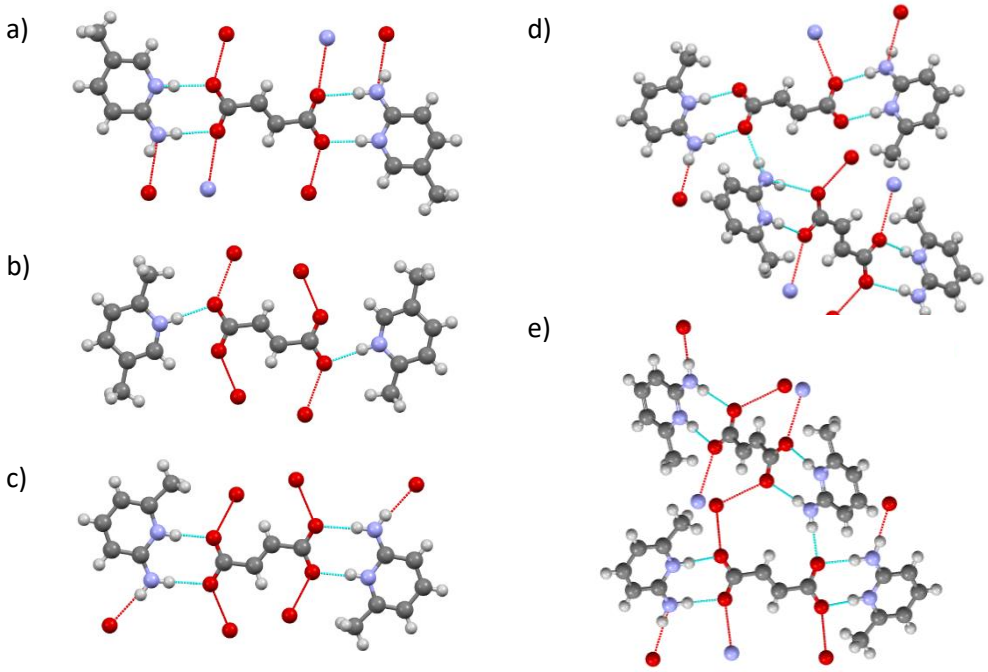


Figure 6.5: The b-a-b units present in (a) 25AMP:$F_{0.5}$:$FA_{0.5}$, (b) 25L:$F_{0.5}$:FA, (c) 26AMP:$F_{0.5}$:$(H_2O)_2$, (d) $26AMP_2$:F:$FA_{0.5}$ and (e) $26AMP_2$:F:H_2O. Hanging contacts (red lines) are shown to illustrate the range of additional H-bonds affecting the b-a-b units geometry.

25L:$F_{0.5}$:FA. The structures of all of the main systems are supported by H-bond networks with only a single weak π-interaction ($^1H\ \Delta\delta_{Crys-Mol} = -1.0$ for 26L:HF) observed. It is the patterns within the H-bond networks that have, therefore, been focussed on.

Five out of the seven systems contain the b-a-b motif with an inversion centre at the mid-point of the fumarate molecule (Fig. 6.5), with the two s-*cis* conformation adopted by the fumarate ion in all but one case, 26AMP:$F_{0.5}$:$(H_2O)_2$, which contains the higher energy two s-*trans* conformation. Another five of the seven also contain direct H-bonding between subsequent acid molecules to form an acid chain (Fig. 6.6), with three containing both. Both

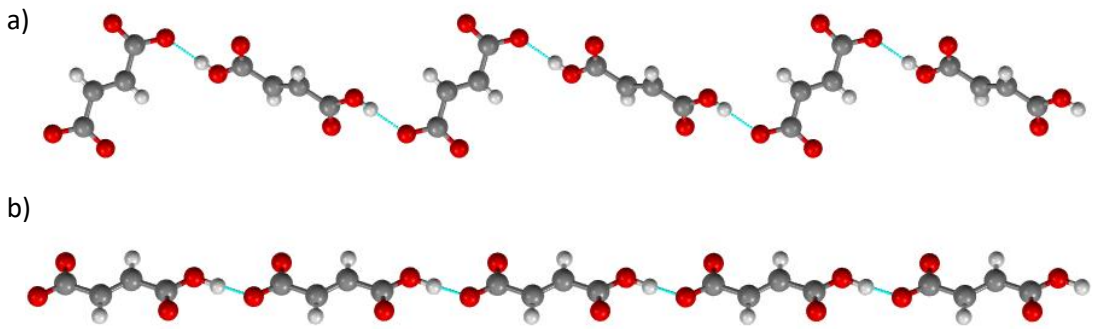


Figure 6.6: The two forms of acid chain exhibited: (a) formed from both fumarate ions and FA molecules (example from 25AMP:$F_{0.5}$:$FA_{0.5}$) and (b) formed from hydrogen fumarate molecules (example from 52AMP:HF).

of these motifs are dependent on the acid having two carboxylic groups to act as both H-bond donors and acceptors, with the b-a-b units also exploiting FA's symmetry.

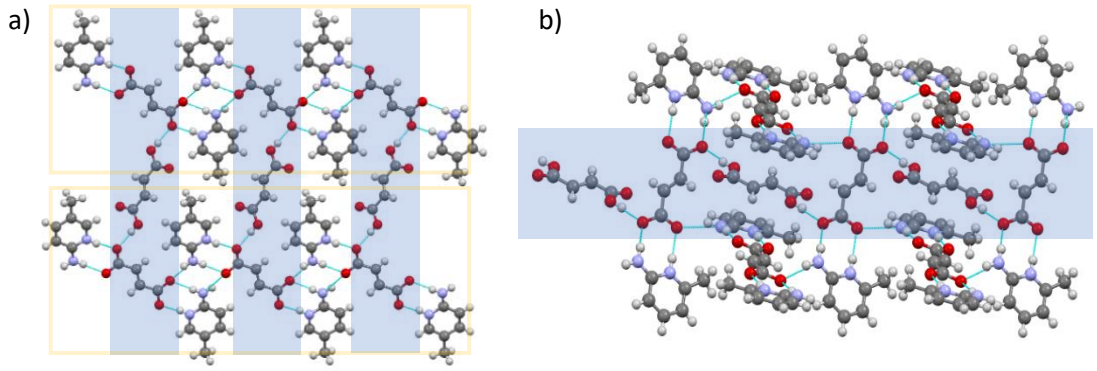


Figure 6.7: Packing of b-a-b units (a) 25AMP:$F_{0.5}$:$FA_{0.5}$, with acid chains running vertically and chains of b-a-b units horizontally; and (b) 26AMP$_2$:F:$FA_{0.5}$, with the acid chain running horizontally.

In most systems, there are a range of patterns that can be defined to describe the structure depending on which interactions are prioritised. Due to the relative strength of the $NH^{+}\cdots O^{-}$ interaction compared to the other NH···O H-bonds from the amino groups (based on the $^{1}H\ \Delta\delta_{Crys-Mol}$ value) the motif containing this was considered the most dominant. The b-a-b unit was therefore selected as a key structural feature, a choice also supported by its common occurrence. The acid chain is similarly highlighted as the OH···O interactions between FA/fumarate molecules consistently have the largest of all $^{1}H\ \Delta\delta_{Crys-Mol}$ values. The three systems that contain both of these structural patterns are the three cocrystal of a salt systems.

25AMP:$F_{0.5}$:$FA_{0.5}$ and 26AMP$_2$:F:$FA_{0.5}$ both have a b-a-b unit strengthened by inclusion of one of the amine protons, forming a second H-bond to each carboxylate ion. In 25AMP:$F_{0.5}$:$FA_{0.5}$, the other amino proton also stabilises/directs the wider structure by formation of a b-a-b-a ring (with graph set notation $R_4^2(8)$) which links the b-a-b units into chains (Fig. 6.7a). FA molecules sit between the fumarate anions of the b-a-b units in neighbouring chains, forming an acid chain orthogonal to the chains of b-a-b units and crosslinking them into a plane on the $(10\bar{1})$ crystal plane. The same alternating acid chain pattern is seen in 26AMP$_2$:F:$FA_{0.5}$ (Fig 6.7b) and a pair of them creates the diamond shaped array in the (010) crystal plane of 25L:$F_{0.5}$:FA (Fig. 6.8a) with a b-a-b unit that sits at each

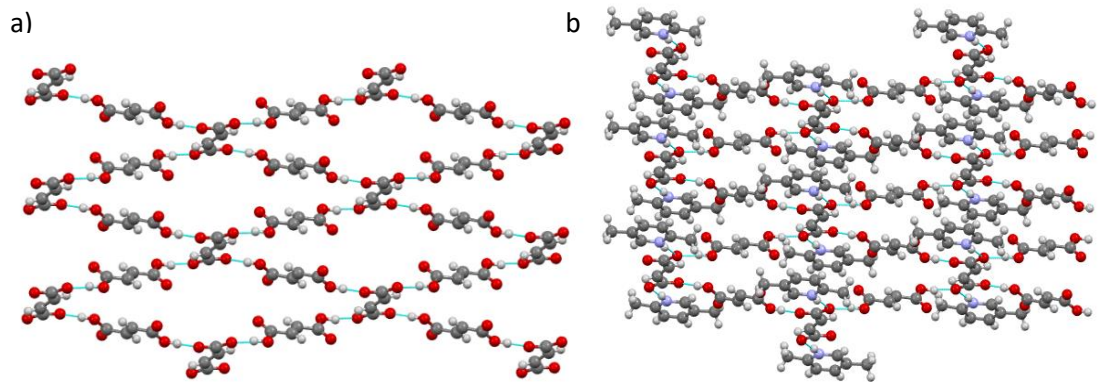


Figure 6.8: Packing of 25L:$F_{0.5}$:FA (a) with base molecules deleted and (b) with base molecules shown.

vertex (Fig. 6.8b). This pair of alternating mixed acid chains likely helps stabilise the b-a-b unit of 25L:$F_{0.5}$:FA which does not have the benefit of the amine group to form a $R_2^2(8)$ motif. All three systems form H-bonded planes that then stack together. In the case of 26AMP_2:F:$FA_{0.5}$ and 25L:$F_{0.5}$:FA, overlap of the pyridines of subsequent layers suggests π-π stacking despite the lack of evidence for this from the isolated molecule calculations.

The two salt hydrates, 26AMP:$F_{0.5}$:$(H_2O)_2$ and 26AMP_2:F:H_2O, also form b-a-b units stabilised by a b-a-b-a ring involving the amino group, as in 26AMP_2:F:$FA_{0.5}$ and

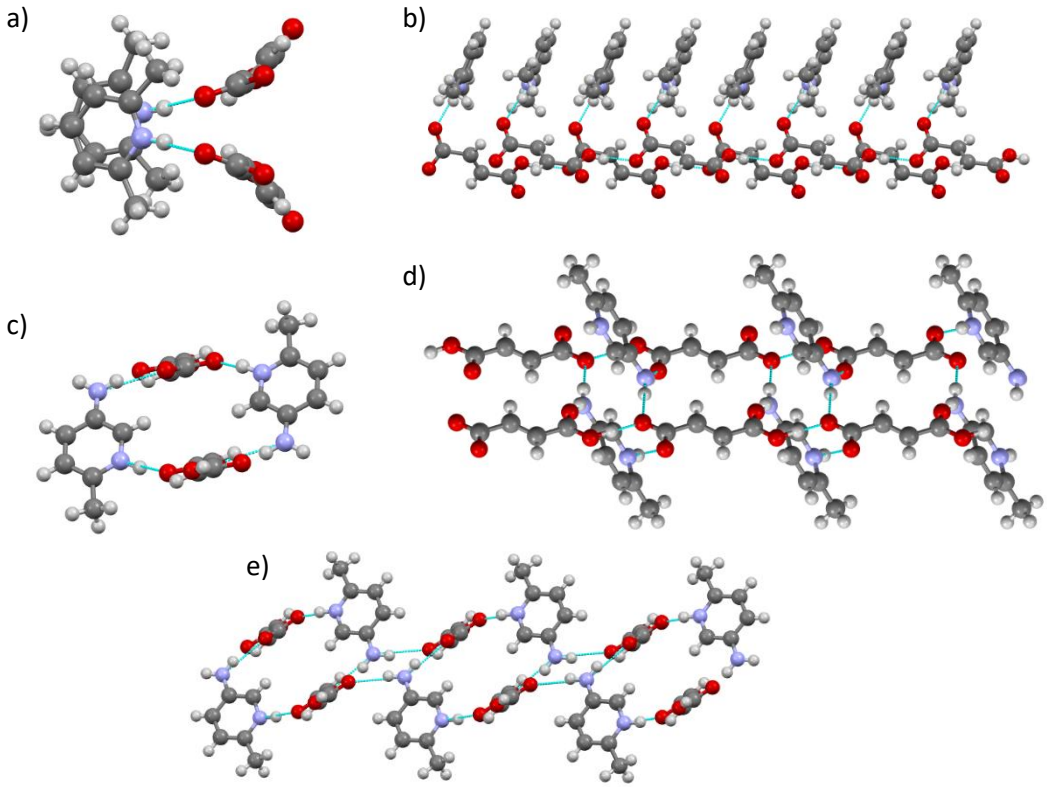


Figure 6.9: Paired acid chains of (a, b) 26L:HF viewed down (a) the c axis and (b) the b axis and of (c, d) 52AMP:HF viewed down (c) the a axis and (d) the c axis. The additional H-bonding in 52AMP:HF (e) between paired chains to form a plane through the structure.

25AMP:$F_{0.5}$:$FA_{0.5}$. 26AMP:$F_{0.5}$:$(H_2O)_2$ is the only structure that shows obvious interactions linking the entire 3D structure rather than just planes throughout it (possibly explaining why it forms large and good quality crystals so readily).

The two anhydrous salt forms both contain hydrogen fumarate anions, rather than doubly deprotonated fumarate anions, which in both cases form into acid chains with the pyridinium NH^+ H-bonded to the single carboxylate group of each anion. In 26L:HF, pairs of acid chains run in the same direction down the c-axis, with their alignment resulting in the stacking of the H-bonded pyridine rings (Fig 6.9a and b) whereas, in 52AMP:HF, paired acid chains run in opposite directions, linked into pairs by b-a-b-a rings through H-bonds from the pyridinium and one amino NH to each O of the carboxylate (Fig. 6.9c and d) with graph set notation $R_4^4(16)$.[70] These paired chains are further linked by a smaller b-a-a-b-a-a ring between the other amine NH and the carboxylic acid O=C (Fig. 6.9e). This crosslinking could also be used to define a b-a-b-a chain through the amine H-bonds and along the length of the hydrogen fumarate, running through the structure orthogonal to the acid chains. Although the position of the pyridinium-carboxylate H-bond suggests that the 52AMP:HF is in the s-*cis*/s-*trans* conformation that is exhibited in 26L:HF, the relative lengths of the carboxylate C-O bonds suggest it is in fact in the two s-*cis* conformation.

Although there is a clear link between the occurrence of hydrogen fumarate and lack of the b-a-b unit (both due to the preference for acid chain formation and the symmetry of the acid being broken), there does not seem to be any direct link to the pK_a difference or the fumaric acid conformation (see Table 1.1 in Chapter 1, section 1.5). This is unsurprising given the small set and the lack of a link between pK_a and the formation of a salt compared to a cocrystal of a salt, as previously noted in Chapter 1.

In summary, the key structural patterns that were identified following analysis of these core systems are:

- acid chains
- b-a-b units

- b-a-b-a- rings
- b-a-b-a chains

6.5. CSD Searches

The occurrence of the motifs identified in section 6.4 throughout the CSD was investigated for both fumarate/FA multicomponent systems and succinate systems/succinic acid systems. Succinic acid (SA) is another pharmaceutically acceptable dicarboxylic acid, related to FA through dehydrogenation. As the pK_a values of SA are higher than those of FA (4.21 and 5.64)[140] it was expected that the incidence of cocrystal systems would correspondingly be higher but it is unknown whether the structural patterns utilised for packing would follow the same trends. The presence of two hydrates within the core set also prompted an analysis of

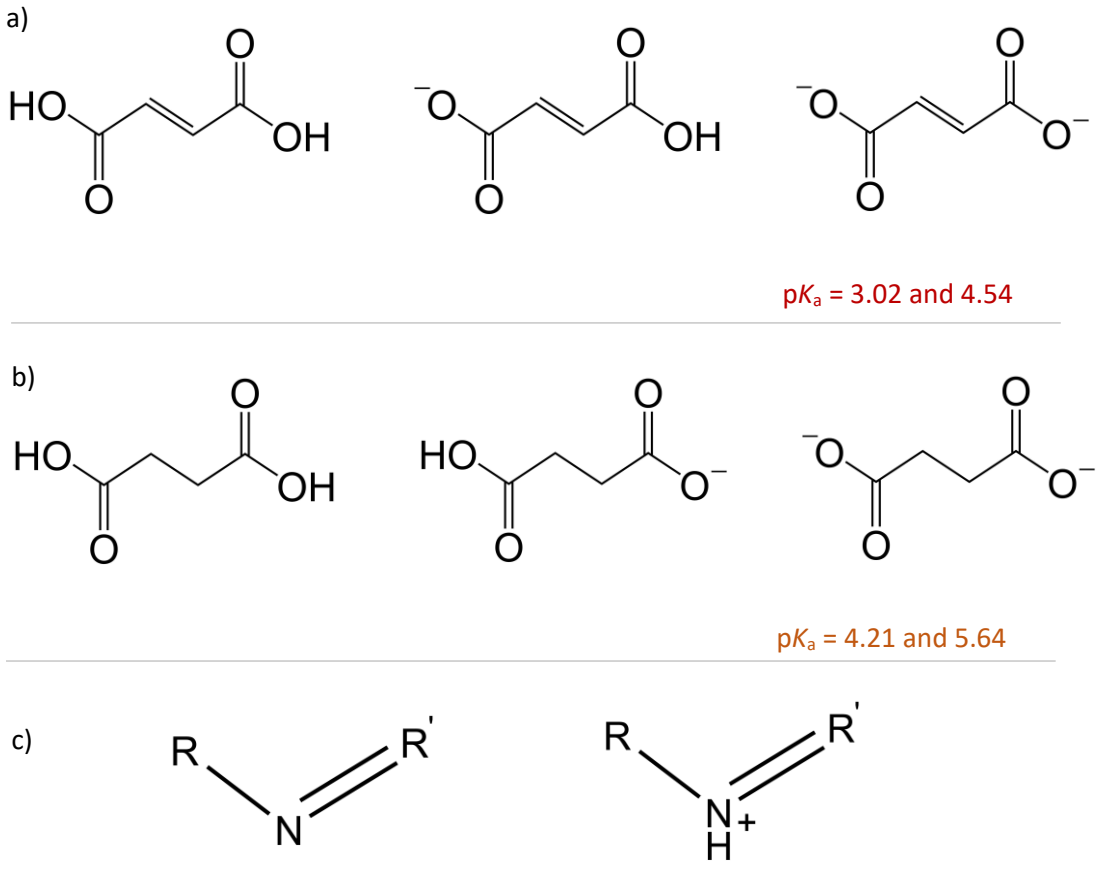


Figure 6.10: The chemical structures that formed the basis of the fragment based ConQuest searches, with each search consisting of one acid form (from either (a) fumarate/FA or (b) succinate/SA) paired with either a neutral or ionic N 'base' fragment (c).

the incidence of hydrate/solvate formation for these co-formers; the occurrence of general disorder was also analysed.

Many variations or alternatives to these packing structures are possible and numerous other interactions are involved in packing and maintaining crystal structure. However, those reliant on H-bonding, particularly those that form local or long-range networks, are the easiest to identify and therefore most convenient to search for within a large set. Other related structural patterns based on H-bonding will be noted and commented on.

ConQuest searches of the CSD (version Mar19) were performed for compound names containing 'fumarate', 'fumaric acid', 'succinate' or 'succinic acid' and also run using a fragment based search for each of the chemical structures of FA, hydrogen fumarate, fumarate, SA, hydrogen succinate or succinate paired with either a neutral N or an ionised NH (Fig. 6.10). The results from each ConQuest search were merged and unrelated results removed, *i.e.* those found not to contain the relevant acid but merely a fragment resembling it. This combination of name and fragment search ensured the set of captured systems containing an analogous acid-base interaction to the set studied in this work but does not completely guarantee all fumarate/FA and succinate/SA multicomponent crystals are included as it is possible that their given names may not have been captured by the relevant search. Structures were also removed from consideration if they did not contain any proton positions or exhibited extreme disorder, such that structural patterns and motifs could not be identified. The numbers excluded from each set were comparable. The smaller size of the succinate/succinic acid set compared to fumarate/FA may indicate that it is less amenable to co-crystallisation but it is also possibly that its higher pK_a, increasing its propensity to cocrystal formation relative to salts, reduced its use prior to the recent interest in the neutral cocrystal systems.

Following this, 535 multicomponent structures remained in the fumarate/FA set, with 210, 321 and 110 containing FA, hydrogen fumarate and fumarate, respectively, and 300 crystal structures remained within the succinate/succinic acid set, with 187, 77 and 75 containing succinic acid, hydrogen succinate and succinate, respectively (Table 6.5, Fig. 6.11). It is clear that for the succinate/SA set, the incidence of the neutral acid relative to either

Table 6.5: Breakdown of the fumarate/FA and succinate/SA sets by crystal form and co-former present.

Fumarate/FA set: 535	Salt: 333	Cocrystal of a Salt: 65	Cocrystal 136
	Fumarate: 110	Hydrogen Fumarate: 321	FA: 210
Succinate/SA set: 300	Salt: 107	Cocrystal of a Salt: 30	Cocrystal 152
	Succinate: 75	Hydrogen succinate: 77	SA: 187

of the ionised forms has risen dramatically compared to the fumarate/FA set. This is reflected in an increased proportion of cocrystals in the succinate/SA set, accounting for just over 50 % of the structures (152), essentially double the fraction seen in the fumarate/FA set (25 %, although the larger set size means this corresponds to 136 systems). Unsurprisingly, given the low pK_a of FA, around three fifths of the fumarate/FA structures (333) are salt forms. The salt form still makes up more than a third of the succinate/SA set (107), indicating that, although

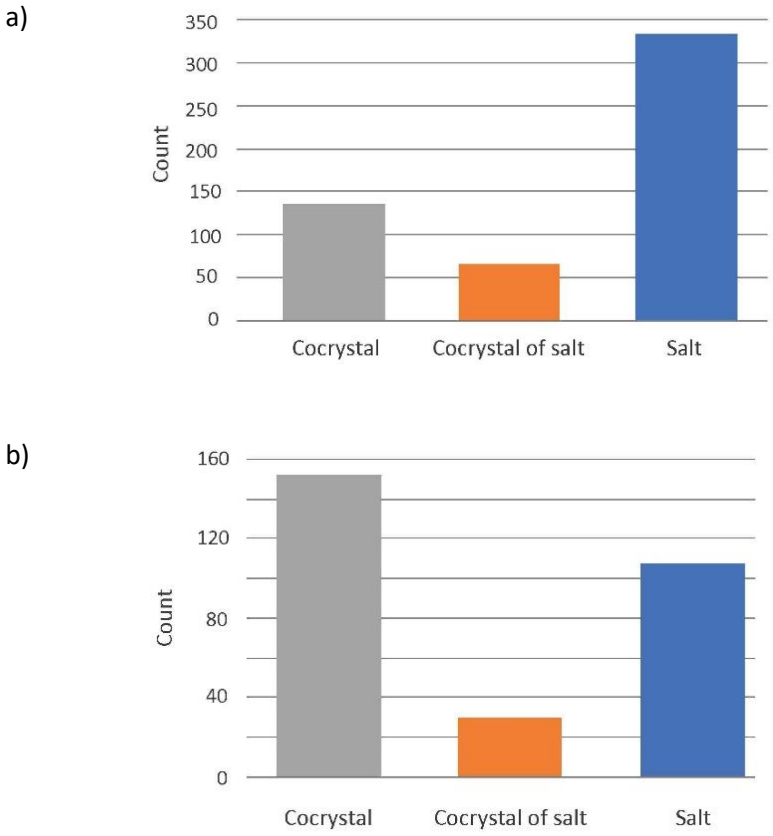


Figure 6.11: Crystal form of (a) the 535 fumarate/FA based multicomponent systems and (b) the 300 succinate/SA based multicomponent systems, as downloaded from the CSD.

the acids pK_a is higher, the ΔpK_a between acid and base is still rarely low enough to guarantee cocrystal formation.

The fraction that take the cocrystal of a salt crystal form is more similar between sets, only reducing from 12 % in fumarate/FA to 10 % in succinate/SA. Given the smaller size of the succinate/SA set, this does result in a relatively small subset of cocrystal of a salt systems (30) making it harder to draw conclusions from the trends seen within this subset, whereas the corresponding fumarate/FA subset still consists of 65 systems. Despite the occurrence of hydrogen fumarate being three times higher than fumarate, it is less likely to be present in the mixed cocrystal of a salt forms than the doubly ionised molecule. This may explain the lack of a trend between ΔpK_a and formation of either a salt or a cocrystal (Section 6.2.1) as single ionisation of every molecule, to produce entirely hydrogen fumarate, and double ionisation one in every two molecules with the other remaining neutral, to produce fumarate and fumaric acid, produce the same overall ionisation, despite resulting in different forms. This is supported by the fact that hydrogen succinate also has a significantly lower incidence in the mixed cocrystal of a salt form than succinate. Only 4 fumarate/FA systems contain both hydrogen fumarate and fumarate molecules, GIGKEO, KIZNIQ, GIGJUD and XONJUI, the latter 2 of which are cocrystal of a salt systems with neutral FA also present.

6.5.1. Disorder and solvation

Only a single fumarate/FA system, YEPJOV, showed proton disorder that prevented the crystal form from being determined (with the structure not removed from the set as the disorder did not prevent packing structure identification as b-a-b) whereas the succinate/SA set has 11 structures that exhibit disorder of the acid-base interaction proton such that the crystal form cannot be classified. Note that these systems are counted in Tables 6.8 and 6.9 as being half salt and half cocrystal. This increase in crystal form disorder was also reflected in the increase in the overal incidence of disorder from 12.1 % in the fumarate/FA set to 15. 7 % in the succinate/SA set.

Table 6.7: The occurrence of each structural pattern within each crystal form for fumarate/FA multicomponent crystals in the CSD, reported as both an absolute value and (in brackets) as a percentage of structures within that form..

Pattern	Salt		Cocrystal of a Salt		Cocrystal		Not classified	Total[a]	
b-a-b	61	(18.3)	35	(53.8)	95	(69.9)	1	**192**	**(35.9)**
b-a-b-a chain	99	(29.7)	10	(15.4)	27	(19.9)	0	**136**	**(25.4)**
b-a-b-a- ring	48	(14.4)	2	(3.1)	5	(3.7)	0	**55**	**(10.3)**
a chain	71	(21.3)	39	(60.0)	1	(0.7)	0	**111**	**(20.7)**
b-a-a-b	1	(0.3)	0		2	(1.5)	0	**3**	**(0.6)**
a-b-a	8	(2.4)	0		1	(0.7)	0	**9**	**(1.7)**
a-b-b-a	7	(2.1)	3	(4.6)	0		0	**10**	**(1.9)**
solvate	15	(4.5)	2	(3.1)	2	(1.5)	0	**19**	**(3.6)**
hydrate	58	(17.4)	8	(12.3)	6	(4.4)	0	**72**	**(13.5)**
disorder	45	(13.5)	9	(13.8)	10	(7.4)	1	**65**	**(12.1)**
Total[b]	**333**		**65**		**136**		**1**	**535**	

[a]The total incidence of each structural pattern across the fumarate/FA set is given by the sum of each row.
[b]Due to the overlap in the incidence of structural patterns, the total number of each crystal form is not given by the sum of the column.

As expected, the percentage incidence of both solvation (not including hydration) and hydration are lower for cocrystals than for either the salt or cocrystal of salt forms, with each of these ionic containing forms contributing more than twice as many solvates or hydrates than the neutral form for fumarate/FA (and more than 4 times as many for succinate/SA) (Tables 6.6 and 6.7). The ratio between hydrates and non-water solvates is similar across

Table 6.6: The occurrence of each structural pattern within each crystal form succinate/succinic acid multicomponent crystals in the CSD, reported as both an absolute value and (in brackets) as a percentage of structures within that form.

Pattern	Salt		Cocrystal of a Salt		Cocrystal		Not classified	Total[a]	
b-a-b	35	(32.7)	23	(76.7)	104	(68.4)	1	**163**	**(54.3)**
baba chain	21	(19.6)	3	(10.0)	42	(27.6)	1	**67**	**(22.3)**
b-a-b-a- ring	6	(5.6)	3	(10.0)	5	(3.3)	1	**15**	**(5.0)**
a chain	41	(38.3)	0		0		1	**42**	**(14.0)**
b-a-a-b	1	(0.9)	0		1	(0.7)	0	**2**	**(0.7)**
a-b-a	1	(0.9)	0		1	(0.7)	0	**2**	**(0.7)**
a-b-b-a	1	(0.9)	0		0		0	**1**	**(0.3)**
solvate	9	(8.4)	3	(10.0)	2	(1.3)	6	**20**	**(6.7)**
hydrate	27	(25.2)	5	(16.7)	5	(3.3)	0	**37**	**(12.3)**
disorder	16	(15.0)	4	(13.3)	16	(10.5)	11	**47**	**(15.7)**
Total[b]	**107**		**30**		**152**		**11**	**300**	

[a]The total incidence of each structural pattern across the fumarate/FA set is given by the sum of each row.
[b]Due to the overlap in the incidence of structural patterns, the total number of each crystal form is not given by the sum of the column.

crystal forms for fumarate/FA systems (~ 4:1 for hydrate : solvate), but varies significantly between succinate/SA forms (ranging from ~ 1.5:1 for cocrystals to 3:1 for salts). The lower incidence of hydrates and non-water solvates within cocrystals corresponds to a smaller percentage containing either FA or SA (Tables 6.8 and 6.9), although the percentage is slightly higher than for cocrystals due to the presence of these molecules within cocrystals of salts as well as neutral cocrystals (*e. g.* hydrates account for only 4.4 % of fumarate/FA cocrystals but are present in 5.7 % of FA containing systems). Although hydration is still present, the findings of Haynes *et al*. were found to hold true for both sets, with a lower incidence of hydrates seen for systems with pyridine containing molecules.[217] This trend was also true for other solvates and, for all the crystal forms together, the prevalence of hydration and solvation reduced by about half for fumarate/FA systems (13.5 % and 3.6 % of all fumarate/FA

Table 6.8: Breakdown of structural pattern seen in the fumarate/FA set by coformer[a] as a percentage of structures containing that coformer.

	Fumarate			Hydrogen Fumarate			FA		
b-a-b	89 (80.9)	**S:**	58.5	9 (2.8)	**S**	4	129 (61.4)	**S**	0.5
		CoS:	30		**CoS**	5		**CoS**	33
		C:	0.5		**C**	0		**C**	95.5
baba chain	13 (11.8)	**S**	8	97 (30.2)	**S**	91	38 (18.1)	**S**	0
		CoS	5		**CoS**	6		**CoS**	11
		C	0		**C**	0		**C**	27
b-a-b-a- ring	3 (2.7)	**S**	3	47 (14.6)	**S**	45	7 (3.3)	**S**	0
		CoS	0		**CoS**	2		**CoS**	2
		C	0		**C**	0		**C**	5
a chain	33 (30.0)	**S**	1	80 (24.9)	**S**	71	38 (18.1)	**S**	0
		CoS	32		**CoS**	9		**CoS**	37
		C	0		**C**	0		**C**	1
solvate	8 (7.3)	**S**	6	10 (3.1)	**S**	9	4 (1.9)	**S**	0
		CoS	2		**CoS**	1		**CoS**	2
		C	0		**C**	0		**C**	2
hydrate	26 (23.6)	**S**	21	42 (13.1)	**S**	38	12 (5.7)	**S**	0
		CoS	5		**CoS**	4		**CoS**	6
		C	0		**C**	0		**C**	6
disorder	18 (16.4)	**S**	12.5	38 (11.8)	**S**	33	17 (8.1)	**S**	0.5
		CoS	5		**CoS**	5		**CoS**	6
		C	0.5		**C**	0		**C**	10.5
Total[b]	**110**			**321**			**210**		

[a]Further separation by crystal form (salt, S, cocrystal of a salt, CoS, and cocrystal, C) is also recorded. The single undetermined system, YEPJOV, is represented as halved between salt and cocrystal

[b]Due to the overlap in the incidence of structural patterns, the total number of each coformer is not given by the sum of the column.

Table 6.9: Breakdown of structural pattern seen in the succinate/SA set by coformer[a] reported both an absolute value and (in brackets) as a percentage of structures containing that coformer.

	Succinate			Hydrogen Succinate			Succinic Acid		
b-a-b	55 (73.3)	**S:**	34	2 (2.6)	**S**	2	124 (66.3)	**S**	0
		CoS:	21		**CoS**	0		**CoS**	19
		C:	0		**C**	0		**C**	105
baba chain	12 (16.0)	**S**	9	13 (16.9)	**S**	13	45 (24.1)	**S**	0
		CoS	3		**CoS**	0		**CoS**	3
		C	0		**C**	0		**C**	42
b-a-b-a- ring	5 (6.7)	**S**	3	4 (5.2)	**S**	4	8 (4.3)	**S**	0
		CoS	2		**CoS**	0		**CoS**	3
		C	0		**C**	0		**C**	5
a chain	0 (0.0)	**S**	0	41 (53.2)	**S**	41	0 (0.0)	**S**	0
		CoS	0		**CoS**	0		**CoS**	0
		C	0		**C**	0		**C**	0
solvate	3 (4.0)	**S**	2	9 (11.7)	**S**	7	4 (2.1)	**S**	0
		CoS	1		**CoS**	2		**CoS**	2
		C	0		**C**	0		**C**	2
hydrate	22 (29.3)	**S**	17	12 (15.6)	**S**	12	8 (4.3)	**S**	0
		CoS	5		**CoS**	0		**CoS**	4
		C	0		**C**	0		**C**	4
disorder	14 (18.7)	**S**	10	8 (10.4)	**S**	7	17 (9.1)	**S**	0
		CoS	4		**CoS**	1		**CoS**	2
		C	0		**C**	0		**C**	15
Total			**75**			**77**			**187**

[a]Further separation by crystal form (salt, S, cocrystal of a salt, CoS, and cocrystal, C) is also recorded.
[b]Due to the overlap in the incidence of structural patterns, the total number of each coformer is not given by the sum of the column.

structures compared to 6.6 % and 1.5 % of pyridine containing systems, for hydration and other solvation, respectively). The difference was smaller within the succinate/SA set, with hydration reducing from 12.3 % to 8.9 % for pyridine systems.

Disorder was also 25% lower within fumarate/FA pyridine containing systems, likely due to the overlap between solvated and disordered structures (with a third of disordered systems identified also being a solvate/hydrate). Comparing occurrence between crystal form, disorder was also nearly twice as prevalent within fumarate/FA salt or cocrystal of a salt forms than in the sets cocrystals, most commonly associated with systems containing doubly ionised fumarate molecules (16.4 %) which also show the highest incidence of hydration and solvation (23.6 % and 7.3 %, respectively). This is mirrored in the succinate/SA set.

6.5.2. Structural motifs

The incidence of each structural pattern is different between sets, both overall and within crystal forms. The motifs' relative incidence within the cocrystal form are most comparable but they differ significantly in the other crystal forms. Across both sets, the b-a-b unit is the most prevalent of the four structural motifs, followed by b-a-b-a chains, acid chains and finally b-a-b-a rings (Fig. 6.12a and 6.13a). The formation of the b-a-b unit is even more dominant across the succinate/SA set, with a significant reduction in the proportion of structures exhibiting the other three motifs. It is by far the most common structural pattern in both FA

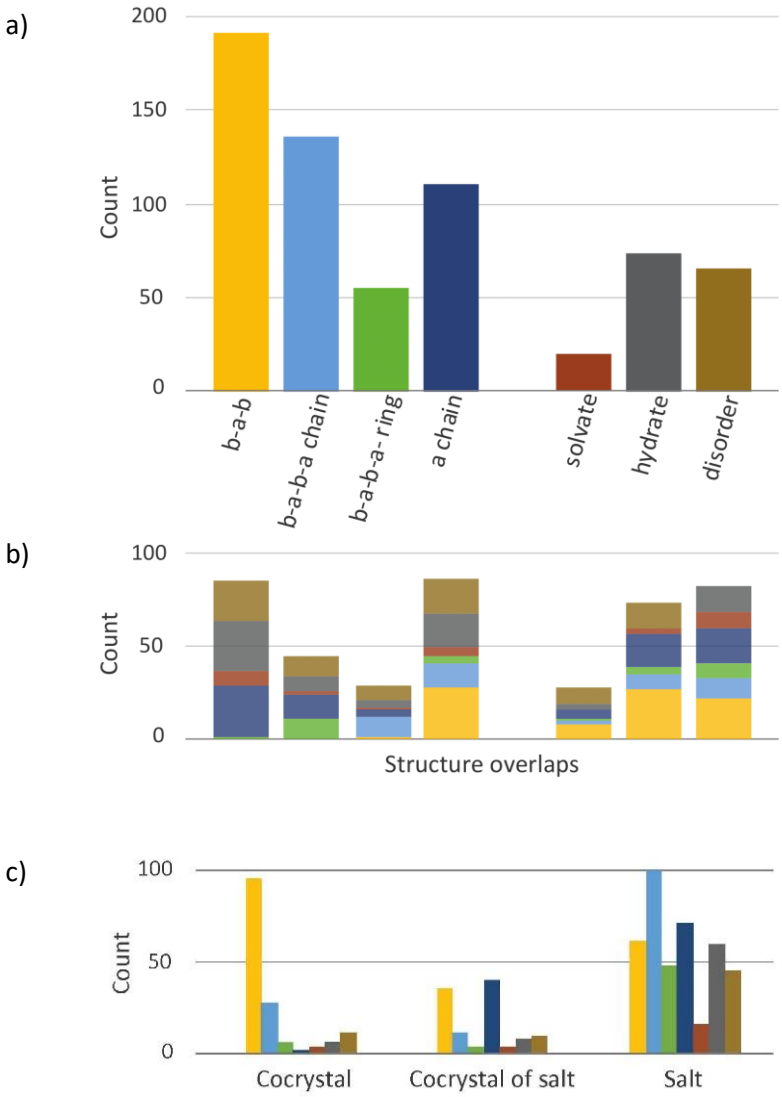


Figure 6.12: Incidence for (a) structural patterns and system solvation/disorder within the fumarate/FA set, (b) overlap in these structural characteristics and (c) their incidence separated by crystal form. The height of each block in (b) is on the same scale as used in (a) but, as there are additional overlaps between the blocks which are not shown, the total height of the stack does not correspond to the total number of systems that exhibit a structural overlap. Colours in (b) and (c)

and SA cocrystals (identified in 69.9% and 68.4 % of FA and SA cocrystals, respectively) and remains notable in both cocrystals of salts (present in 53.8 % and 76.7 %) and salts (present in 18.3 % and 32.7 %) (Fig. 6.12c and Fig 6.13c, Tables 6.6 and 6.8). The significantly lower incidence seen for salts within both sets is not surprising given the prevalence of hydrogen fumarate and hydrogen succinate as the acid molecules within this form, rather than fumarate or succinate (Tables 6.7 and 6.9).

As the singly ionised acids lack the symmetry of either fumarate/succinate or FA/SA, they do not form b-a-b units but instead lend themselves to either the related patterns, b-a-b-a

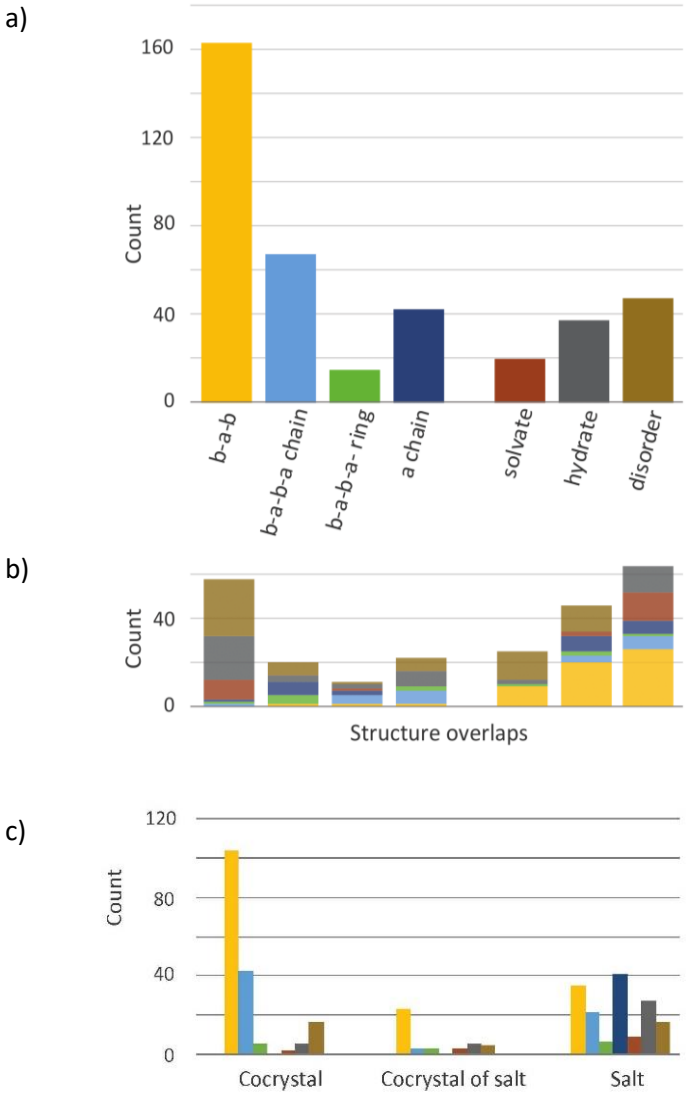


Figure 6.13: Incidence for (a) structural patterns and system solvation/disorder within the succinate/succinic acid set, (b) overlap in these structural characteristics and (c) their incidence separated by crystal form. The height of each block in (b) is on the same scale as used in (a) but, as there are additional overlaps between the blocks which are not shown, the total height of the stack does not correspond to the total number of systems that exhibit a structural overlap. Colours in (b) and (c) correspond to the colours in (a).

chain and ring structures, or acid chain formation. In b-a-b-a chains and rings, where the donors/acceptors on the base molecule are very similar (with at least two such moieties required for the formation of either), the acid molecule tends to be either FA/SA or fumarate/succinate, instead of the corresponding hydrogen form. When, as is more commonly the case, the acceptors/donors on the base differ, the acid is singly deprotonated and the base molecules are joined by hydrogen fumarate molecules with alternating neutral and ionic interactions. The incidence of b-a-b-a rings is less than half that of the b-a-b-a chains in both sets (less than a quarter in the succinate/SA set) due to the need for the base acceptors/donors to sit on the same side of one of the molecule's 3D conformation, as shown for HEQZOT, rather than different sides, as shown for YEPCOM (Fig. 6.14).

While, in the fumarate/FA set, acid chains are the the most prevalent in the cocrystal of a salt form, occurring in 60.0 % (Fig. 6.12c, Table 6.6), they were not evident at all in this form within the succinate/SA set, only forming from hydrogen succinate molecules in the salt form, where it is the most common strucutral pattern (38.3 % compared to second mst prevalent in the fumarate/FA set at 21.3 %). This may be related to the different conformational landscape available to each of the acids, with the lack of the central C=C bond in succinic acid contributing to a far more complex and varied set of possible conformations

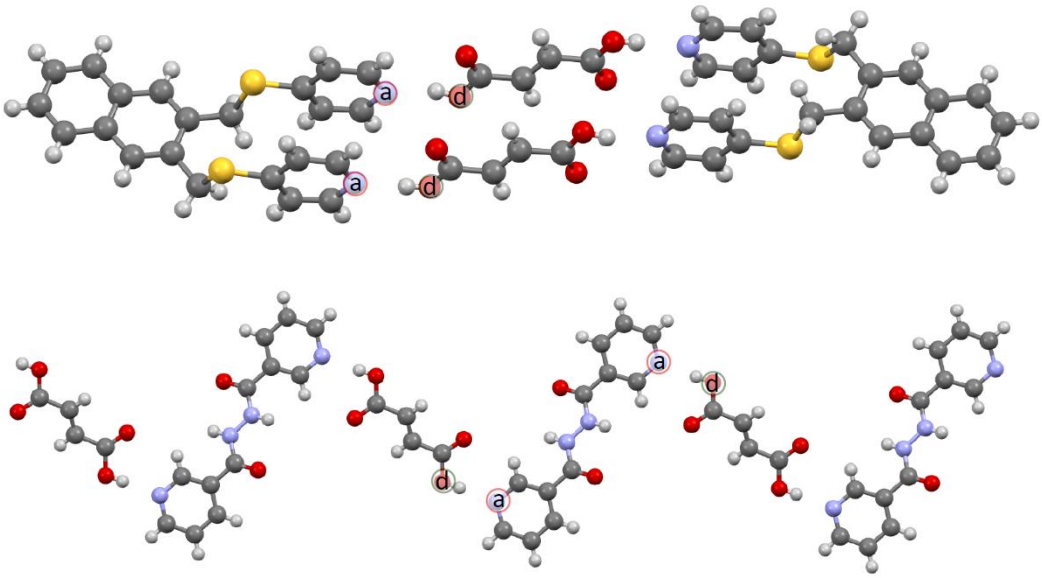


Figure 6.14:Example of cocrystals containing a b-a-b-a ring (top) and a b-a-b-a chain (bottom). The b-a-b-a ring in HEQZOT has both acceptor atoms on the same face of the base molecule whereas they are on opposite sides of the base molecule in the b-a-b-a chain in YEPCOM.

possessing similar energies.[235] Purely hydrogen fumarate/hydrogen succinate based systems account for all but one of the acid chains within the salts (this example occurs in the fumarate/FA set) as fumarate/succinate have no donors with which to form H-bonds to themselves so require either hydrogen fumarate/hydrogen succinate or FA/SA to be present as well. The exception in the fumarate salts, KIZNIQ, contains both hydrogen fumarate and fumarate anions. Alternating doubly ionised and neutral molecules form the basis of most of the acid chains in fumarate/FA cocrystals of salts. It is the absence of these mixed chains in the succinate/SA set that eliminates the occurrence of acid chains within the cocrystal of a salt form. As FA and SA can satisfy all acceptor/donor interactions when they form neutral acid chains, which is the basis of the pure crystalline form of FA, it is not surprising that it is an uncommon motif within the cocrystal systems as its disruption is what allows a multicomponent crystal to form. The sole occurrence of a neutral acid chain is the FA cocrystal TOQKER, which has one carboxylic acid group involved in the chain and the second H-bonding to the base.

Within the salt form, it is the occurrence of b-a-b-a chains and rings that show the most dramatic change between sets, accounting for the core structural pattern in only 19.6 % and 5.6 % of succinate salt systems compared to 29.7 % and 14.4 % in the fumarate/FA set, for b-a-b-a chains and rings, respectively. This reduction saw a corresponding increase in both b-a-b units and acid chains, the occurrence of which rose by 14.4 % and 17.0 %, respectively, in the succinate salts. As b-a-b-a chains and b-a-b units are closely related, their overlap is minimal across all the structures, but both occur within the same systems as b-a-b-a rings, if rarely. The acid chains present simultaneously with all of the other three motifs, as was evident within the small local set of structures.

Analysis of the fumarate/FA set suggests that cocrystals tend to be dominated by one structural pattern, with no overlap identified in the occurrence of the four motifs considered. Within the set of salts, it was also rare to observe a combination of the these patterns, with complex H-bond networks that allow classification of both b-a-b-a chains and b-a-b-a rings being the most common, presenting in ~3 % of structures. More than 90 % of structures were

found to possess only a single H-bond motif. The crystal form which shows the highest mixing of structural patterns, possibly due to the increased variety of donor and acceptor groups resulting from both the neutral and ionic forms of the same molecule, is the cocrystal of a salt form, with nearly half of the structures classified as containing two or more of the motifs, accounting for most of the structural overlap shown in Fig. 6.12b. This is not seen in the succinate/SA set, where there is no overlap in the structural patterns exhibited in the cocrystal of a salt crystal form. As most of the mixing was combinations of one of the two b-a-b-a alternating motifs with an acid chain, this result is likely due to the lack of acid chains for this succinic crystal form. The reason why succinic acid does not readily form chains of doubly ionised and neutral molecules may be linked to its higher pK_a. The limited co-occurrence of the structural patterns seen for succinic multicomponent systems is largely within the salt form and similar to the level exhibited by fumaric systems.

As no isolated molecule GIPAW calculations have been conducted and the H-bond parameters have not been measured for each of the structures, it is not possible to determine the relative importance of the structural motifs to forming and maintaining the systems where multiple structural patterns are present. Within the small set considered in chapters 3-5, both isolated molecule calculations and H-bond parameters (Table 6.4) suggest that acid chains are the most stable structures as they have both the strongest H-bonds (largest $^{1}H\ \Delta\delta_{Crys-Mol}$) with the most optimal geometries. However, as all systems are reliant on the acid-base interaction, it is unclear which of the interactions drive the initial packing.

Other acid-base structural patterns, of the same type as the b-a-b unit and b-a-b-a chains and rings, were identified but they were far less prevalent variations of those discussed above (see Tables 6.6 and 6.7). There were 3 systems that contained b-a-a-b units, 8 with symmetrical base molecules that allowed formations of an a-b-a unit (3 of which were present alongside acid chains) and a further 10 systems with non-symmetrical bases that were classified as a-b-b-a units (with 3 seen in combination with b-a-b-a rings).

6.6. Summary

In the absence of crystal form classification, or when validation is required, GIPAW calculations suggest that the ^{14}N shift of cyclic amines provides a useful measure of its interaction state. Ionised and neutral systems fall in distinct shift ranges with a separation (at 14.1 T) of 126 ppm, which rises to 250 ppm when outliers for which the calculations are less reliable are removed. Salt interactions lie below 0 ppm and cocrystal interactions from around 200 ppm up. Although there is overlap, it is also seen that cyclic amines which are non-interacting have a ^{14}N shift above 600 ppm whereas interacting systems are generally lower. The validity of this method is supported by the good agreement between experiment and GIPAW calculation seen for the ^{14}N shifts of the systems considered in chapters 3-5.

Throughout this work, discrepancies have been identified between GIPAW calculated and experimental chemical shifts for two specific chemical environments, with significantly larger differences shown than expected. The affected environments are the ^{1}H in an OH···O hydrogen bond, which is observed at a lower chemical shift than calculated, and the quaternary ^{13}C which is covalently bound to both a pyridinium nitrogen and an amino nitrogen, which is observed at a higher chemical shift than calcualted.

Within the same small set, two key structural patterns were identified, b-a-b units and acid chains. The prevalence of these, alongside several variations and the relative propensity towards each crystal form, was investigated within the CSD for sets of both fumarate/FA and succinate/succinic acid containing systems. Structures derived from FA are more likely to contain ionic interactions than neutral ones, as expected due to its low pK_a values, whereas those derived from succinic acid tend towards cocrystal structures. Within the set of systems considered in chapters 3-5, no relationship is seen between the ΔpK_a and whether a salt or a cocrystal of a salt is formed. The CSD searches show that the incidence of this mixed form does not actually change significantly between fumarate/FA and succinate/succinic acid sets, despite the dramatic drop in the proportion of salt forms. However, a significantly higher proportion of the succinate/succinic acid set were unclassified due to disorder of the proton

within the key acid base interaction. There are systems reported in the literature as continuums, where the proton is considered to be shared more equally between the molecules and free to move between them, often with a dependence on temperature.[236-239]

The shared symmetry of FA and succinic acid means that overall they both favour the formation of b-a-b units above any other structural pattern, although this varies with crystal form and the two related structural patterns are seen instead depending on the structure of the base molecule. The nature of the acid chains exhibited differs between the two sets. Although the salt form of both sets is dominated by the single component (hydrogen fumarate or hydrogen succinate) chains, the cocrystal of a salts form show a prevalence of mixed chains in the fumarate/FA and none in the succinate/succinic acid. It would be interesting to note whether any correlations exist between each structural pattern and the acid conformation, particularly for fumaric acid where it is more easily identified and has fewer variations, reducing the invesitigation's complexity.

It would be useful to have physical properties recorded and available as an easily accessible parameter within the CSD to allow trends between either crystal form or structural patterns and, for example, the thermal stability to be easily investigated. The melting point is a parameter that can already be recorded within CIF files but it is not common practice to do so (of 61 CIFs specifically checked, only 9 recorded it). As it stands, therefore, no direct property-structure links can be probed in the wider CSD and only hydration/solvation and disorder are readily available as measures of the preferability of each structural pattern for form design and selection.

Chapter 7. Summary and Outlook

The work presented in this thesis demonstrates the power of employing a multitechnique appoach to the characterisation of organic multicomponent forms, focusing particularly on integrating the complementary data provided by XRD, solid-state MAS NMR and DFT calculations. A detailed characterisation of seven related pyridinium-fumarate based multicomponent systems was undertaken, specifically investigating their structures and stabilities using both single crystal and powder XRD at a range of temperatures, ^{1}H and ^{13}C CP MAS 1D spectra and ^{1}H-^{1}H, ^{1}H-^{13}C and ^{14}N-^{1}H 2D correlation solid state NMR experiments, TGA and DSC. DFT-based geometry optimisations were performed to investigate the structure convergence and GIPAW calculations of the NMR parameters facilitated both the analysis of the experimental solid-state NMR and highlighted key the intermolecular interactions present in each system. Alongside this, larger sets of crystal strucutres were downloaded from the CSD, firstly to identify a trend in ^{14}N shift to help distinguish between ionic and nuetral crystal forms and secondly to probe the variety and conservation of structural patterns within fumarate/FA and succinate/SA systems.

Within both chapters 3 and 4, a loss of base molecules was seen resulting in a transformation to single component fumaric acid and a multicomponent anhydrous cocrystal of a salt for 26L:HF and $26AMP_2$:F:H_2O, respectively. In both cases, this requires the transfer of a proton from the base molecule to one of the acid molecules that remains in the resultant system. While the thermal analysis was predominantly conducted by TGA for 26L:HF, it was expanded in chapter 4 to include VT-XRD and VT-NMR. Given the changing hydration of the 26AMP systems, these methods were of particuar use to develop an understanding of the transition pathway from the dihydrate, 26AMP:$F_{0.5}$:$(H_2O)_2$, to the anhydrous forms, $26AMP_2$:F:$FA_{0.5}$ and the unknown Form IV.

Chapter 5 highlighted the significant impact that small changes to the chemical structure can have on the crystal packing. The three 2,5 amino and methyl substituted

pyridines, exhibited a distinct change in both packing and form, not merely from inclusion of an amino group as seen for the 2,6 substitutions, but also dependent on its substitution position. The newly solved structure, 52AMP:HF, forms a hydrogen fumarate salt with acid chains similar to 26L:HF, while the related 25AMP and 25L systems both form cocrystals of salts with strong b-a-b units, as seen for all of the 26AMP systems, which are linked by fumarate/FA acid chains. This seemingly structurally sensitive variation in packing contrasts with a high consevation of key patterns, such as the b-a-b unit and two forms of acid chains, leading to the broad searches of the CSD in chapter 6 to determine their wider incidence. Although the trend of crystal form versus ΔpK_a was as expected for salts and cocrystals, the key difference between the fumarate/FA and succinate/SA sets, the nature of the acid chain, was unanticipated given the small difference between coformer molecules. The mixed ionic/neutral form of the acid chain is integral to the fumarate/FA cocrystals of salts but are not seen at all within the succinate/SA set.

The recent work within NMR crystallography focuses on continually improving the depth and detail that can be determined about a system, from structure to dynamics, but also pushing the limits of structure solution in the absence of SXRD to allow analysis of increasingly complex systems. This is aided by ongoing work to increase resolution within the experimental solid-state NMR, particularly for ^{1}H linewidths, to allow a full assignment of resonances and structural structural characterisation for large, fully protonated systems. The development of ultrafast MAS (up to 111 kHz) has done a great deal to facilitate this but continual pulse sequence development also plays a significant role.[240] As well as improved ^{1}H-^{1}H decoupling strategies, there is also a move to develop recoupling methods to allow fine measurements of dipolar couplings and distances.[236, 241, 242] This allows more precise constraints to be determined, aiding structure solution from PXRD,[243, 244] informing CSP methods[245] and distinguishing between crystal forms[236] or desmotrophs.[241]

In some instances, XRD is not feasible, making the precision of the parameters and constraints determined from complementary methods of even greater importance to the quality of the structural model that can be developed. A range of strategies can be adopted, depending

on the system, with the application of ^{14}N NMR shifts, as discussed in chapter 6, one method to qualitatively determine crystal form for amine chemical environments. In instances of dealing with nanocrystalline structures, electron diffraction has been employed, with an NMR Crystallography approach applied to confirm atomic assignments between similar weight atoms (C/N/O) as well as place protons within the structure.[246] Alternatively, no diffraction can be used but instead DFT calculations can be utilised to link experimental NMR parameters to trial structures, developed through either CSP methods or chemical intuition.[245]

As computational methods provide the link between complementary techniques, the accuracy of the calculations is an integral factor in the characterisation. Attempting to understand and resolve instances where the computational approach is insufficient or innacurate is therefore key to progress. In the case of the OH···O hydrogen bond ^{1}H chemical shift discrepancy identified throughout this thesis, it would be interesting to investigate the temperature dependence[226-233] to see if there are marked differences for the fumaric acid OH ^{1}H resonances as compared to the NH^{+} ^{1}H resonances. The inclusion of temperature effects into calculations presents a number of challenges,[99] particularly for more complex systems, but they are integral for probing phase transitions and dynamics, which may be key to understanding the source of a systems properties.[247]

Application of a combined approach to structural characterisation is progressing in terms of both its power to elucidate the structure of complex systems and the insight provided into their formation and its relation to their physical properties, allowing for corresponding development in crystal engineering.

References

1. S. M. Berge, L. D. Bighley and D. C. Monkhouse, *J. Pharm. Sci.*, 1977, **66**, 1-19.
2. O. Almarsson and M. J. Zaworotko, *Chem. Commun.*, 2004, 1889-1896.
3. C. B. Aakeroy and D. J. Salmon, *CrystEngComm*, 2005, **7**, 439-448.
4. P. Vishweshwar, J. A. McMahon, J. A. Bis and M. J. Zaworotko, *J. Pharm. Sci*, 2006, **95**, 499-516.
5. A. V. Trask, *Mol. Pharm.*, 2007, **4**, 301-309.
6. D. J. Berry and J. W. Steed, *Adv. Drug Delivery Rev.*, 2017, **117**, 3-24.
7. U. S. D. o. H. a. H. Services, F. a. D. Administration and C. f. D. E. a. R. (CDER), Regulatory Classification of Pharmaceutical Co-Crystlas Guidance for Industry, *FDA* 2016.
8. EMA/CHMP/CVMP/QWP, *Reflection paper on the use of cocrystals of active substances in medicinal products*, EMA, London, U. K., 2015.
9. P. C. Vioglio, M. R. Chierotti and R. Gobetto, *Adv. Drug Delivery Rev.*, 2017, **117**, 86-110.
10. L. A. E. Sousa, S. M. Reutzel-Edens, G. A. Stephenson and L. S. Taylor, *Mol. Pharm.*, 2015, **12**, 484-495.
11. L. A. E. Sousa, S. M. Reutzel-Edens, G. A. Stephenson and L. S. Taylor, *Cryst. Growth Des.*, 2016, **16**, 737-748.
12. J. M. Miller, B. M. Collman, L. R. Greene, D. J. W. Grant and A. C. Blackburn, *Pharm. Dev. Technol.*, 2005, **10**, 291-297.
13. M. Hoffman and J. A. Lindeman, in *Pharmaceutical Salts and Co-crystals*, eds. J. Wouters and L. Quere, Royal Society of Chemistry, Cambridge, 2011, pp. 318-329.
14. A. Tilborg, B. Norberg and J. Wouters, *Eur. J. Med. Chem.*, 2014, **74**, 411-426.
15. M. L. Cheney, N. Shan, E. R. Healey, M. Hanna, L. Wojtas, M. J. Zaworotko, V. Sava, S. J. Song and J. R. Sanchez-Ramos, *Cryst. Growth Des.*, 2010, **10**, 394-405.
16. N. Shan, M. L. Perry, D. R. Weyna and M. J. Zaworotko, *Expert Opin. Drug Met.*, 2014, **10**, 1255-1271.
17. IUPAC, *Compendium of Chemical Terminology*, Blackwell Scientific Publications, Oxford, 2nd edn., 1997.
18. S. L. Childs, G. P. Stahly and A. Park, *Mol. Pharm.*, 2007, **4**, 323-338.
19. E. Grothe, H. Meekes, E. Vlieg, J. H. ter Horst and R. de Gelder, *Cryst. Growth Des.*, 2016, **16**, 3237-3243.
20. A. M. Healy, Z. A. Worku, D. Kumar and A. M. Madi, *Adv. Drug Delivery Rev.*, 2017, **117**, 25-46.
21. N. K. Duggirala, M. L. Perry, O. Almarsson and M. J. Zaworotko, *Chem. Commun.*, 2016, **52**, 640-655.
22. H. G. Brittain, *J. Pharm. Sci.*, 2013, **102**, 311-317.
23. U. J. Griesser, in *Polymorphism: in the Pharmaceutical Industry*, ed. R. Hilfiker, Wiley - VCH Verlang GmbH & Co., Weinheim, 2006.
24. P. A. Thiel and T. E. Madey, *Surf. Sci. Rep.*, 1987, **7**, 211-385.

25. C. H. Gorbitz and E. Gundersen, *Acta Chem. Scand.*, 1996, **50**, 537-543.
26. C. H. Gorbitz and E. Gundersen, *Acta Crystallogr. Sect. C*, 1996, **52**, 1764-1767.
27. C. H. Gorbitz and E. Torgersen, *Acta Crystallogr. Sect. B*, 1999, **55**, 104-113.
28. L. Infantes and S. Motherwell, *CrystEngComm*, 2002, **4**, 454-461.
29. B. C. R. Sansam, K. M. Anderson and J. W. Steed, *Cryst. Growth Des.*, 2007, **7**, 2649-2653.
30. C. H. Gorbitz and H. P. Hersleth, *Acta Crystallogr. Sect. B*, 2000, **56**, 526-534.
31. G. A. Jeffrey and H. Maluszynska, *Acta Crystallogr. Sect. B*, 1990, **46**, 546-549.
32. A. V. Trask, W. D. S. Motherwell and W. Jones, *Cryst. Growth Des.*, 2005, **5**, 1013-1021.
33. J. Rantanen, H. Wikstrom, F. E. Rhea and L. S. Taylor, *Appl. Spectrosc.*, 2005, **59**, 942-951.
34. L. Yu, *Adv. Drug Delivery Rev.*, 2001, **48**, 27-42.
35. R. A. Bellantone, in *Amorphous Solid Dispersions Theory and Practice*, eds. N. Shah, H. Sandhu, D. S. Choi, H. Chokshi and A. W. Malick, Springer, 2014, pp. 3-34.
36. C. B. Aakeroy, M. E. Fasulo and J. Desper, *Mol. Pharm.*, 2007, **4**, 317-322.
37. C. B. Aakeroy, *Acta Crystallogr. Sect. B*, 2015, **71**, 387-391.
38. S. Mohamed, D. A. Tocher and S. L. Price, *Int. J. Pharm.*, 2011, **418**, 187-198.
39. D. Ludeker and G. Brunklaus, *Solid State Nucl. Mag.*, 2015, **65**, 29-40.
40. D. Luedeker, R. Gossmann, K. Langer and G. Brunklaus, *Cryst. Growth Des.*, 2016, **16**, 3087-3100.
41. R. Lee, A. J. Firbank, M. R. Probert and J. W. Steed, *Cryst. Growth Des.*, 2016, **16**, 4005-4011.
42. S. M. Pratik and A. Datta, *J. Phys. Chem. B*, 2016, **120**, 7606-7613.
43. A. M. Reilly, R. I. Cooper, C. S. Adjiman, S. Bhattacharya, A. D. Boese, J. G. Brandenburg, P. J. Bygrave, R. Bylsma, J. E. Campbell, R. Car, D. H. Case, R. Chadha, J. C. Cole, K. Cosburn, H. M. Cuppen, F. Curtis, G. M. Day, R. A. DiStasio, A. Dzyabchenko, B. P. van Eijck, D. M. Elking, J. A. van den Ende, J. C. Facelli, M. B. Ferraro, L. Fusti-Molnar, C. A. Gatsiou, T. S. Gee, R. de Gelder, L. M. Ghiringhelli, H. Goto, S. Grimme, R. Guo, D. W. M. Hofmann, J. Hoja, R. K. Hylton, L. Iuzzolino, W. Jankiewicz, D. T. de Jong, J. Kendrick, N. J. J. de Klerk, H. Y. Ko, L. N. Kuleshova, X. Y. Li, S. Lohani, F. J. J. Leusen, A. M. Lund, J. Lv, Y. M. Ma, N. Marom, A. E. Masunov, P. McCabe, D. P. McMahon, H. Meekes, M. P. Metz, A. J. Misquitta, S. Mohamed, B. Monserrat, R. J. Needs, M. A. Neumann, J. Nyman, S. Obata, H. Oberhofer, A. R. Oganov, A. M. Orendt, G. I. Pagola, C. C. Pantelides, C. J. Pickard, R. Podeszwa, L. S. Price, S. L. Price, A. Pulido, M. G. Read, K. Reuter, E. Schneider, C. Schober, G. P. Shields, P. Singh, I. J. Sugden, K. Szalewicz, C. R. Taylor, A. Tkatchenko, M. E. Tuckerman, F. Vacarro, M. Vasileiadis, A. Vazquez-Mayagoitia, L. Vogt, Y. C. Wang, R. E. Watson, G. A. de Wijs, J. Yang, Q. Zhu and C. R. Groom, *Acta Crystallogr. Sect. B*, 2016, **72**, 439-459.
44. S. L. Price, *Chem. Soc. Rev.*, 2014, **43**, 2098-2111.
45. A. J. Cruz Cabeza, G. M. Day, W. D. S. Motherwell and W. Jones, *J. Am. Chem. Soc.*, 2006, **128**, 14466-14467.
46. C. B. Aakeroy, A. M. Beatty and B. A. Helfrich, *Angew. Chem. Int. Ed.*, 2001, **40**, 3240-+.

47. D. A. Adsmond, A. S. Sinha, U. B. R. Khandavilli, A. R. Maguire and S. E. Lawrence, *Cryst. Growth Des.*, 2016, **16**, 59-69.
48. A. Otero-de-la-Roza, B. H. Cao, I. K. Price, J. E. Hein and E. R. Johnson, *Angew. Chem. Int. Ed.*, 2014, **53**, 7879-7882.
49. W. Deng, Y. You, A. Zhang and A. D. l. o. o. p. Zhang, in *Recent Advances in Smart Self-Healing Polymers and Composites*, Woodhead Publishing, Cambridge, 2015, pp. 181-210.
50. G. R. Desiraju and T. Steiner, *The Weak Hydrogen Bond in Structural Chemistry and Biology*, Oxford University Press, Oxford, 2001.
51. D. K. Bucar, G. M. Day, I. Halasz, G. G. Z. Zhang, J. R. G. Sander, D. G. Reid, L. R. MacGillivray, M. J. Duer and W. Jones, *Chem. Sci.*, 2013, **4**, 4417-4425.
52. C. J. Aspley, C. Boxwell, M. L. Buil, C. L. Higgitt, C. Long and R. N. Perutz, *Chem. Commun.*, 1999, 1027-1028.
53. J. C. Ma and D. A. Dougherty, *Chem. Rev.*, 1997, **97**, 1303-1324.
54. G. Cavallo, P. Metrangolo, R. Milani, T. Pilati, A. Priimagi, G. Resnati and G. Terraneo, *Chem. Rev.*, 2016, **116**, 2478-2601.
55. H. Schmidbaur and A. Schier, *Chem. Soc. Rev.*, 2012, **41**, 370-412.
56. J. Starbuck, N. C. Norman and A. G. Orpen, *New J. Chem.*, 1999, **23**, 969-972.
57. G. D. Andretti, R. Ungaro and A. Pochini, *J. Chem. Soc.-Chem. Comm.*, 1979, 1005-1007.
58. K. M. Steed and J. W. Steed, *Chem. Rev.*, 2015, **115**, 2895-2933.
59. G. R. Desiraju, *Angew. Chem. Int. Ed.*, 1995, **34**, 2311-2327.
60. N. Blagden, M. de Matas, P. T. Gavan and P. York, *Adv. Drug Delivery Rev.*, 2007, **59**, 617-630.
61. M. L. Cheney, D. R. Weyna, N. Shan, M. Hanna, L. Wojtas and M. J. Zaworotko, *Cryst. Growth Des.*, 2010, **10**, 4401-4413.
62. N. Shan and M. J. Zaworotko, *Drug Discovery Today*, 2008, **13**, 440-446.
63. G. P. Stahly, *Cryst. Growth Des.*, 2009, **9**, 4212-4229.
64. N. J. Babu and A. Nangia, *Cryst. Growth Des.*, 2011, **11**, 2662-2679.
65. B. Swapna, D. Maddileti and A. Nangia, *Cryst. Growth Des.*, 2014, **14**, 5991-6005.
66. J. D. Wuest, *Nat. Chem.*, 2012, **4**, 74-75.
67. J. A. Bis, P. Vishweshwar, D. R. Weyna and M. J. Zaworotko, *Mol. Pharm.*, 2007, **4**, 401-416.
68. M. C. Etter, *Accounts Chem. Res.*, 1990, **23**, 120-126.
69. M. R. Edwards, W. Jones and W. D. S. Motherwell, *Cryst. Eng.*, 2002, **5**, 25-36.
70. J. Bernstein, R. E. Davis, L. Shimoni and N. L. Chang, *Angew. Chem. Int. Ed.*, 1995, **34**, 1555-1573.
71. M. C. Etter, J. C. Macdonald and J. Bernstein, *Acta Crystallogr. Sect. B*, 1990, **46**, 256-262.
72. D. P. McNamara, S. L. Childs, J. Giordano, A. Iarriccio, J. Cassidy, M. S. Shet, R. Mannion, E. O'Donnell and A. Park, *Pharm. Res.*, 2006, **23**, 1888-1897.
73. C. Sowa, R. E. Gold, T. Chiodo, and R. Vogel, *Co-crystals of Cyprodinil and Dithianon*, Google Patents, 2012.
74. S. A. Ross, D. A. Lamprou and D. Douroumis, *Chem. Commun.*, 2016, **52**, 8772-8786.

75. R. S. Dhumal, A. L. Kelly, P. York, P. D. Coates and A. Paradkar, *Pharm. Res.*, 2010, **27**, 2725-2733.
76. N. Schultheiss and A. Newman, *Cryst. Growth Des.*, 2009, **9**, 2950-2967.
77. Y. Chen, L. Li, J. Yao, Y. Y. Ma, J. M. Chen and T. B. Lu, *Cryst. Growth Des.*, 2016, **16**, 2923-2930.
78. S. Jain, *Pharm. Sci. Technol. To.*, 1999, **2**, 20-31.
79. E. N. Hiestand, *J. Pharm. Sci.*, 1985, **74**, 768-770.
80. E. N. Hiestand, *Int. J. Pharm*, 1991, **67**, 217-229.
81. E. N. Hiestand and D. P. Smith, *Int. J. Pharm*, 1991, **67**, 231-246.
82. G. Bolla and A. Nangia, *Chem. Commun.*, 2016, **52**, 8342-8360.
83. N. Blagden, S. J. Colesb and D. J. Berry, *CrystEngComm*, 2014, **16**, 5753-5761.
84. C. C. Sun and H. Hou, *Cryst. Growth Des.*, 2008, **8**, 1575-1579.
85. S. L. James, C. J. Adams, C. Bolm, D. Braga, P. Collier, T. Friscic, F. Grepioni, K. D. M. Harris, G. Hyett, W. Jones, A. Krebs, J. Mack, L. Maini, A. G. Orpen, I. P. Parkin, W. C. Shearouse, J. W. Steed and D. C. Waddell, *Chem. Soc. Rev.*, 2012, **41**, 413-447.
86. D. Braga, L. Maini and F. Grepioni, *Chem. Soc. Rev.*, 2013, **42**, 7638-7648.
87. D. J. Berry, C. C. Seaton, W. Clegg, R. W. Harrington, S. J. Coles, P. N. Horton, M. B. Hursthouse, R. Storey, W. Jones, T. Friscic and N. Blagden, *Cryst. Growth Des.*, 2008, **8**, 1697-1712.
88. D. Braga, F. Grepioni, G. I. Lampronti, L. Maini, K. Rubini, A. Turrina and F. Zorzi, *Crystengcomm*, 2014, **16**, 7452-7458.
89. K. D. M. Harris, R. L. Johnston, B. M. Kariuki, E. Tedesco, G. W. Turner, R. Delhez and E. J. Mittemeijer, *Epdic 7: European Powder Diffraction, Pts 1 and 2*, 2001, **378-3**, 38-46.
90. K. D. M. Harris and E. Y. Cheung, *Chem. Soc. Rev.*, 2004, **33**, 526-538.
91. K. Shankland, M. J. Spillman, E. A. Kabova, D. S. Edgeley and N. Shankland, *Acta Crystallogr. Sect. C*, 2013, **69**, 1251-1259.
92. A. Altomare, C. Cuocci, C. Giacovazzo, A. Moliterni, R. Rizzi, N. Corriero and A. Falcicchio, *J. Appl. Crystallogr.*, 2013, **46**, 1231-1235.
93. D. V. Dudenko, P. A. Williams, C. E. Hughes, O. N. Antzutkin, S. P. Velaga, S. P. Brown and K. D. M. Harris, *J. Phys. Chem. C*, 2013, **117**, 12258-12265.
94. D. H. Brouwer and K. P. Langendoen, *Crystengcomm*, 2013, **15**, 8748-8762.
95. C. E. Hughes, G. N. M. Reddy, S. Masiero, S. P. Brown, P. A. Williams and K. D. M. Harris, *Chem. Sci.*, 2017, **8**, 3971-3979.
96. I. J. Lowe, *Phys. Rev. Lett.*, 1959, **2**, 285-287.
97. E. R. Andrew, A. Bradbury and R. G. Eades, *Nature*, 1958, **182**, 1659.
98. E. R. Andrew, A. Bradbury and R. G. Eades, *Nature*, 1959, **183**, 1802-1803.
99. S. E. Ashbrook and D. McKay, *Chem. Commun.*, 2016, **52**, 7186-7204.
100. B. Elena, G. Pintacuda, N. Mifsud and L. Emsley, *J. Am. Chem. Soc.*, 2006, **128**, 9555-9560.
101. R. K. Harris, *Solid State Sci.*, 2004, **6**, 1025-1037.

102. C. Bonhomme, C. Gervais, F. Babonneau, C. Coelho, F. Pourpoint, T. Azais, S. E. Ashbrook, J. M. Griffin, J. R. Yates, F. Mauri and C. J. Pickard, *Chem. Rev.*, 2012, **112**, 5733-5779.
103. T. Charpentier, *Solid State Nucl. Magn.*, 2011, **40**, 1-20.
104. D. L. Bryce, *Iucrj*, 2017, **4**, 350-359.
105. M. R. Chierotti and R. Gobetto, *CrystEngComm*, 2013, **15**, 8599-8612.
106. S. P. Brown, *Solid State Nucl. Magn.*, 2012, **41**, 1-27.
107. C. J. Pickard and F. Mauri, *Phys. Rev. B*, 2001, **63**, 245101.
108. C. M. Widdifield and D. L. Bryce, *Phys.Chem. Chem. Phys.*, 2009, **11**, 7120-7122.
109. C. M. Widdifield, H. Robson and P. Hodgkinson, *Chem. Commun.*, 2016, **52**, 6685-6688.
110. M. T. Weller, P. F. Henry, V. P. Ting and C. C. Wilson, *Chem. Commun.*, 2009, 2973-2989.
111. C. Bilton, J. A. K. Howard, N. N. L. Madhavi, G. R. Desiraju, F. H. Allen and C. C. Wilson, *Acta Crystallogr. Sect. B*, 2001, **57**, 560-566.
112. L. H. Thomas, A. J. Florence and C. C. Wilson, *New J. Chem.*, 2009, **33**, 2486-2490.
113. K. Fucke and J. W. Steed, *Water*, 2010, **2**, 333-350.
114. D. V. Dudenko, J. R. Yates, K. D. M. Harris and S. P. Brown, *CrystEngComm*, 2013, **15**, 8797-8807.
115. C. M. Gowda, F. Vasconcelos, E. Schwartz, E. R. H. van Eck, M. Marsman, J. J. L. M. Cornelissen, A. E. Rowan, G. A. de Wijs and A. P. M. Kentgens, *Phys.Chem. Chem. Phys.*, 2011, **13**, 13082-13095.
116. P. Paluch, T. Pawlak, M. Oszajca, W. Lasocha and M. J. Potrzebowski, *Solid State Nucl.Magn.*, 2015, **65**, 2-11.
117. G. N. M. Reddy, D. S. Cook, D. Iuga, R. I. Walton, A. Marsh and S. P. Brown, *Solid State Nucl.r Magn.*, 2015, **65**, 41-48.
118. J. R. Yates, T. N. Pham, C. J. Pickard, F. Mauri, A. M. Amado, A. M. Gil and S. P. Brown, *J. Am. Chem. Soc.*, 2005, **127**, 10216-10220.
119. S. P. Brown and H. W. Spiess, *Chem. Rev.*, 2001, **101**, 4125-4155.
120. A. C. Poppler, E. K. Corlett, H. Pearce, M. P. Seymour, M. Reid, M. G. Montgomery and S. P. Brown, *Acta Crystallogr. Sect. C*, 2017, **73**, 149-156.
121. L. Mafra, S. M. Santos, R. Siegel, I. Alves, F. A. A. Paz, D. Dudenko and H. W. Spiess, *J. Am. Chem. Soc.*, 2012, **134**, 71-74.
122. M. Zilka, S. Sturniolo, S. P. Brown and J. R. Yates, *J. Chem. Phys.*, 2017, **147**, 144203.
123. J. Schmidt, A. Hoffmann, H. W. Spiess and D. Sebastiani, *J. Phys. Chem. B*, 2006, **110**, 23204-23210.
124. C. Gervais, M. Profeta, V. Lafond, C. Bonhomme, T. Azaïs, H. Mutin, C. J. Pickard, F. Mauri and F. Babonneau, *Magn. Reson. Chem.*, 2004, **42**, 445-452.
125. A. C. Uldry, J. M. Griffin, J. R. Yates, M. Perez-Torralba, M. D. Maria, A. L. Webber, M. L. L. Beaumont, A. Samoson, R. M. Claramunt, C. J. Pickard and S. P. Brown, *J. Am. Chem. Soc.*, 2008, **130**, 945-954.

126. *British National Formulary*, The National Institute for Health and Care Excellence (NICE), http://bnf.nice.org.uk/, 2018.
127. N. Vogt, M. A. Abaev and N. M. Karasev, *J. Mol. Struct.*, 2011, **987**, 199-205.
128. E. M. S. Macoas, R. Fausto, J. Lundell, M. Pettersson, L. Khriachtchev and M. Rasanen, *J. Phys. Chem. A*, 2001, **105**, 3922-3933.
129. https://www.sigmaaldrich.com/catalog/product/mm/800269?lang=en&region=GB&cm_sp=Insite-_-prodRecCold_xviews-_-prodRecCold10-2, (accessed 04/08/2019, 2019).
130. *CRC Handbook of Chemistry and Physics*, ed. W. M. Haynes, 91st edn., 2010, ch. 8, pp. 42-51.
131. https://www.sigmaaldrich.com/catalog/product/aldrich/a75706?lang=en&region=GB, (accessed 30/09/2019, 2019).
132. https://www.chemicalbook.com/ChemicalProductProperty_EN_CB3853241.htm (accessed 25/8/2019, 2019).
133. https://www.sigmaaldrich.com/catalog/product/aldrich/a75684?lang=en&region=GB&cm_sp=Insite-_-prodRecCold_xviews-_-prodRecCold10-5, (accessed 29/09/2019, 2019).
134. https://www.chemicalbook.com/ChemicalProductProperty_EN_CB0485858.htm, (accessed 25/08/2019, 2019).
135. https://www.sigmaaldrich.com/catalog/product/aldrich/662704?lang=en&region=GB&cm_sp=Insite-_-prodRecCold_xviews-_-prodRecCold10-5, (accessed 30/09/2019, 2019).
136. https://www.sigmaaldrich.com/catalog/product/aldrich/151351?lang=en&region=GB, (accessed 30/09/2019, 2019).
137. https://www.sigmaaldrich.com/catalog/product/aldrich/336106?lang=en&region=GB&gclid=Cj0KCQjwz8bsBRC6ARIsAEyNnvr94zW82Qu5WiRcWk2p8PlyHVwkSIJJo_QS8Ft_hkRwzYbrzxcn67MaAr1SEALw_wcB, (accessed 04/08/2019, 2019).
138. C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, *Acta Crystallogr. Sect. B*, 2016, **72**, 171-179.
139. Y. J. Pan, Z. M. Jin, C. R. Sun and C. W. Jiang, *Chem. Lett.*, 2001, **30**, 1008-1009.
140. D. A. Haynes, W. Jones and W. D. S. Motherwell, *CrystEngComm*, 2006, **8**, 830-840.
141. M. Hemamalini and H. K. Fun, *Acta Crystallogr. Sect. E*, 2010, **66**, O2093-U1847.
142. S. Selyani and M. Dincer, *Mol. Cryst. Liq. Cryst.*, 2018, **666**, 65-78.
143. *X-ray Crystallography*, Oxford University Press, Oxford, 2nd edn., 2015.
144. C. Giacovazzo, H. L. Monaco, D. Viterbo, F. Scordari, G. Gilli, G. Zanotti and M. Catti, *Fundamentals of Crystallography*, Oxford University Press, Oxford, Second edn., 1992.
145. *Structure from Diffraction Methods*, Wiley, Chichester, 2014.
146. B. H. Toby, *Powder Diffraction*, 2006, **21**, 67-70.
147. Agilent, *CrysAlisPRO*, Agilent Technologies Ltd., 2014.
148. G. M. Sheldrick, *Acta Crystallogr. Sect. A*, 2008, **64**, 112-122.
149. O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. K. Howard and H. Puschmann, *J. Appl. Crystallogr.*, 2009, **42**, 339-341.
150. A. A. Coehlo, *TOPAS-Academic V6*, Coelho Software, 2016.

151. A. Le Bail, H. Duroy and J. L. Fourquet, *Matt. Res. Bull.*, 1988, **23**, 447-452.
152. H. M. Rietveld, *J. Appl. Crystallogr.*, 1969, **2**, 65-71.
153. P. J. Hore, J. A. Jones and S. Wimperis, *NMR: The Toolkit*, Oxford University Press, Oxford, 2000.
154. M. J. Duer, *Introduction to Solid-State NMR Spectroscopy*, Blackwell, Oxford, 2004.
155. M. H. Levitt, *Spin Dynamics*, Wiley, Chichester, 2001.
156. J. Keeler, *Understanding NMR Spectroscopy*, Wiley, Chichester, 2010.
157. U. Haeberlen, *Adv. Magn. Reson.*, 1976.
158. D. J. States, R. A. Haberkorn and D. J. Ruben, *J. Magn. Reson.*, 1982, **48**, 286-292.
159. D. Marion, M. Ikura, R. Tschudin and A. Bax, *J. Magn. Reson.*, 1989, **85**, 393-399.
160. B. M. Fung, A. K. Khitrin and K. Ermolaev, *J. Magn. Reson.*, 2000, **142**, 97-101.
161. D. Sakellariou, A. Lesage, P. Hodgkinson and L. Emsley, *Chem. Phys. Lett.*, 2000, **319**, 253-260.
162. B. Elena, G. de Paepe and L. Emsley, *Chem. Phys. Lett.*, 2004, **398**, 532-538.
163. S. R. Hartmann and E. L. Hahn, *Phys. Rev.*, 1962, **128**, 2042-2053.
164. G. Metz, X. L. Wu and S. O. Smith, *J. Magn. Reson. Ser. A*, 1994, **110**, 219-227.
165. W. Sommer, J. Gottwald, D. E. Demco and H. W. Spiess, *J. Magn. Reson. Ser. A*, 1995, **113**, 131-134.
166. I. Schnell, A. Lupulescu, S. Hafner, D. E. Demco and H. W. Spiess, *J. Magn. Reson.*, 1998, **133**, 61-69.
167. J. P. Bradley, C. Tripon, C. Filip and S. P. Brown, *Phys. Chem. Chem. Phys.*, 2009, **11**, 6941-6952.
168. J. P. Bradley, C. J. Pickard, J. C. Burley, D. R. Martin, L. P. Hughes, S. D. Cosgrove and S. P. Brown, *J. Pharm. Sci.*, 2012, **101**, 1821-1830.
169. S. Cavadini, S. Antonijevic, A. Lupulescu and G. Bodenhausen, *J. Magn. Reson.*, 2006, **182**, 168-172.
170. S. Cavadini, *Prog. Nucl. Mag. Res. Sp.*, 2010, **56**, 46-77.
171. A. S. Tatton, J. P. Bradley, D. Iuga and S. P. Brown, *Z. Phys. Chem.*, 2012, **226**, 1187-1203.
172. Z. H. Gan, J. P. Amoureux and J. Trebosc, *Chem. Phys. Lett.*, 2007, **435**, 163-169.
173. M. H. Levitt, T. G. Oas and R. G. Griffin, *Is. J. Chem.* 1988, **28**, 271-282.
174. T. G. Oas, R. G. Griffin and M. H. Levitt, *J. Chem. Phys.*, 1988, **89**, 692-695.
175. S. Hayashi and K. Hayamizu, *Bull. Chem. Soc. Jpn.*, 1991, **64**, 685-687.
176. C. R. Morcombe and K. W. Zilm, *J. Magn. Reson.*, 2003, **162**, 479-486.
177. S. Hayashi and K. Hayamizu, *Bull. Chem. Soc. Jpn*, 1991, **64**, 688-690.
178. P. Hohenberg and W. Kohn, *Phys. Rev. B*, 1964, **136**, 864-871.
179. W. Kohn and L. J. Sham, *Phys. Rev. A*, 1965, **140**, 1133-1138.
180. J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865-3868.
181. J. Perdew, S. Kurth, A. Zupan and P. Blaha, *Phys. Rev. Lett.*, 1999, **82**, 2544-2547.
182. J. M. Tao, J. P. Perdew, V. N. Staroverov and G. E. Scuseria, *Phys. Rev. Lett.*, 2003, **91**, 146401.

183. P. J. Stephens, K. J. Jalkanen, F. J. Devlin and C. F. Chabalowski, *J. Phys. Chem.*, 1993, **97**, 6107-6110.
184. C. Adamo and V. Barone, *Chem. Phys. Lett.*, 1998, **298**, 113-119.
185. Z. G. Wu and R. E. Cohen, *Phys. Rev. B*, 2006, **73** 235116.
186. J. Klimes and A. Michaelides, *J. Chem. Phys.*, 2012, **137**, 12708.
187. S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.
188. A. Tkatchenko and M. Scheffler, *Phys. Rev. Lett.*, 2009, **102**, 073005.
189. G. Roman-Perez and J. M. Soler, *Phys. Rev. Lett.*, 2009, **103**, 096102.
190. F. Furche and T. Van Voorhis, *J. Chem. Phys.*, 2005, **122**, 164106.
191. T. Schwabe and S. Grimme, *Phys. Chem. Chem. Phys.*, 2007, **9**, 3397-3406.
192. A. Tkatchenko, R. DiStasio, R. Car and M. Scheffler, *Phys. Rev. Lett.*, 2012, **108**, 236402.
193. M. W. Cole, D. Velegol, H. Y. Kim and A. A. Lucas, *Mol. Sim.*, 2009, **35**, 849-866.
194. S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson and M. C. Payne, *Z. Kristallogr.*, 2005, **220**, 567-570.
195. I. J. Bruno, J. C. Cole, M. Kessler, J. Luo, W. D. S. Motherwell, L. H. Purkis, B. R. Smith, R. Taylor, R. I. Cooper, S. E. Harris and A. G. Orpen, *J. Chem. Inf. Comp. Sci.*, 2004, **44**, 2133-2144.
196. A. C. Poppler, D. Walker and S. P. Brown, *CrystEngComm*, 2017, **19**, 224-236.
197. R. K. Harris, P. Hodgkinson, C. J. Pickard, J. R. Yates and V. Zorin, *Magn. Reson. Chem.*, 2007, **45**, S174-S186.
198. T. Venâncio, L. M. Oliveira, J. Ellena, N. Boechat and S. P. Brown, *Solid State Nuclear Magn. Reson.*, 2017, **87**, 73-79.
199. F. M. Paruzzo, A. Hofstetter, F. Musil, S. De, M. Ceriotti and L. Emsley, *Nat. Commun.*, 2018, **9**, 4501.
200. J. D. Hartman, R. A. Kudla, G. M. Day, L. J. Mueller and G. J. O. Beran, *Phys. Chem. Chem. Phys.*, 2016, **18**, 21686-21709.
201. E. K. Corlett, H. Blade, L. P. Hughes, P. J. Sidebottom, D. Walker, R. I. Walton and S. P. Brown, *CrystEngComm*, 2019, **21**, 3502-3516.
202. R. K. Harris, P. Y. Ghi, H. Puschmann, D. C. Apperley, U. J. Griesser, R. B. Hammond, C. Y. Ma, K. J. Roberts, G. J. Pearce, J. R. Yates and C. J. Pickard, *Org. Proc. Res. Dev.*, 2005, **9**, 902-910.
203. J. R. Yates, S. E. Dobbins, C. J. Pickard, F. Mauri, P. Y. Ghi and R. K. Harris, *Phys. Chem. Chem. Phys.*, 2005, **7**, 1402-1407.
204. A. L. Webber, L. Emsley, R. M. Claramunt and S. P. Brown, *J. Phys. Chem. A*, 2010, **114**, 10435-10442.
205. X. Z. Li, L. Tapmeyer, M. Bolte and J. van de Streek, *ChemPhysChem*, 2016, **17**, 2496-2502.
206. S. P. Brown, *Prog. Nucl. Mag. Res. Sp.*, 2007, **50**, 199-251.
207. S. P. Brown, in *Modern Methods in Solid-state NMR: A Practitioner's Guide*, ed. P. Hodgkinson, Royal Society of Chemistry, London, 2018, ch. 2, pp. 39-74.

208. R. K. Harris, P. Jackson, L. H. Merwin, B. J. Say and G. Hagele, *J. Chem. Soc.*, 1988, **84**, 3649-3672.
209. L. H. Merwin and S. D. Ross, *Magn. Reson. Chem.*, 1992, **30**, 440-448.
210. A. L. Bednowitz and B. Post, *Acta Crystallogr.*, 1966, **21**, 566-571.
211. https://www.sigmaaldrich.com/catalog/product/aldrich/336106?lang=en&region=GB&gclid=Cj0KCQjwz8bsBRC6ARIsAEyNnvr94zW82Qu5WiRcWk2p8PlyHVwkSIJJo_QS8Ft_hkRwzYbrzxcn67MaAr1SEALw_wcB, (accessed 04/08/2019, 2019).
212. O. Almarsson, M. B. Hickey, M. L. Peterson, S. L. Morissette, S. Soukasene, C. McNulty, M. Tawa, J. M. MacPhee and J. F. Remenar, *Cryst. Growth Des.*, 2003, **3**, 927-933.
213. C. R. Gardner, O. Almarsson, H. M. Chen, S. Morissette, M. Peterson, Z. Zhang, S. Wang, A. Lemmo, J. Gonzalez-Zugasti, J. Monagle, J. Marchionna, S. Ellis, C. McNulty, A. Johnson, D. Levinson and M. Cima, *Comput. Chem. Eng.*, 2004, **28**, 943-953.
214. E. Losev and E. Boldyreva, *Acta Crystallogr. Sect. C*, 2018, **74**, 177-185.
215. G. P. Stahly, *Cryst. Growth Des.*, 2007, **7**, 1007-1026.
216. S. Boothroyd, A. Kerridge, A. Broo, D. Buttar and J. Anwar, *Cryst. Growth Des.*, 2018, **18**, 1903-1908.
217. D. A. Haynes, W. Jones and W. D. S. Motherwell, *CrystEngComm*, 2005, **7**, 342-345.
218. P. W. Borthwick, *Acta Crystallogr. Sect. B*, 1980, **36**, 628-632.
219. V. Semenov, D. Samultsev and L. Krivdin, *Magn. Reson. Chem.*, 2015, **53**, 433-441.
220. R. Marek, A. Lycka, E. Kolehmainen, E. Sievanen and J. Tousek, *Curr. Org. Chem.*, 2007, **11**, 1154-1205.
221. R. Gobetto, C. Nervi, E. Valfre, M. R. Chierotti, D. Braga, L. Maini, F. Grepioni, R. K. Harris and P. Ghi, *Chem. Mater.*, 2005, **17**, 1457-1466.
222. R. Gobetto, C. Nervi, M. R. Chierotti, D. Braga, L. Maini, F. Grepioni, R. K. Harris and P. Hodgkinson, *Chem. -Eur. J.*, 2005, **11**, 7461-7471.
223. Z. J. Li, Y. Abramov, J. Bordner, J. Leonard, A. Medek and A. V. Trask, *J. Am. Chem. Soc.*, 2006, **128**, 8199-8210.
224. J. Stevens, S. Byard, C. Muryn and S. Schroeder, *J. Phys. Chem. B*, 2010, **114**, 13961-13969.
225. V. M. S. Gil and J. N. Murrell, *T. Faraday Soc.*, 1964, **60**, 248-255.
226. S. P. Brown, I. Schnell, J. D. Brand, K. Mullen and H. W. Spiess, *Phys. Chem. Chem. Phys.*, 2000, **2**, 1735-1745.
227. N. Muller and R. C. Reiter, *J. Chem. Phys.*, 1965, **42**, 3265-3269.
228. J. T. Arnold and M. E. Packard, *J. Chem. Phys.*, 1951, **19**, 1608-1609.
229. U. Liddel and N. F. Ramsey, *J. Chem. Phys.*, 1951, **19**, 1608.
230. K. Modig and B. Halle, *J. Am. Chem. Soc.*, 2002, **124**, 12031-12041.
231. Y. Li, Y. M. Jia, Z. W. Wang, X. H. Li, W. Feng, P. C. Deng and L. H. Yuan, *Rsc Advances*, 2014, **4**, 29702-29714.

232. A. L. Webber, B. Elena, J. M. Griffin, J. R. Yates, T. N. Pham, F. Mauri, C. J. Pickard, A. M. Gil, R. Stein, A. Lesage, L. Emsley and S. P. Brown, *Phys. Chem. Chem. Phys.*, 2010, **12**, 6970-6983.
233. J. N. Dumez and C. J. Pickard, *J. Chem. Phys.*, 2009, **130**, 104701.
234. R. K. Harris, S. A. Joyce, C. J. Pickard, S. Cadars and L. Emsley, *Phys. Chem. Chem. Phys.*, 2006, **8**, 137-143.
235. N. Vogt, M. A. Abaev, A. N. Rykov and I. F. Shishkov, *J. Mol. Struct.*, 2011, **996**, 120-127.
236. L. Rajput, M. Banik, J. R. Yarava, S. Joseph, M. K. Pandey, Y. Nishiyama and G. R. Desiraju, *IUCrJ*, 2017, **4**, 466-475.
237. T. Steiner, I. Majerz and C. C. Wilson, *Angew. Chem. Int. Ed.*, 2001, **40**, 2651-2654.
238. Z. Malarski, I. Majerz and T. Lis, *J. Mol. Struct.*, 1987, **158**, 369-377.
239. M. G. Siskos, M. I. Choudhary and I. P. Gerothanassis, *Org. Biomol. Chem.*, 2017, **15**, 4655-4666.
240. F. M. Paruzzo, B. J. Walder and L. Emsley, *J. Magn. Reson.*, 2019, **305**, 131-137.
241. M. G. Jain, K. R. Mote and P. K. Madhu, *Crystals*, 2019, **9**, 131-137.
242. P. Thureau, S. Sturniolo, M. Zilka, F. Ziarelli, S. Viel, J. R. Yates and G. Mollica, *Magn. Reson. Chem.*, 2019, **57**, 256-264.
243. A. J. Florence, *International Tables of Crystallography*, 2019, **H**, 433-441.
244. M. Husak, A. Jegorov, J. Czernek, J. Rohlicek, S. Zizkova, P. Vraspir, P. Kolesa, A. Fitch and J. Brus, *Cryst. Growth Des.*, 2019, **19**, 4625-4631.
245. A. Hofstetter, M. Balodis, F. M. Paruzzo, C. M. Widdifield, G. Steyanato, A. C. Pinon, P. J. Bygrave, G. M. Day and L. Emsley, *J. Am. Chem.l Soc.*, 2019, **141**, 16624-16634.
246. C. Guzman-Afonso, Y. L. Hong, H. Colaux, H. Iijima, A. Saitow, T. Fukumura, Y. Aoyama, S. Motoki, T. Oikawa, T. Yamazaki, K. Yonekura and Y. Nishiyama, *Nat. Commun.*, 2019, **10**, 3537.
247. B. L. Dale, N. R. Halcovitch, M. J. G. Peach and J. M. Griffin, *Magn. Reson. Chem.*, 2019, **57**, 230-242.

Appendix

A1 Investigating the Structure of 2,6-Lutidinium Hydrogen Fumarate

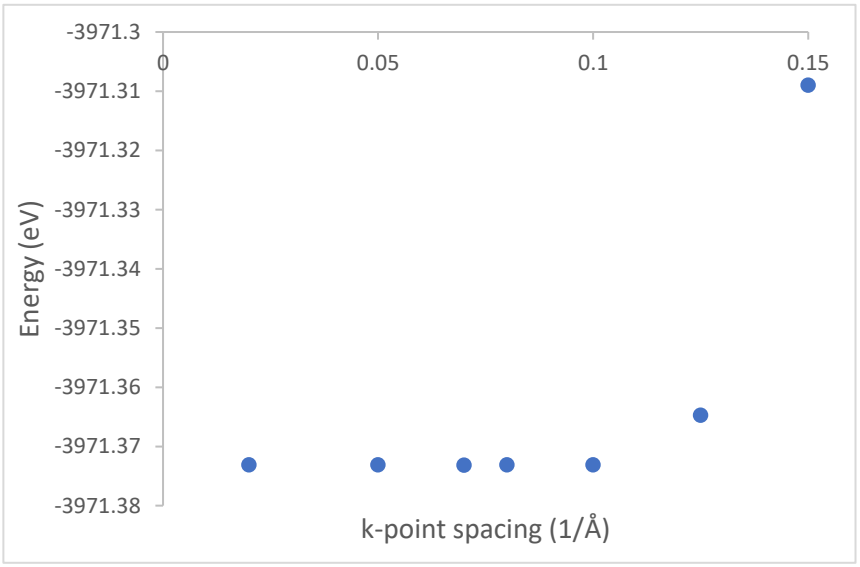


Figure A1.1: Convergence of energy with decreasing k-point spacing in the DFT (CASTEP) calculations for the literature structure (181445).

Table A1.1: Comparison of the fractional coordinates of the atoms in the asymmetric unit of the 26L:HF structure, for the literature structure (181445) and the Rietveld refined positions for the high-resolution synchrotron structure. All atoms have an occupancy of 1.

	181445[a]			MAC[b] RT			
Atom	**x**	**y**	**z**	**x**	**y**	**z**	**U_{iso} (Å^2)**
O1	0.30136(14)	0.86276(9)	0.01990(17)	0.2978(3)	0.86374(19)	0.0178(4)	1.94(10)
O2	0.08188(14)	0.81554(9)	−0.02052(18)	0.0775(3)	0.81443(19)	−0.0284(4)	2.16(10)
O3	0.26795(17)	0.85651(10)	0.66876(19)	0.2681(3)	0.85867(19)	0.6752(4)	2.11(10)
O4	0.4098(2)	0.96976(12)	0.6903(2)	0.4088(4)	0.9681(2)	0.6951(4)	4.72(14)
C8	0.20194(19)	0.84136(12)	0.0809(3)	0.1995(5)	0.8439(3)	0.0910(7)	0.50(15)
C9	0.22464(19)	0.84836(12)	0.2866(3)	0.2125(5)	0.8436(3)	0.3025(7)	0.79(14)
C10	0.3156(2)	0.90267(12)	0.3964(3)	0.3197(4)	0.9004(3)	0.3671(6)	0.63(15)
C11	0.3358(2)	0.91336(13)	0.6002(3)	0.3371(5)	0.9148(4)	0.5902(8)	3.4(2)
N	−0.15263(17)	0.77978(10)	0.0572(2)	−0.1491(4)	0.7833(4)	0.0584(5)	1.21(12)
C1	−0.0946(2)	0.62644(14)	0.0555(3)	−0.0866(4)	0.6209(3)	0.0576(5)	0.84(14)
C2	−0.1951(2)	0.69710(13)	0.0642(3)	−0.1890(6)	0.6934(4)	0.0580(5)	1.03(14)
C3	−0.3288(2)	0.68343(15)	0.0810(3)	−0.3259(6)	0.6775(3)	0.0848(5)	1.06(14)
C4	−0.4131(2)	0.75330(16)	0.0901(3)	−0.4180(4)	0.7533(4)	0.0897(5)	0.23(15)
C5	−0.3655(2)	0.83706(15)	0.0812(3)	−0.3703(6)	0.8387(4)	0.0726(6)	2.82(16)
C6	−0.2335(2)	0.85013(13)	0.0628(3)	−0.2317(7)	0.8548(4)	0.0597(6)	2.97(18)
C7	−0.1739(3)	0.93811(14)	0.0482(4)	−0.1618(4)	0.9432(3)	0.0620(5)	0.58(14)

[a] Structure determined by Pan *et al.*[139]
[b] Multi-Analysing Crystals, high resolution synchrotron scan

Table A1.2: Crystallographic data, instrumental parameters and final residuals for Rietveld refined model of 26L:HF against high-resolution synchrotron data recorded at 300 K.

Formula (Z = 4)	$C_{11}H_{13}O_4N$			
***a* (Å)**	***b* (Å)**	***c* (Å)**	***β* (°)**	***V* (Å^3)**
9.906752(18)	15.35923(3)	7. 501959(12)	107.78973(12)	1086.917(4)

Number of parameters refined	96	**Space group**	$P2_1/c$
Scale factor	0.0000698(2)	**Zero point**	−0.00237(2)
PseudoVoigt peak width parameters (°^{2})	U 0.00229(8)	V −0.00007(2)	W 0.00250(4)
R_{wp}	10.11 %	R_p	7.66 %
R_{exp}	3.07 %	R_{bragg}	5.59 %

Table A1.3: Longer-range C···H proximities between 1.2 and 2.8 Å for 26L:HF.

C	δ_{iso}^{exp} (ppm)	H	δ_{iso}^{exp} (ppm)	Separation[a] (Å)
C1	152.6	H10	17.7	2.11
		H4	2.1	2.12
		H6	2.1	2.13
		H1	6.3	2.15
		H5	2.1	2.15
C2	123.8	H2	7.0	2.16
		H6	2.1	2.69
C3	146.4	H3	7.9	2.17
		H1	6.3	2.17
C4	128.2	H2	7.0	2.14
		H9	2.1	2.64
C5	152.6	H10	17.7	2.09
		H9	2.1	2.13
		H7	2.1	2.14
		H3	7.9	2.15
		H8	2.1	2.15
C6	19.6	H10	17.7	2.60
		H1	6.3	2.77
C7	19.6	H10	17.7	2.57
		H3	7.9	2.77
C8	173.4	H11	7.9	2.21
		H13	*15.8*	*2.38*
		H10	*17.7*	*2.53*
		H12	7.0	2.75
C9	142.7	H12	7.0	2.13
C10	132.4	H11	7.9	2.12
C11	169.9	H13	15.8	2.00
		H12	7.0	2.19
		H11	7.9	2.73
		H2	*7.0*	*2.79*

[a] H-H distances are taken from the DFT (CASTEP) optimised structure. Intermolecular proximities are denoted using italic font.

Table A1.4: The fractional coordinates of the atoms in the asymmetric unit of the 26L:HF structural model refined against high-resolution synchrotron data recorded at 100 K (see Fig. 3.6).

Atom	x	y	z	U_{iso} (Å^2)
O1	0.2015(7)	0.3633(4)	0.9938(8)	1.6(2)
O2	0.4325(7)	0.3129(4)	1.0401(8)	0.50(19)
O3	0.2306(6)	0.3553(4)	0.3237(8)	0.54(18)
O4	0.0971(7)	0.4692(4)	0.3150(8)	0.9(2)
C8	0.3048(12)	0.3378(6)	0.9306(16)	–0.5(3)
C9	0.2836(10)	0.3391(6)	0.7157(14)	0.1(3)
C10	0.1838(10)	0.4028(7)	0.6286(14)	1.3(3)
C11	0.1764(11)	0.4102(9)	0.4447(19)	4.6(4)
N	0.6530(8)	0.2762(7)	0.9528(10)	1.1(2)
C1	0.6990(13)	0.1914(8)	0.9569(13)	2.5(3)
C2	0.8357(13)	0.1713(6)	0.9261(11)	1.7(3)
C3	0.9130(9)	0.2484(8)	0.9087(12)	0.3(3)
C4	0.8719(11)	0.3396(7)	0.9228(12)	–0.1(3)
C5	0.7304(13)	0.3594(7)	0.9383(13)	2.0(3)
C6	0.5933(7)	0.1099(5)	0.9585(9)	–2.3(2)
C7	0.6668(9)	0.4436(6)	0.9566(10)	–0.5(3)

Table A1.5: Crystallographic data, instrumental parameters and final residuals for Rietveld refined model of 26L:HF against high-resolution synchrotron data recorded at 100 K.

Formula (Z = 4)	$C_{11}H_{13}O_4N$			
a **(Å)**	***b*** **(Å)**	***c*** **(Å)**	***β*** **(°)**	***V*** **(Å^3)**
9.83062(5)	15.16849(9)	7.48175(3)	108.8958(4)	1055.524(9)

Number of parameters refined	96	**Space group**	$P2_1/c$
Scale factor	0.00000729(5)	**Zero point**	–0.00232(6)
PseudoVoigt peak width parameters (°^{2})	U 0.0037(3)	V –0.00018(6)	W 0.0025(2)
R_{wp}	19.37 %	R_p	14.95 %
R_{exp}	18.88 %	R_{bragg}	4.92 %

Table A1.6: GIPAW calculated chemical shifts for the CCDC structure 181445 of 26L:HF and 1876100 and the difference between them.

Atom	181445 δ_{iso}^{calc} (ppm)	1876100 δ_{iso}^{calc} (ppm)	$\Delta\delta_{iso}^{calc}$ (ppm)
H4/H5/H6	2.1	2.1	0.0
H7/H8/H9	2.1	2.1	0.0
H1	6.3	6.4	0.0
H2	6.9	6.8	0.0
H12	7.5	7.5	0.0
H3	7.9	7.9	0.0
H11	8.0	8.0	0.1
H13	17.7	17.6	−0.1
H10	17.7	17.9	0.2
C7	14.3	14.6	0.3
C6	15.5	16.1	0.6
C2	124.3	124.7	0.3
C4	129.1	129.3	0.1
C10	134.7	135.1	0.3
C9	145.3	144.9	−0.4
C3	145.8	145.5	−0.3
C5	152.4	152.1	−0.3
C1	152.6	152.6	0.1
C11	172.3	172.5	0.2
C8	175.2	175.4	0.2

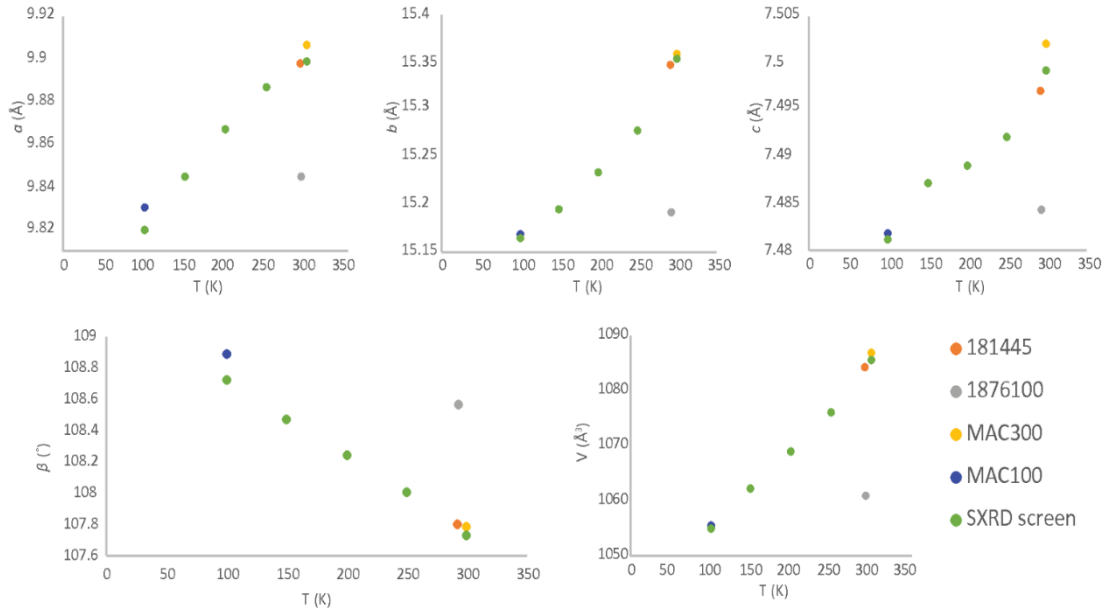


Figure A1.2: Graphs showing the change in the value of the unit cell parameters of 2,6-lutidinium hydrogen fumarate with temperature. MAC refers to high resolution synchrotron PXRD data recorded with a Multi-Analysing Crystal detector.

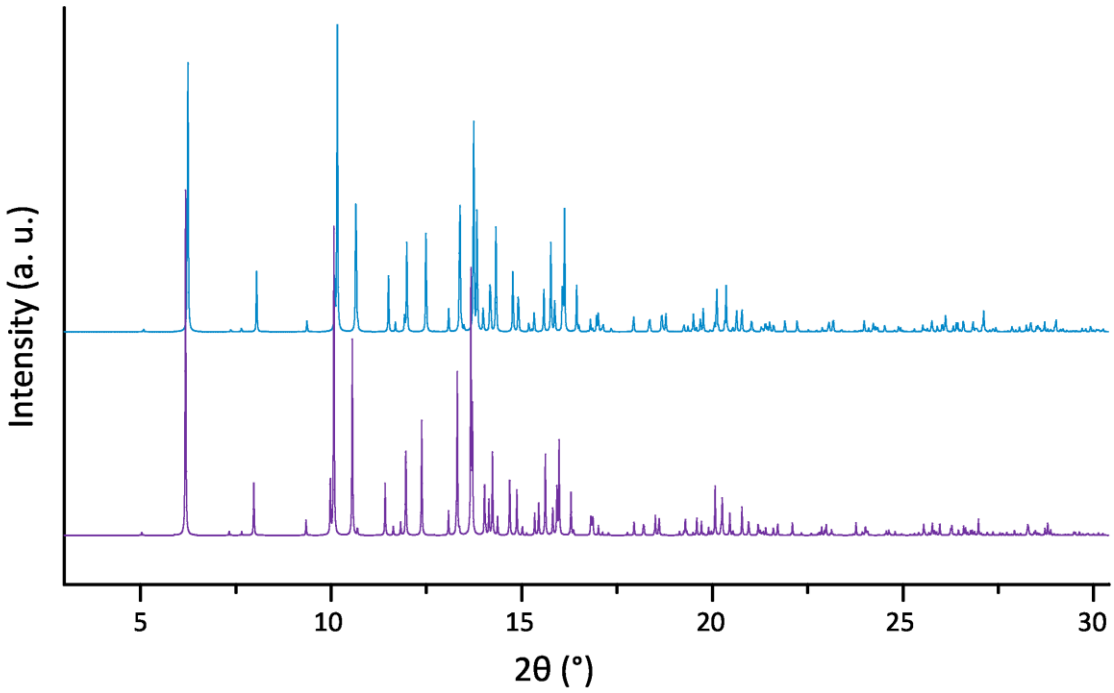


Figure A1.3: Simulated PXRD patterns of 2,6-lutidinium hydrogen fumarate for the CCDC structures 181445 (purple) and 1876100 (blue).

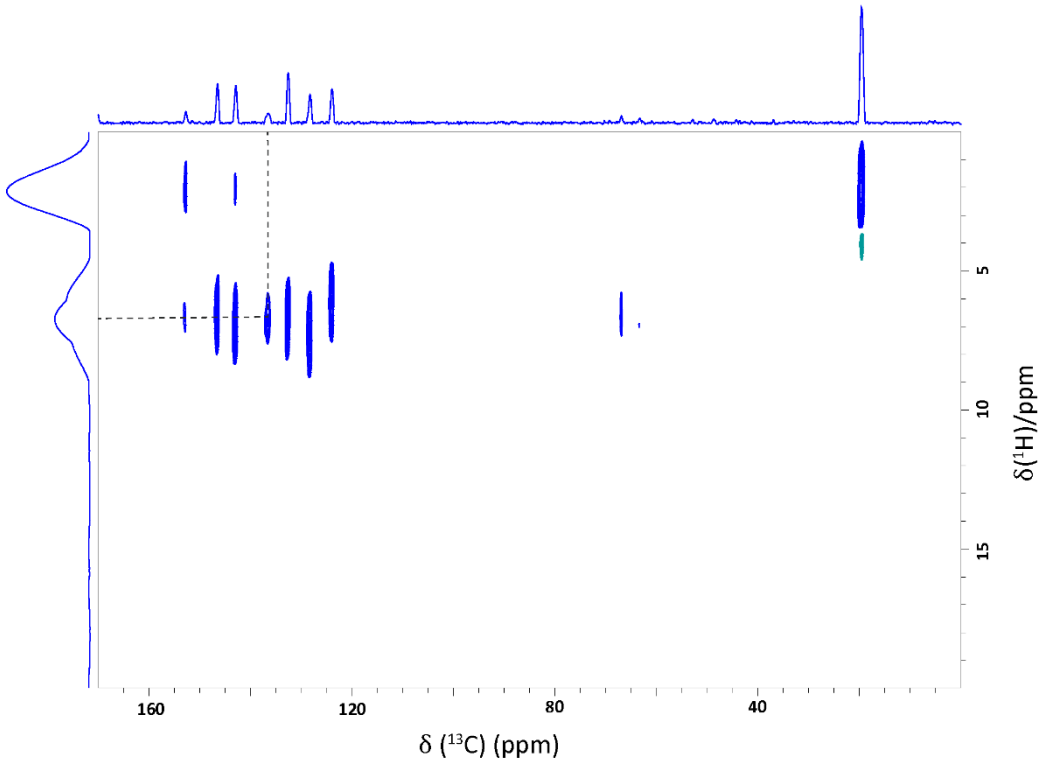


Figure A1.4: A ^{1}H (500 MHz)-^{13}C CP (200 μs) HETCOR MAS (12.5 kHz) NMR spectrum of 2,6-lutidinium hydrogen fumarate with fumaric acid present. The dashed grey lines show the CH correlation for the fumaric acid backbone. The base contour level is at 6.2% of the maximum peak height.

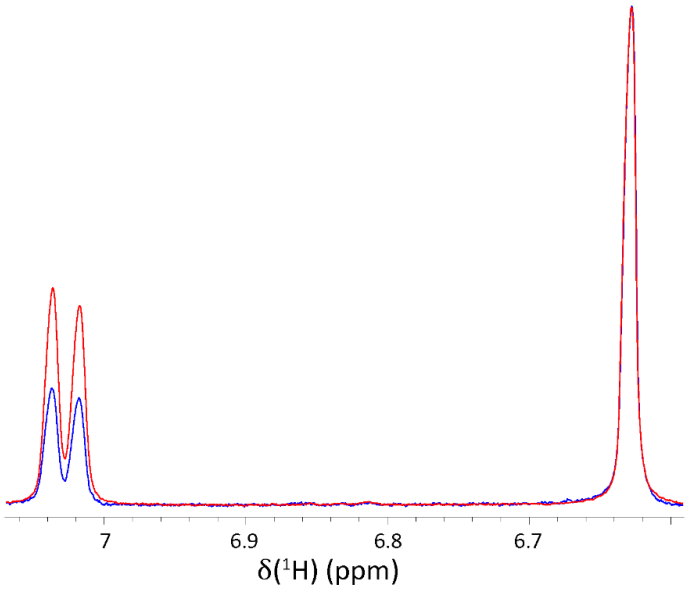


Figure A1.5: Solution-state ^{1}H (400 MHz, one pulse) NMR spectra of 2,6-lutidinium hydrogen fumarate dissolved in d6-DMSO. A sample made up freshly from a single crystal is shown in red and a sample made up from a powder that had been stored under ambient conditions for more than 2 weeks is shown in blue. The doublet at 7.01 ppm corresponds to H1 and H3 on 2,6-lutidine and the singlet at 6.61 ppm corresponds to H11 and H12 on fumaric acid.

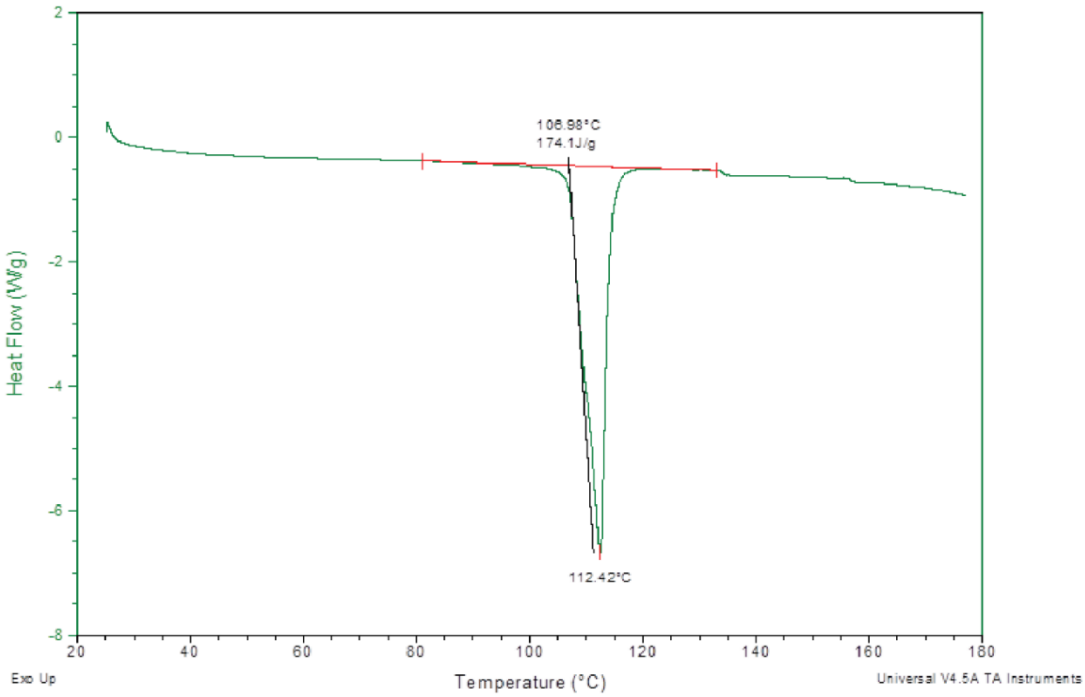


Figure A1.6:DSC of small plate crystals of 2,6-lutidinium hydrogen fumarate recorded on a TA Instruments DSC 2000 with a ramp of 10°C/min from 20-180°C.

A2 Exploring the Polymorphism of 2-Amino-6-Methylpyridine Salts

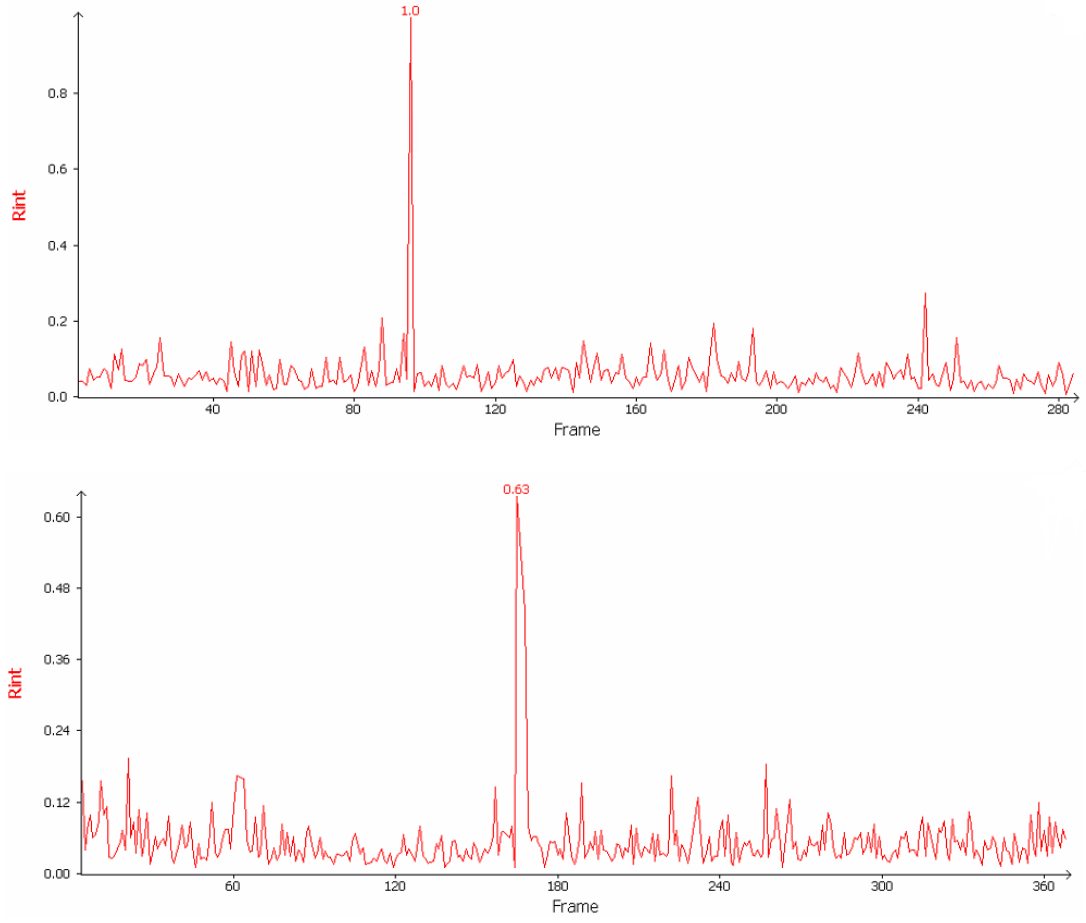


Figure A2.1: Variation in R_{int} throughout SXRD data collection for a single crystal, initially 26AMP:$F_{0.5}$:$(H_2O)_2$, when fmeasured at 67 °C (top) and at 77 °C (bottom).

Table A2.1: Comparison of the fractional coordinates of the atoms in the asymmetric unit of 26AMP$_2$:F:FA$_{0.5}$, 26AMP$_2$:F:H$_2$O and 26AMP:F$_{0.5}$:(H$_2$O)$_2$ and the Rietveld refined O and N atomic positions for the RT PXRD, alongside thermal parameters following refinement.

	26AMP$_2$:F:FA$_{0.5}$			**Refined model**			
Atom	**x**	**y**	**z**	**x**	**y**	**z**	**Uiso (Å^2)**
O1	0.37952(12)	0.11147(9)	0.83926(7)	0.38(5)	0.11(5)	0.81(5)	10(2)
O2	0.67553(12)	0.02884(10)	0.76120(7)	0.68(10)	0.03(5)	0.77(6)	11(2)
C1	0.54482(16)	0.04213(12)	0.84075(9)	-	-	-	14(5)
C2	0.56072(16)	−0.02188(12)	0.95115(10)	-	-	-	5(2)
O5	0.19028(13)	0.36447(11)	0.12566(8)	0.18(10)	0.38(10)	0.14(6)	16(3)
O6	0.38911(13)	0.37955(11)	−0.03399(8)	0.3(1)	0.4(1)	0.22(8)	16(4)
N1	0.66166(14)	0.24883(10)	0.05658(8)	0.66(10)	0.26(6)	0.05(7)	3(2)
N2	0.46015(15)	0.24424(12)	0.22839(9)	0.47(10)	0.22(6)	0.22(8)	12(3)
C9	0.62806(17)	0.21972(12)	0.16798(10)	-	-	-	17(6)
C5	0.23273(16)	0.40386(12)	0.02579(10)	-	-	-	10(5)
C6	0.08578(17)	0.48536(12)	−0.02574(10)	-	-	-	17(6)
C10	0.77339(19)	0.16306(15)	0.21417(11)	-	-	-	18(6)
C11	0.9419(2)	0.13692(16)	0.14529(13)	-	-	-	9(3)
C12	0.82896(18)	0.22285(13)	−0.01318(11)	-	-	-	8(4)
C13	0.97145(19)	0.16613(15)	0.02955(12)	-	-	-	9(3)
C14	0.8427(2)	0.26011(17)	−0.13444(11)	-	-	-	5(3)
O9	0.39381(14)	0.25840(9)	0.47192(7)	0.39(7)	0.25(5)	0.46(5)	9(2)
O10	0.33652(14)	0.23856(9)	0.65853(7)	0.29(8)	0.26(6)	0.66(5)	15(3)
N5	0.23035(15)	0.54019(10)	0.50389(9)	0.22(1)	0.55(6)	0.49(7)	6(3)
N6	0.05704(18)	0.49777(12)	0.67492(10)	0.08(10)	0.50(9)	0.67(5)	9(2)
C21	0.40103(18)	0.18833(12)	0.55826(10)	-	-	-	7(2)
C22	0.48788(19)	0.03523(12)	0.54531(10)	-	-	-	12(3)
C25	0.09224(18)	0.58663(13)	0.59881(11)	-	-	-	3(3)
C26	0.28134(19)	0.62337(14)	0.42207(11)	-	-	-	9(4)
C27	0.1853(2)	0.76089(15)	0.43469(13)	-	-	-	0.05(18)
C28	−0.0096(2)	0.72865(14)	0.61205(12)	-	-	-	10(4)
C29	0.0370(2)	0.81219(14)	0.53035(14)	-	-	-	7(3)
C30	0.4403(2)	0.55465(16)	0.32465(12)	-	-	-	−3(2)
	26AMP$_2$:F:H$_2$O						
O1	0.34922(15)	0.73502(10)	0.56790(7)	0.35(8)	0.69(7)	0.55(5)	-2(3)
O2	0.29036(16)	0.77200(10)	0.39158(7)	0.37(12)	0.73(9)	0.42(9)	8(4)
O5	0.20760(15)	0.30579(12)	−0.05622(7)	0.13(11)	0.26(8)	-0.11(7)	4(3)
O9	0.0383(2)	0.89385(13)	0.24187(10)	0.03(12)	0.88(1)	0.26(8)	7(4)
N1	0.23302(15)	0.45753(11)	0.50131(8)	0.24(24)	0.49(7)	0.50(12)	19(7)
O6	0.24558(19)	0.38997(14)	0.11610(8)	0.32(2)	0.40(9)	0.11(9)	9(4)
N5	0.52348(15)	0.21258(11)	0.02084(8)	0.57(9)	0.23(9)	0.05(11)	2(4)
N6	0.48721(19)	0.21860(14)	0.19408(9)	0.4(3)	0.26(5)	0.17(3)	12(6)
N2	0.2710(2)	0.49478(15)	0.33153(10)	0.27(6)	0.47(11)	0.32(9)	5(4)
C1	0.36265(18)	0.81125(13)	0.49392(9)	-	-	-	-10(3)
C21	0.58248(18)	0.18941(13)	0.12784(10)	-	-	-	-9(3)
C5	0.17158(19)	0.38017(14)	0.01407(10)	-	-	-	-5(3)
C9	0.22885(18)	0.40633(14)	0.39914(10)	-	-	-	19(10)
C10	0.19386(18)	0.37673(14)	0.57774(11)	-	-	-	-10(2)
C22	0.60609(19)	0.18248(13)	−0.05577(11)	-	-	-	-2(6)
C2	0.47632(19)	0.95706(13)	0.53315(10)	-	-	-	-1(3)
C6	0.0299(2)	0.46171(15)	−0.03059(11)	-	-	-	-5(3)
C11	0.1809(2)	0.26156(15)	0.37111(12)	-	-	-	-9(3)
C12	0.2072(3)	0.45100(16)	0.68632(12)	-	-	-	-6(3)
C23	0.7399(2)	0.13253(16)	0.16106(13)	-	-	-	2(3)
C13	0.1472(2)	0.23661(15)	0.55138(13)	-	-	-	19(9)
C24	0.7566(2)	0.12533(16)	−0.02446(14)	-	-	-	-4(5)
C25	0.5237(3)	0.21729(19)	−0.16911(12)	-	-	-	5(4)

C26	0.8233(2)	0.10152(17)	0.08593(15)	-	-	-	11(5)
C14	0.1406(2)	0.17956(15)	0.44641(14)	-	-	-	19(11)
	26AMP:$F_{0.5}$:$(H_2O)_2$						
O1	0.10830(9)	0.35585(5)	0.14824(13)	0.19(9)	0.34(6)	0.11(9)	-0.5(3)
O2	0.24638(9)	0.46942(6)	0.26205(13)	0.27(5)	0.45(8)	0.39(12)	11(5)
O9	0.25642(10)	0.63425(7)	0.46103(13)	0.28(6)	0.40(8)	0.37(6)	16(9)
N1	0.29824(9)	0.22537(7)	0.25868(12)	0.23(5)	0.29(11)	0.2(2)	4(5)
O13	0.16466(13)	0.79709(9)	0.29005(17)	0.2(1)	0.80(7)	0.4(1)	7(4)
N2	0.45656(12)	0.33602(8)	0.37018(17)	0.5(2)	0.55(10)	0.5(2)	4(9)
C1	0.14377(10)	0.43984(7)	0.16165(14)	-	-	-	0.3(4)
C2	0.05823(11)	0.50981(7)	0.05070(15)	-	-	-	3.2(5)
C9	0.42618(11)	0.24758(8)	0.33792(14)	-	-	-	1(6)
C10	0.25538(12)	0.13717(8)	0.21788(16)	-	-	-	10(7)
C11	0.52100(14)	0.17430(10)	0.38084(18)	-	-	-	-8(3)
C12	0.11154(15)	0.12691(12)	0.1289(2)	-	-	-	-7(4)
C13	0.34675(15)	0.06594(9)	0.2593(2)	-	-	-	11(9)
C14	0.48034(15)	0.08606(10)	0.34004(19)	-	-	-	14(8)

Table A2.1: A comparison of GIPAW calcualted ^1H shifts (in ppm) for crystal structures of AF26-H0.5 that were recorded at a range of temperatures, alongside a calculation with the water molecules removed.*

Atom	**$\delta_{LT\ Crystal}$ (150 K)**	**$\delta_{CT\ Crystal}$ (250 K)**	**$\delta_{Crystal}$ (290 K)**	**$\delta_{HT\ Crystal}$ (340 K)**	**$\delta_{Dehydrated}$ (290 K)**	**$\Delta\delta_{Crystal\ -\ Dehydrated}$**
H9/H10/H11	1.3	1.3	1.4	1.4	1.7	-0.1
H27/H28/H29	1.4	1.4	1.4	1.3	1.5	-0.3
H5	4.4	4.3	4.2	4.3	-	-
H6	4.6	4.5	4.5	4.4	-	-
H3	5.2	5.2	5.3	5.2	5.2	0.1
H31	5.3	5.3	5.4	5.4	5.3	0.1
H14	5.5	5.5	5.5	5.4	6.4	-0.9
H1	5.9	5.9	5.9	5.8	6.0	-0.1
H16	6.2	6.2	6.3	6.3	5.1	1.2
H33	6.4	6.3	6.4	6.5	5.4	1.0
H13	6.8	6.8	7.0	6.8	5.2	1.8
H35	7.3	7.3	7.4	7.3	6.4	1.0
H17	8.5	8.5	8.7	8.4	8.8	-0.1
H34	8.6	8.5	8.7	8.6	9.4	-0.7
H15	9.5	9.6	9.5	9.8	10.4	-0.9
H32	10.3	10.2	10.5	10.1	10.5	0.0
H12	14.9	14.9	15.0	14.8	15.0	0.0
H30	15.1	15.1	15.0	15.0	16.6	-1.6

Table A2.2: Crystallographic data, instrumental parameters and final residuals for multiphase Rietveld refined model of 26AMP$_2$:F:FA$_{0.5}$, 26AMP$_2$:F:H$_2$O and 26AMP:F$_{0.5}$:(H$_2$O)$_2$ against RT PXRD.

AFF26 74 %	**Formula** (Z = 3)	$C_{18}H_{22}N_4O_6$		**Scale factor**		0.0009(4)	
	***a* (Å)**	***b* (Å)**	***c* (Å)**	***α* (°)**	***β* (°)**	***γ* (°)**	***V* (Å^3)**
	8.521(2)	10.645(2)	12.383(4)	84.74(3)	72.77(3)	67.02(3)	987.4(5)
AF26-H0.5 17.5 %	**Formula** (Z = 2)	$C_{16}H_{22}N_4O_5$		**Scale factor**		0.0002(2)	
	***a* (Å)**	***b* (Å)**	***c* (Å)**	***α* (°)**	***β* (°)**	***γ* (°)**	***V* (Å^3)**
	7.689(5)	10.022(8)	12.726(9)	93.62(7)	107.23(6)	102.62(5)	905(1)
AF26-H2 8.5 %	**Formula** (Z = 4)	$C_8H_{14}N_2O_4$		**Scale factor**		0.0001(1)	
	***a* (Å)**	***b* (Å)**	***c* (Å)**	***β* (°)**		***V* (Å^3)**	
	9.759(10)	14.30(3)	7.452(8)	96.97 (10)		1032(2)	
Number of parameters refined	165	**Space groups**	$P\bar{1}$ and $P2_1/c$	**Zero point**	0.058(2)		
PseudoVoigt peak width parameters (°^{2})	U 0.03(23)	V 0.1(1)	W 0.2(2)				
***R*$_{wp}$**	9.33 %	***R*$_{bragg}$ (AFF26)**	1.48 %				
***R*$_{exp}$**	3.62 %	***R*$_{bragg}$ (AF26-H0.5)**	3.66 %				
***R*$_p$**	6.72 %	***R*$_{bragg}$ (AF26-H2)**	1.81 %				

A3 5-Amino-2-Methylpyridine and Related Systems

Table A3.1: Crystallographic data, instrumental parameters and final residuals for Rietveld refined model of 52AMP:HF against RT PXRD

Formula (Z = 2)			$C_{10}H_{12}O_4N_2$			
a **(Å)**	***b*** **(Å)**	***c*** **(Å)**	***α*** **(°)**	***β*** **(°)**	***γ*** **(°)**	***V*** **(Å^3)**
8.0604(7)	8.24925(10)	9.563(1)	110.971(7)	92.736(7)	111.320(8)	1086.917(4)
Number of parameters refined		68	**Space group**		$P\bar{1}$	
Scale factor		0.0082(3)	**Zero point**		0.1(1)	
PseudoVoigt peak width parameters (°^{2})		U 0.02(49)	V 0.4(1)		W 0.1(5)	
R_{wp}		8.83 %	R_p		6.47 %	
R_{exp}		2.42 %	R_{bragg}		1.39 %	

Table A3.2: Comparison of the fractional coordinates of the atoms in the asymmetric unit of the SXRD structure of 52AMP:HF and the Rietveld refined O and N atomic positions for the RT PXRD, alongside thermal parameters following refinement.

	52AMP:HF			Refined model			
Atom	**x**	**y**	**z**	**x**	**y**	**z**	**Uiso (Å²)**
O1	0.41329(15)	0.35567(19)	0.10896(12)	0.398(3)	0.344(2)	0.107(3)	2(1)
O2	0.22842(14)	0.19182(17)	−0.12002(13)	0.238(5)	0.194(4)	−0.112(4)	11(1)
O3	0.81193(15)	0.12732(19)	−0.23797(12)	0.814(3)	0.136(3)	−0.243(3)	2.8(8)
O4	0.98512(15)	0.2864(2)	−0.00397(13)	0.983(4)	0.286(3)	−0.015(3)	5(1)
C8	0.52100(19)	0.1962(2)	−0.09969(16)	-	-	-	−3(1)
C9	0.69197(19)	0.2759(2)	−0.03153(16)	-	-	-	3(1)
C10	0.83433(18)	0.2208(2)	−0.10287(15)	-	-	-	0.1(1.4)
C7	0.37744(17)	0.2529(2)	−0.02931(16)	-	-	-	5(2)
N1	0.23152(16)	0.40445(16)	0.33420(13)	0.210(5)	0.364(5)	0.302(3)	6(1)
N2	−0.13463(18)	0.08599(19)	0.44077(17)	−0.127(4)	0.089(5)	0.453(6)	7(1)
C1	0.01787(18)	0.23958(19)	0.45211(15)	-	-	-	10(2)
C2	0.08745(19)	0.2487(2)	0.32356(16)	-	-	-	2(2)
C3	0.1096(2)	0.3965(2)	0.59130(16)	-	-	-	4(1)
C4	0.2562(2)	0.5534(2)	0.59659(16)	-	-	-	0.8(9)
C5	0.31887(18)	0.55926(19)	0.46530(16)	-	-	-	−5.1(9)
C6	0.4732(2)	0.7249(2)	0.4588(2)	-	-	-	−5(1)

Table A3.3: Crystallographic data, instrumental parameters and final residuals for Rietveld refined model of 25L:$F_{0.5}$:FA against RT PXRD

25L:$F_{0.5}$: FA 88.5 %	**Formula** (Z = 2)		$C_{13}H_{15}O_6N$		**Scale factor**		0.00553(8)
	***a* (Å)**	***b* (Å)**	***c* (Å)**	***α* (°)**	***β* (°)**	***γ* (°)**	***V* (Å³)**
	7.3548(2)	9.1894(5)	11.6668(10)	84.375(4)	71.932(4)	89.210(3)	745.91(8)
FA 11.5 %	**Formula** (Z = 1)		$C_4H_4O_4$		**Scale factor**		0.0002(2)
	***a* (Å)**	***b* (Å)**	***c* (Å)**	***α* (°)**	***β* (°)**	***γ* (°)**	***V* (Å³)**
	5.248(2)	7.346(2)	4.531(1)	107.21(2)	85.87(1)	134.13(2)	117.44(7)
Number of parameters refined		106	**Space group**		$P\bar{1}$	**Zero point**	−0.007(54)
PseudoVoigt peak width parameters (°²)			U 0.07(21)		V 0.19(4)		W 0.04(21)
R_{wp}		10.33 %		***R_{bragg}* (25L:$F_{0.5}$:FA)**		2.51 %	
R_{exp}		2.05 %					
R_p		7.98 %		***R_{bragg}* (FA)**		1.69 %	

Table A3.4: Comparison of the fractional coordinates of the atoms in the asymmetric unit of 25AMP:$F_{0.5}$:$FA_{0.5}$ and FA, for the literature structures (RESGEC and FUMAAC01), and the Rietveld refined O and N atomic positions for the RT PXRD, alongside thermal parameters following refinement.

Atom	25L:$F_{0.5}$:FA [a]			Refined model			
	x	y	z	x	y	z	Uiso ($Å^2$)
N1	0.23324(16)	0.46707(12)	0.01520(11)	0.231(4)	0.440(4)	0.047(3)	12(2)
C1	0.1393(2)	0.35702(15)	0.09588(15)	-	-	-	16(3)
C2	0.0825(2)	0.23655(16)	0.04993(18)	-	-	-	0.3(9)
C3	0.1283(2)	0.23059(16)	−0.07278(17)	-	-	-	16(2)
C4	0.2315(2)	0.34413(16)	−0.15410(16)	-	-	-	2(1)
C5	0.2796(2)	0.46406(15)	−0.10516(13)	-	-	-	−0.6(9)
C6	0.1069(3)	0.3706(2)	0.22629(16)	-	-	-	−1.2(8)
C7	0.2923(3)	0.3384(2)	−0.28885(19)	-	-	-	−0.2(9)
O3	0.6212(2)	0.27440(14)	0.34108(11)	0.603(3)	0.283(3)	0.357(2)	14(1)
C11	0.52934(19)	0.06933(16)	0.48610(11)	-	-	-	−0.5(9)
C10	0.5554(2)	0.15112(16)	0.36562(12)	-	-	-	6(2)
O4	0.50276(18)	0.07742(13)	0.28897(9)	0.523(4)	0.086(3)	0.283(4)	6.4(7)
O2	0.55446(15)	0.23639(10)	0.08458(8)	0.547(3)	0.227(3)	0.092(3)	11(1)
C8	0.60715(17)	0.19655(13)	−0.01985(10)	-	-	-	6(2)
O1	0.69539(14)	0.28229(10)	−0.11184(8)	0.714(2)	0.298(3)	−0.112(2)	6.2(7)
C9	0.56341(18)	0.04391(13)	−0.04050(11)	-	-	-	9(1)
O5	0.81194(15)	0.21625(11)	−0.33431(8)	0.815(3)	0.224(2)	−0.334(2)	4.7(6)
C12	0.91716(19)	0.09677(15)	−0.35453(11)	-	-	-	0.5(10)
O6	0.97031(16)	0.02247(12)	−0.27860(9)	0.987(3)	0.034(2)	−0.293(2)	9.8(8)
C13	0.96267(18)	0.06362(16)	−0.48273(11)	-	-	-	12(2)
	FA						
C1	0.4079	0.3804	0.4003	-	-	-	13(5)
C2	0.4541	0.2185	0.4495	-	-	-	17(4)
O1	0.2716	−0.0045	0.2567	0.23(2)	−0.01(2)	−0.08(2)	12(4)
O2	0.6842	0.3255	0.7026	0.65(3)	0.31(2)	0.79(2)	11(5)
C1A	0.5921	0.6196	0.5997	-	-	-	13(5)
C2A	0.5459	0.7815	0.5505	-	-	-	17(5)
O1A	0.7284	1.0045	0.7433	0.85(2)	1.19(1)	0.78(2)	11(5)
O2A	0.3158	0.6745	0.2974	0.32(3)	0.72(2)	0.23(1)	−0.4(30)

[a] Structure determined by Haynes *et al.*[140]

Table A3.5: Crystallographic data, instrumental parameters and final residuals for Rietveld refined modela of 25AMP:$F_{0.5}$:$FA_{0.5}$ against RT PXRD

Formula (Z = 2)			$C_{10}H_{12}O_4N_2$			
a **(Å)**	***b*** **(Å)**	***c*** **(Å)**	***α*** **(°)**	***β*** **(°)**	***γ*** **(°)**	***V*** **($Å^3$)**
4.2170(2)	9.3347(8)	13.972(1)	94.150(5)	96.074(5)	90.421(7)	545.44(8)
Number of parameters refined		68	**Space group**		$P\bar{1}$	
Scale factor		0.0115 (3)	**Zero point**		0.11(7)	
PseudoVoigt peak width parameters ($°^2$)		U 0.04(32)	V 0.11(7)		W 0.04(32)	
R_{wp}		13.28 %	R_p		8.82 %	
R_{exp}		1.59 %	R_{bragg}		5.00 %	

Table A3.6: Comparison of the fractional coordinates of the atoms in the asymmetric unit of 25AMP:$F_{0.5}$:$FA_{0.5}$ structure, for the literature structure (DUTNUC) and the Rietveld refined O and N atomic positions for the RT PXRD

	25AMP:$F_{0.5}$:$FA_{0.5}$[a]			RT PXRD			
Atom	**x**	**y**	**z**	**x**	**y**	**z**	**U_{iso} (Å^2)**
O1	0.27081(16)	0.19332(6)	0.39352(4)	0.286(7)	0.287(2)	0.426(1)	12(1)
O2	0.49662(14)	0.03168(5)	0.29314(3)	0.479(8)	0.056(3)	0.294(2)	13(1)
C13	0.42204(15)	0.08387(6)	0.37847(4)	-	-	-	14(2)
C14	0.54504(15)	−0.00956(6)	0.45556(4)	-	-	-	13(1)
O3	0.10076(14)	0.26911(5)	0.02053(3)	0.124(4)	0.280(2)	0.024(1)	−0.09(72)
O4	0.34405(13)	0.19338(5)	0.15605(3)	0.372(4)	0.173(2)	0.171(1)	5.8(9)
C17	0.17514(14)	0.17185(6)	0.07482(4)	-	-	-	−1(1)
C18	0.06922(15)	0.01973(6)	0.04396(4)	-	-	-	6(1)
N1	0.64925(12)	0.45378(5)	0.20997(3)	0.695(7)	0.333(3)	0.240(2)	13(1)
N2	0.31601(12)	0.55122(5)	0.09009(3)	0.307(7)	0.623(3)	0.08(2)	11(1)
C1	0.54215(13)	0.56996(5)	0.16485(4)	-	-	-	1(1)
C2	0.67623(13)	0.70674(5)	0.20110(4)	-	-	-	−5.8(9)
C3	0.89669(13)	0.71642(6)	0.28125(4)	-	-	-	5(1)
C4	0.99507(13)	0.59304(6)	0.32951(4)	-	-	-	12(1)
C5	0.86767(13)	0.46307(6)	0.29027(4)	-	-	-	12(1)
C6	1.21690(15)	0.60555(7)	0.42149(4)	-	-	-	3(1)

[a] Structure determined by Hemamalini *et al.*[141]

Table A3.7: GIPAW calculated ^{1}H chemical shifts

	δ^{iso}_{calc}(^{1}H) (ppm)		
	52AMP:HF	**25L:$F_{0.5}$:FA**	**25AMP:$F_{0.5}$:$FA_{0.5}$**
Me	0.9	1.6	0.4
	-	1.6	-
CH	5.8	5.6	5.5
	6.0	6.2	6.0
	6.2	6.3	6.1
	6.3	6.7	6.4
	6.9	7.0	7.1
	-	8.2	-
NH_2	5.2	-	6.7
	7.0	-	9.6
NH	15.1	13.9	14.2
OH	18.4	14.7	16.0
		14.8	-

Table A3.8: GIPAW calculated ^{13}C chemical shifts

	$\delta_{calc}^{iso}(^{13}C)$ (ppm)		
	52AMP:HF	**25L:F$_{0.5}$:FA**	**25AMP:F$_{0.5}$:FA$_{0.5}$**
Me	12.9	11.8	15.8
	-	12.1	-
CH/C	124.5	127.5	114.5
	125.4	137.4	122.1
	131.7	138.6	134.7
	135.4	138.9	138.4
	135.9	139.1	138.4
	143.4	141.3	138.6
	146.3	146.0	138.6
	-	152.5	142.8
	-	-	147.9
COOH/$^-$	172.7	170.0	168.2
	173.2	170.2	172.9
	-	173.6	-

Table A3.9: GIPAW calculated ^{15}N chemical shifts, quadrupolar parameters and calculated ^{14}N shifts for each geometry optimised crystal structure, alongside the experimental ^{14}N shifts.

Structure	**NH$^+$**	**^{15}N δ_{iso}^{calc} (ppm)**	**η_Q**	**C_Q (MHz)***	**P_Q (MHz)**	**δ_{iso}^{Q} (ppm)**	**^{14}N δ^{calc} (ppm)**	**^{14}N δ^{expt} (ppm)**
52AMP:HF	N1	−170.2	0.8	−0.9	−1.0	43.6	−126.6	−126.1
	N2	−311.4	0.4	−3.6	−3.7	548.9	237.5	-
25AMP:F$_{0.5}$:FA$_{0.5}$	N1	−216.5	0.8	1.6	1.8	127.9	−88.5	−103.8
	N2	−289.6	0.5	−3..0	−3.1	382.2	92.6	87.1
25L:F$_{0.5}$:FA	N1	−175.9	0.7	0.9	0.9	34.1	−141.7	−150.0

*A C_Q scaling factor of 0.95 is employed.

Table A3.10: Longer-range C···H proximities between 1.2 and 3.3 Å for 52AMP:HF

C	δ_{iso}^{exp} (ppm)	H	δ_{iso}^{exp} (ppm)	Separation[a] (Å)
C1	170.6	H3	16.8	1.99
		H2	6.1	2.20
		H1	6.1	2.70
		H8	*6.1*	*2.99*
		H10	*6.1*	*3.16*
		H7	*7.1*	*3.29*
C2	171.7	H1	6.1	2.21
		H3	*16.8*	*2.30*
		H12	*15.0*	*2.63*
		H2	6.1	2.71
		H11	*7.1*	*2.84*
C3	132.5	H1	6.1	2.11
		H9	*6.1*	*2.98*
		H8	*6.1*	*3.29*
		H3	16.8	3.29
C4	139.7	H2	6.1	2.11
		H11	*7.1*	*2.98*
		H13/H14/H15	*1.0*	*3.10*
C9	147.5	H10	6.1	2.04
		H11	7.1	2.04
		H8	6.1	2.16
		H7	7.1	2.19
C10	138.2	H12	15.0	2.08
		H13/H14/H15	1.0	2.13
		H9	6.1	2.14
C11	127.0	H9	6.1	2.15
		H10	6.1	2.62
		H2	*6.1*	*2.90*
		H13/H14/H15	1.0	3.04
C12	127.0	H12	15.0	2.05
		H11	7.1	2.58
		H13/H14/H15	1.0	3.17
		H10	6.1	3.30
C13	129.9	H8	6.1	2.15
		H13/H14/H15	1.0	2.71
		H2	*6.1*	*2.92*
		H12	15.0	3.30
C14	12.9	H12	15.0	2.60
		H9	6.1	2.77
		H10	6.1	3.01
		H13/H14/H15	*1.0*	*3.17*
		H1	*6.1*	*3.17*
		H11	*7.1*	*3.25*

[a] H-H distances are taken from the DFT (CASTEP) optimised structure. Intermolecular proximities are denoted using italic font.

Table A3.11: Longer-range C···H proximities between 1.2 and 2.8 Å for 25AMP:$F_{0.5}$:$FA_{0.5}$.

C	δ_{iso}^{exp} (ppm)	H	δ_{iso}^{exp} (ppm)	Separation[a] (Å)
C1	153.8	H3	6.7	2.05
		H2	8.8	2.06
		H1	14.0	2.08
		H4	6.1	2.16
C2	116.3	H5	6.4	2.12
		H3	6.7	2.66
C3	144.2	H4	6.1	2.15
		H7/H8/H9	0.8	2.79
C4	121.5	H7/H8/H9	0.8	2.14
		H6	7.1	2.17
		H5	6.4	2.17
C5	134.7	H1	14.0	2.08
		H7/H8/H9	0.8	2.62
C6	18.7	H5	6.4	2.76
		H6	7.1	2.78
C13	166.6	H19	*14.7*	1.97
		H20	6.1	2.22
C14	136.1	H20	6.1	2.10
C17	172.7	H23	6.1	2.22
		H19	*14.7*	*2.55*
		H1	*14.0*	*2.57*
		H2	*8.8*	*2.60*
C18	136.1	H23	6.1	2.11

[a] H-H distances are taken from the DFT (CASTEP) optimised structure. Intermolecular proximities are denoted using italic font.

Table A3.12: Longer-range C···H proximities between 1.2 and 2.9 Å for 25L:$F_{0.5}$:FA.

C	δ_{iso}^{exp} (ppm)	H	δ_{iso}^{exp} (ppm)	Separation[a] (Å)
C1	152.4	H10	14.3	2.07
		H4/H5/H6	1.8	2.12
		H1	6.4	2.14
C2	128.1	H2	7.1	2.16
		H4/H5/H6	1.8	2.80
C3	146.9	H1	6.4	2.17
		H7/H8/H9	1.8	2.83
C4	136.0	H7/H8/H9	1.8	2.13
		H2	7.1	2.15
		H3	8.1	2.19
C5	142.5	H10	14.3	2.07
		H7/H8/H9	1.8	2.61
C7	16.8	H2	7.1	2.76
		H3	8.1	2.78
C8	173.5	H11	5.6	2.20
		H10	14.3	2.36
		H12	13.4	2.50
		H14	13.4	2.51
C9	136.0	H11	5.6	2.11
		H14	13.7	2.83
C10	169.8	H12	13.4	1.95
		H13	6.4	2.17
C11	136.0	H7/H8/H9	1.8	2.90
C12	169.8	H14	13.7	1.96
		H15	6.4	2.21
C13	136.0	H15	6.4	2.11
		H13	6.4	2.85

*^{1}H correlations written in italics correspond to those seen in both the short and long range experiments

A4 Comparisons Identifying Crystal Forms and Structural Patterns

Table A4.1: CSD codes of the additional 56 structures downloaded from the CCDC for determination of ^{14}N GIPAW calculated shifts.

ABAQEB	CUVDUT	LOHJAU	RESHIH
ABEKUN03	DIFNUC	LOHLIE	RETZEW
ABELAU01	DIGCUS	MOGWAI	TOYGOF02
ABUDEI	EKEPEP	ONAPID	UMININ
ACACEO	FICTAO	RAKQAV04	WIDDES
APALIN	GAHFAX	RESFIF	WOPZUW
AXUFUV	GODNAO	RESFOL	WOQBOT
BANCID	GODNES	RESFUR	YEPCOM
BATYAZ01	GODNIW	RESGAY	ZITZAD
BATYIH01	GOJQAY	RESGIG	CIRXAD
BEBTIB	GUKWIT	RESGOM	EXUQUJ
BEVXEH	HOGFOA	RESGUS	GANYAW
BICQAH	LADDEB	RESHAZ	XEJXIU
BIDWAN	LATSUW	RESHED	XEJXAM

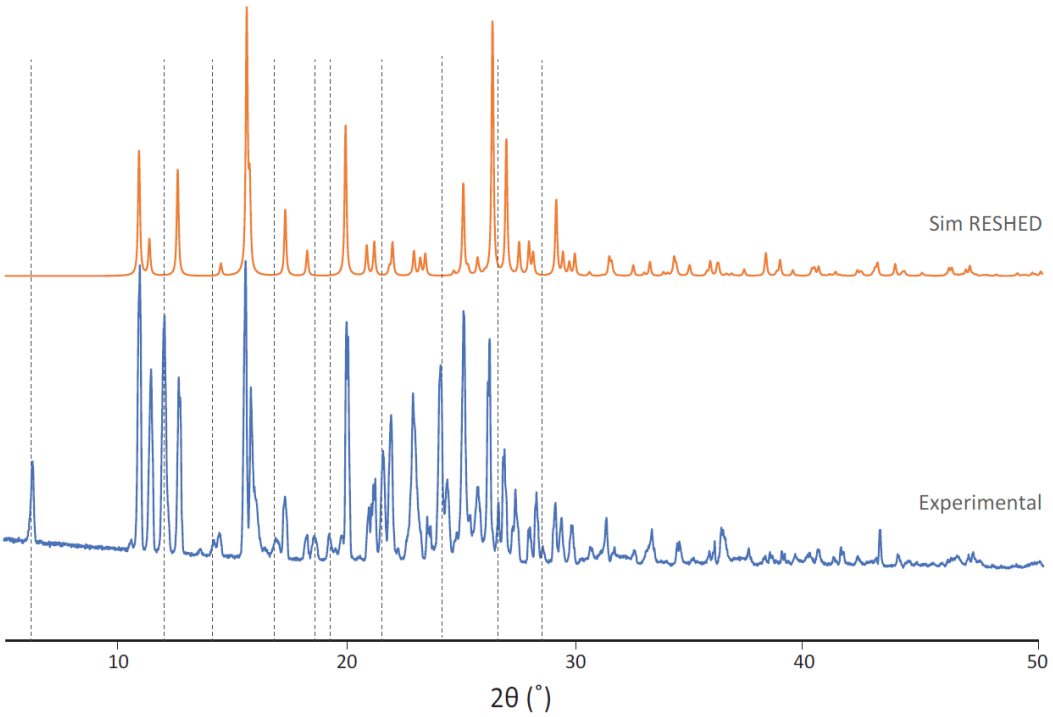


Figure A4.1: Experimental PXRD pattern of P35L:F presented alongside the simulated powder pattern for RESHED.

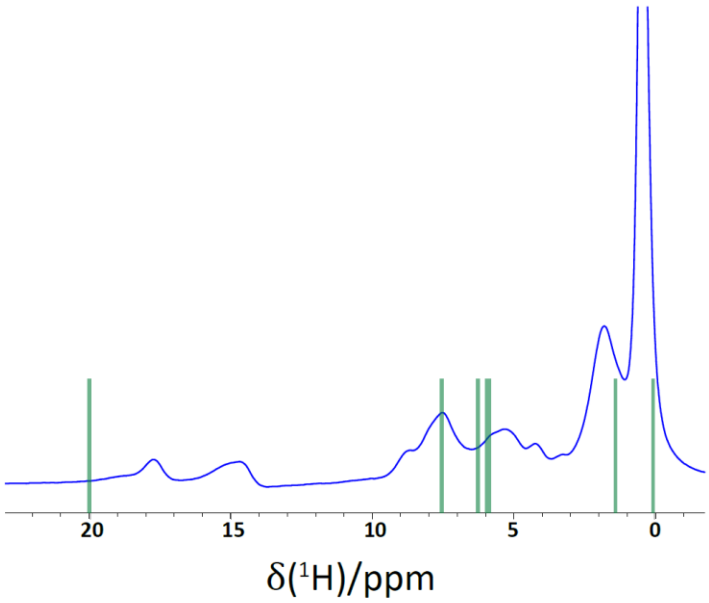


Figure A4.2: ^{1}H (600 MHz) one-pulse MAS (60 kHz) spectrum of P35L:F with a stick spectrum corresponding to GIPAW calculated chemical shifts for the geometry optimised crystal structure of RESHED

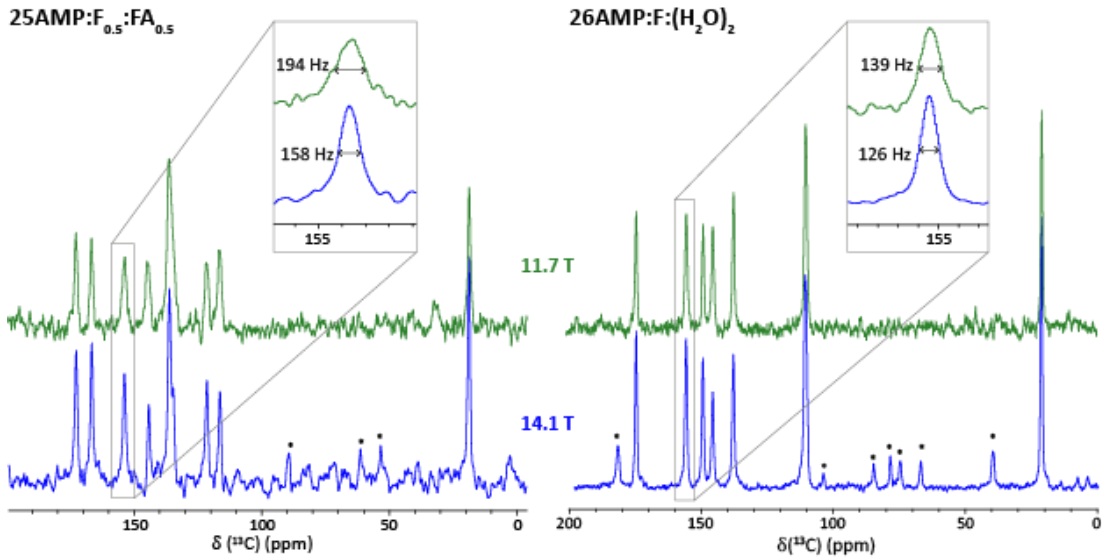
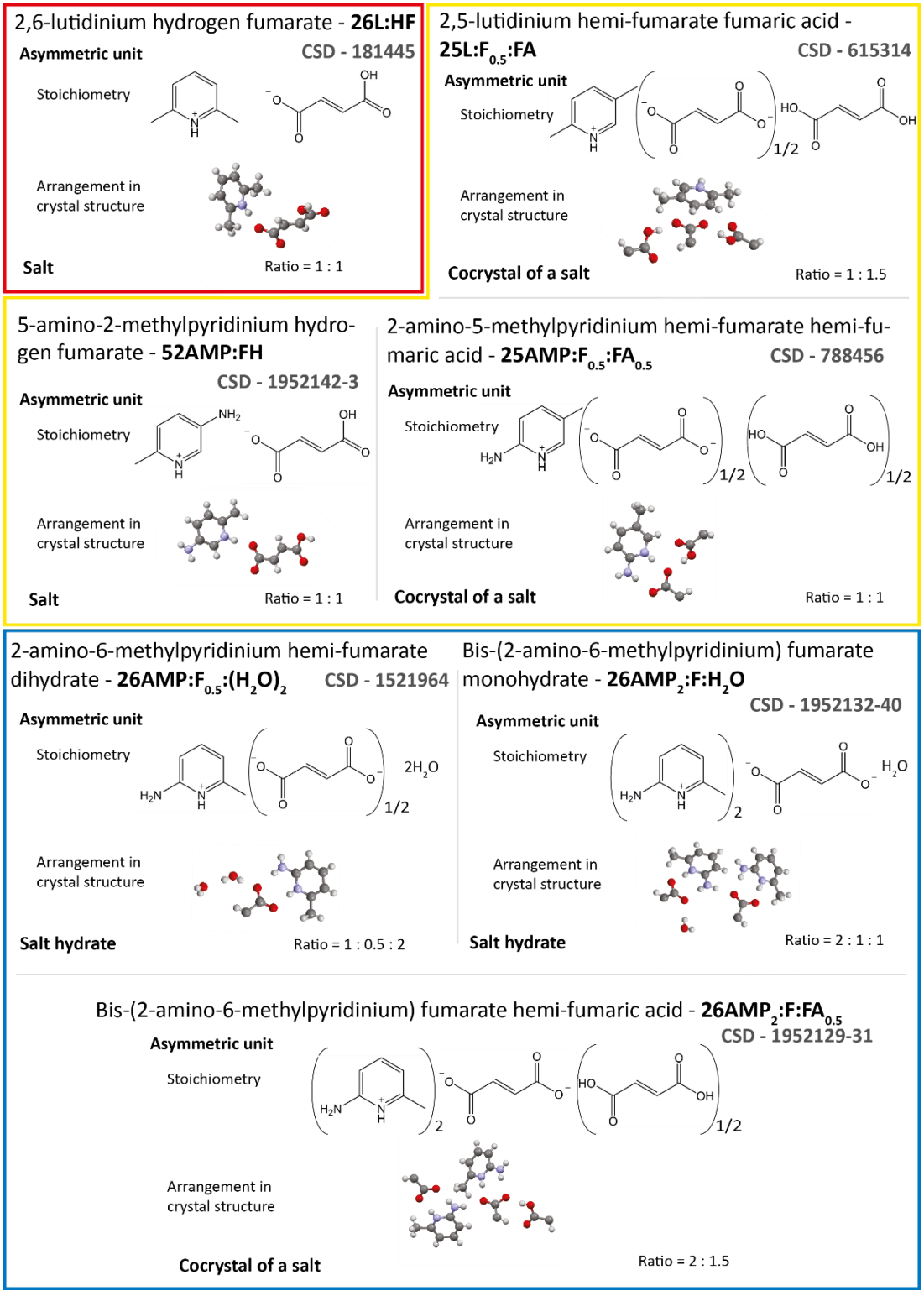


Figure A4.3: Comparison of ^{1}H-^{13}C CP-MAS (12.5 kHz) spectra of 25AMP:$F_{0.5}$:$FA_{0.5}$ (left) and 26AMP:$F_{0.5}$:$(H_2O)_2$ (right) recorded at 11.7 T (top) and 14.1 T (bottom), corresponding to a ^{1}H Larmor frequency of 500 MHz and 600 MHz, respectively. Insets are zoomed regions for the peaks discussed in the main text.



Summary of the multicomponent crystal systems studied within this work. All names are given relative to a single base molecule whereas the stoichiometric ratio is here given in terms of the asymmetric unit. Coloured boxes indicate the groups in which the structures were introduced in Chapter 3 (red), Chapter 4 (blue) and Chapter 5 (yellow).