

University of  
**Strathclyde**  
**Glasgow**

**Structural investigations with high  
pressure techniques and  
multicomponent systems**

Lauren Evelyn Connor

Thesis submitted to the Strathclyde Institute of Pharmacy and Biomedical Sciences  
at the University of Strathclyde in accordance with the requirements for the degree  
of Doctor of Philosophy

2018

**Lauren Connor**

## **Declaration of Author's Rights**

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

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

Signed: 

Date: 29<sup>th</sup> of May 2019

### Abstract

This thesis illustrates the use of high pressure crystallography techniques for the discovery and investigation of solid-state forms and probes the relationship between molecular structure and compression of both single and multicomponent systems. As well as investigating a data-driven approach to directing experimental co-crystallisation attempts.

Single crystal X-ray diffraction techniques are a highlight in all areas of this study, as well as computational approaches which were used in the evaluation of the interactions of small molecule systems. Data-mining of the Cambridge Structural Database made the comparison of the compression studies richer.

The pharmaceutical co-crystal, indomethacin and saccharin was analysed with respect to increasing pressure. The system is an example of a homomolecular synthon co-crystal allowing investigation of the component dimers free of strong interaction with surrounding molecules. The ambient pressure structure remains stable but investigation showed that the saccharin dimer sits in a pocket made by indomethacin allowing the dimer to lie further apart than in the pure compound.

To follow, a structural compression study of the single component saccharin using synchrotron radiation lead to the structural characterisation of the first new polymorph of saccharin. The hydrogen bonding pattern of the new phase remains consistent however Pixel calculations revealed that the biggest difference in packing arises due to the reduction of an interlayer distance.

To further explore multicomponent systems, two stoichiometric ratios of benzoic acid and isonicotinamide (2:1 & 1:1) were investigated. The rate of compression in these systems are almost identical despite the different molecular packing in each of the stoichiometric ratios. Through the investigation of materials in these initial chapters, the rate of compression in particular supramolecular synthons, e.g. amide-dimers, is demonstrated to be consistent despite the difference in the molecular make-up of the materials under study and their packing arrangements.

Lastly, a data-driven approach was applied in directing the discovery of a new solid-state entity. Following previous failed attempts, machine learning was employed to direct experimental co-crystallisations which led to a new co-crystal of Artemisinin and 1-Naphthol. Pixel calculations revealed that the largest contribution to crystal stabilisation comes from dispersion energy and enabled the identification of dominant intermolecular interactions in the crystal structures.

**Lauren Connor**

**Dedication**

*For JP, the mischievous one  
And  
'Auld Annie', the giggling one.*

### Acknowledgements

Firstly, I am greatly indebted to my supervisor, Dr. Iain Oswald. Iain, your unending support over the many years has been appreciated more than you could know, I could not have wished for a better supervisor. Your encouragement has been second to none and I could not have gotten anywhere near where I am without it.

Thank you to Prof. Zaher Judeh and all of the team in Singapore for welcoming me into the lab. To the solid state and crystallisation teams in AstraZeneca, thank you for all of the help and advice as well as a very enjoyable placement. To the I19 beamline scientists at Diamond, thank you for making trips to Diamond so interesting and rewarding.

For his enthusiasm and explanations of all things machine learning, my sincerest gratitude goes to Dr. Tony Vassileiou, you are a fantastic teacher and you really sparked my interest in machine learning. Working with you over the last few months of my Ph.D. has been enjoyable as well as beneficial.

To the members of the Oswald group, past and present, I am so grateful to have met each of you and worked alongside you, Dr. Amit Delori, Abdullah Al Balushi and Lloyd Farquar. For putting up with my many questions at the start and for all of the training you provided, thank you to Dr. Ian Hutchison. Thank you to Dr. Martin Ward, you never fail to offer brilliant solutions as well as enthusiasm whether the topic is work related or a group go-karting trip – one day! To Suse Bebiano, a warm thanks goes to you for every time we ended up sitting on the floor at work or in our flat to ‘think clearly’ and got lost in one of our scientific discussions. Thank you for always being an active member of the late night working crew along with Bilal Ahmed and Carlota Mendez. You all made the long days enjoyable and motivating. To Vaclav Svoboda for his constant support and those pesky 2:1 co-crystals, thank you. To all of those who climb the walls, especially Alice Turner, thank you for the escapism. A most warm and sincere thank you to Clarissa Forbes, you have been an incredible friend, fellow troll and shoulder. Here’s to our next chapter, Caba.

Thank you to all of my family and friends outside of Strathclyde, especially Kerry, Annie, Heather and Danielle for their encouragement but also for the occasional escapism.

To my parents, it is all for you both. Your support and encouragement is, as always, endless. I cannot express my gratitude enough, especially for the gallons of tea you brought me during the write-up of this thesis! You have both made this thesis and all my years of study possible.

Lastly, to Nils, thank you for putting up with my stress and for helping me in every way you could, at any opportunity. I will always appreciate your encouragement and genuine interest in my work, I even appreciate the unending questions on every piece of work I try to explain. Thank you for being so patient over the years... the long distance is over! Finalement!

## Contents

<b>1. INTRODUCTION .....</b>	<b>1</b>
1.1. SYNOPSIS .....	1
1.2. SUPRAMOLECULAR CHEMISTRY AND THE CRYSTALLINE FORM.....	1
1.3. POLYMORPHISM .....	1
1.3.1. <i>Importance of polymorphism</i> .....	4
1.4. CRYSTAL ENGINEERING.....	4
1.5. MACHINE LEARNING .....	7
1.6. CO-CRYSTALS .....	8
1.7. SOLUTION BASED METHODS .....	9
1.7.1. <i>Pharmaceutical co-crystals</i> .....	10
1.8. X-RAY DIFFRACTION.....	12
1.9. HIGH PRESSURE CRYSTALLOGRAPHY .....	12
1.10. THESIS OUTLINE .....	14
<b>2. MATERIALS AND METHODS .....</b>	<b>16</b>
2.1. CRYSTALLOGRAPHY.....	16
2.1.1. <i>Producing X-rays</i> .....	16
2.1.2. <i>X-ray crystallography</i> .....	16
2.1.3. <i>Crystal systems</i> .....	18
2.2. HIGH PRESSURE CRYSTALLOGRAPHY.....	20
2.2.1. <i>Collection and issues with high pressure experiments</i> .....	21
2.2.2. <i>Synchrotron radiation</i> .....	22
2.3. PIXEL.....	23
2.3.1. <i>The Pixel method</i> .....	23
<b>3. HIGH PRESSURE STRUCTURAL INVESTIGATION OF INDOMETHACIN AND SACCHARIN .....</b>	<b>33</b>
3.1. SYNOPSIS .....	33
3.2. INTRODUCTION .....	33
3.3. EXPERIMENTAL.....	35
3.3.1. <i>General procedure</i> .....	35
3.3.2. <i>Co-crystal formation</i> .....	35
3.3.3. <i>Diamond anvil cell preparation</i> .....	35
3.3.4. <i>Ambient crystal collection, processing and refinement</i> .....	35
3.3.5. <i>Compression study collection, processing and refinement</i> .....	36
3.3.6. <i>H-bonding length distribution in CSD</i> .....	36
3.4. RESULTS AND DISCUSSION .....	37
3.4.1. <i>Structure</i> .....	37
3.4.2. <i>Pixel calculations</i> .....	42
3.4.3. <i>Compression study</i> .....	42
3.4.4. <i>Compression behaviour of the amide dimer</i> .....	45
3.5. CONCLUSION.....	51
<b>4. POLYMORPHIC PHASE TRANSITION OF SACCHARIN DURING COMPRESSION.....</b>	<b>56</b>
4.1. SYNOPSIS .....	56
4.2. INTRODUCTION .....	56
4.2.1. <i>Saccharin in the food industry</i> .....	56
4.2.2. <i>Saccharin in the pharmaceutical industry</i> .....	57
4.3. EXPERIMENTAL.....	58
4.3.1. <i>General procedure</i> .....	58

4.3.2. Crystal .....	59
4.3.3. Diamond anvil cell preparation.....	59
4.3.4. Compression study collection, processing and refinement.....	59
4.4. RESULTS AND DISCUSSION .....	62
4.4.1. Structure.....	62
4.4.2. Morphology predictions and face indexing.....	64
4.4.3. Compression of saccharin.....	67
4.4.4. Phase transformation .....	69
4.4.5. Void space analysis .....	73
4.4.6. Compression behaviour of amide dimer .....	73
4.4.7. Increasing density in phase changes at pressure.....	75
4.5. CONCLUSION.....	76
<b>5. COMPRESSION OF 1:1 AND 2:1 CO-CRYSTALS OF BENZOIC ACID AND ISONICOTINAMIDE .....</b>	<b>82</b>
5.1. SYNOPSIS .....	82
5.2. INTRODUCTION .....	82
5.3. EXPERIMENTAL.....	85
5.3.1. The 1:1 co-crystal of benzoic acid and isonicotinamide .....	85
5.3.2. The 2:1 co-crystal of benzoic acid and isonicotinamide .....	91
5.4. RESULTS AND DISCUSSION .....	96
5.4.1. Analysis of the 1:1 co-crystal of benzoic acid and isonicotinamide .....	96
5.4.2. Compression behaviour of the amide dimer .....	103
5.4.3. Analysis of the 2:1 co-crystal of benzoic acid and isonicotinamide .....	104
5.4.4. Compression comparison of the 1:1 and 2:1 co-crystal.....	111
5.5. CONCLUSION.....	112
<b>6. DATA-DRIVEN APPROACH TO DIRECT THE CO-CRYSTALLISATION OF ANTI-MALARIAL ARTEMISININ .....</b>	<b>117</b>
6.1. SYNOPSIS .....	117
6.2. INTRODUCTION .....	117
6.2.1. Artemisinin co-crystals .....	117
6.2.2. Computational screening .....	118
6.2.3. Machine learning.....	119
6.3. COMPUTATIONAL SCREENING METHODS.....	121
6.3.1. Molecular complementarity .....	121
6.3.2. COSMOtherm based on excess enthalpy.....	121
6.3.3. Machine learning.....	121
6.4. EXPERIMENTAL.....	122
6.4.1. Ball mill.....	122
6.4.2. X-ray Powder Diffraction (XRPD).....	122
6.4.3. Differential Scanning Calorimetry .....	123
6.4.4. Single crystal growth.....	123
6.4.5. SCXRD.....	123
6.5. RESULTS AND DISCUSSION .....	125
6.5.1. Negative experimental results from co-crystal computational methods .....	125
6.5.2. Machine learning success.....	127
6.5.3. Analysis of ArtNap.....	129
6.6. CONCLUSION.....	135
<b>7. CONCLUDING REMARKS AND FURTHER WORK.....</b>	<b>139</b>

## List of figures

<b>Figure 1.1</b> Schematic of five commonly encountered supramolecular synthons. ....	5
<b>Figure 1.2</b> Example of a ternary phase diagram with Coformers 1 and 3 at 100% at the bottom two apexes. 100% solvent at 4 in the top apex. Regions 1 and 3 within the triangle show regions where the co-formers 1 and 3 would be produced in solution, respectively. Regions 1 and 3 within the triangle show regions where the co-formers 1 and 3 would be produced in solution, respectively. Region 2 shows the region for formation of a co-crystal. ....	9
<b>Figure 2.1</b> X-ray diffraction from crystal lattice planes (Bragg's law). ....	17
<b>Figure 2.2</b> The four types of crystal lattice. ....	19
<b>Figure 2.3</b> Expanded view of an open DAC with components, b) bird's eye view of a Merrill-Basset DAC and c) labelled sample chamber between two diamonds. ....	20
<b>Figure 2.4</b> View of the sample chamber of a DAC containing a single crystal and many small ruby spheres. ....	21
<b>Figure 3.1</b> The asymmetric unit of INSA with numbering scheme (hydrogen atoms omitted for clarity). The colours of the elements are consistent throughout unless otherwise stated. (carbon, grey; oxygen, red; sulphur, yellow; nitrogen, blue; chlorine, green; hydrogen, white) ....	37
<b>Figure 3.2</b> A section of the structure of indomethacin and saccharin with hydrogen bonding contacts shown. Depicted a and b indicate a strong carboxylic acid dimer formed by indomethacin molecules and the dimer between saccharin molecules, respectively. ....	38
<b>Figure 3.3</b> Section of two chains running parallel along the b-axis in the ab plane interlocking through a $\pi/\pi$ interaction between indomethacin and saccharin (i) and a CH- $\pi$ interaction between indomethacin molecules (ii) shown in the blackened molecules. ....	38
<b>Figure 3.4</b> a) Unit cell volume of INSA as a function of pressure (EosFit 7), (b) Normalised unit cell lengths as a function of pressure. The errors on each of the measurements is smaller than the symbol. ....	43
<b>Figure 3.5</b> The structure of the unit cell down the a-axis (a). The void space at 0.0001 GPa (b), 3.33 GPa (c), and 5.06 GPa (d) obtained by using probe radius of 0.5Å with high definition of 0.1Å. ....	44
<b>Figure 3.6</b> Histograms showing the distribution of O-H $\cdots$ O and N-H $\cdots$ O distances reported in the CSD with markers for the distances reported in the compression of INSA at ambient and the highest pressure dataset refined. ....	46
<b>Figure 3.7</b> Plots showing how the length of piracetam (form II: half symbol & form V: filled symbol), INSA, tolazamide, and salicylamide's N-H $\cdots$ O amide dimer respond to pressure. ....	50
<b>Figure 4.1</b> The asymmetric unit of saccharin with numbering scheme at ambient pressure. The colours of the elements are consistent throughout (carbon, grey; oxygen, red; sulphur, yellow; nitrogen, blue; hydrogen, white). ....	62
<b>Figure 4.2</b> a) View of the structure down the b-axis showing the hydrogen bonded dimers; the layers of dimers are numbered 1 & 2. b) View down the b-axis highlighting the layers in colour, corresponding to the numbers in a) purple, layer 1; green, layer 2. The dispersive interactions between the layers are weaker than the dispersive interactions in the layers. ....	63
<b>Figure 4.3</b> Dimers of saccharin molecules interacting through $\pi$ - $\pi$ forces in a layer (Interaction 1; 4.562 Å) and between layers (interaction 2; 5.777 Å). Colours shown corresponds with layers in Figure 4.2 b). ....	64
<b>Figure 4.4</b> a) The BFDH predicted morphology of saccharin, viewed down the a-axis, b) the BFDH predicted morphology perpendicular to the main face. c) observed morphology of saccharin grown at ambient pressure from ethanol. d) face indexed crystal of saccharin with the main crystal faces labelled. ....	65
<b>Figure 4.5</b> Potential slip plane directions in saccharin shown by the dotted lines. ....	66
<b>Figure 4.6</b> Pixel plot of the main forces in the crystal structure of saccharin. Interactions are numbered by total strength and labelled by direction of interaction. ....	67

**Figure 4.7** Plot of linear strain showing principal axes of strain (maximum strain, red; medium strain; blue). The minimum strain is along the b-axis (into the page) as required by symmetry. .... 68

**Figure 4.8** View of the unit cell of saccharin down the c-axis showing the covalent and hydrogen bonding which propagates along the b–axis direction and is responsible for the least strain experienced in this direction..... 68

**Figure 4.9** Unit cell lengths and volume of saccharin as a function of pressure. The majority of the errors are smaller than the symbol. Half filled symbols denote the new phase of saccharin. .... 69

**Figure 4.10** Arrows depict the decrease in layer distances. Distance 1 decreases from 5.537 Å at 0.08 GPa to 5.384 Å in the new form. Distance 2 decreases from 6.305 Å at 0.08 GPa to 6.000 Å in the new form. Hydrogen bonding shown between dimers. .... 70

**Figure 4.11** Dispersive molecule to molecule interactions are indicated by numbers which correspond to the numbers in Figure 4.12. The dashed line at interaction 4 indicates that the Pixel interaction is directly into the plane of the page, with another molecule of saccharin directly behind the centre molecule when viewed in this orientation..... 71

**Figure 4.12** Calculated Pixel energies of the five main interactions in β-saccharin. Coulombic, polarisation, dispersion, repulsion and total energies are plotted for each of the interactions 1-5, corresponding to Figure 4.11. The interaction directions are written as a function of the unit cell a) ambient pressure polymorph (at 0.08 GPa) in full colour; graph 1 b) high pressure polymorph (1.32 GPa) with pattern; graph 2. .... 72

**Figure 4.13** Void space analysis with a probe of 0.6Å radius and 0.2Å grid spacing at 0.08 GPa before the phase transition and 1.32 GPa after the phase transition. .... 73

**Figure 4.14** Plots showing how the length of piracetam (form II: half symbol & form V: filled symbol), INSA, tolazamide, and salicylamide’s N-H···O amide dimer respond to pressure. The compression of α and β saccharin is plotted. β-saccharin contains amide dimers at two different lengths..... 75

**Figure 4.15** Selected compression studies from the CSD showing a polymorphic phase transition in the compression series by an empty symbol<sup>37,38,43,69–72</sup> ..... 76

**Figure 5.1** Ternary phase diagrams for the formation of benzoic acid and isonicotinamide in a) water<sup>10</sup> b) ethanol<sup>10</sup> and c) methanol<sup>1</sup> ..... 84

**Figure 5.2** Ternary phase diagram showing regions of the co-crystal formation in ethanol as produced by Svoboda et al. .... 86

**Figure 5.3** Ternary phase diagram showing regions of the co-crystal formation in water as produced by Svoboda et al. .... 91

**Figure 5.4** The asymmetric unit of benzoic acid and isonicotinamide with numbering scheme, at ambient pressure. The colours of the elements are consistent throughout (carbon, grey; oxygen, red; nitrogen, blue; hydrogen, white). .... 96

**Figure 5.5** Tetramer unit with the hydrogen bonding interactions of the 1:1..... 97

**Figure 5.6** Hydrogen bonded sheets of 1:1 benzoic acid and isonicotinamide in colour, made up of the base tetramer unit of the 1:1 a) complete tetramer shown in the middle layer in black, which is side bonded to another tetramer at an angle of 87.26° (arrows signify the lines of molecules which are in and out of plane) b) coloured layers of the structure in the a-axis showing a ‘wine-rack-type’ structure. .... 98

**Figure 5.7** The aromatic stacking directions of the 1:1 relative to the unit cell axes, with ring to ring centroid distances noted. .... 99

**Figure 5.8** a) BFDH predicted morphology of the 1:1 co-crystal of benzoic acid and isonicotinamide viewed down the c axis of the unit cell in agreement with the experimentally found morphology b). .... 100

**Figure 5.9** Unit cell lengths as a function of pressure. The errors on each of the measurements is smaller than the symbol. Volume of the unit cell as a function of pressure (EoS fit 7). .... 101

**Figure 5.10** Unit cell and packing of the co-crystal viewed down the b-axis. Void spacing of the unit cell shown at (a) ambient, (b) 3.05 GPa and (c) 5.32 GPa. .... 102

<b>Figure 5.11</b> Molecules of the 1:1 shown with stick images superimposed to enable easier visualisation of the ‘hinge’ mechanism. At ambient pressure the two molecules are ‘open’ with an angle of $87.26^\circ$ and at 5.32 GPa the hinge mechanism ‘closes’ with an angle of $74.07^\circ$ between the two molecules. ....	<b>103</b>
<b>Figure 5.12</b> Plots showing how the length of piracetam (form II: half symbol & form V: filled symbol), INSA (1:1 co-crystal of indomethacin and saccharin), tolazamide, salicylamide, saccharin and the 1:1 co-crystal of benzoic acid and isonicotinamide’s N-H··O amide dimer responds to pressure.....	<b>104</b>
<b>Figure 5.13</b> Polarised light microscope image of the 2:1 co-crystals. ....	<b>105</b>
<b>Figure 5.14</b> Two trimers of the 2:1 co-crystal with hydrogen bonding motif shown. ....	<b>106</b>
<b>Figure 5.15</b> Hydrogen bonded layers of 2:1 benzoic acid and isonicotinamide in colour, the coloured layers of the structure in the c-axis display the same ‘wine-rack-type’ structure present in the 1:1 co-crystal.....	<b>107</b>
<b>Figure 5.16</b> The aromatic stacking directions of the 2:1 relative to the unit cell axes, with ring to ring centroid distances depicted.....	<b>108</b>
<b>Figure 5.17</b> Unit cell lengths as a function of pressure. The errors on each of the measurements is smaller than the symbol. Volume of the unit cell as a function of pressure (EoS fit 7). ....	<b>109</b>
<b>Figure 5.18</b> Unit cell and packing of the co-crystal viewed down the b-axis. Void spacing of the unit cell shown at (a) ambient, (b) 2.73 GPa and (c) 5.99 GPa. ....	<b>110</b>
<b>Figure 5.19</b> The 2:1 co-crystal, viewed down the b-axis at ambient and 5.99 GPa. The benzoic acid molecules display a wine-rack-type compression behaviour similar to that of the 1:1 co-crystal. ....	<b>111</b>
<b>Figure 5.20</b> Molecular volume comparison over the course of the compression study for the 1:1 and 2:1 co-crystal of benzoic acid and isonicotinamide. ....	<b>111</b>
<b>Figure 6.1</b> Artemisinin molecule.....	<b>117</b>
<b>Figure 6.2</b> Representation of a decision tree based on co-crystallisation with responses ‘yes’ for successful co-crystallisation and ‘no’ for unsuccessful co-crystallisation. A RF is composed of many decision trees.....	<b>121</b>
<b>Figure 6.3</b> XRPD patterns of Artemisinin, 1-naphthol and the ball milled sample revealing new solid phase ArtNap. ....	<b>129</b>
<b>Figure 6.4</b> Thermogram of the ArtNap co-crystal. ....	<b>130</b>
<b>Figure 6.5</b> Thermogram of artemisinin. ....	<b>131</b>
<b>Figure 6.6</b> Thermogram of 1-naphthol.....	<b>131</b>
<b>Figure 6.7</b> The labelled co-crystal of artemisinin and 1-naphthol, only the hydrogen atom involved in the supramolecular synthon is labelled for clarity. The colours of the elements are consistent throughout (carbon, grey; oxygen, red; hydrogen, white). ....	<b>132</b>
<b>Figure 6.8</b> View down the b axis showing 1:1 molecular units of the co-crystal and the hydrogen bonding between components. ....	<b>132</b>
<b>Figure 6.9</b> Artemisinin is hydrogen bonded to 1-naphthol in 1:1 stoichiometric ratio as well as interacting with a neighbouring 1-naphthol molecule via dispersive interactions. Two main interactions labelled to match Table 6.5.....	<b>133</b>
<b>Figure 6.10</b> BFDH predicted morphology viewed down the a) c-axis and b) b-axis. c) face indexing confirms the largest face to be the (0 0 1) face, consistent with BFDH morphology predictions.....	<b>134</b>
<b>Figure 6.11</b> View of the ArtNap down the b axis showing the layers of interactive components. ....	<b>135</b>

### List of abbreviations

Active Pharmaceutical Ingredient	API
Artemisinin and 1-Naphthol co-crystal	ArtNap
Bravais Friedel Donnay Harker	BFDH
Cambridge Structural Database	CSD
Charge Coupled Device	CCD
Diamond Anvil Cell	DAC
Differential Scanning Calorimetry	DSC
Equation of State	EoS
Generally Regarded As Safe	GRAS
Indomethacin and Saccharin co-crystal	INSA
Liquid Assisted Grinding	LAG
Neat Grinding	NG
Non-Steroidal Anti-Inflammatory	NSAID
Pressure Transmitting Medium	PTM
Random Forest	RF
Single Crystal X-Ray Diffraction	SCXRD
X-Ray Powder Diffraction	XRPD

***-Chapter One-***

***Introduction***

## 1. Introduction

### 1.1. Synopsis

An overview of the theory and literature surrounding the topics within this thesis are presented in the introduction and materials and methods chapter. Beginning with a broad view of supramolecular chemistry, specific terms of relevance such as the definition of a crystal, polymorph and co-crystal are discussed. The backgrounds of certain topics of interest to the work within this thesis are covered such as crystal engineering, machine learning, co-crystals in the pharmaceutical industry and high pressure crystallography.

### 1.2. Supramolecular chemistry and the crystalline form

Fundamental properties of a compound, such as the simple melting point, could not be understood without considering the chemistry beyond the single molecule. Considering the grander scale beyond the single molecule, we begin to understand the physicochemical properties of materials. Many physical attributes depend on the assembly of multiple molecules rather than just the single molecule. Supramolecular chemistry aids in explaining certain physicochemical properties. One of the principle ideas of supramolecular chemistry concerns the overall network of molecules rather than the network components i.e., the molecules.

The specifications of the crystalline form are confirmed at a supramolecular level. Firstly, the solid must be comprised of specifically packed molecules which form a repeating unit and show internal regularity. A crystal exhibits long range order and as a result will exhibit Bragg diffraction when probed by X-rays and should be homogenous. The unit cell is the fundamental building block applied by crystallographers as a means to understand the overall crystalline lattice. The parameters of the motif are defined and identically repeated creating a structured crystalline lattice. The crystalline lattice can be thought of as a set of theoretical points, a mathematical concept. By understanding the unit cell and its symmetry, it is possible to investigate the physical and physicochemical properties displayed by the molecule.

### 1.3. Polymorphism

Polymorphism comes from Greek meaning many (polus) shape (morph). In materials science the term is used to describe different crystalline forms that a chemical entity may adopt. These different forms are identical in the liquid and vapour states and the

## Lauren Connor

term polymorph comes into play when we consider the solid state. If a chemically identical compound can exist in at least two different solid crystalline phases, then such a compound is said to be polymorphic. Differences in the three dimensional packing arrangements of the same compound may give rise to different physicochemical properties.<sup>1</sup> Friedrich Wöhler and Justus von Liebig reported the first reported case of polymorphism. In 1832, they investigated a boiling solution of benzamide and observed needle shaped crystals slowly transforming into rhombic shaped crystals. 175 years later Thun *et al.*<sup>2</sup> revisited the work of Wöhler and Liebig to solve the structure of the polymorph and a year later discovered another new form.<sup>3</sup> It was noted that the three forms of benzamide have the same hydrogen bonding interactions but different  $\pi$ - $\pi$  interactions. A difference in intermolecular interactions as well as changes in arrangement of the molecules can influence many physicochemical properties between polymorphic forms. A range of properties such as melting or sublimation temperature, heat capacity, solubility, density and refractive index may differ between polymorphs of a given compound. For this reason, the phenomenon has been the focus of research for many years in materials science<sup>4-7</sup> and in the pharmaceutical industry.<sup>8-17</sup>

Microscopist Walter McCrone famously noted:

“Those who study polymorphism are rapidly reaching the conclusion that all compounds, organic or inorganic, can crystallize in different forms or polymorphs. In fact, the more diligently any system is studied, the larger the number of polymorphs discovered.”<sup>18</sup>

McCrone's statement may be one of the most well-known and used quotes by a scientist in relation to polymorphism. It suggests that the phenomenon of polymorphism may be the rule rather than the exception. As more methods of crystallisation are explored and the number of experiments rises as well as the recording of experiments, there has been a rise in the number of polymorphs published. However, despite the rising numbers and despite McCrone's statement, there are still only 28,703 of the current >900,000 structures in the Cambridge Structural Database (CSD)<sup>19</sup> with the word 'polymorph' explicitly stated. The number of reported polymorphs is low, at only ~3% of the whole CSD. Such a low percentage of reported polymorphs could indicate that the experimental space routinely covered needs expanding. Another possible reason for a low percentage of polymorphs may simply be due to the rapid expansion of the CSD in general, the number of new primary structures has been rapidly increasing over the years due to vast

## Lauren Connor

improvements in X-ray diffraction. 5-methyl-2-[(2-nitrophenyl)amino]-3-thiophenecarbonitrile, commonly known as ROY for the colours of its polymorphs, had been known as the compound with the most structurally solved (seven) polymorphs from 2005. That was until 2012, when López-Mejías *et al.*<sup>20</sup> solved six structures of a non-steroidal anti-inflammatory, flufenamic acid, adding to the two previously solved structures. López-Mejías *et al.* noted that there was a ninth structure of flufenamic acid that as of yet had only been confirmed by X-ray powder diffraction (XRPD) but had not been structurally characterised. Flufenamic acid remains the compound with the most structurally characterised forms but there have been reports of up to ten polymorphs of 1,4-diazabicyclo[2.2.2]octane hydroiodide however only six of these have been structurally characterised.<sup>21</sup>

There are different categories of polymorphism; packing, conformational, synthon, tautomeric and concomitant. Packing polymorphism arises from differences in the packing arrangement of the molecules in a crystal structure. Conformational polymorphism is mostly observed to occur in larger, more flexible molecules. Larger, flexible molecules allow for rotation around covalent bonds and as a result will usually have multiple molecular shapes within a small energy window. As a result of the conformational flexibility of the molecule, multiple crystal structures can be obtained with the same chemical entity. Ritonavir exhibited conformational polymorphism around one bond and due to the change in rotation around the bond, the molecule changed in shape, lead to different packing arrangements.<sup>13</sup> It is common for polymorphs to fall under more than one polymorphism category. Synthon polymorphism is the rarest form of polymorphism observed. Changes in the primary synthon and therefore pattern of bonding between molecules are observed in this form. Pogoda *et al.*<sup>22</sup> investigated a case of synthon polymorphism in 5-nitrofurazone. Form  $\beta$  exhibited an amide synthon of a primary and secondary amide whereas the  $\gamma$ -form had an unusual N-H $\cdots$ O synthon involving a primary amide and a ternary cyclic amide as well as an instance of a carbonyl accepting two hydrogens. The  $\alpha$ -form involved 3 different synthon patterns, all of which were different to those in the other two forms. Tautomeric polymorphism occurs when an acidic proton is able to move to another position on the molecule. The anti-ulcer drug, omeprazole, has two different tautomeric polymorphs. The acidic proton can either sit on the 5 or 6 position of the methoxy- ring in the structure. The word concomitant can be used to describe polymorphs which crystallise under the same experimental conditions. Three distinct polymorphs of CICSSN, a diathiadiazolyl all crystallise via sublimation.<sup>23</sup>

## Lauren Connor

Polymorphism can be induced using a variety of different crystallisation techniques.<sup>22,24–28</sup> Crystallisation is one of the staple techniques used by scientists everywhere to purify substances and in the pharmaceutical industry it is a pivotal step in the formulation development of a drug.

### 1.3.1. Importance of polymorphism

Polymorphism is of interest across many different industries. In the food industry for example, manufacturing chocolate is a delicate process due to the six different polymorphic forms of cocoa butter of which one particular polymorph (form V) has the desired characteristics.<sup>28</sup> The process of tempering chocolate is fragile as the six different forms of chocolate can appear over temperatures 289-309 K.

The pharmaceutical industry also has great interest in polymorphism as lack of due care towards screening and isolating the intended form can be costly. The drug discovery timeline is a lengthy and expensive process upon which companies can spend somewhere around the value of £834 million.<sup>29</sup> Different polymorphs have different physicochemical properties and physicochemical properties are the key factors affecting the development of appropriate dosage forms. The costs of unexpected polymorphic transitions can impact efficacy for the patient and impact processing as well as being costly in the monetary sense of the word. The classic example of polymorphic cost is ritonavir, the Active Pharmaceutical Ingredient (API) in Norvir®, an antiretroviral medication targeting HIV. Ritonavir exhibited conformational polymorphism and transitioned into a new thermodynamically more stable form two years after its release. Inherently, the new stable polymorph had much lower solubility

(-50 %) and bioavailability. The product had to be retracted by its producers (Abbott) and reformulated at considerable cost to the company.<sup>13</sup>

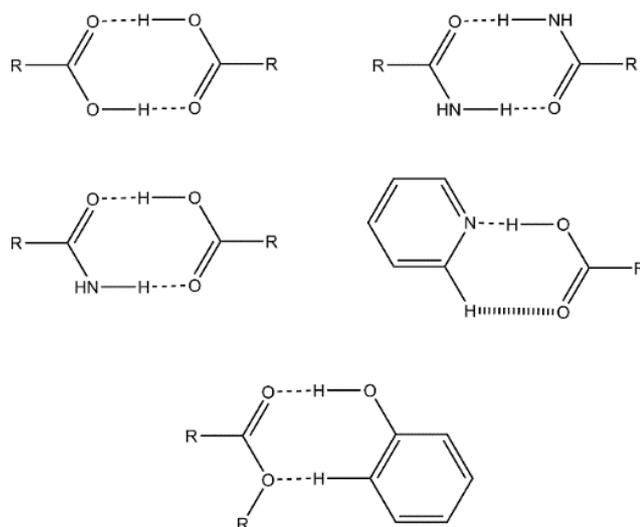
### 1.4. Crystal engineering

Crystal engineering was introduced by Pipensky, however, the first mention of the term crystal engineering is thought to have been made by Gerard Schmidt in a meeting of the American Physical Society, in 1955. Schmidt's explanation of the term included the three main components considered in the definition of crystal engineering today: analysis, design and function (relating to structure). The term crystal engineering suggests that we can direct and control the design of the structure for specific scientific needs; for application or fundamental purposes.<sup>30</sup>

## Lauren Connor

The key to crystal engineering is noncovalent bonding. Much of the initial work on organic systems focused on the use of hydrogen bonds.<sup>31</sup> However, halogen bonds<sup>32</sup> and weaker dispersive forces such as  $\pi\cdots\pi$  and  $\text{CH}\cdots\pi$  interactions may be considered when designing a functional crystal. Yao *et al.*<sup>33</sup> recently described introducing a bulky substituent group to pentacene to enable an improvement in the  $\pi$ - $\pi$  overlap of the system by modification of the packing arrangement. The herringbone packing arrangement in pentacene had poor  $\pi$ - $\pi$  overlap whereas the brick layer stacking of TIPS-pentacene (a derivative of pentacene) showed improved overlap and solubility.

Organic molecular self-assembly or recognition typically involves an interaction between moieties with hydrogen bond donor or acceptor capabilities. Hydrogen bond directed crystal engineering approaches are the most widely explored due to our familiarity with the supramolecular synthon. The supramolecular synthon provides the link between the molecular components. Five of the most common supramolecular synthons are shown in Figure 1.1.



**Figure 1.1 Schematic of five commonly encountered supramolecular synthons.**

In applying analysis of known structures to functional designs we usually consider the supramolecular synthon approach. Attempts to direct the bonding of the molecular units may begin by utilising knowledge of previous synthon successes. Majumde *et al.*<sup>34</sup> made use of their years of previous work on the 1D supramolecular synthons of low molecular weight gelators to design a salt-based supramolecular gelator of non-steroidal anti-inflammatory, naproxen. The gelator was converted into a topical gel and proven to be efficacious in treating skin inflammation in mice. This eliminated the

## Lauren Connor

need for drug loading into a gel matrix by simply using the supramolecular synthon approach to convert the drug into a gel itself.

Crystal engineering approaches can also be used to effectively target the morphology of crystalline systems. Morphology and crystal structure can impact the dissolution rate, bioavailability, downstream processing, and ultimately, the efficacy of a crystalline drug. Experimental variables such as pressure, temperature, pH, media composition and concentration may be adapted during crystal growth of active drug compounds in order to dictate parameters such as crystal shape and form.<sup>35</sup> Mirza *et al.*<sup>36</sup> demonstrated using an additive to control the crystal habit of erythromycin in order to improve tableting performance. A common pharmaceutical excipient, hydroxypropyl cellulose, was simply added in various concentrations to the crystallisation of erythromycin which aided in producing more equidimensional crystals with significantly improved compaction performance.

Agnew *et al.*<sup>37</sup> demonstrated their use of crystal engineering to control form. Production of an elusive metastable polymorph of paracetamol was made possible by using a templating approach where a 4-halobenzoic acid derivative was added to cooling crystallisations of paracetamol.

A more computational approach is becoming common with tools to analyse structures and gather data from previous structures published in the Cambridge Structural Database (CSD).<sup>19</sup> A bank of almost 960,000 structures is present in the CSD. Established in 1965 in Cambridge University by Kennard, it has rapidly expanded in line with the boom of “big data”. The CSD contains results of X-ray and neutron diffraction studies of organics, organometallics and complexes of metals. The database stores bibliographic information, crystallographic data and chemical connectivity information for each entry which is named as a refcode. The CSD consists of many components but two of which are of interest in this thesis 1) ConQuest: allows for the searching of information and retrieving it. 2) Mercury: helps in visually looking at a structure. The growth of the CSD along with its ever expanding wealth of tools to aid in data mining and programming make it a perfect tool for the crystal engineering and crystallography fields of study. Data collected from this software helps in understanding the supramolecular synthons that could be formed between functional groups.

### 1.5. Machine learning

Machine learning is a data-driven approach which is coming to the fore in crystal engineering. It is an application of Artificial Intelligence (AI) and is the science of a computer functioning without being explicitly programmed. It can be used to make a prediction or determination surrounding the data it learns from. In an age of a data boom, machine learning has become a topic of huge interest. It has the ability to programme self-driving cars<sup>38</sup>, speech recognition<sup>39</sup>, face recognition<sup>40</sup> and effective internet searches.<sup>41</sup> We may use devices and applications which are programmed by machine learning many times a day without even realising. The breakthrough of machine learning with such force was instigated by a realisation, made by a pioneer of AI, Arthur Samuel in 1959.<sup>42</sup> Samuel realised that it may be worthwhile allowing computers to learn for themselves rather than to teach them everything we know. The relatively recent emergence of the internet and the boom of digital data generation and storage was a great enabler in the machine learning breakthrough.

In machine learning, a research question is posed and the dataset contains responses to the question. Some datasets can then be classified into both positive and negative data if the answer to the question is a categorical answer rather than regression based. Data that meets the machine learning research question posed is classified as positive data and data that does not meet the research question is classified as negative data. One of the main problems in applying machine learning approaches is access to data. Machine learning requires large consistent sets of data to train the computer model on and with more consistent data the models become more accurate. Access to banks of data like this has become more normalised<sup>43</sup> in the past few years however for specific problems access to negative data may be an issue. Wicker *et al.*<sup>44</sup> encountered the negative data issue when using a Support Vector Machine (SVM) approach to co-crystallisation. Wicker *et al.* negated the issue by generating a dataset of positive and negative experimental results in-house. The large dataset was the first of its size and kind published in the field of co-crystallisation. Some researchers from this group then went on to target the negative data issue directly by publishing a crystallisation data set. Pillong *et al.* produced the data set to enable choice of solvent for crystallisation. Raccuglia *et al.*<sup>45</sup> took a different approach to the problem of negative data from the outset. Raccuglia *et al.* set up a web-accessible database to facilitate the entry of negative data. Their recent study directly uses the data from failed experiments to successfully predict conditions for new inorganic-organic compounds such as metal-organic frameworks (MOFs) and perovskites.

## Lauren Connor

In crystal engineering the first step is analysis of previous cases. By employing a machine to apply analysis based on known cases, diverse data can be analysed where human understanding may fall short due to the inability to consider many independent variables at once or spot any kind of pattern. The second and third steps in crystal engineering are design and function. By controlling the test data fed into the model the design and function steps are also managed. Whilst machine learning may negate the need for the in depth understanding of the problem, the analysis of the previous cases is present and the concept does fit well with crystal engineering for functional design. Johnston *et al.*<sup>46</sup> demonstrated using analysis of previous cases in the search for solvates of a drug. Johnston *et al.* used the results of a previous experimental study on the classification of solvents to target the crystallisation of novel carbamazepine solvates.

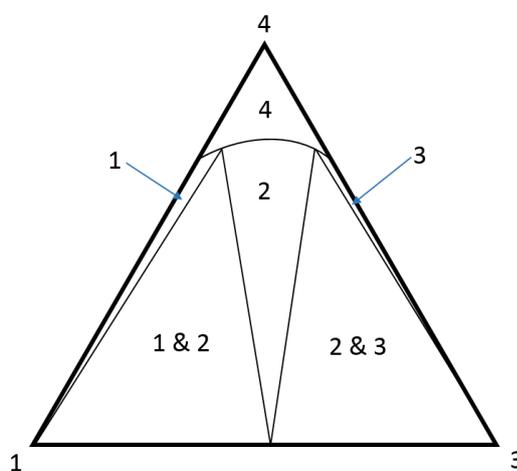
### 1.6. Co-crystals

Both the word co-crystal (hyphen/no hyphen) and the definition of the term have been a point of topical debate since the time of their consideration as a form of modification within solid state chemistry. To define the co-crystal we apply supramolecular chemistry as well as crystal engineering: "a co-crystal is the consequence of a molecular recognition event between different molecular species."<sup>47</sup> As with polymorphism, the term co-crystal only applies to the solid state:

Co-crystals are "solids that are crystalline single phase materials composed of two or more different molecular or ionic compounds generally in a stoichiometric ratio which are neither solvates nor simple salts".<sup>48</sup> This definition is generally accepted and clearly defines the co-crystal as non-covalently interacting, separating it from chemical reactions. As well as containing only components which are solids in their singular forms, separating it from solvates and hydrates under the umbrella of multicomponents.<sup>25,49,50</sup> The earliest considered co-crystal was studied by German chemist Friedrich Wöhler in 1844, the co-crystal was a 1:1 molar multicomponent system called quinhydrone, and the components of the system were quinone and hydroquinone. This co-crystal was studied for years after its discovery.<sup>51</sup> Over 170 years later, co-crystals are still discovered by various screening techniques. There are many ways to co-crystallise a compound; solution based techniques<sup>52</sup>, neat grinding<sup>44</sup>, liquid assisted grinding<sup>53</sup> and high pressure<sup>54</sup> to name a few.

### 1.7. Solution based methods

Co-crystallisation by solution based methods has an added complication when considering the relative solubilities of the two components. It is only useful for a combination of components whose co-crystallisation would be thermodynamically favoured when considering how complimentary the functional groups are, however, the co-crystallisation may not be kinetically favoured.<sup>55</sup> There are ways to tamper with the solubility in the crystal engineer's favour during co-crystallisation to shift the equilibrium towards co-crystallisation. Binary (two component) and ternary (three component) phase diagrams can be produced for multicomponent systems to show the relative environments for the production of each phase. The relationships between the two components can be quantitatively described. A molar excess of one component can result in a co-crystal due to the decrease in solubility of the component. Ternary phase diagrams can indicate regions of different stoichiometric formations. Ternary phase diagrams are often displayed as a triangle and show the compositions at a specific temperature. They can show that, for example, evaporation of solvent would lead to the movement from one region to another (away from the solvent apex) and potentially into the co-crystal phase (2 in Figure 1.2).



**Figure 1.2 Example of a ternary phase diagram with Coformers 1 and 3 at 100% at the bottom two apexes. 100% solvent at 4 in the top apex. Regions 1 and 3 within the triangle show regions where the co-formers 1 and 3 would be produced in solution, respectively. Regions 1 and 3 within the triangle show regions where the co-formers 1 and 3 would be produced in solution, respectively. Region 2 shows the region for formation of a co-crystal.**

## Lauren Connor

From a crystal engineering perspective the intermolecular interactions of a compound can be assessed and manipulated to alter the physical properties of a material. A recent paper by Sandhu *et al.*<sup>56</sup> used a co-crystallisation approach to alter the solubility and lower the hygroscopicity of urea by adding in a less soluble co-former. As a result urea presented improved physicochemical properties for use as an agrochemical. As yet, this process of fine tuning intermolecular interactions to control the outcome is not precise but examples like that of Sandhu *et al.* show that chemical intuition can show progress in the field.

### 1.7.1. Pharmaceutical co-crystals

The pharmaceutical industry aims to produce a safe, efficacious drug to either maintain or improve the life of those taking it. Approximately ten years are spent on research and development of the API. The aim of a pharmaceutical co-crystal is to achieve better physicochemical properties than those of the API in its pure form without altering the chemical or pharmacological properties of the API. The role of the co-former is to optimize the physical properties of the API. In terms of pharmaceutical compounds, most APIs rely on their dissolution in the gastrointestinal tract to be able to reach the site of action within the body. Therefore, the most problematic physical property of an API tends to be its solubility. Solubility can also affect the ability of an API to be administered in appropriate correct form. Sometimes if the drug is very poorly soluble then excess amounts must be administered for therapeutic effect. Such as was the case with anti-epileptic, carbamazepine (>100mg/day for therapeutic effect). Oral administration of carbamazepine was problematic with its low water solubility and limited bioavailability. When carbamazepine was co-crystallised in a 1:1 stoichiometric ratio with saccharin, the suspension stability as well as dissolution properties improved.<sup>57</sup>

Typically the improvement in physicochemical properties of a drug is measured in terms of the neat co-crystal dissolution profile. However, the 2013 paper by Childs *et al.* highlights co-crystals are new chemical entities and it is important to treat co-crystals as new APIs would be. Childs *et al.* not only found that the dissolution of a poorly soluble form of Danazol is improved 1.7 times by co-crystallisation with vanillin. The co-crystal was then tested in a formulation with typical excipients and found to be ten times more soluble than the initial API form. The supersaturated formulation was key to creating conditions for co-crystal dissolution, further improving the dissolution of the co-crystal.<sup>58</sup>

## Lauren Connor

Co-crystallisation offers a route to multicomponent forms that may be essential for some APIs that are non-ionisable and therefore cannot be improved by means of pharmaceutical salt formation. In 2006, McNamara *et al.*<sup>59</sup> investigated a sodium ion blocker API due to its low solubility and high permeability (Biopharmaceutical drug Classification Scheme class II). McNamara *et al.* found that by forming co-crystals with carboxylic acids the dissolution rate of the API improved 18 fold. A complexity can arise in the stoichiometry of the pharmaceutical co-crystal. The stoichiometric ratio does influence the physical properties of the API but it is not clear yet whether there is a direct correlation between the addition of a co-former in increased stoichiometric ratios and the improvement in solubility.<sup>60</sup> In a 2016 paper by Li *et al.*<sup>61</sup>, carbamazepine and para-amino benzoic acid were investigated in 1:1, 1:2 and 1:4 ratios. There was no correlation noted between the addition of the more soluble co-former and the solubility.

There is a larger scope of components available to use as a counter component in co-crystals than there are suitable counter ions in salt formation. The area of pharmaceutical co-crystals has boomed. Many APIs present multiple functional groups with hydrogen bonding making them attractive for co-crystal formation. Increased hydrogen bonding availability can also be a good indicator of the risk of polymorphism in the co-crystal. Polymorphism in co-crystals can open up further avenues for suitable drug molecules as is the case with polymorphism in singular components. Two co-crystals of a ethenzamide and gentisic acid co-crystal were investigated by Sokal *et al.* in 2017.<sup>62</sup> The two components were compared to ethenzamide and to one another. Their compression behaviour and dissolution was studied to determine which would be most suitable for tableting and pharmaceutical formulation. Once more, the excipients played a major role, having an effect on the compression and dissolution. Addition of excipients on compression prevented phase transition from the metastable co-crystal to the stable co-crystal and the excipients greatly influenced the dissolution profiles of both co-crystals.

It was previously thought that co-crystallisation could be a method to avoid polymorphism<sup>55,63-65</sup> however a recent paper by Cruz Cabeza *et al.*<sup>66</sup> states that polymorphism is just as likely with multicomponent lattices. Holding true to Walter McCrone's statement on the likelihood of polymorphism, there may, as of yet, be a lack of time spent on co-crystals to really explore polymorphism of multicomponents.

## Lauren Connor

### 1.8. X-ray diffraction

To solve the structures of the crystalline form X-ray diffraction is employed. Max von Laue was awarded the Nobel Prize in 1914 for his successful diffraction of X-rays. However, it was the Braggs, father and son duo, who homed in on X-rays as a crystallographic tool. They were able to conclude that X-rays were electromagnetic waves. Their research verified that X-rays have a wavelength of atomic dimensions, enabling them to interfere with crystalline lattices and provide information on the crystal structure. This was the first tool to yield images from which atomic structures of the solid state could be determined. The Braggs were able to look at simple structures such as NaCl. The Braggs' breakthrough in the field meant that the work of Niels Bohr on the atomic model was able to be investigated and proven which lead to the concept of chemical bonding.

Single Crystal X-ray Diffraction (SCXRD) is responsible for most widely accepted atomic models of molecular structure. The X-ray diffractometers required for data collection on single crystals are widely available, and the breadth of chemical samples now solved by single crystal X-ray methods is impressive, ranging from small organic molecules<sup>67</sup> and inorganic coordination complexes<sup>68</sup> to some of the largest discrete molecules in biological systems, such as the 70S ribosome macromolecular complex.<sup>69,70</sup> Pure compounds, from small molecules to organometallic complexes, proteins, and polymers, solidify into crystals under the proper conditions. The disadvantage of single crystal X-ray methods is the requirement for a crystal of sufficient size and quality. Many molecules cannot form crystals that are sufficiently well ordered to diffract to "atomic resolution" with issues such as disorder, a problem in many systems.

### 1.9. High pressure crystallography

High pressure diffraction was first attempted soon after the application of X-rays to determine crystal structures. Cohn et al. carried out high pressure powder diffraction in 1933.<sup>71</sup> The first experiment was conducted using a steel pressure bomb compressing gas and the sample and reaching pressures of ~0.1 GPa. Two beryllium windows in the cell allowed for the incident and diffracted beams to pass through. The first high pressure experiments were limited as the equipment was too large and noisy background from the pressure cell meant that reflections were often not strong enough to record.<sup>72</sup> High pressure equipment improved after the concept of the Diamond Anvil Cell (DAC) was introduced by Bridgman and improved by Jamieson et

## Lauren Connor

*al.*<sup>73</sup> Bridgman constructed an apparatus with opposing anvils that were capable of applying and withstanding pressures of up to 10 GPa in the 1950s.<sup>74</sup> However the anvils of Bridgman's design were bulky and problematic. Jamieson *et al.* proposed a change of anvil material from tungsten carbide cemented in cobalt to diamond, in 1959. This proposal stuck and diamond remains the material in the modern day DAC. Today's DAC is suitable for any electromagnetic radiation to pass through its diamond windows and is small enough to fit into a diffractometer casing for SCXRD. The modern, elegant DACs are used in many branches of high-pressure science such as geophysics,<sup>75</sup> biology,<sup>76</sup> condensed matter physics and chemistry.<sup>77,78</sup>

Pressure is a fundamental thermodynamic variable that can be explored to a much greater extent than other variables such as temperature. The temperature we can exert on any given structure is limited to a few thousand degrees. Whereas, in high pressure research we are able to reach pressures in the terapascal region,  $10^7$  orders of magnitude from atmospheric pressure.<sup>79</sup> In varying the pressure of a system we are potentially able to encounter new forms of compounds and possibly bring the new forms back to ambient pressures. Fabbiani *et al.* investigated  $\gamma$ -aminobutyric acid (GABA), an inhibitory neurotransmitter studied for the treatment of anxiety and Parkinson's diseases, at high pressure. Fabbiani *et al.*<sup>80</sup> crystallised GABA in its monohydrate form in the 0.4-0.8 GPa pressure range and were able to recover this form back to ambient pressure. The small sample size of a loaded DAC can be inhibiting if a form (which is recoverable to ambient) is of interest. However, like the case of GABA by Fabbiani *et al.*, the small sample can be used to seed reactions to obtain the form of interest in larger quantities.

High pressure research does not only concern itself with the search for new polymorphs at high pressure to bring back to ambient. Compression studies have been used to gain detailed knowledge of how intra- and intermolecular forces respond to high pressure environments. Both the direct compression studies of materials and crystallisation at high pressure from solution have been used in many studies to explore interactions.<sup>7,15,81-85</sup> Hydrogen bonding and van der Waals forces are sensitive to compression in the pressure range usually explored for organic materials (up to ~10 GPa). Patterns relating these forces to pressure have been made. For example, a reasonable observation from many high pressure studies is that hydrogen bonded chains withstand compression along the direction of the chain.<sup>86</sup> Where hydrogen bonded chains are stacked with an inter-planar 'slip plane' (a plane across which there is little interaction) the compression will be greatest perpendicular to the

## Lauren Connor

slip plane. Some studies have also reported that hydrogen bonding can act like a spring<sup>82</sup> or a wine-rack<sup>87,88</sup> if it is zig zagged along one direction. Another reasonable conclusion from high pressure trends is that, in general, reduction of void space occurs during compression<sup>11</sup> although there are exceptions in molecular organic frameworks (MOF). Moggach *et al.*<sup>89</sup> found that the pore size of ZIF-8 ( $\text{Zn}(\text{MeIM})_2$ , MeIM = 2-methylimidazolate) increased with compression due to the entry of the hydrostatic medium into the structure, increasing the unit cell.

High pressure research is of interest to the pharmaceutical industry as often APIs are subjected to mechanical stress and pressure during tableting and milling. Although relatively small pressures in comparison to most studies performed in a DAC (<0.1 GPa), phase transitions can occur. Paracetamol is known to change form depending on the crystallisation technique. Compression and high pressure recrystallization can impact on paracetamol.<sup>90,91</sup>

### 1.10. Thesis outline

This thesis aims to explore small molecule crystal structures at high pressure using a DAC and elucidate their compression behaviour. Structural investigations by the use of SCXRD is key in this thesis and the combination of experimental and computational methods allow for in-depth analysis throughout. After a description of the methods used in this thesis, for example, crystallography, high pressure crystallography, synchrotron radiation and Pixel calculations in the following chapter, three chapters will contain structural studies on single and multicomponent crystals at high pressure. In a structural informatics approach, the systems are analysed using tools in Mercury.<sup>92</sup> Structures in the first two chapters are further understood using Pixel methods. The behaviour of all of the systems under compression will be compared to other compression studies by means of data-mining the CSD. The final experimental chapter of this thesis aims to explain the use of a machine learning approach to co-crystallisation of a challenging API. Following failed screening attempts as part of a placement at Nanyang Technological University in Singapore, the final experimental chapter encompasses crystal engineering methods with a data driven approach to co-crystallisation.

***-Chapter Two-***  
***Materials and methods***

### 2. Materials and methods

All materials and solvents used in this research were commercially available and were purchased from Sigma Aldrich or TCI (Tokyo Chemical Industry). Materials were used as received or recrystallised as specified in the body of this thesis.

#### 2.1. Crystallography

##### 2.1.1. Producing X-rays

X-rays are short-wavelength and high frequency beams of electromagnetic radiation. X-rays are produced using an X-ray tube. Inside the sealed tube, a heated filament (cathode) emits electrons. The electrons are accelerated towards a metal target (anode) due to a high voltage across the tube. On collision with the metal target, an electron of sufficient energy will displace electrons from the inner shell of the metal atom creating holes in the inner shell. The subsequent cascade of electrons from the outer shell of the atom to the inner shell emits radiation in the form of X-rays. Three main wavelengths of radiation are produced;  $K\alpha_1$ ,  $K\alpha_2$  and  $K\beta$ . The metal target can be changed and is dependent on the X-ray wavelength required for a particular experiment. The most commonly used metals are copper ( $\lambda = 1.5405 \text{ \AA}$ ), molybdenum ( $\lambda = 0.7107 \text{ \AA}$ ) and silver ( $\lambda = 0.4559 \text{ \AA}$ ). In high pressure diffraction studies, molybdenum or silver is typically used to minimise the effects of the shading from the sample environment (Diamond anvil Cell). The short wavelength of the radiation allows access to a larger range of reciprocal space.

##### 2.1.2. X-ray crystallography

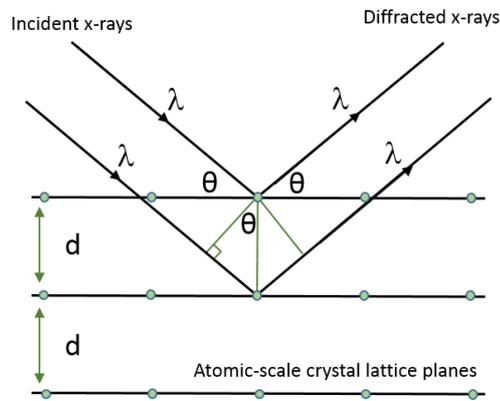
X-rays were discovered in 1895 by Wilhelm Conrad Röntgen who later received a Nobel prize. He serendipitously discovered these 'new rays' whilst investigating Crookes discharge tubes. He saw the importance of the discovery to medical applications when he took a picture of his wife's hand with X-rays. Following on from his discovery, Max von Laue was a pioneer and Nobel Prize winner who used X-rays to investigate crystals. It was William Henry Bragg and his son, William Lawrence Bragg who later proposed Bragg diffraction; a mathematical construct that relates direct and reciprocal space. The wavelength ( $\lambda$ ) of an X-ray is of the same order of magnitude (1-100  $\text{\AA}$ ) as the d spacing between planes in a crystal hence can give rise to constructive and destructive interference. D-spacing is the distance between

planes of atoms that give rise to diffraction peaks. For a material to be considered as a crystal it must obey Bragg's law:

$$n\lambda = 2d\sin\theta \quad \text{Equation 2.1}$$

Where  $\lambda$  is the radiation wavelength,  $d$  is the inter-planar spacing and  $\theta$  is the angle between the incident or diffracted X-ray with the plane.

Bragg's law is used to explain the diffraction experiment. For the intensities to be observed constructive interference must occur. This relies on the spacing between the planes to be at specific values such that the diffracted X-rays must be in phase (Figure 2.1).



**Figure 2.1 X-ray diffraction from crystal lattice planes (Bragg's law).**

A reflection is observed when a set of lattice planes (hkl) satisfies Bragg's law. The relative intensities of the peaks and the peak positions give information on the contents of the unit cell, the cell parameters and lattice system. Firstly, the data are indexed so that each measured peak is assigned (hkl) values. The unit cell parameters are calculated using the d-spacing of the (hkl) reflection (which is calculated using Bragg's law). The Bravais lattice type and space group are determined by investigation of absent reflections in the diffraction pattern due to symmetry related atoms. There are two types of absence, a general absence relating to Bravais lattice type and a systematic absence relating to translational space group symmetry. Depending on the symmetry of the crystal the absences will affect different zones of data. For example, in a body centred lattice, those reflections whose  $h + k + l$  is odd will be absent. For glide planes, planes of data will be affected e.g. (h 0 l) will be absent when  $h+l$  is odd for a n-glide or if h is odd then an a-glide is present. A  $2_1$  screw axis in the a direction will have a systematic absence at (h, 0, 0) where h is odd and absent.

## Lauren Connor

Each peak in the diffraction pattern has a structure factor,  $F_{hkl}$  which is related to the relative intensity and determines the phase and amplitude of the diffracted beams.

$$F_{hkl} = \sum_{j=1}^N f_j e^{-2\pi i(hx_j + ky_j + lz_j)} \quad \text{Equation 2.2}$$

The structure factors are related to the electron density which is required for structure solution. The reverse Fourier transform allows calculation of electron density. Electron density,  $\rho(xyz)$  is obtained by summing all of the structure factors and phase angles. Every part of a crystal structure contributes towards every reflection present in a diffraction pattern.

$$\rho(xyz) = \frac{1}{V} \sum_{h,k,l} [F(hkl)] \exp[i\phi(hkl)] \exp[-2\pi i(hx + ky + lz)] \quad \text{Equation 2.3}$$

Before the electron density map can be obtained, the phase problem must be considered. The phases cannot be recorded in a diffraction pattern, only the intensities. A method to overcoming this is by using direct methods in which possible phase angles are estimated. Patterson methods are also available that may be useful in structures with heavy atoms present. A successful solution then undergoes refinement against the experimental data using least-squares procedures to ensure that the structure is correct. The measure of a 'correct' structure is one in which the  $F_{\text{calculated}} |F_c|$  is equal to  $F_{\text{observed}} |F_o|$ ; the R-factor. The R-factor and the weighted  $R^2$ , in which each reflection is assigned its own weighted,  $w$ , are shown below. As the accuracy of the model improves, the R-factor reduces towards zero.

$$R = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|} \quad \text{Equation 2.4}$$

$$R = \sqrt{\frac{\sum w(|F_o|^2 - |F_c|^2)^2}{\sum w(|F_o|^2)^2}} \quad \text{Equation 2.5}$$

### 2.1.3. Crystal systems

There are seven different crystal systems with various restrictions on the unit cell parameters due to symmetry (Table 2.1). These seven systems vary from being low symmetry (triclinic) to high symmetry (cubic). For most organic systems, they reside in the lower symmetry systems due to the fact that on a molecular level they have low symmetry. Further classification can be made with four different lattice types;

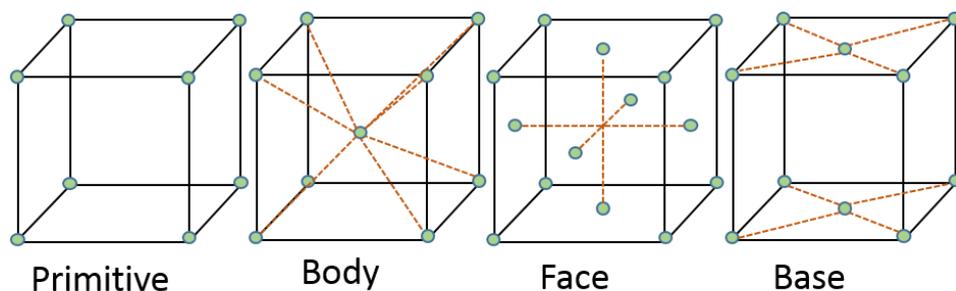
## Lauren Connor

primitive, body-centred, face-centred and base centred (sides A, B or C can be the 'base') shown in Figure 2.2. These lattice types can be identified from the diffraction experiment and the absence (or not) of reflections.

**Table 2.1 The seven crystal systems**

Crystal system	Unit cell restrictions	
	axes	angles
<b>Triclinic</b>	$a \neq b \neq c$	$\alpha \neq \beta \neq \gamma$
<b>Monoclinic</b>	$a \neq b \neq c$	$\alpha = \gamma, \beta \neq 90^\circ$
<b>Orthorhombic</b>	$a \neq b \neq c$	$\alpha = \beta = \gamma = 90^\circ$
<b>Tetragonal</b>	$a = b \neq c$	$\alpha = \beta = \gamma = 90^\circ$
<b>Trigonal</b>	$a = b = c$	$\alpha = \beta = \gamma \neq 90^\circ$
<b>Hexagonal</b>	$a = b \neq c$	$\alpha = \beta = 90^\circ; \gamma = 120^\circ$
<b>Cubic</b>	$a = b = c$	$\alpha = \beta = \gamma = 90^\circ$

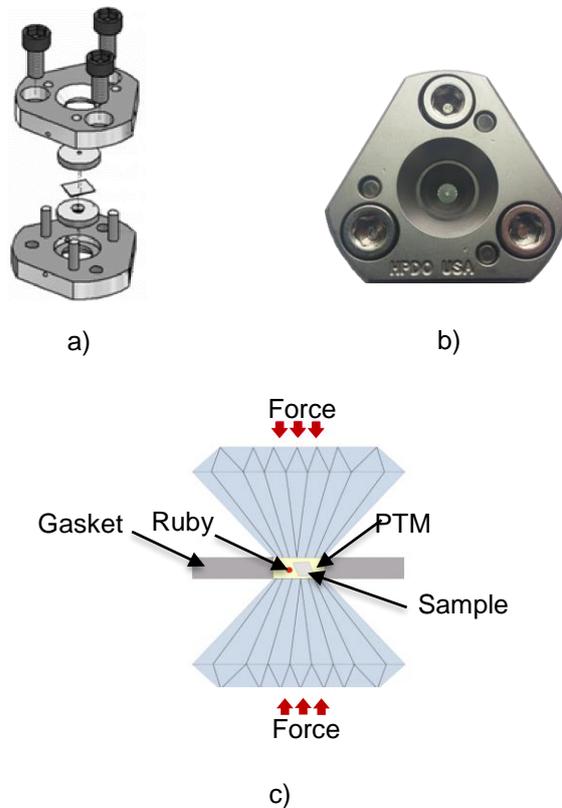
A crystal system is classified by a space group which represents all of the symmetry elements within the system. Symmetry elements include inversion, reflection, rotation, screw axes and glide planes. The combination of the symmetry elements, seven crystal systems and the fourteen possible Bravais lattices amount to 230 potential space groups.



**Figure 2.2 The four types of crystal lattice.**

## 2.2. High pressure crystallography

A Merrill-Bassett DAC is used in high pressure crystallography to compress samples in the GPa scale and analyse the sample by X-ray diffraction methods (Figure 2.3).

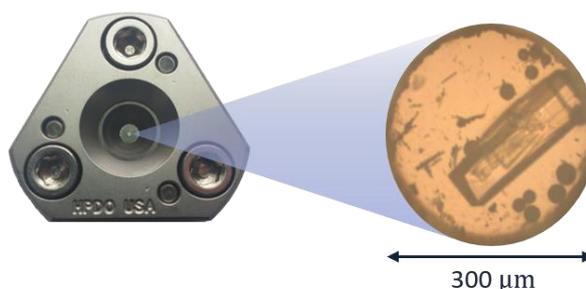


**Figure 2.3 Expanded view of an open DAC with components, b) bird's eye view of a Merrill-Basset DAC and c) labelled sample chamber between two diamonds.**

The cell is composed of two steel plates which have a conical hole at the centre (to allow radiation to pass through). The plates house tungsten carbide backing anvil seats which hold an opposing diamond each. A thin sheet ( $\sim 200\mu\text{m}$ ) of metal such as stainless steel or tungsten is then indented (to  $\sim 100\mu\text{m}$ ) by the diamonds to create a well. A hole of  $\sim 300\mu\text{m}$  diameter is drilled directly through the centre of the well to create the sample chamber. The gasket is placed flush onto the diamond culet and the sample is loaded into the chamber with a ruby and pressure-transmitting medium to ensure that the pressure is applied hydrostatically and that it can be measured. The pressure inside the sample chamber is increased by forcing the diamond faces closer together and is achieved by tightening the three Allen screws of the steel plates. The sample within the chamber (be it single crystal or powder) is surrounded by a Pressure Transmitting Medium (PTM) which maintains hydrostatic pressure within the chamber when the opposing diamonds push together. The PTM also acts as a safe guard against the diamonds touching and breaking under such impressive force. It is

## Lauren Connor

possible to load other media into the chamber. The work carried out in this thesis uses only liquid petroleum ether as a PTM. Liquid media are simply pipetted into the chamber. To enable the determination of pressure inside the chamber a ruby chip is loaded with the sample and the PTM. Figure 2.4 displays ruby spheres surrounding a loaded crystal.



**Figure 2.4 View of the sample chamber of a DAC containing a single crystal and many small ruby spheres.**

The fluorescence of ruby is dependent on the pressure the ruby is subjected to. Rubies of ambient pressure act as the calibrant in the well-established ruby fluorescence method which uses a Raman spectrometer for measurement.<sup>93</sup> A calculation using the fluorescence of the ruby inside the cell and the fluorescence of the ambient ruby provides the pressure reading.

### 2.2.1. Collection and issues with high pressure experiments

The high-pressure diffraction experiments described in this thesis were either carried out on a Bruker APEX-II diffractometer with a CCD detector, or on station I19 using EH2 with a Newport 4-circle goniometer. In both cases a modified data collection strategy is required to account for the shading of the body of the cell and to optimise data collection times. Modified scan runs ensure the angle between the incident beam and the cell axis is less than 40°. One of the problems with high pressure crystallography is the completeness of the data due to the cell half opening angle which can affect the processing of the data. Firstly, in order to index the sample, the harvested reflections need to be carefully selected to eliminate powder rings diffracted from the tungsten gasket and diamond reflections which are very intense compared with the diffraction pattern of the sample.

During indexing care must be taken to eliminate areas of the frames collected which are shaded out by the bulk of the DAC. Reflections which are poorly measured due to the shading of the DAC must be omitted. Reflections which appear behind a strong

## Lauren Connor

diamond reflection are also omitted but can cause problems during the structural solution stage, if present in the reduced data. The program SHADE automatically omits reflections with a poor profile correlation coefficient that are within  $2^\circ$  from the cell opening angle (i.e. between  $38-42^\circ$ ).

With well diffracting, high-symmetry systems the data quality obtained from high-pressure data can be comparable to data obtained under ambient conditions. Cases where data quality still proves challenging are low-symmetry systems and weakly diffracting samples i.e. organic materials with only carbon, nitrogen, oxygen and hydrogen.

### 2.2.2. Synchrotron radiation

Due to the casing and restrictions of the diamond anvil cell data collection in house may be an issue for challenging crystals. The use of synchrotron radiation with shorter wavelengths can be employed.

A synchrotron creates extremely bright light which is approximately a million times brighter than sunlight. A synchrotron is an extremely powerful source of x-rays. It uses large, powerful 'bending' magnets and radio frequency waves to accelerate and direct the electrons. The electrons are generated in the electron gun, they are then accelerated in a series of three structures. Firstly, the magnets and radio frequency waves accelerate negatively charged electrons along a stainless steel tube (the LINAC/ linear accelerator). The electrons are accelerated in the booster ring to almost the speed of light in extremely high vacuum. The electron beam is then injected into the storage ring. The storage ring is a polygon of straight edges and powerful bending magnets called undulators. An undulators is a periodic array of dipole magnets which are used to 'wobble' the electrons around the ring. When the path of the high speed, high energy electrons is deflected or 'bent' by the magnets inside the storage ring, an extremely brilliant light is produced and directed into the beamline for use at a specific wavelength.

One of the benefits of synchrotron radiation in high pressure crystallography is the improved data completeness. The DACs opening angle of  $40^\circ$  restricts the data that can be collected however experiments done at the synchrotron compresses the diffraction pattern into a smaller volume of reciprocal space due to the use of shorter wavelengths. A collection in-house may take 18 hours, yet only 30 minutes at a synchrotron beamline. The increased incident flux of synchrotron radiation enables

## Lauren Connor

faster collections of samples which are weakly diffracting. This is especially beneficial for high pressure as absorption from the diamonds and backing plates is a problem.

### 2.3. Pixel

The semi-classical density sums (Pixel) method was developed by Gavezotti<sup>94</sup> in 2005 and can be used to calculate molecule to molecule interaction energies. Interactions can be grouped into electrostatic forces or non-bonding van der Waals forces. Pixel was developed as *in silico* approach to quantify molecule to molecule interactions that are non-bonding. Non-bonded interactions occur due to a balance between attractive and repulsive terms.

**Table 2.2 Types of intermolecular interaction and their dependence on intermolecular distance, r.**

Force	Cause	Dependence on r
Coulombic	Charge- charge interaction	$r^{-1}$
Polarisation	Permanent dipole-dipole	$r^{-3}$
Dispersive	Induced dipole-dipole	$r^{-6}$
Repulsive	Repulsion between atom spins	0 (overlap of electron cloud)

Pixel is a useful tool to gain insight into interactions that may be stabilising in a crystal structure.

#### 2.3.1. The Pixel method

The method quantifies the total energy in a lattice or of an interaction between two molecules and decomposes the energy into four components; coulombic, polarisation, dispersion and repulsion. Firstly, the hydrogen, X...H distances are normalised to 1.08 Å for C-H and 1.00Å for O-H and N-H. Pixel then takes the geometry of the molecule from the observed crystal structure but does not optimise it. The electron densities for the molecule are then calculated in an external quantum mechanical programme such as Gaussian<sup>95</sup>. A standard MP2/6-31G\*\* level of theory calculation is performed in Gaussian. The output from Gaussian is an electron map for each atom on a 3D grid of step size 0.08-0.1 Å. The electron density map output by Gaussian may contain over 10<sup>6</sup> pixels which is far too large so the map is modified

## Lauren Connor

in three stages to reduce computational expense. Electron density files produced by Gaussian are condensed into super-pixels. The size of the super pixels is a multiple of the original pixels and is dependent on the condensation level. The default size is a super pixel of 4 x 4 x 4 original pixels however a condensation value of 3 can be used for a more accurate calculation or 6 for a faster but less accurate calculation. Following the condensation step, the pixels are screened by charge. Pixels with a charge below a certain threshold are removed (default to  $10^{-6}$  electrons). This does not significantly affect the electron density but reduces computational expense.

Next, Pixel generates the required molecules, a cluster is required for a lattice calculation whereas a dimer calculation only requires the two molecules. In lattice calculations, the space group symmetry is utilised to produce the molecules around a central reference molecule, until a crystal cut off radius (default is 18 Å). The intermolecular energy between the central reference molecule and the others in the cluster are determined by pair-wise evaluation of pixel to pixel energies between symmetry related molecules. The energies are split into the four main types.

Pixel is a fast method to quantify and understand molecule to molecule interactions and its decomposition to the four energetic terms is very beneficial. Pixel calculations on a full compression study can take as little as a day on a standard desktop computer which is relatively short compared to some *ab initio* methods. Maschio *et al.*<sup>96</sup> tested the Pixel method against dispersion corrected periodic density functional theory (DFT-D) calculations for sixty organic compounds and found close agreement. The Pixel lattice energies were compared to DFT-D calculated lattice energies at the B3LYP-D\*/6-31G\*\* level and the results show that Pixel was in good agreement (correlation coefficient of 0.95) with the higher level computational methods. It has many applications including understanding how pressure affects crystal structures<sup>24,97</sup> as well as justifying crystal morphology predictions, as is shown in this thesis.

There are limitations of Pixel. The user should be wary of its treatment of strongly hydrogen bonded systems. Pixels rigid treatment of electron densities mean that the method does not account for electronic rearrangement that could occur after molecular contact. This is a problem mainly for strongly hydrogen bonded systems where charge transfer may occur.

## References

1. Dunitz, J. D. Phase transitions in molecular crystals from a chemical viewpoint. *Pure Appl. Chem.* **63**, 177–185 (1991).
2. Thun, J., Seyfarth, L., Senker, J., Dinnebier, R. E. & Brey, J. Polymorphism in benzamide: Solving a 175-year-old riddle. *Angew. Chemie - Int. Ed.* **46**, 6729–6731 (2007).
3. Thun, J., Seyfarth, L., Butterhof, C., Senker, J., Dinnebier, R. E. & Brey, J. Wohler and Liebig Revisited: 176 Years of Polymorphism in Benzamide - and the Story Still Continues! *Cryst. Growth Des.* **9**, 3–9 (2009).
4. Dittrich, B., Bergmann, J., Roloff, P., Crystals, G. R. - & 2018, undefined. New Polymorphs of the Phase-Change Material Sodium Acetate. *Mdpi.Com* (2018). doi:10.3390/cryst8050213
5. Madusanka, N., Eddleston, M. D., Arhangelskis, M. & Jones, W. Polymorphs, hydrates and solvates of a co-crystal of caffeine with anthranilic acid. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **70**, 72–80 (2014).
6. Threlfall, T. Crystallisation of Polymorphs: Thermodynamic Insight into the Role of Solvent. *Org. Process Res. & Dev.* **4**, 384–390 (2000).
7. Oswald, I. D. H. & Urquhart, A. J. Polymorphism and polymerisation of acrylic and methacrylic acid at high pressure. *CrystEngComm* **13**, 4503–4507 (2011).
8. Ayala, A. P. Polymorphism in drugs investigated by low wavenumber Raman scattering. *Vib. Spectrosc.* **45**, 112–116 (2007).
9. Pallipurath, A. R., Skelton, J. M., Warren, M. R., Kamali, N., McArdle, P. & Erxleben, A. Sulfamerazine: Understanding the influence of slip planes in the polymorphic phase transformation through x-ray crystallographic studies and ab initio lattice dynamics. *Mol. Pharm.* **12**, 3735–3748 (2015).
10. Okumura, T., Ishida, M., Takayama, K. & Otsuka, M. Polymorphic Transformation of Indomethacin under High Pressures. *J. Pharm. Sci.* **95**, 689–700 (2005).
11. Johnstone, R. D. L., Lennie, A. R., Parker, S. F., Parsons, S., Pidcock, E., Richardson, P. R., Warren, J. E. & Wood, P. A. High-pressure polymorphism in salicylamide. *CrystEngComm* **12**, 1065–1078 (2010).
12. Miller, J. M., Collman, B. M., Greene, L. R., Grant, D. J. W. & Blackburn, A. C. Identifying the stable polymorph early in the drug discovery-development process. *Pharm. Dev. Technol.* **10**, 291–297 (2005).
13. Bauer, J., Spanton, S., Henry, R., Quick, J., Dziki, W., Porter, W. & Morris, J. Ritonavir: an extraordinary example of conformational polymorphism. *Pharm. Res.* **18**, 859–866 (2001).
14. Carmo, H., Brulport, M., Hermes, M., Oesch, F., Silva, R., Ferreira, L. M., Branco, P. S., Boer, D. de, Remião, F., Carvalho, F., Schön, M. R., Krebsfaenger, N., Doehmer, J., Bastos, M. de L. & Hengstler, J. G. Influence of CYP2D6 polymorphism on 3,4-methylenedioxymethamphetamine ('Ecstasy') cytotoxicity. *Pharmacogenet. Genomics* **16**, 789–799 (2006).
15. Patyk, E., Podsiadło, M. & Katrusiak, A. CH···N Bonds and Dynamics in Isostructural Pyrimidine Polymorphs. *Cryst. Growth Des.* **15**, 4039–4044 (2015).
16. Horosanskaia, E., Seidel-Morgenstern, A. & Lorenz, H. Investigation of drug polymorphism: Case of artemisinin. *Thermochim. Acta* **578**, 74–81 (2014).

17. Rietveld, I. B. & Ceolin, R. Rotigotine: Unexpected Polymorphism with Predictable Overall Monotropic Behavior. *J. Pharm. Sci.* **104**, 4117–4122 (2015).
18. McCrone, W. Application of Fusion Methods in Chemical Microscopy. (1957).
19. Groom, C. R., Bruno, I. J., Lightfoot, M. P. & Ward, S. C. The Cambridge Structural Database. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **B72**, 171–179 (2016).
20. López-Mejías, V., Kampf, J. W. & Matzger, A. J. Nonamorphism in flufenamic acid and a new record for a polymorphic compound with solved structures. *J. Am. Chem. Soc.* **134**, 9872–9875 (2012).
21. Olejniczak, A., Katrusiak, A. & Szafranski, M. Ten polymorphs of NH<sup>+</sup>...N hydrogen-bonded 1,4-diazabicyclo[2.2.2]octane complexes: Supramolecular origin of giant anisotropic dielectric response in polymorph v. *Cryst. Growth Des.* **10**, 3537–3546 (2010).
22. Pogoda, D., Janczak, J. & Videnova-Adrabinska, V. New polymorphs of an old drug: Conformational and synthon polymorphism of 5-nitrofurazone. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **72**, 263–273 (2016).
23. Bond, A. D., Haynes, D. A., Pask, C. M. & Rawson, J. M. Concomitant polymorphs: structural studies on the trimorphic dithiadiazolyl radical, CICNSSN. *J. Chem. Soc. Dalton Trans.* 2522 (2002). doi:10.1039/b110922g
24. Moggach, S. A., Parsons, S. & Wood, P. A. High-pressure polymorphism in amino acids. *Crystallography Reviews* **14**, (2008).
25. Trask, A. V, Samuel, W. D. & Jones, W. Solvent-drop grinding: green polymorph control of cocrystallisation. *Chem. Commun.* 890–891 (2004).
26. Zakharov, B. a., Losev, E. a. & Boldyreva, E. V. Polymorphism of “glycine–glutaric acid” co-crystals: the same phase at low temperatures and high pressures. *CrystEngComm* **15**, 1693 (2013).
27. Thomas, L. H., Wales, C., Zhao, L. & Wilson, C. C. Paracetamol form II: An elusive polymorph through facile multicomponent crystallization routes. *Cryst. Growth Des.* **11**, 1450–1452 (2011).
28. Loisel, C., Keller, G., Lecq, G., Bourgaux, C. & Ollivon, M. Phase transitions and polymorphism of cocoa butter. *JAOCS, J. Am. Oil Chem. Soc.* **75**, 425–439 (1998).
29. Light, D. W. & Lexchin, J. R. Pharmaceutical research and development: what do we get for all that money? *Bmj* **345**, e4348–e4348 (2012).
30. Desiraju, G. R., Vittal, J. J. & Ramanan, A. *Crystal engineering: a textbook. Crystallography Reviews* (2011). doi:10.1080/0889311X.2012.724405
31. Aakeroy, C. B. & Seddon, K. R. The hydrogen bond and crystal engineering. *Chem. Soc. Rev.* **22**, 397–407 (1993).
32. Mukherjee, A., Tothadi, S. & Desiraju, G. R. Halogen bonds in crystal engineering: Like hydrogen bonds yet different. *Acc. Chem. Res.* **47**, 2514–2524 (2014).
33. Yao, Z. F., Wang, J. Y. & Pei, J. Control of π-π Stacking via Crystal Engineering in Organic Conjugated Small Molecule Crystals. *Cryst. Growth Des.* **18**, 7–15 (2018).

34. Majumder, J., Deb, J., Das, M. R., Jana, S. S. & Dastidar, P. Designing a simple organic salt-based supramolecular topical gel capable of displaying in vivo self-delivery application. *Chem. Commun.* **50**, 1671–1674 (2014).
35. Thompson, C. Investigating the fundamentals of drug crystal growth using Atomic Force Microscopy. (The University of Nottingham, 2003).
36. Mirza, S., Miroshnyk, I., Heinämäki, J., Antikainen, O., Rantanen, J., Vuorela, P., Vuorela, H. & Yliruusi, J. Crystal Morphology Engineering of Pharmaceutical Solids: Tableting Performance Enhancement. *AAPS PharmSciTech* **10**, 113–119 (2009).
37. Agnew, L. R., Cruickshank, D. L., McGlone, T. & Wilson, C. C. Controlled production of the elusive metastable form II of acetaminophen (paracetamol): A fully scalable templating approach in a cooling environment. *Chem. Commun.* **52**, 7368–7371 (2016).
38. Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., Jackel, L. D., Monfort, M., Muller, U., Zhang, J., Zhang, X., Zhao, J. & Zieba, K. End to End Learning for Self-Driving Cars. 1–9 (2016). doi:10.1109/IJCNN.2005.1556090
39. Graves, A., Mohamed, A.-R. & Hinton, G. Speech recognition with deep recurrent neural networks. *2013 IEEE Int. Conf. Acoust. Speech Signal Process.* 6645–6649 (2013). doi:10.1109/ICASSP.2013.6638947
40. Gao, Y., Ma, J. & Yuille, A. L. Semi-Supervised Sparse Representation Based Classification for Face Recognition with Insufficient Labeled Samples. *IEEE Trans. Image Process.* **26**, 2545–2560 (2017).
41. Kumar, S., Gao, X., Welch, I. & Mansoori, M. A Machine Learning Based Web Spam Filtering Approach. *2016 IEEE 30th Int. Conf. Adv. Inf. Netw. Appl.* 973–980 (2016). doi:10.1109/AINA.2016.177
42. Samuel, A. L. Some Studies in Machine Learning Using the Game of Checkers. *IBM J. Res. Dev.* **3**, 210–229 (1959).
43. Obermeyer, Z. & Ezekiel, J. E. Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. *N. Engl. J. Med.* **375**, 1216–1219 (2016).
44. Wicker, J. G. P., Crowley, L. M., Robshaw, O., Little, E. J., Stokes, S. P., Cooper, R. I. & Lawrence, S. E. Will they co-crystallize? *CrystEngComm* **19**, 5336–5340 (2017).
45. Raccuglia, P., Elbert, K. C., Adler, P. D. F., Falk, C., Wenny, M. B., Mollo, A., Zeller, M., Friedler, S. A., Schrier, J. & Norquist, A. J. Machine-learning-assisted materials discovery using failed experiments. *Nature* **533**, 73–76 (2016).
46. Johnston, A., Johnston, B. F., Kennedy, A. R. & Florence, A. J. Targeted crystallisation of novel carbamazepine solvates based on a retrospective Random Forest classification. *CrystEngComm* **10**, 23–25 (2008).
47. Crystals, M. & Wright, B. J. D. Molecular Crystals and Organic Solids Zeolites. *Angew. Chem. Int. Ed. Engl.* 1987–1988 (1987).
48. Aitipamula, S., Banerjee, R., Bansal, A. K., Biradha, K., Cheney, M. L., Choudhury, A. R., Desiraju, G. R., Dikundwar, A. G., Dubey, R., Duggirala, N., Ghogale, P. P., Ghosh, S., Goswami, P. K., Goud, N. R., Jetti, R. R. K. R., Karpinski, P., Kaushik, P., Kumar, D., Kumar, V., Moulton, B., Mukherjee, A., Mukherjee, G., Myerson, A. S., Puri, V., Ramanan, A., Rajamannar, T., Reddy, C. M., Rodriguez-Hornedo, N., Rogers, R. D., Row, T. N. G., Sanphui, P., Shan, N., Shete, G., Singh, A., Sun, C. C., Swift, J. A., Thaimattam, R., Thakur, T. S., Kumar Thaper, R., Thomas, S. P., Tothadi, S., Vangala, V. R., Variankaval, N.,

Vishweshwar, P., Weyna, D. R. & Zaworotko, M. J. Polymorphs, salts, and cocrystals: What's in a name? *Cryst. Growth Des.* **12**, 2147–2152 (2012).

49. Elder, D. P., Holm, R. & De Diego, H. L. Use of pharmaceutical salts and cocrystals to address the issue of poor solubility. *Int. J. Pharm.* **453**, 88–100 (2013).

50. Fasulo, M. E. & Desper, J. articles Cocrystal or Salt : Does It Really Matter ? *Mol. Pharm.* **4**, 317–322 (2007).

51. Yu, L. & Stingelin, N. Controlling the solid-state microstructure of organic semiconducting materials by molecular compound formation. *J. Org. Semicond.* **1**, 16–21 (2013).

52. Basavoju, S., Boström, D. & Velaga, S. P. Indomethacin-saccharin cocrystal: Design, synthesis and preliminary pharmaceutical characterization. *Pharm. Res.* **25**, 530–541 (2008).

53. Karki, S., Friiċ, T., Fábíán, L. & Jones, W. New solid forms of artemisinin obtained through cocrystallisation. *CrystEngComm* **12**, 4038–4041 (2010).

54. Oswald, I. D. H. & Pulham, C. R. Co-crystallisation at high pressure—an additional tool for the preparation and study of co-crystals. *CrystEngComm* **10**, 1114 (2008).

55. Vishweshwar, P., A. McMahon, J., A. Bis, J. & J. Zaworotko, M. Pharmaceutical Cocrystals. *Journal of pharmaceutical sciences* **95**, 499–516 (2006).

56. Sandhu, B., Sinha, A. S., Desper, J. & Aakeröy, C. B. Modulating the physical properties of solid forms of urea using co-crystallization technology. *Chem. Commun.* **54**, 4657–4660 (2018).

57. Pagire, S. K., Jadav, N., Vangala, V. R., Whiteside, B. & Paradkar, A. Thermodynamic Investigation of Carbamazepine-Saccharin Co-Crystal Polymorphs. *J. Pharm. Sci.* **106**, 2009–2014 (2017).

58. Childs, S. L., Kandi, P. & Lingireddy, S. R. Formulation of a danazol cocrystal with controlled supersaturation plays an essential role in improving bioavailability. *Mol. Pharm.* **10**, 3112–3127 (2013).

59. McNamara, D. P., Childs, S. L., Giordano, J., Iarriccio, A., Cassidy, J., Shet, M. S., Mannion, R., O'Donnell, E. & Park, A. Use of a glutaric acid cocrystal to improve oral bioavailability of a low solubility API. *Pharm. Res.* **23**, 1888–1897 (2006).

60. Eddleston, M. D., Sivachelvam, S. & Jones, W. Screening for polymorphs of cocrystals: a case study. *CrystEngComm* **15**, 175–181 (2013).

61. Li, Z. & Matzger, A. J. Influence of Coformer Stoichiometric Ratio on Pharmaceutical Cocrystal Dissolution: Three Cocrystals of Carbamazepine/4-Aminobenzoic Acid. *Mol. Pharm.* **13**, 990–995 (2016).

62. Sokal, A., Pindelska, E., Szeleszczuk, L. & Kolodziejski, W. Pharmaceutical properties of two ethenzamide-gentisic acid cocrystal polymorphs: Drug release profiles, spectroscopic studies and theoretical calculations. *Int. J. Pharm.* **522**, 80–89 (2017).

63. Almarsson, Ö. & Zaworotko, M. J. Crystal engineering of the composition of pharmaceutical phases. Do pharmaceutical co-crystals represent a new path to improved medicines? *Chem. Commun.* 1889–1896 (2004). doi:10.1039/b402150a

64. Babu, N. J., Reddy, L. S. & Nangia, A. Amide N-oxide heterosynthon and amide dimer homosynthon in cocrystals of carboxamide drugs and pyridine N-oxides. *Mol. Pharm.* **4**, 417–434 (2007).

65. McMahon, J. A., Bis, J. A., Vishweshwar, P., Shattock, T. R., McLaughlin, O. L. & Zaworotko, M. J. Crystal engineering of the composition of pharmaceutical phases. Primary amide supramolecular heterosynthons and their role in the design of pharmaceutical co-crystals. *Zeitschrift für Krist.* **220**, 340–350 (2005).
66. Cruz-Cabeza, A. J., Reutzel-Edens, S. M. & Bernstein, J. Facts and fictions about polymorphism. *Chem. Soc. Rev.* **44**, 8619–8635 (2015).
67. Ward, H. K. An X-Ray Study of the Structure of Liquid Benzene, Cyclohexane and Their Mixtures. *J. Chem. Phys.* **2**, 153–159 (1934).
68. Olmstead, M. M., Power, P. P. & Shoner, S. C. Three-Coordinate Iron Complexes: X-ray Structural Characterization of the Amide-Bridged Dimers  $[\text{Fe}(\text{NR}_2)_2]_2$  ( $\text{R} = \text{SiMe}_3, \text{C}_6\text{H}_5$ ) and the Adduct  $\text{Fe}[\text{N}(\text{SiMe}_3)_2]_2(\text{THF})$  and Determination of the Association Energy of the Monomer  $\text{Fe}\{\text{N}(\text{SiMe}_3)_2\}_2$  in Solution. *Inorg. Chem.* **30**, 2547–2551 (1991).
69. Shi, Y. A glimpse of structural biology through X-ray crystallography. *Cell* **159**, 995–1014 (2014).
70. Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N. & Noller, H. F. X-ray Crystal Structures of 70 S Ribosome Functional Complexes. *Science (80- )*. **285**, 2095–2104 (1999).
71. Jacobs, R. B. X-Ray Diffraction of Substances under High Pressures. *Phys. Rev.* **54**, 325–331 (1938).
72. Hobday, C. L. Effect of guest uptake and high pressure on zn- and zr- metal-organic frameworks. *PQDT - UK & Ireland* (2017).
73. Jamieson, J. C., Lawson, A. W. & Nachtrieb, N. D. New device for obtaining X-ray diffraction patterns from substances exposed to high pressure. *Rev. Sci. Instrum.* **30**, 1016–1019 (1959).
74. Bridgman, P. The resistance of 72 elements, alloys and compounds to 100,000 Kg/Cm<sup>2</sup>. *Proc. Am. Acad. Arts Sci.* **81**, 165 (1952).
75. Boehler, R. High-pressure experiments and the phase diagram of lower mantle and core materials. *Rev. Geophys.* **38**, 221–245 (2000).
76. Winter, R. & Köhling, R. Static and time-resolved synchrotron small-angle x-ray scattering studies of lyotropic lipid mesophases, model biomembranes and proteins in solution. *J. Phys. Condens. Matter* **16**, 160–184 (2004).
77. David, R., William, I. F., Alistair, J., Lennie, A. R., Parsons, S., Pulham, C. R. & Warren, J. E. High-Pressure Studies of Pharmaceuticals : An Exploration of the Behavior of Piracetam 2007. (2007).
78. Boldyreva, E. V. Multicomponent organic crystals at high pressure. *Zeitschrift für Krist.* **229**, 236–245 (2014).
79. Nagao, H., Nakamura, K. G., Kondo, K., Ozaki, N., Takamatsu, K., Ono, T., Shiota, T., Ichinose, D., Tanaka, K. A., Wakabayashi, K., Okada, K., Yoshida, M., Nakai, M., Nagai, K., Shigemori, K., Sakaiya, T. & Otani, K. Hugoniot measurement of diamond under laser shock compression up to 2 TPa. *Phys. Plasmas* **13**, 0–5 (2006).
80. Fabbiani, F. P. A., Buth, G., Levendis, D. C. & Cruz-Cabeza, A. J. Pharmaceutical hydrates under ambient conditions from high-pressure seeds: A case study of GABA monohydrate. *Chem. Commun.* **50**, 1817–1819 (2014).

81. Oswald, I. D. H. & Pulham, C. R. Co-crystallisation at high pressure—an additional tool for the preparation and study of co-crystals. *CrystEngComm* **10**, 1114 (2008).
82. Boldyreva, E. V. High-pressure diffraction studies of molecular organic solids . A personal view. *Acta Crystallogr. A* **64**, 218–231 (2008).
83. Kolesov, B. a, Mikhailenko, M. a & Boldyreva, E. V. Dynamics of the intermolecular hydrogen bonds in the polymorphs of paracetamol in relation to crystal packing and conformational transitions: a variable-temperature polarized Raman spectroscopy study. *Phys. Chem. Chem. Phys.* **13**, 14243–14253 (2011).
84. Boldyreva, E. V. High-pressure-induced structural changes in molecular crystals preserving the space group symmetry: anisotropic distortion/isosymmetric polymorphism. *Cryst. Eng.* **6**, 235–254 (2003).
85. Connor, L. E., Delori, A., Hutchison, I. B., Nic Daeid, N., Sutcliffe, O. B. & Oswald, I. D. H. The ecstasy and the agony; compression studies of 3,4-methylenedioxymethamphetamine (MDMA). *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **71**, 3–9 (2015).
86. Fabbiani, F. P. A., Allan, D. R., David, W. I. F., Davidson, A. J., Lennie, A. R., Parsons, S., Pulham, C. R. & Warren, J. E. High-pressure studies of pharmaceuticals: An exploration of the behavior of piracetam. *Cryst. Growth Des.* **7**, 1115–1124 (2007).
87. Qiao, Y., Wang, K., Yuan, H., Yang, K. & Zou, B. Negative linear compressibility in organic mineral ammonium oxalate monohydrate with hydrogen bonding wine-rack motifs. *J. Phys. Chem. Lett.* **6**, 2755–2760 (2015).
88. Harty, E. L., Ha, A. R., Warren, M. R., Thompson, A. L., Allan, D. R., Goodwin, A. L. & Funnell, N. P. Reversible piezochromism in a molecular wine-rack. *Chem. Commun.* **51**, 10608–10611 (2015).
89. Moggach, S. A., Bennett, T. D. & Cheetham, A. K. The effect of pressure on ZIF-8: Increasing pore size with pressure and the formation of a high-pressure phase at 1.47 GPa. *Angew. Chemie - Int. Ed.* **48**, 7087–7089 (2009).
90. Joiris, E., Di Martino, P., Berneron, C., Guyot-Hermann, A. M. & Guyot, J. C. Compression behavior of orthorhombic paracetamol. *Pharm. Res.* **15**, 1122–1130 (1998).
91. Boldyreva, E. V., Shakhtshneider, T. P., Ahsbahs, H., Sowa, H. & Uchtmann, H. Effect of high pressure on the polymorphs of paracetamol. *J. Therm. Anal. Calorim.* **68**, 437–452 (2002).
92. Macrae, J., Edgington, C. F., McCabe, P. ., Pidcock, P., Shields, E., Taylor, G. ., Towler, R. & Van de Streek, J. MERCURY. *J. Appl. Crystallogr.* **39**, 453–457 (2006).
93. Piermarini, G. J., Block, S., Barnett, J. D. & Forman, R. a. Calibration of the Pressure Dependence of the R//1 Ruby Fluorescence Line To 195 Kbar. *J. Appl. Phys.* **46**, 2774–2780 (1975).
94. Gavezzotti, A. Calculation of Intermolecular Interaction Energies by Direct Numerical Integration over Electron Densities. 2. An Improved Polarization Model and the Evaluation of Dispersion and Repulsion Energies. *J. Phys. Chem. B* **107**, 2344–2353 (2003).
95. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E. ., Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V. . M., B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H., P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M. ., Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M. . N., T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., J. ., Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd,

## Lauren Connor

J. J.; Brothers, E. . K., K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K. ., Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M. . R., N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V. ., Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O. ., Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L. ., Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P. ., Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O. . & Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09. (2009).

96. Maschio, L., Civalleri, B., Ugliengo, P. & Gavezzotti, A. Intermolecular Interaction Energies in Molecular Crystals: Comparison and Agreement of Localized Møller-Plesset 2, Dispersion-Corrected Density Functional, and Classical Empirical Two-Body Calculations. *J. Phys. Chem.* **115**, 11179–11186 (2011).

97. Cai, W., Marciniak, J., Andrzejewski, M. & Katrusiak, A. Pressure effect on d,l-Mandelic acid racemate crystallization. *J. Phys. Chem. C* **117**, 7279–7285 (2013).

***-Chapter Three-***

***High pressure structural investigation of indomethacin and  
saccharin***

### 3. High pressure structural investigation of indomethacin and saccharin

#### 3.1. Synopsis

The co-crystalline structure of the non-steroidal, anti-inflammatory indomethacin with the non-toxic, Generally Regarded As Safe (GRAS) sweetener component saccharin was investigated up to 6.33GPa using a Diamond Anvil Cell (DAC). Single crystal X-ray diffraction (SCXRD) measurements show that the co-crystal remains in the same triclinic, P-1, phase throughout the compression with a significant reduction in void space (155.69 to 55.61Å<sup>3</sup>). Information on the response of different types of intermolecular interactions has been enabled by the use of a co-crystal. We have rationalised that the 'intermediate' length of the saccharin amide dimer in the co-crystal is caused by the dimer sitting in a 'pocket' surrounded by the indomethacin framework. This framework reduces the effects of molecular packing on the dimer allowing for an ideal hydrogen bonding geometry. Calculations are underway to support this conclusion.

#### 3.2. Introduction

Co-crystallisation is a branch of solid state modification that has been of growing interest in the energetic material, ferroelectric and pharmaceutical industry in recent years.<sup>1-6</sup> Much akin to its fellow branch of solid state modification, salt formation, co-crystallisation is the incorporation of two or more species in the crystalline lattice. Co-crystallisation has a much greater scope due to the use of neutral co-formers and the lack of requirement for an ionisable group.<sup>7</sup> Crystal engineering modifications are sought to explore a few avenues of interest but one common interest spanning across all industries is the alteration of the physicochemical properties.<sup>8</sup> A change in the physicochemical properties could also mean a change in the mechanochemical properties and functionality of the singular component material on the larger, potentially tangible, product scale. To fundamental scientists, co-crystallisation offers something of value on the molecular scale; the potential to investigate intermolecular interactions which may not have occurred within the lattice of a singular component. By adding another component to the crystalline lattice there are new synthons to analyse which open up a wealth of new molecular opportunities. In understanding these competing interactions we can begin to understand how the molecules self-assemble into the co-crystalline form. Co-crystals may contain components that are sterically largely different and understanding how these lattices are held together is

usually done by looking at the intermolecular interactions. Whilst in many systems hydrogen bonding interactions are the linkage of the two components in the lattice, weaker dispersive forces play a significant role in the overall energy of the structure. By understanding the basic fundamentals of the structure, the properties can be explained and possibly controlled. However, even when we have understood a system invariably polymorphism arises. A recent study by Cruz-Cabeza *et al.*<sup>9</sup> has shown through database mining that polymorphism in co-crystalline phases are equally as common as with pure components. The initial concept of co-crystallisation was to try to inhibit or reduce the propensity for polymorphism but this has proved to be not the case.<sup>10-12</sup> Considering this, we have set about to use high pressure to evaluate effects of pressure on co-crystalline systems. Using pressure we can: i) survey the thermodynamic stability of ambient pressure forms; ii) access to metastable polymorphs; and iii) investigate the response of hydrogen bonds with respect to pressure.<sup>9,13</sup>

Indomethacin is a Non-Steroidal Anti-Inflammatory (NSAID) that is known for its ability to reduce fever, pain, stiffness and swelling and can be prescribed for migraines. There are 10 different multi-component forms found in the Cambridge Structural Database (CSD)<sup>14</sup> In its pure form Indomethacin has been well-characterised under ambient and high pressure conditions. Under ambient conditions there are three polymorphs  $\alpha$ - $\gamma$ ; a fourth  $\delta$ -form is a hydrate. At ambient pressure, the  $\gamma$ -form is the most stable form which was demonstrated by Okumura *et al.* where they slurried the  $\alpha$ -form of indomethacin in ethanol over the course of four days and observed that transformed to the  $\gamma$ -form. At high pressure, however, the behaviour of the slurry changes. At a modest pressure of 0.4 GPa, the transformation of the  $\alpha$ -form to the  $\gamma$ -form is reversed indicating that whilst metastable at ambient pressure, the  $\alpha$ -form is more stable at higher pressures.<sup>15</sup> A compression of the  $\gamma$ -form to 0.4 GPa without ethanol solvent did not reveal any polymorphic transformation which reveals the significance of the solvent in mediating these pressure-induced phase transitions.<sup>16</sup>

Given the rich behaviour of Indomethacin, the co-crystal of indomethacin and saccharin (INSA) was chosen as a compound to investigate under high pressure conditions. It is a well-characterised system known to have altered physicochemical properties compared with indomethacin itself such as a higher solubility and bioavailability.<sup>17,18</sup> There is no indication of any polymorphism in the previous studies. In the co-crystal, the indomethacin and saccharin molecules hydrogen bond with

## Lauren Connor

themselves allowing both types of hydrogen bond (carboxylic acid dimer & amide dimer) to be assessed under a range of pressure conditions.

### 3.3. Experimental

#### 3.3.1. General procedure

A Merrill-Bassett DAC with a half-opening angle of  $40^\circ$  was used for the experiments. The opposing diamonds possessed  $600\mu\text{m}$  culets and sat in Boehler-Almax designed tungsten carbide backing seats. The DAC sample chamber was a  $300\mu\text{m}$  diameter hole drilled into a tungsten  $10\text{mm} \times 10\text{mm}$  square gasket pre-indented to a thickness of  $90\mu\text{m}$  (Almax EasyLab  $\mu$ Driller). A ruby sphere of approximately  $15\mu\text{m}$  diameter was loaded into the chamber along with the sample to measure the pressure *in situ* via the ruby fluorescence technique.<sup>19</sup> A Horiba XplorA Raman spectrometer equipped with a  $532\text{nm}$  laser was used to measure the ruby fluorescence.

#### 3.3.2. Co-crystal formation

A 1:1 stoichiometric co-crystal of indomethacin and saccharin (INSA) was grown from ethyl acetate via small batch evaporative co-crystallisation ( $35.7\text{mg}$  of indomethacin to  $18.32\text{mg}$  of saccharin in  $1\text{ml}$ ). A crystal of dimensions  $0.3\text{mm} \times 0.05\text{mm} \times 0.05\text{mm}$  was used to determine the crystal structure at ambient pressure.

#### 3.3.3. Diamond anvil cell preparation

For the compression study a prismatic co-crystal was loaded into the DAC with petroleum ether (40:60) to allow for hydrostatic compression before the sample chamber was sealed; the initial sealing of the cell introduced a pressure of  $0.55\text{GPa}$ . Single crystal X-ray diffraction data were collected with the DAC in a single orientation throughout the study, attempts to improve on the data completeness by collecting a second orientation and merging the data with the initial dataset failed to provide a significant improvement. Due to a pressure jump from  $0.55\text{GPa}$  to  $3.33\text{GPa}$ , another cell was loaded in the same manner with another crystal. A dataset was taken at  $1.66\text{GPa}$  to fill the gap in the compression study.

#### 3.3.4. Ambient crystal collection, processing and refinement

X-ray diffraction intensities were collected from a sequence of six scans at  $293\text{K}$  using a Bruker APEX II CCD diffractometer with an IncoTec  $I\mu\text{S}$  microsource ( $\text{Mo } K\alpha$

## Lauren Connor

radiation,  $\lambda = 0.71073 \text{ \AA}$ ). Data were indexed and integrated using SAINT as incorporated in the APEX II software.<sup>20</sup> An absorption correction was applied using SADABS.<sup>21</sup> The crystal structure at ambient pressure was refined within *Olex2*<sup>22</sup> ( $F^2$ ) using the atomic coordinates from the CSD. All heavy atoms were anisotropically refined. Hydrogen atoms were placed in calculated positions and constrained to ride their parent atom as per procedure in *Olex2*.

### 3.3.5. Compression study collection, processing and refinement

The high-pressure data were collected and treated in a similar manner to the ambient procedure described above. X-ray diffraction intensities were collected from a sequence of twelve scans at 293K following the procedure described by Dawson *et al.* Data were indexed and integrated using SAINT in conjunction with dynamic masks.<sup>23</sup> An absorption correction was applied in two stages using both SHADE and SADABS, respectively.<sup>21,24</sup> Known coordinates from a crystal structure of INSA found in the CSD were used as a starting model for the refinement. For high-pressure datasets the heavier atoms (S-35 and Cl-1, Figure 3.1) were refined anisotropically and the remainder of the heavy atoms were refined isotropically due to the limited data and large number of atoms. Additionally, distance restraints were applied to the bonds of the high pressure structures of INSA using distances taken from a Mogul geometry search of the Cambridge Structural Database (CSD) as implemented in Materials Mercury.<sup>25</sup> The AFIX66 command was used to constrain the phenyl rings. Hydrogen atoms were placed in calculated positions and constrained to ride their parent atoms. The refined model from each pressure point was used as a starting model for each subsequent model.

### 3.3.6. H-bonding length distribution in CSD

The CSD version 5.39 (March 2018 + 3 updates until May 2018) was used along with Conquest Version 1.21 to search for H-bond distributions.<sup>14</sup> The first search for crystal structures with an intermolecular (carboxylic acid) O-H $\cdots$ O dimer contact length between 2.45 $\text{\AA}$  and 2.75 $\text{\AA}$  produced 4977 refcodes. The second search for crystal structures with an intermolecular (amide) N-H $\cdots$ O dimer contact length between 2.65 $\text{\AA}$  and 3.1 $\text{\AA}$  produced 7600 refcodes. Both searches excluded all errors, polymeric, ion and powder structures and considered only organic molecules with 3D coordinates determined. An additional constraint of no alkali metals was used since original searches contained molecules with these species.

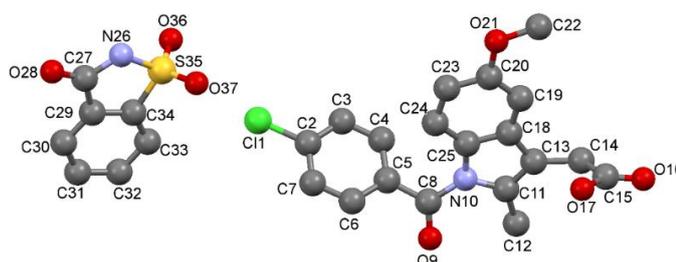
## Lauren Connor

A script was written for CSD Python API to produce a restricted CSD entry list of compounds whose data were collected at high pressure (2454 CSD refcodes). The script only takes into account refcodes with pressure explicitly stated in the experimental conditions and converts the pressure units to GPa. The search parameters previously described were used for consistency with the restriction of the subset of high pressure refcodes. The restriction narrowed the results of the carboxylic acid dimer search to 45 refcodes which equates to 12 compounds and the amide dimer search to 30 refcodes which equates to 5 compounds.

### 3.4. Results and Discussion

#### 3.4.1. Structure

The crystals grown from ethyl acetate were of similar size and consistent crystal habit. The unit cell information was consistent with those found in the Cambridge Structural Database (CSD) for the 1:1 stoichiometric co-crystal of INSA (CSD refcode: UFERED) (Figure 3.1).<sup>14,17</sup> INSA crystallises in triclinic spacegroup  $\bar{P}1$  with one molecule of each component in the asymmetric unit.



**Figure 3.1 The asymmetric unit of INSA with numbering scheme (hydrogen atoms omitted for clarity). The colours of the elements are consistent throughout unless otherwise stated. (carbon, grey; oxygen, red; sulphur, yellow; nitrogen, blue; chlorine, green; hydrogen, white)**

Indomethacin is an indole derivative with a chlorobenzoyl ring, methoxy group and carboxylic acid presenting opportunities for hydrogen bonding with suitable donors. Saccharin, the artificial sweetener, is a dioxobenzothiazol and is considerably smaller than indomethacin. The crystal structure of INSA consists of dimers of each component linked through close interaction between the NH of saccharin and C=O of the indomethacin. Utilising this close interaction, the two components form a chain along the *b*-axis as shown in Figure 3.2. Neighbouring chains run parallel to each

other interlocking through non-directional CH- $\pi$  interactions between the chlorobenzoyl rings (Figure 3.2).

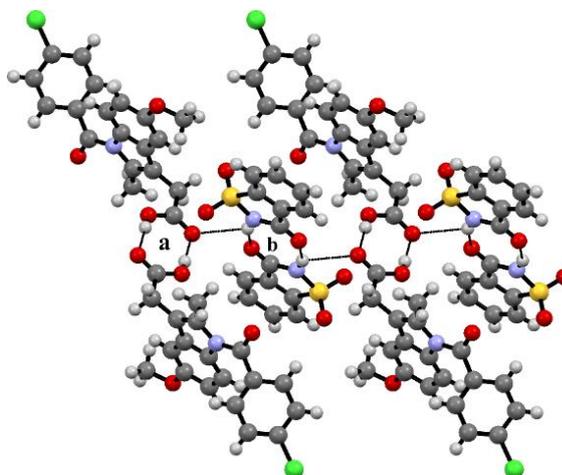


Figure 3.2 A section of the structure of indomethacin and saccharin with hydrogen bonding contacts shown. Depicted a and b indicate a strong carboxylic acid dimer formed by indomethacin molecules and the dimer between saccharin molecules, respectively.

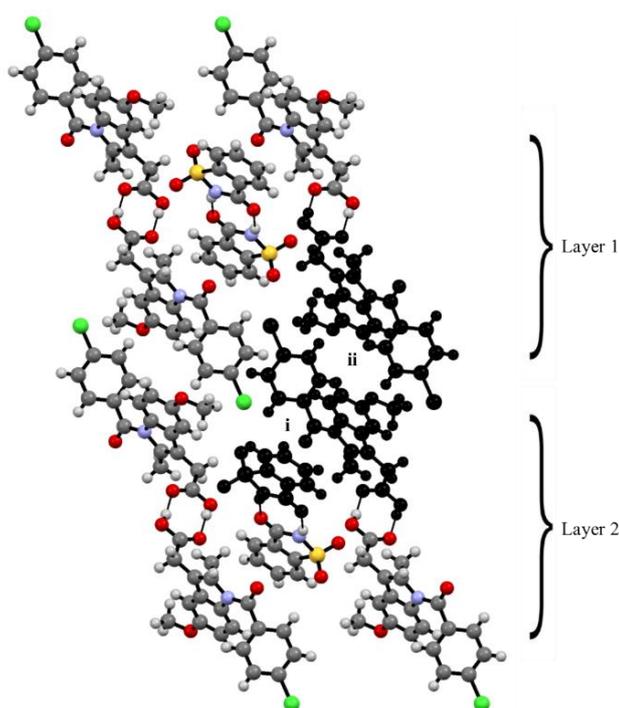


Figure 3.3 Section of two chains running parallel along the b-axis in the ab plane interlocking through a  $\pi/\pi$  interaction between indomethacin and saccharin (i) and a CH- $\pi$  interaction between indomethacin molecules (ii) shown in the blackened molecules.

**Table 3.1 Experimental details from ambient to 1.66 GPa**

For all structures:  $C_{19}H_{16}ClNO_4 \cdot C_7H_5NO_3S$ ,  $M_r = 540.96$ , triclinic,  $\bar{P}1$ ,  $Z = 2$ . Experiments were carried out with Mo  $K\alpha$  radiation using a Bruker SMART APEX2 area detector. H-atom parameters were constrained.

**Table 3.2 Experimental details from ambient to 1.66 GPa**

	Ambient pressure, 293K	0.55 GPa, 297 K	1.66 GPa, 297 K
$a, b, c$ (Å)	7.1314 (2), 10.4809 (2), 16.7109 (4)	7.0133 (6), 10.155 (2), 16.607 (3)	6.9301 (4), 9.9420 (7), 16.514 (3)
$\alpha, \beta, \gamma$ (°)	80.5779 (12), 86.4941 (12), 79.2821 (12)	79.683 (14), 85.829 (11), 79.844 (14)	79.013 (11), 85.441 (10), 80.339 (5)
$V$ (Å <sup>3</sup> )	1210.08 (5)	1144.3 (3)	1099.8 (2)
$\mu$ (mm <sup>-1</sup> )	0.30	0.31	0.33
Crystal size (mm)	0.3 × 0.05 × 0.05	0.25 × 0.05 × 0.04	0.24 × 0.04 × 0.04
Absorption correction	Multi-scan SADABS (Siemens, 1996)	Multi-scan SADABS2016/2 (Bruker,2016/2) was used for absorption correction. $wR2(int)$ was 0.1025 before and 0.0509 after correction. The Ratio of minimum to maximum transmission is 0.8711.	Multi-scan SADABS2016/2 (Bruker,2016/2) was used for absorption correction. $wR2(int)$ was 0.0861 before and 0.0545 after correction. The Ratio of minimum to maximum transmission is 0.7871.
$T_{min}, T_{max}$	0.91, 0.99	0.649, 0.745	0.586, 0.745
No. of measured, independent and observed [ $I > 2.0\sigma(I)$ ] reflections	26677, 7292, 4554	3840, 1015, 682	5514, 1044, 761
$R_{int}$	0.033	0.073	0.081
$\theta_{max}$ (°)	30.5	23.2	23.3
$(\sin \theta/\lambda)_{max}$ (Å <sup>-1</sup> )	0.714	0.555	0.556
$R[F^2 > 2\sigma(F^2)], wR(F^2), S$	0.047, 0.038, 1.13	0.066, 0.166, 1.06	0.059, 0.149, 1.09
No. of reflections	3972	1015	1044
No. of parameters	342	126	126
No. of restraints	6	20	24
$\Delta\rho_{max}, \Delta\rho_{min}$ (e Å <sup>-3</sup> )	0.27, -0.45	0.28, -0.29	0.25, -0.22

Computer programs: SAINT v8.37A<sup>26</sup> SHELXT<sup>27</sup> XL<sup>28</sup> and Olex2<sup>22</sup>

Table 3.1 Contd. Experimental details for 3.33 GPa and 4.06 GPa

	3.33 GPa, 297K	4.06 GPa, 297 K
$a, b, c$ (Å)	6.7888 (3), 9.6045 (11), 16.3703 (16)	6.7561 (3), 9.5308 (11), 16.3385 (16)
$\alpha, \beta, \gamma$ (°)	77.764 (8), 84.858 (6), 81.297 (8)	77.508 (8), 84.710 (6), 81.556 (8)
$V$ (Å <sup>3</sup> )	1029.30 (17)	1014.02 (16)
$\mu$ (mm <sup>-1</sup> )	0.35	0.35
Crystal size (mm)	0.25 × 0.05 × 0.04	0.25 × 0.05 × 0.04
Absorption correction	Multi-scan <i>SADABS2014/5</i> (Bruker,2014/5) was used for absorption correction. $wR2(int)$ was 0.0729 before and 0.0488 after correction. The Ratio of minimum to maximum transmission is 0.9098.	Multi-scan <i>SADABS2014/5</i> (Bruker,2014/5) was used for absorption correction. $wR2(int)$ was 0.0696 before and 0.0508 after correction. The Ratio of minimum to maximum transmission is 0.9264.
$T_{min}, T_{max}$	0.678, 0.745	0.690, 0.745
No. of measured, independent and observed [ $I > 2.0\sigma(I)$ ] reflections	4793, 941, 747	4707, 921, 731
$R_{int}$	0.050	0.054
$\theta_{max}$ (°)	23.3	23.3
$(\sin \theta/\lambda)_{max}$ (Å <sup>-1</sup> )	0.557	0.556
$R[F^2 > 2\sigma(F^2)], wR(F^2),$ $S$	0.064, 0.163, 1.09	0.064, 0.157, 1.07
No. of reflections	941	921
No. of parameters	126	126
No. of restraints	24	24
$\Delta\rho_{max}, \Delta\rho_{min}$ (e Å <sup>-3</sup> )	0.27, -0.24	0.25, -0.31

Computer programs: *SAINT* v8.37A<sup>26</sup> *SHELXT*<sup>27</sup> *XL*<sup>28</sup> and *Olex2*<sup>22</sup>

Table 3.1 Contd. Experimental details for 4.34 GPa and 5.06 GPa

	4.34 GPa, 297K	5.06 GPa, 297 K
<i>a</i> , <i>b</i> , <i>c</i> (Å)	6.7109 (4), 9.4367 (12), 16.2988 (17)	6.6623 (3), 9.3453 (11), 16.2637 (15)
$\alpha$ , $\beta$ , $\gamma$ (°)	77.132 (9), 84.534 (7), 81.962 (9)	76.764 (8), 84.331 (6), 82.375 (8)
<i>V</i> (Å <sup>3</sup> )	994.17 (18)	974.52 (16)
$\mu$ (mm <sup>-1</sup> )	0.36	0.37
Crystal size (mm)	0.25 × 0.05 × 0.04	0.25 × 0.05 × 0.04
Absorption correction	Multi-scan <i>SADABS2014/5</i> (Bruker,2014/5) was used for absorption correction. <i>wR2(int)</i> was 0.0702 before and 0.0505 after correction. The Ratio of minimum to maximum transmission is 0.9046.	Multi-scan <i>SADABS2016/2</i> (Bruker,2016/2) was used for absorption correction. <i>wR2(int)</i> was 0.0674 before and 0.0484 after correction. The Ratio of minimum to maximum transmission is 0.8882.
<i>T</i> <sub>min</sub> , <i>T</i> <sub>max</sub>	0.674, 0.745	0.662, 0.745
No. of measured, independent and observed [ <i>I</i> > 2.0 $\sigma$ ( <i>I</i> )] reflections	4363, 898, 715	4517, 894, 694
<i>R</i> <sub>int</sub>	0.050	0.053
$\theta$ <sub>max</sub> (°)	23.2	23.4
( <i>sin</i> $\theta$ / $\lambda$ ) <sub>max</sub> (Å <sup>-1</sup> )	0.555	0.559
<i>R</i> [ <i>F</i> <sup>2</sup> > 2 $\sigma$ ( <i>F</i> <sup>2</sup> )], <i>wR</i> ( <i>F</i> <sup>2</sup> ), <i>S</i>	0.065, 0.153, 1.05	0.067, 0.185, 1.08
No. of reflections	898	894
No. of parameters	126	126
No. of restraints	24	24
$\Delta\rho$ <sub>max</sub> , $\Delta\rho$ <sub>min</sub> (e Å <sup>-3</sup> )	0.33, -0.25	0.33, -0.33

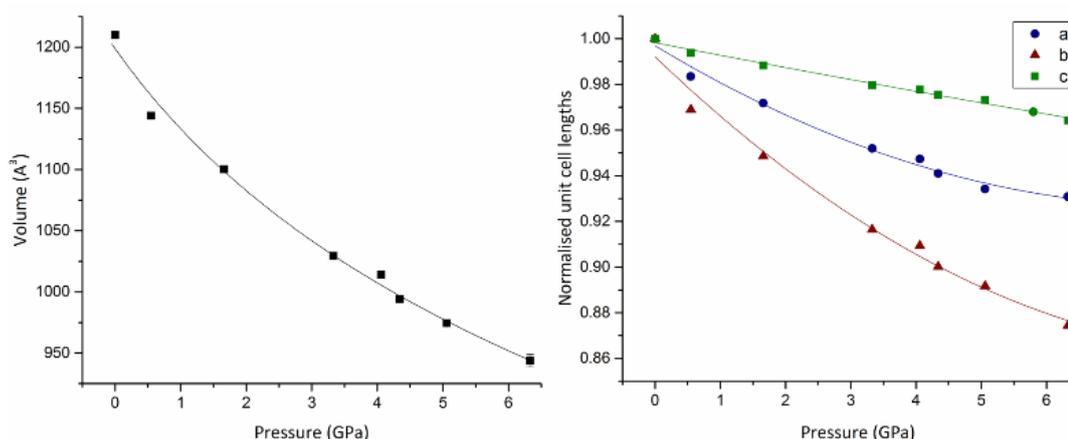
Computer programs: *SAINT* v8.37A<sup>26</sup> *SHELXT*<sup>27</sup> *XL*<sup>28</sup> and *Olex2*<sup>22</sup>

### 3.4.2. Pixel calculations

Pixel calculations<sup>29,30</sup> of the INSA co-crystal, reveal that the strongest molecule to molecule interactions are those associated with the hydrogen bonded dimers; the carboxylic acid dimer between indomethacin molecules (total 83.6 kJmol<sup>-1</sup>; 2.677(3) Å O16···O17') and the sulphonamide dimer between the saccharin components (total 60.2 kJmol<sup>-1</sup>; 2.884(3) Å N26···O28'). The third and fourth notable interaction can be attributed to a  $\pi/\pi$  interaction (indomethacin to saccharin, total -50.3 kJmol<sup>-1</sup> dispersion -51.1 kJmol<sup>-1</sup>) and a CH/ $\pi$  interaction (indomethacin to indomethacin, total -46.1 kJmol<sup>-1</sup>; dispersion -53.6 kJmol<sup>-1</sup>) respectively. These latter two molecule-molecule interactions are surprisingly strong given that they are predominantly comprised of dispersive forces. The bicyclic moieties of the indomethacin and saccharin are well-aligned which contributes significantly to the attractive dispersive forces. In this instance, the total energy of the interaction is nearing the typical energies for hydrogen bonded interactions. The bicyclic indole moiety of the indomethacin interacts through a CH··· $\pi$  interaction where the chlorobenzoyl of the second indomethacin is in an edge-to-face arrangement with the symmetry equivalent molecule (-x, -y, -z).

#### Compression study

The co-crystal undergoes a 20% decrease in volume from 1210.1 Å<sup>3</sup> at ambient pressure to 974.5 Å<sup>3</sup> at 5.06 GPa (Figure 3.4). The 3<sup>rd</sup> order Birch-Murnaghan Equation of State gives a bulk modulus ( $K_0$ ) of 15.46 (13) GPa  $V_0 = 1199.35$  (4) Å<sup>3</sup>,  $K' = 4.49$  (2). A further dataset taken at 6.33 GPa was of such poor quality that no structural information could be derived even after collecting data in a different orientation, however, the unit cell information was used in the determination of the bulk modulus.



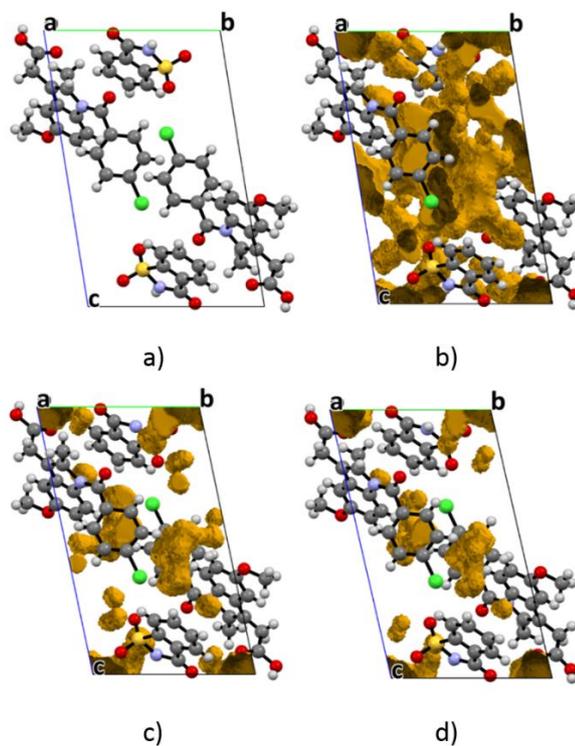
**Figure 3.4 a) Unit cell volume of INSA as a function of pressure (EosFit 7), (b) Normalised unit cell lengths as a function of pressure. The errors on each of the measurements is smaller than the symbol.**

The unit cell axes display a smooth compression with decreases of 6.5%, 10.8% and 2.7% in the a-, b- and c-axis respectively (Figure 3.4). The large decrease in the b-axis can be explained by the analysis of the void space in the structure and the absence of strong interactions between the indomethacin and saccharin dimers. Figure 3.5 shows the voids present in the structure at ambient pressure with large pockets of space between the chlorobenzoyl rings of neighbouring indomethacin molecules translated along the same chain. Over the course of the compression these voids decrease from 17.8% of the unit cell to 4.7% at 5.06 GPa (Figure 3.5; Table 3.3). Even at the highest pressure of 6.33 GPa, the unit cell parameters that could be observed from the deteriorating diffraction pattern indicated the retention of this phase hence the void space will have compressed even further. It is apparent that the voids are three dimensional throughout the lattice and could be held partly accountable for the resistance to any phase transformation in the system i.e. in every direction compression can take place without necessary rearrangement of atoms to relieve repulsion between atoms. The small pocket of void space is ‘protected’ from compression to some extent due to the size and bulkiness of the indomethacin molecule with any reduction in the pocket necessitating compression of molecular groups substantially inside the sum of the van der Waals radii.

One of the significant developments over the compression is the change in the interaction energy between the indomethacin molecule and the saccharin molecule ( $x-1, 1+y, 1+z$ ). The close contact, and only real interaction between the molecules, is between the hydroxyl oxygen and carbonyl oxygen atoms indomethacin and saccharin respectively. The interaction between these molecules becomes highly

## Lauren Connor

repulsive increasing from  $1.2 \text{ kJmol}^{-1}$  at ambient pressure to  $13.3 \text{ kJmol}^{-1}$  at 5.06 GPa. There are very little attractive forces between the molecules hence the repulsion term dominates the molecular interaction.



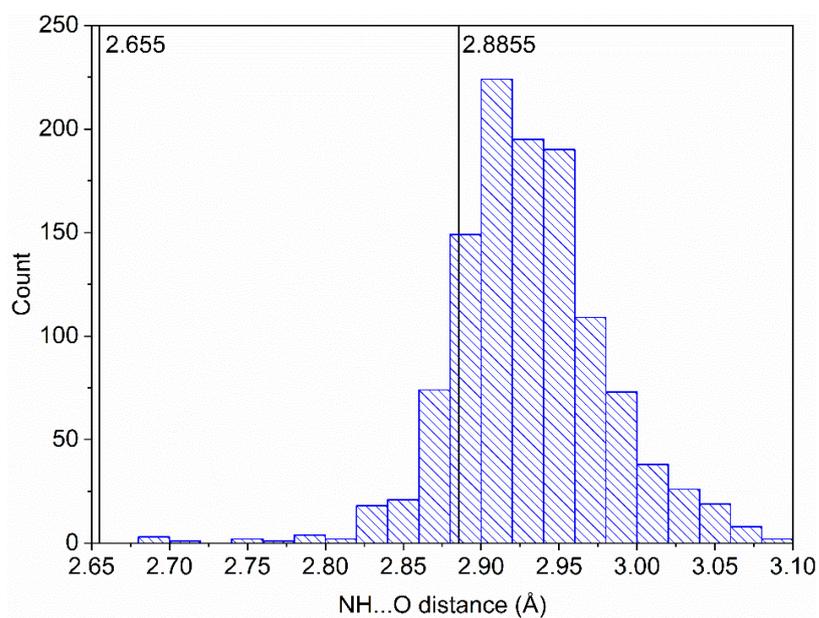
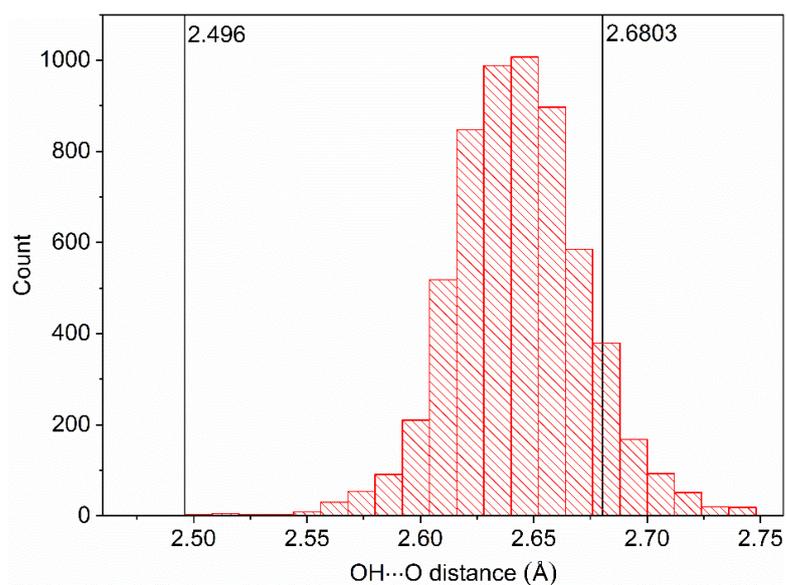
**Figure 3.5** The structure of the unit cell down the a-axis (a). The void space at 0.0001 GPa (b), 3.33 GPa (c), and 5.06 GPa (d) obtained by using probe radius of  $0.5 \text{ \AA}$  with high definition of  $0.1 \text{ \AA}$ .

**Table 3.3** Percentage of the unit cell occupied by void space at each pressure. The probe radius was set to  $0.5 \text{ \AA}$  with 0.1 grid spacing.

Pressure (GPa)	Percentage of unit cell (%)
0.0001	17.8
0.55	12.7
1.66	11.3
3.33	6.9
4.06	6.6
4.34	5.5
5.06	4.7

### 3.4.3. Compression behaviour of the amide dimer

One of the key aspects to this study was to investigate how intermolecular interactions change as a function of pressure. There have been a large number of studies of materials under pressure but the types of hydrogen bond that have been probed is not particularly diverse nor have there been a significant number of studies of particular hydrogen bonds except for, perhaps, ammonium-carboxylate interaction, due to studies of amino acids.<sup>31–36</sup> INSA exists as a co-crystal with only homomeric hydrogen bonding interactions (carboxylic acid and amide dimers). We probed the Cambridge Structural Database for similar hydrogen bonding interactions to those in our system. Analysis of hydrogen bonding distributions of similar dimer interactions are presented in Figure 3.6. Discussion of hydrogen bonding distances will concern only Donor⋯Acceptor (D⋯A) due to the limited data from the high-pressure experiment. In terms of the carboxylic acid dimer, the mean distance is 2.643Å. At ambient pressure, the carboxylic acid dimer of INSA (2.6803(18) Å) is slightly above this value, however on compression to 5.06 GPa the dimer reduces by 7% to 2.496Å which is far below those observed in the database. The majority of the database structures are at ambient pressure but the lower end of the distribution has been used in previous work to show that a phase transition may be imminent during compression studies.<sup>37</sup> Perhaps a more instructive comparison would be with other pressure studies to see if, despite changes to molecular connectivity, the hydrogen bonds are compressed to a similar level. Restricting the database to those refcodes that have pressure explicitly stated in their experimental conditions, there are 2454 refcodes reported in the CSD. 45 of these refcodes contain a carboxylic acid dimer contact. R,S-Ibuprofen is the only observation of a carboxylic acid dimer compression study in the database with the other entries being mostly individual collections at particular pressures hence do not give information on how hydrogen bonds change with respect to pressure.<sup>38</sup> The hydrogen bonded dimer in ibuprofen compresses to a similar value to our INSA example and at 4 GPa it is observed at 2.549Å (cf. 2.529(10) Å for INSA at 4.06GPa) (Table 4). Further compression of INSA to 5.06 GPa reduces the hydrogen bond further to 2.498(11) Å to be amongst the shortest reported in the database.



**Figure 3.6** Histograms showing the distribution of O-H...O and N-H...O distances reported in the CSD with markers for the distances reported in the compression of INSA at ambient and the highest pressure dataset refined.

Table 3.4 Selected hydrogen-bond parameters at pressure

$D-H\cdots A$	$D-H$ (Å)	$H\cdots A$ (Å)	$D\cdots A$ (Å)	$D-H\cdots A$ (°)
Ambient Pressure				
O16—H161...O17 <sup>i</sup>	0.84	1.84	2.6803(18)	176.2
N26—H261...O28 <sup>ii</sup>	0.88	2.04	2.8855(19)	159.8
0.55GPa				
O16—H161...O17 <sup>i</sup>	0.82	1.86	2.639(9)	158.2
N26—H261...O28 <sup>ii</sup>	0.86	2.04	2.860(10)	159.4
1.66GPa				
O16—H161...O17 <sup>i</sup>	0.82	1.79	2.604(15)	169.7
N26—H261...O28 <sup>ii</sup>	0.86	1.96	2.796(12)	162.5
3.33GPa				
O16—H161...O17 <sup>i</sup>	0.82	1.73	2.545(10)	173.8
N26—H261...O28 <sup>ii</sup>	0.86	1.89	2.723(13)	161.5
4.06GPa				
O16—H161...O17 <sup>i</sup>	0.82	1.73	2.529(10)	163.5
N26—H261...O28 <sup>ii</sup>	0.86	1.87	2.702(9)	161.8
4.34GPa				
O16—H161...O17 <sup>i</sup>	0.82	1.7	2.513(10)	170.8
N26—H261...O28 <sup>ii</sup>	0.86	1.85	2.681(10)	162.4
5.06GPa				
O16—H161...O17 <sup>i</sup>	0.82	1.68	2.496(11)	174.7
N26—H261...O28 <sup>ii</sup>	0.86	1.82	2.655(10)	162.2

Symmetry code(s): (i)  $-x-1, -y+2, -z+2$ ; (ii)  $-x+1, -y+1, -z$

The distribution of the amide dimers in the CSD are shown in Figure 3.6 compared with the values observed in our study. The mean distance of the amide dimer in the CSD is 2.931 Å with our model showing a dimer distance of 2.8855(19) Å. As a comparator, pure saccharin has a dimer contact of 2.818 Å. On compression to 5.06 GPa the hydrogen bond reduces by 8% to 2.655(10) Å again approaching the lowest values observed in the CSD. By restricting the CSD to show structures under pressure conditions, there are only thirty refcodes observed pertaining to five molecules. Two different polymorphs of 2-oxopyrrolidineacetamide (piracetam),<sup>39</sup> 2-hydroxybenzamide (salicylamide)<sup>37</sup> and a compression of two polymorphs of anti-diabetic, tolazamide.<sup>40</sup>

Figure 3.7 shows the changes in the hydrogen bonding in INSA compared with the dimer distances from other pressure studies of amides. The relative compression in each of these systems is comparable with a suggestion that the dimers in INSA and tolazamide are compressed a little faster. The reason behind this slight change in behaviour is potentially a combination of both the packing of the compounds and respective positioning of the hydrogen bonds in the structure. Both amide dimer contacts for piracetam<sup>39</sup> and salicylamide<sup>37</sup> are involved in a second hydrogen bonding interaction in an opposing direction within the structure which will aid in stabilising the amide dimer against shortening as a result of compression. Piracetam dimers form a hydrogen bond to the carbonyl of another neighbouring dimer unit in each direction, stabilising the central dimer. Piracetam undergoes a subtle phase transition after 0.45 GPa from Form II to Form V which is the reason for the change in compression behaviour but the hydrogen bonding remains the same. The driving force for this transition is thought to be the reduction of voids in the structure. Salicylamide packs to form a channel of hydrogen bonding due to a screw axes in this molecules packing arrangement. Each molecule in the dimer bonds perpendicularly bonds to another molecule, forming channels of hydrogen bonding. Piracetam and salicylamide also possess a layered structure where the hydrogen bonding is situated within the layer. The principal axis of strain was calculated to be approximately perpendicular to these layers in each case signifying that the main compression of the system comes from reducing the spacing between them hence the hydrogen bonding (in the layer) will be less impacted.

Tolazamide<sup>40</sup> polymorph II shows a similar rate of compression to INSA. The tolazamide structure exists in layers as hydrogen-bonded dimers with  $\pi$ - $\pi$  interactions

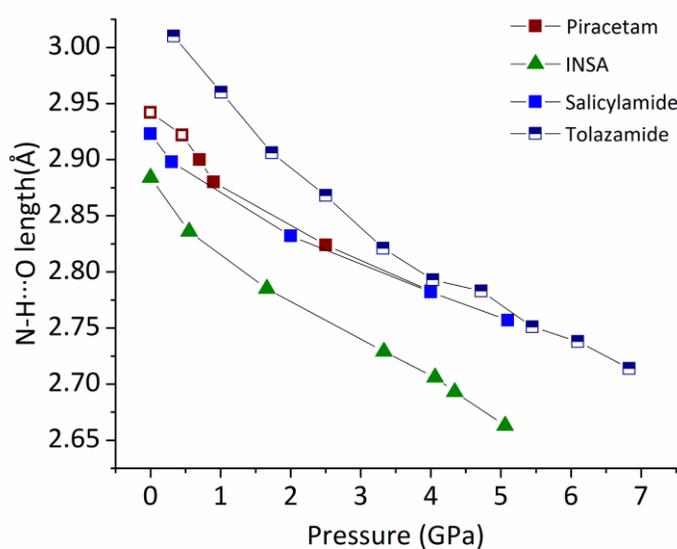
## Lauren Connor

between the layers. Like INSA, the amide dimer contact is not involved in any other hydrogen bonding therefor may be more susceptible to compression.

Comparatively, the dimer observed in INSA is much closer at all pressure points than previous high pressure studies of piracetam, salicylamide, and Tolazamide but is close to the average distance as observed in the database. The dimer in saccharin itself lies at  $\sim 2.80$  Å depending on the temperature of the study. The other compounds under study possess longer hydrogen bonding parameters at ambient pressure suggesting that the packing of the molecules (and other hydrogen bonds) affects the hydrogen bonding parameters more than the saccharin dimer in INSA. The packing indexes for each compound were calculated in Platon<sup>41</sup> and found to be 73.0 %, 71.0 %, 70.5 %, 68.3 % and 68.2 % for Piracetam, INSA, saccharin, tolzamide and salicylamide, respectively. Piracetam packs more efficiently than the rest but the nitrogen of the amide dimer is involved in another N-H $\cdots$ O amide contact using the second hydrogen of the amide in the opposite direction which may serve to weaken and lengthen the dimer h-bond. Tolzamide and salicylamide have comparable lower packing indexes. Tolzamide is a relatively bulky structure that is more than double the molecular mass (311.4 g/mol) of salicylamide (137.1 g/mol). In the compression comparison Tolzamide was the only other structure with an amide dimer not involved in other hydrogen bonding allowing for a faster rate of compression. At ambient pressure, however, it is thought that sterics will affect the packing of this structure, with a 105° angle at the sulphonamide part of the molecule. In contrast, the low packing index for salicylamide is not due to sterics but as a result of the channelled hydrogen bonding and layered structure. INSA and saccharin itself possess an amide dimer that is free from other hydrogen bonding interactions. The pure substance is also a layered structure with a layer of hydrogen bonding and then a slip plane. The hydrogen bonding is supported by the close packing of the rings within the layer providing a stable environment for a short dimer contact. A potential reason for the saccharin dimer being slightly longer in INSA is that the packing of the solid in INSA will be dominated by the packing of the indomethacin molecules due to its size. The saccharin dimers are in a pocket bookended by two indomethacin molecules; the interaction energy between the molecules is  $-15$  kJmol<sup>-1</sup>. The saccharin dimer is not bound by the close packing of similar molecules (in the case of the pure compound) hence may be lying at an ideal h-bonding distance for a gas-phase dimer within this pocket. Calculations are underway to calculate the potential energy well of this dimer in the gas phase to establish whether this hypothesis is correct. Unfortunately, the

## Lauren Connor

timescales for completion of the calculations were beyond the completion of this thesis. As pressure is applied, the saccharin molecules are pushed together by the indomethacin framework. INSA's void space distribution is situated throughout the structure and allows for compression in multiple directions without necessarily compressing in the direction of the hydrogen bonding.



**Figure 3.7** Plots showing how the length of piracetam (form II: half symbol & form V: filled symbol), INSA, tolazamide, and salicylamide's N-H...O amide dimer respond to pressure.

### 3.5. Conclusion

The investigation of the continuous compression of the co-crystal of indomethacin and saccharin showed no phase transition however there were some interesting points to note from the exploration of the structure on compression. In this compression study, the co-crystalline lattice was investigated with a focus on the co-crystals intermolecular interactions. Surprisingly, there are few compression studies involving acid or amide dimers in the CSD. A comparative study between INSA and other compression studies in the database with an amide dimer motif revealed that the amide bond compression was dependant on the other hydrogen bonding in the system. The compression rate was slower for systems with amide dimers which were involved in a second hydrogen bonding interaction in an opposing direction. The discrepancy in ambient amide bond lengths were considered to be a mixture of the sterics, packing and hydrogen bonding patterns of each system. In comparison, INSA possesses a shorter amide dimer contact length and has a fast rate of compression. The pure compound, saccharin, possesses a shorter ambient amide dimer bonding contact length than in co-crystalline INSA. As a result, chapter four targets the compression of pure saccharin as a point of interest. The 'intermediate' hydrogen bonding observed in the ambient saccharin dimer of INSA is due to the dimer sitting in a 'pocket' with the indomethacin framework at an ideal hydrogen bonding distance. Calculations are underway to provide theoretical evidence for this observation.

## References

1. Landenberger, K. B. & Matzger, A. J. Cocrystal engineering of a prototype energetic material: Supramolecular chemistry of 2,4,6-trinitrotoluene. *Cryst. Growth Des.* **10**, 5341–5347 (2010).
2. Yang, Z., Li, H., Zhou, X., Zhang, C., Huang, H., Li, J. & Nie, F. Characterization and properties of a novel energetic-energetic cocrystal explosive composed of HNIW and BTF. *Cryst. Growth Des.* **12**, 5155–5158 (2012).
3. Jona, F., Shirane, G. & Pepinsky, R. Dielectric, X-Ray and Optical Study of Ferroelectric Cd<sub>2</sub>Nb<sub>2</sub>O<sub>7</sub> and Related Compounds. *Phys. Rev.* **98**, (1955).
4. Bolla, G., Dong, H., Zhen, Y., Wang, Z. & Hu, W. Organic cocrystals: the development of ferroelectric properties. *Sci. China Mater.* **59**, 523–530 (2016).
5. Karki, S., Friscic, T., Fabiani, L. & Jones, W. New solid forms of artemisinin obtained through cocrystallisation. *CrystEngComm* **12**, 4038–4041 (2010).
6. Kumar, S. S., Thakuria, R. & Nangia, A. Pharmaceutical cocrystals and a nitrate salt of voriconazole. *CrystEngComm* 4722–4731 (2014). doi:10.1039/c3ce42324g
7. Steed, J. W. The role of co-crystals in pharmaceutical design. *Trends in Pharmacological sciences* **34**, (Durham University, 2013).
8. Schultheiss, N. & Newman, A. Pharmaceutical cocrystals and their physicochemical properties. *Cryst. Growth Des.* **9**, 2950–2967 (2009).
9. Cruz-Cabeza, A. J., Reutzel-Edens, S. M. & Bernstein, J. Facts and fictions about polymorphism. *Chem. Soc. Rev.* **44**, 8619–8635 (2015).
10. Almarsson, Ö. & Zaworotko, M. J. Crystal engineering of the composition of pharmaceutical phases. Do pharmaceutical co-crystals represent a new path to improved medicines? *Chem. Commun.* 1889–1896 (2004). doi:10.1039/b402150a
11. Vishweshwar, P., McMahon, J., A. Bis, J. & J. Zaworotko, M. Pharmaceutical Cocrystals. *Journal of pharmaceutical sciences* **95**, 499–516 (2006).
12. Reddy, L. S., Babu, N. J. & Nangia, A. Carboxamide-pyridine N-oxide heterosynthons for crystal engineering and pharmaceutical cocrystals. *Chem. Commun.* 1369–1371 (2006). doi:10.1039/b515510j
13. Boldyreva, E. V. Multicomponent organic crystals at high pressure. *Zeitschrift für Krist.* **229**, 236–245 (2014).
14. Groom, C. R., Bruno, I. J., Lightfoot, M. P. & Ward, S. C. The Cambridge Structural Database. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **B72**, 171–179 (2016).
15. Okumura, T., Ishida, M., Takayama, K. & Otsuka, M. Polymorphic transformation of indomethacin under high pressures. *J. Pharm. Sci.* **95**, 689–700 (2006).
16. Fabbiani, F. P. a, Buth, G., Levendis, D. C. & Cruz-Cabeza, A. J. Pharmaceutical hydrates under ambient conditions from high-pressure seeds: a case study of GABA monohydrate. *Chem. Commun. (Camb)*. **50**, 1817–9 (2014).
17. Basavoju, S., Boström, D. & Velaga, S. P. Indomethacin-saccharin cocrystal: Design, synthesis and preliminary pharmaceutical characterization. *Pharm. Res.* **25**, 530–541 (2008).

18. Jung, M., Kim, J., Kim, M., Alhalaweh, A., Cho, W., Hwang, S. & Velaga, S. P. Bioavailability of indomethacin-saccharin cocrystals. *J. Pharm. Pharmacology* **62**, 1560–1568 (2010).
19. Piermarini, G. J., Block, S., Barnett, J. D. & Forman, R. a. Calibration of the Pressure Dependence of the R//1 Ruby Fluorescence Line To 195 Kbar. *J. Appl. Phys.* **46**, 2774–2780 (1975).
20. Bruker. APEX II. *APEX II* Bruker AXS Inc., Madison, Wisconsin, USA. (2013).
21. Sheldrick, G. M. SADABS, Programs for Scaling and Absorption Correction of Area Detector Data. *SADABS, Programs Scaling Absorpt. Correct. Area Detect. Data* University of Göttingen: Göttingen, Germany (2008).
22. Dolomanov, O. V., Bourhis, L. J., Gildea, R. J., Howard, J. A. K. & Puschmann, H. OLEX2: a complete structure solution, refinement and analysis program. 339–341 (2009).
23. Dawson, A., Allan, D. R., Parsons, S. & Ruf, M. Use of a CCD diffractometer in crystal structure determinations at high pressure. *J. Appl. Crystallogr.* **37**, 410–416 (2004).
24. Parsons, S. SHADE. Program for Empirical Absorption Corrections to High Pressure Data. (2004).
25. Macrae, J., Edgington, C. F., McCabe, P. ., Pidcock, P., Shields, E., Taylor, G. ., Towler, R. & Van de Streek, J. MERCURY. *J. Appl. Crystallogr.* **39**, 453–457 (2006).
26. Bruker. APEX3, SAINT and SADABS. (2015).
27. Sheldrick, G. M. SHELXT - Integrated space-group and crystal-structure determination. *Acta Crystallogr. Sect. A Found. Crystallogr.* **71**, 3–8 (2015).
28. Sheldrick, G. M. Crystal structure refinement with SHELXL. *Acta Crystallogr. Sect. C Struct. Chem.* **71**, 3–8 (2015).
29. Gavezzotti, A. Calculation of Intermolecular Interaction Energies by Direct Numerical Integration over Electron Densities . 2 . An Improved Polarization Model and the Evaluation of Dispersion and Repulsion Energies. *J. Phys. Chem.* **107**, 2344–2353 (2003).
30. Maschio, L., Civalleri, B., Ugliengo, P. & Gavezzotti, A. Intermolecular Interaction Energies in Molecular Crystals: Comparison and Agreement of Localized Møller-Plesset 2, Dispersion-Corrected Density Functional, and Classical Empirical Two-Body Calculations. *J. Phys. Chem.* **115**, 11179–11186 (2011).
31. Momany, F. A., McGuire, R. F., Burgess, A. W. & Scheraga, H. A. Energy parameters in polypeptides. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids. *J. Phys. Chem.* **79**, 2361–2381 (1975).
32. Pauling, L., Corey, R. B. & Branson, H. R. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci.* **37**, 205–211 (1951).
33. Bernstein, J., Davis, R. E., Shimon, L. & Chang, N. -L. Patterns in Hydrogen Bonding: Functionality and Graph Set Analysis in Crystals. *Angew. Chemie Int. Ed. English* **34**, 1555–1573 (1995).
34. Zakharov, B. A. & Boldyreva, E. V. A high-pressure single-crystal to single-crystal phase transition in dl-alaninium semi-oxalate mono-hydrate with switching-over hydrogen bonds. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **69**, 271–280 (2013).

## Lauren Connor

35. Moggach, S. A., Parsons, S. & Wood, P. A. High-pressure polymorphism in amino acids. *Crystallography Reviews* **14**, (2008).
36. Görbitz, C. H. Crystal structures of amino acids: From bond lengths in glycine to metal complexes and high-pressure polymorphs. *Crystallogr. Rev.* **21**, 160–212 (2015).
37. Johnstone, R. D. L., Lennie, A. R., Parker, S. F., Parsons, S., Pidcock, E., Richardson, P. R., Warren, J. E. & Wood, P. A. High-pressure polymorphism in salicylamide. *CrystEngComm* **12**, 983–1316 (2010).
38. Ostrowska, K. & Katrusiak, A. High-Pressure Crystallization and Structural Transformations in Compressed R , S - Ibuprofen. *Cryst. Growth Des.* **15**, 1512–1517 (2015).
39. Fabbiani, F. P. A., Allan, D. R., David, W. I. F., Davidson, A. J., Lennie, A. R., Parsons, S., Pulham, C. R. & Warren, J. E. High-Pressure Studies of Pharmaceuticals : An Exploration of the Behavior of Piracetam. *Cryst. Growth Des.* **7**, 1115–1124 (2007).
40. Fedorov, A. Y., Rychkov, D. A., Losev, E. A., Zakharov, B. A., Stare, J. & Boldyreva, E. V. Effect of pressure on two polymorphs of tolazamide: Why no interconversion? *CrystEngComm* **19**, 2243–2252 (2017).
41. Spek, A. L. Structure validation in chemical crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **65**, 148–155 (2009).

***-Chapter Four-***

***Polymorphic phase transition of saccharin during  
compression***

### 4. Polymorphic phase transition of saccharin during compression

#### 4.1. Synopsis

The work, herein, follows on from the study in chapter three on the compression of a co-crystal of indomethacin and saccharin to probe the intermolecular interactions in the pure system and how they respond to hydrostatic compression. Study of saccharin in the co-crystal lead to a conclusion that the saccharin dimer may exist in an unperturbed environment, with little interaction from other components. This study attempts to understand how this structure reacts to compression when in a homomolecular crystal structure. Interestingly, saccharin possesses strong dispersive forces which we were able to quantify and explore using Pixel calculations. Layers of saccharin are dominated by amide dimer interactions with  $\pi$ - $\pi$  stacking between the layers. Saccharin undergoes a polymorphic phase transition at 1.32 GPa from a monoclinic system to a triclinic system. The hydrogen bonding pattern of the structure does not change but the dispersive forces dominate the main changes in the cell as the intra- and inter-layer distances shorten and accommodate the phase change upon compression.

#### 4.2. Introduction

##### 4.2.1. Saccharin in the food industry

Saccharin is an attractive substitute to sucrose in the food industry as it is 200-700 times sweeter than sucrose and has zero calories. Sugar (specifically sucrose) has increasingly come under fire for its detrimental effects to our bodies in terms of weight gain and tooth decay.<sup>1,2</sup> In an effort to reduce our sugar intake, many countries of the world have imposed a tax on drinks which are high in sugar. To date, 28 countries have introduced a sugar tax on food and drinks, with other countries considering similar proposals. In 2016, the World Health Organisation (WHO) supported the concept of taxing sugary drinks by 20% or more to help reduce sugar consumption. The UK is one of the most recent to introduce the tax in 2018.<sup>3,4</sup> In an effort to maintain sales and profits, some large companies have changed their recipes to avoid the tax which has led to increased use of alternative sweeteners in the food industry such as aspartame, sorbitol, xylitol and saccharin. Saccharin is one of the most commonly used artificial sweeteners. It is generally regarded as safe (GRAS) after a study in the 1970's on the link between saccharin and bladder cancer in rats was found to be

## Lauren Connor

irrelevant to humans.<sup>5</sup> Its classification as non-carcinogenic in 1998 allows its consumption as food or as part of a pharmaceutical formulation.

### 4.2.2. Saccharin in the pharmaceutical industry

The molecule itself is of no known therapeutic benefit but it has been investigated as an excipient, counter ion<sup>6</sup> and co-former<sup>7-10</sup> by the pharmaceutical industry. The most common sweeteners used as excipients in final drug formulations are sucrose (53.4%) and saccharin (38.3 %).<sup>11</sup> Saccharin can potentially mask a bitter taste in pharmaceutical products which may lead to enhanced patient compliance, this is especially of interest in paediatric oral drug delivery. In the early stages of drug development, saccharin can be introduced as one component in a multicomponent form of a drug. Multicomponent forms are investigated as a means to improve the physicochemical properties of an active pharmaceutical ingredient (API). Saccharin can act as a counterion to APIs in pharmaceutical salts or, in its neutral form, as a co-former in co-crystalline materials.<sup>8,12</sup> The saccharin anion in saccharinate salts is known to have a high affinity for water. The prospect of improved solubility has led to studies of saccharin with many APIs such as quinine, haloperidol, mirtazapine, pseudophedrine, lamivudine, risperidone, sertraline, venlafaxine, zolpidem and amlodipine.<sup>6</sup> There are also studies on saccharin as a co-crystal with adefovir dipivoxil, Indomethacin and Carbamazepine and it is commonly considered as a valuable co-former.<sup>8-10,12,13</sup>

Despite numerous studies of saccharin, there is only one known polymorph reported in the Cambridge Structural Database (CSD) since the structure was first solved in 1968.<sup>14</sup> It has been thought for many years that the time and effort expended on any given structure is proportional to the amount of polymorphs found<sup>15-17</sup> however, to date, saccharin defies this idea. Polymorphs of a component can be isolated serendipitously from studies into multicomponent systems<sup>18,19</sup> as was the case with benzidine. The formation of four new polymorphs of benzidine during co-crystallisation attempts with diphenyl sulfoxide was reported by Rafilovich et al.<sup>20</sup> Rafilovich et al. concluded that the changes to the crystallisation medium, invoked by the introduction of co-formers to the solvent, led to the discovery of the four new forms of benzidine. This was considered by the authors as a route to new polymorphs that, whilst serendipitous, may be a useful process for polymorph screening. There are 28 multi-component forms of saccharin<sup>1,8,27-35,12,13,21-26</sup> in the Cambridge Structural Database (CSD) that have exposed saccharin to various solvent systems (chloroform,

## Lauren Connor

methanol, water, DMSO, ethanol, DCM, ethyl acetate and benzene) as well as heat and mechanical stress by neat grinding. Despite the numerous experiments on saccharin, no new forms have been reported, but one new polymorph of a co-crystal has been observed (CSD refcode: UNEZAO).<sup>36</sup>

One area of experimental space that has remained unexplored in saccharin's crystal structure landscape is high pressure. High pressure analysis is a useful route of exploration as saccharin's journey into our possession for consumption can include harsh processing conditions. Production of tablets for artificial sweeteners or final drug products involve mechanical stress such as particle-size reduction by milling or through compression using the tableting press. It is advisable to fully understand how this structure behaves under pressure to gain insight into its structure-property relationship. High pressure as a means to explore experimental space has been employed for many years.<sup>15,37,46–50,38–45</sup> However, as equipment limits screening to one compound at a time as well as being expensive in time and money the food and pharmaceutical industry do not routinely explore a wide pressure range in the search for other polymorphs.

The purpose of this chapter is to understand how the crystal structure of pure saccharin behaves as a function of pressure. Following from the previous chapter, on the co-crystal of indomethacin and saccharin, this study will aid in the understanding of how intermolecular interactions respond to pressure. Through the use of X-ray diffraction and Pixel energy calculations we will be able to gain a deeper energetic understanding of how the structure compresses.

### 4.3. Experimental

#### 4.3.1. General procedure

Merrill-Bassett diamond-anvil cell (DAC)<sup>51</sup> with a half-opening angle of 40° was used as the pressure device. Opposing diamonds with 600µm culets and sat in Boehler-Almax designed tungsten carbide backing seats. The DAC sample chamber was a 300µm diameter hole drilled into a tungsten 10mm x 10mm gasket with a thickness of 90µm (Almax EasyLab µDriller). A ruby chip was loaded into the chamber along with the sample to measure the pressure *in situ* via the ruby fluorescence technique. A Horiba XplorA Raman spectrometer equipped with a 532nm laser was used to measure the ruby fluorescence.<sup>52</sup>

## Lauren Connor

### 4.3.2. Crystal

A crystal of dimensions 0.25mm x 0.18mm x 0.08mm was grown in ethanol by evaporation at room temperature (~293 K) over 12 hours.

### 4.3.3. Diamond anvil cell preparation

A well-defined crystal with a plate-like morphology was loaded into the DAC<sup>51</sup> with petroleum ether (40:60)<sup>53</sup> to allow for hydrostatic compression. The initial sealing of the cell introduced a pressure of 0.08GPa. Single crystal X-ray diffraction data were collected at regular pressure intervals with the DAC in a single orientation throughout the study on I19 at Diamond Light Source.

### 4.3.4. Compression study collection, processing and refinement

X-ray diffraction intensities were collected at station I19 at Diamond Light Source, Harwell Science and Innovation Campus. A sequence of ten scans were taken, at 293K, using a Newport IS4CCD (4 circle) diffractometer ( $\lambda = 0.4859 \text{ \AA}$ ) with a Pilatus 300 K detector. Details of the beamline are detailed in Nowell *et al.*<sup>54</sup> Data were indexed and integrated using Xia2<sup>55</sup> which provided accurate unit cell parameters. The diffractometer requires two instrument models for the two positions of the 2-theta circle due to small differences in the beamcentre position at positive and negative 2-theta due to the sag in the detector arm at positive 2-theta. Xia2 uses new beam centres for each run hence is able to provide accurate unit cell parameters; commercial programs only use one beam centre per integration hence the unit cells are distorted from their true values. The integrated intensities used in this study were taken from data reduction using CrysAlisPro<sup>56</sup> due to the ability to process split crystals and the success with weaker diffracting data over Xia2. The initial crystal structure of saccharin, at 0.08 GPa, was solved by Intrinsic Phasing (XT) in *Olex2* (F<sup>2</sup>).<sup>57</sup> The starting point for the high pressure datasets was the 0.08 GPa dataset, each subsequent refinement started with the refined model from the previous pressure. All heavy atoms were refined anisotropically and enhanced rigid body restraints (RIGU)<sup>58</sup> were applied to each pressure structure. The crystal split at 1.98 GPa and RIGU restraints with a standard deviation of  $0.002 \text{ \AA}^2$  applied to refinements at this pressure and higher pressures. Hydrogen atoms were placed in calculated positions and constrained to ride their parent atom as per procedure in *Olex2*.

**Table 4.1 Experimental details from 0.08 GPa to 1.32 GPa**

For all structures:  $C_7H_5NO_3S$ ,  $M_r = 183.18$ . Experiments were carried out with Synchrotron radiation, ( $\lambda = 0.4859 \text{ \AA}$ ) using a Newport detector. H-atom parameters were constrained.

Pressure (GPa)	0.08 GPa, 297 K	0.68 GPa, 297 K	1.32 GPa, 297 K
Crystal system, space group	Monoclinic, $P2_1/c$	Monoclinic, $P2_1/c$	Triclinic, $P-1$
$a, b, c$ (Å)	9.578 (6), 6.9303 (5), 11.8524 (10)	9.320 (7), 6.8779 (5), 11.5359 (9)	6.8364 (8), 9.183 (11), 11.3365 (10)
$\alpha, \beta, \gamma$ (°)	103.81 (3)	102.63 (3)	78.03 (4), 89.921 (10), 86.45 (3)
$V$ (Å <sup>3</sup> )	764.0 (5)	721.5 (5)	694.9 (9)
Z	4	4	4
$\mu$ (mm <sup>-1</sup> )	0.19	0.20	0.16
Crystal size (mm)	0.25 × 0.18 × 0.08	0.25 × 0.18 × 0.08	0.25 × 0.18 × 0.08
Absorption correction	Multi-scan <i>CrysAlis PRO</i> 1.171.38.46 (Rigaku Oxford Diffraction, 2015)	Multi-scan <i>CrysAlis PRO</i> 1.171.38.41 (Rigaku Oxford Diffraction, 2015)	Multi-scan <i>CrysAlis PRO</i> 1.171.38.41 (Rigaku Oxford Diffraction, 2015)
$T_{\min}, T_{\max}$	0.156, 1.000	0.428, 1.000	0.859, 1.000
No. of measured, independent and observed [ $I > 2\sigma(I)$ ] reflections	3862, 661, 571	4305, 598, 552	4005, 993, 856
$R_{\text{int}}$	0.054	0.041	0.051
$(\sin \theta/\lambda)_{\text{max}}$ (Å <sup>-1</sup> )	0.627	0.627	0.627
$R[F > 2\sigma(F)]$ , $wR(F)$ , S	0.057, 0.186, 1.13	0.030, 0.085, 1.07	0.059, 0.166, 1.13
No. of reflections	661	598	993
No. of parameters	109	109	217
$\Delta\rho_{\text{max}}, \Delta\rho_{\text{min}}$ (e Å <sup>-3</sup> )	0.27, -0.34	0.12, -0.18	0.46, -0.27

Computer programs: *CrysAlis PRO* 1.171.38.41<sup>56</sup>, XT<sup>59</sup>, XL<sup>60</sup>, Olex2<sup>57</sup>

Table 4.1 Contd. Experimental details from 1.98 GPa to 2.811 GPa

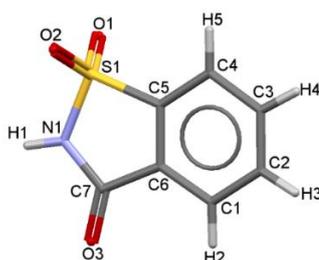
Pressure (GPa)	1.98 GPa, 297 K	2.45 GPa, 297 K	2.811 GPa, 297 K
Crystal system, space group	Triclinic, <i>P</i> -1	Triclinic, <i>P</i> -1	Triclinic, <i>P</i> -1
<i>a</i> , <i>b</i> , <i>c</i> (Å)	6.8196 (8), 9.008 (12), 11.1890 (11)	6.7986 (6), 8.921 (8), 11.0906 (8)	6.7891 (6), 8.869 (9), 11.0208 (8)
$\alpha$ , $\beta$ , $\gamma$ (°)	78.49 (4), 89.844 (7), 85.382 (12)	78.85 (2), 89.800 (8), 84.714 (18)	79.10 (3), 89.751 (8), 84.60 (2)
<i>V</i> (Å <sup>3</sup> )	671.6 (9)	657.1 (6)	648.7 (7)
<i>Z</i>	4	4	4
$\mu$ (mm <sup>-1</sup> )	0.16	0.22	0.17
Crystal size (mm)	0.25 × 0.18 × 0.08	0.25 × 0.18 × 0.08	0.25 × 0.18 × 0.08
Absorption correction	Multi-scan <i>CrysAlis PRO</i> 1.171.39.46 (Rigaku Oxford Diffraction, 2018)	Multi-scan <i>CrysAlis PRO</i> 1.171.38.41 (Rigaku Oxford Diffraction, 2015)	Multi-scan <i>CrysAlis PRO</i> 1.171.38.46 (Rigaku Oxford Diffraction, 2015)
<i>T</i> <sub>min</sub> , <i>T</i> <sub>max</sub>	0.172, 1.000	0.471, 1.000	0.435, 1.000
No. of measured, independent and observed [ <i>I</i> > 2σ( <i>I</i> )] reflections	3705, 935, 795	1876, 635, 539	3380, 893, 743
<i>R</i> <sub>int</sub>	0.075	0.054	0.070
(sin θ/λ) <sub>max</sub> (Å <sup>-1</sup> )	0.628	0.556	0.626
<i>R</i> [ <i>F</i> <sup>2</sup> > 2σ( <i>F</i> <sup>2</sup> )], <i>wR</i> ( <i>F</i> <sup>2</sup> ), <i>S</i>	0.075, 0.225, 1.19	0.070, 0.200, 1.06	0.064, 0.182, 1.06
No. of reflections	935	635	893
No. of parameters	217	223	217
Δρ <sub>max</sub> , Δρ <sub>min</sub> (e Å <sup>-3</sup> )	0.54, -0.45	0.32, -0.32	0.46, -0.35

Computer programs: *CrysAlis PRO* 1.171.38.41<sup>56</sup>, *XT*<sup>59</sup>, *XL*<sup>60</sup>, *Olex2*<sup>57</sup>

## 4.4. Results and Discussion

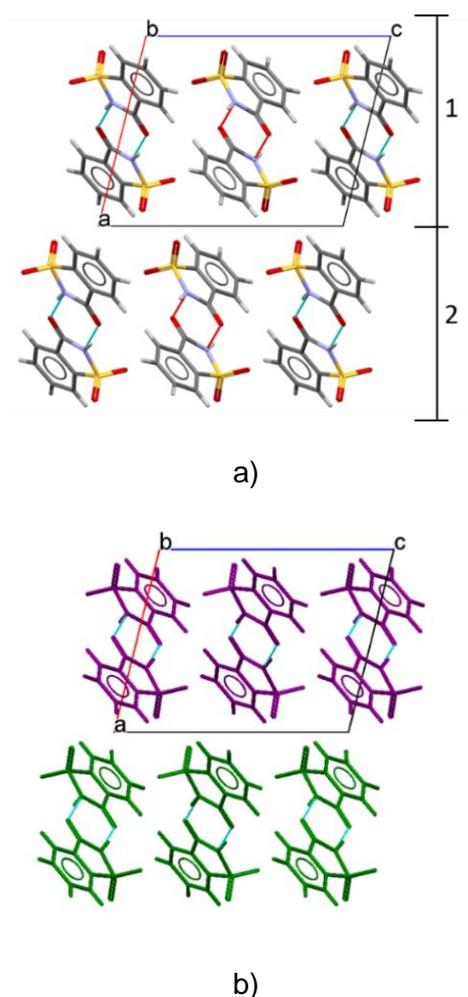
### 4.4.1. Structure

The crystal structure of saccharin at ambient pressure has been published and discussed elsewhere.<sup>9,61–63</sup> However, it is useful to provide an introduction to certain structural aspects and to compare and understand structural changes observed with hydrostatic pressure. The data from our crystal corresponded with the data found in the Cambridge Structural Database (CSD) for saccharin (CSD refcode: SCCHRN03).



**Figure 4.1** The asymmetric unit of saccharin with numbering scheme at ambient pressure. The colours of the elements are consistent throughout (carbon, grey; oxygen, red; sulphur, yellow; nitrogen, blue; hydrogen, white).

Saccharin is a dioxobenzothiazol and it crystallises in monoclinic space group  $P2_1/c$ , with  $Z'$  of 1. The crystal structure of saccharin consists of dimers formed via sulphonamide synthons with no other hydrogen bonding present. The saccharin dimers propagate in layers shown in Figure 4.2 a). At  $c = 0, 1$  etc. in Figure 4.2a) the dimers are pointing out of the plane and those at  $c = \frac{1}{2}$  are pointing into the plane.

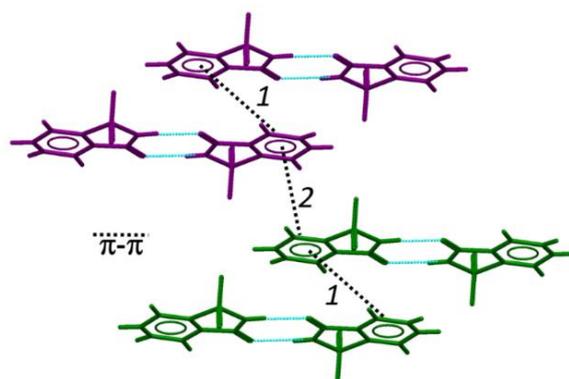


**Figure 4.2 a) View of the structure down the b-axis showing the hydrogen bonded dimers; the layers of dimers are numbered 1 & 2. b) View down the b-axis highlighting the layers in colour, corresponding to the numbers in a) purple, layer 1; green, layer 2. The dispersive interactions between the layers are weaker than the dispersive interactions in the layers.**

*In silico* calculations of molecule to molecule interactions in the structure were performed using Pixel.<sup>64,65</sup> The calculations confirm the strongest molecule to molecule interaction is associated with the sulphonamide dimer interaction in the *ab* direction (total,  $-57.4 \text{ kJ mol}^{-1}$ ; coulomb,  $-79.5 \text{ kJ mol}^{-1}$ ; intercentroid distance,  $6.865 \text{ \AA}$ ) which is to be expected. Reinforcing the layer of hydrogen bonded dimers is a strong dispersive force (Figure 4.3). The strongest dispersive force (denoted as 1 in Figure 4.3) exists in the layer and is associated with a  $\pi$ - $\pi$  interaction in the *a*-axis direction between neighbouring dimer pairs (total energy,  $-26.7 \text{ kJ mol}^{-1}$ ; dispersive,  $-32.2 \text{ kJ mol}^{-1}$ ; intercentroid distance,  $4.562 \text{ \AA}$ ). There is a larger intercentroid distance between interlayer saccharin molecules and the dispersive force is slightly less

## Lauren Connor

energetic (total energy,  $-25 \text{ kJ mol}^{-1}$ ; dispersive,  $-17.8 \text{ kJ mol}^{-1}$ ; intercentroid distance,  $5.777 \text{ \AA}$ ). The alternating  $4.562 \text{ \AA}$  and  $5.777 \text{ \AA}$  centroid to centroid  $\pi$ - $\pi$  stacking distance can be considered bonding. Centroid to centroid distances of up to  $6 \text{ \AA}$  are bonding, recent research has suggested that even distances above  $7 \text{ \AA}$  may be considered bonding.<sup>66</sup>

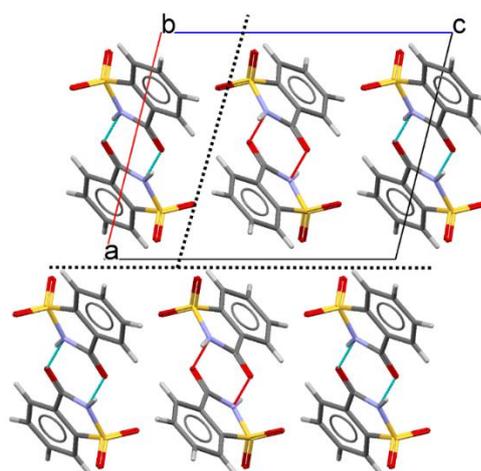


**Figure 4.3 Dimers of saccharin molecules interacting through  $\pi$ - $\pi$  forces in a layer (Interaction 1;  $4.562 \text{ \AA}$ ) and between layers (interaction 2;  $5.777 \text{ \AA}$ ). Colours shown corresponds with layers in Figure 4.2 b).**

### 4.4.2. Morphology predictions and face indexing

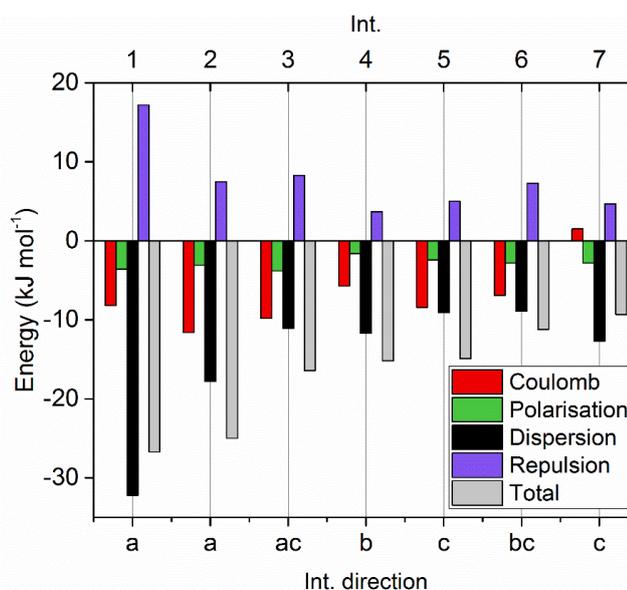
An understanding of the intermolecular interactions is complemented with Bravais Friedel Donnay Harker (BFDH) predicted morphology analysis. BFDH morphology predictions are based only on geometry and crystal lattice symmetry. The predictions do not consider the growth medium of the crystals and intermolecular interactions and forces in the structure are not modelled. It is useful to use Pixel calculations to justify BFDH models when the predicted morphology matches the observed. The predicted morphology of saccharin matches well with the observed morphology shown in Figure 4.4. The largest face is the  $(1\ 0\ 0)$  which is due to the layers acting as slip planes in this direction ( $a$ -axis). Slip planes may be considered as a specific direction in the crystal which is void of strong bonding and with the least intermolecular interaction of the axes directions. Figure 4.4 illustrates two potential slip planes in saccharin based on hydrogen bonding. Dispersive forces exist in the  $a$ -axis direction, as discussed previously. However, the observed and predicted morphology show the  $(1\ 0\ 0)$  face as the largest and therefore slowest growing face.





**Figure 4.5 Potential slip plane directions in saccharin shown by the dotted lines.**

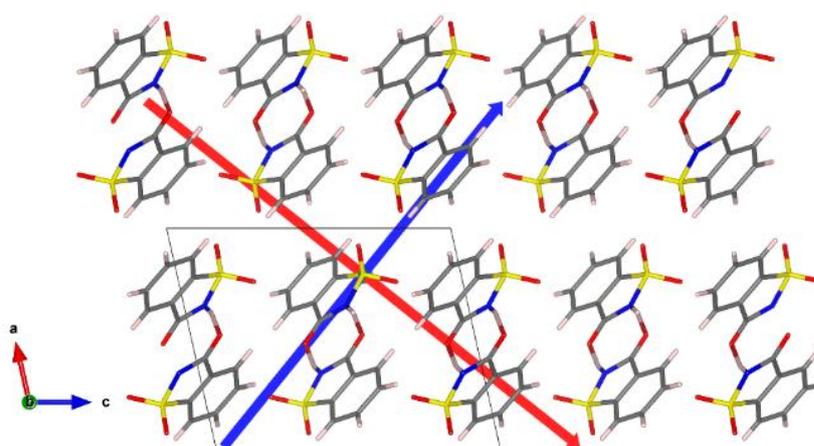
The highest strength interactions exist in the *a*-direction. However, we collated the total energies per direction (minus the strongest interaction in the dimer layer). The energies in each axes direction were *a*,  $-17.8 \text{ kJ mol}^{-1}$ ; *b*,  $-20.6 \text{ kJ mol}^{-1}$ ; *c*,  $-41.8 \text{ kJ mol}^{-1}$ . There are more dispersive forces in the *c*-axis direction than any other, justifying the slip plane in the *a*-axis direction rather than the *c*-axis direction. The slip plane in *a*-direction is responsible for the observed largest face  $(-1 \ 0 \ 0)$ . One aspect of the BFDH predicted morphology which differs slightly from the observed morphology is the smaller  $(0 \ 0 \ -1)$  face (Figure 4.4). One crystal is almost completely rhombohedral due to the fast growing  $(0 \ 0 \ 1)$  face. The accumulation of dispersive energy in the *c*-direction is responsible for the smaller  $(0 \ 0 \ 1)$  face. The dispersive energy promotes faster crystal growth in this direction.



**Figure 4.6 Pixel plot of the main forces in the crystal structure of saccharin. Interactions are numbered by total strength and labelled by direction of interaction.**

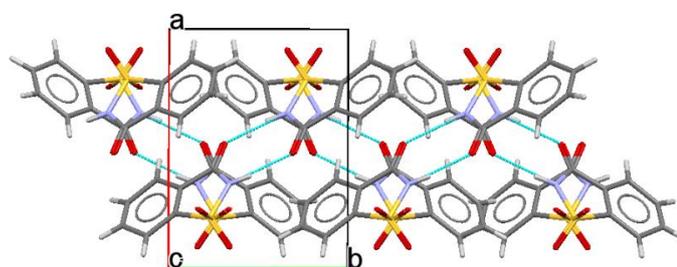
#### 4.4.3. Compression of saccharin

Saccharin compresses to 0.69 GPa without any phase transformation. Over this pressure range the unit cell is compressed by 2.69%, 0.76%, 2.67% in the *a*-, *b*-, and *c*-directions respectively. The *a*-direction and *c*-direction compresses to the same extent, the *a*-axis direction compressing marginally more. The measure of the principal axes of strain are invaluable to understand how low symmetry solids compress as the principal axes of strain do not necessarily have to be aligned to the unit cell axes. In this system, both the maximum (red) and medium (blue) strain are observed along the primarily along the *ac* direction and the minimum strain is calculated along the *b*-axis direction.



**Figure 4.7 Plot of linear strain showing principal axes of strain (maximum strain, red; medium strain; blue). The minimum strain is along the b-axis (into the page) as required by symmetry.**

Analysis of the unit cell axes provide insight into the movement and rearrangement of this structure upon compression. The extent of compression along the *a*- and *c*-axes are consistent with expectations of this structures' energetic profile. The major element of the dimer shape is along the *b*-axis, which, unlike the *a*- and *c*- lacks layering (Figure 4.8). The *b*-direction is reinforced by hydrogen bonding making it marginally more incompressible.

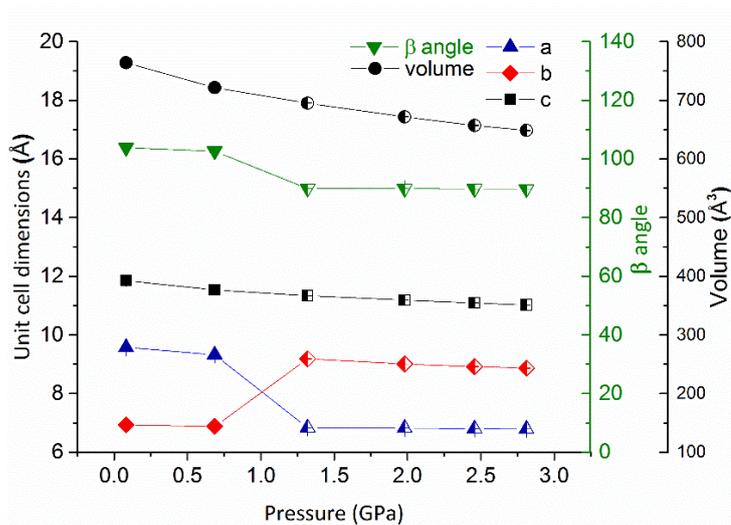


**Figure 4.8 View of the unit cell of saccharin down the c-axis showing the covalent and hydrogen bonding which propagates along the b-axis direction and is responsible for the least strain experienced in this direction.**

The interactions discussed and their layered structure could be held partly accountable for the willingness of the molecules to change position. Hence, saccharin undergoes a phase transition when under great stress and strain from the pressure it is subjected to in the *ac* direction only. The volume decrease of the initial compression of saccharin is prominent with a 5.56 % decrease between the first two points.

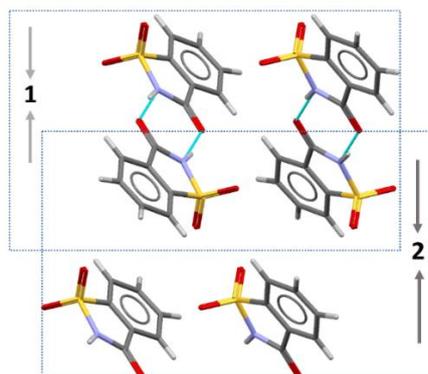
#### 4.4.4. Phase transformation

Between 0.69 GPa and 1.32 GPa, saccharin changes phase to a new high pressure polymorph. The dataset at 1.32 GPa revealed a new polymorph of saccharin which will be referred to as  $\beta$ -saccharin, herein. The cell setting of saccharin changes with the phase transition which switches cell lengths  $a$  and  $b$ . The notable decrease in the  $\beta$  angle signifies a polymorphic phase transition. For clarity, the cell lengths are plotted together alongside angle and unit cell volume data in Figure 4.9.  $\beta$ -saccharin packs in lower symmetry triclinic space group P-1 with  $Z' = 2$ . Over this transition the crystal split but the diffraction was of good enough quality to determine the structure.



**Figure 4.9 Unit cell lengths and volume of saccharin as a function of pressure. The majority of the errors are smaller than the symbol. Half filled symbols denote the new phase of saccharin.**

$\beta$ -saccharin is inherently more tightly packed than the lower pressure polymorph but there is little difference in the packing arrangement of the molecules. The most notable difference is a shift in the layer distances in the new form. The distances in the saccharin dimer layer and the interlayer distance has decreased by 2.76 % and 5.08 % respectively. The distances within the dimer layer (1 in Figure 4.11) and the interlayer distances (2 in Figure 4.11) were measured between centroids calculated from an average of 6 molecules in each of the layers.



**Figure 4.10** Arrows depict the decrease in layer distances. Distance 1 decreases from 5.537 Å at 0.08 GPa to 5.384 Å in the new form. Distance 2 decreases from 6.305 Å at 0.08 GPa to 6.000 Å in the new form. Hydrogen bonding shown between dimers.

The hydrogen bonding of saccharin remains consistent with the hydrogen bonded sulphonamide dimer motif in the ambient pressure phase of saccharin. Supramolecular synthons other than this cannot be expected from such a small molecule with limited donor/acceptor capability. *In silico* Pixel calculations reveal that the most energetic molecule to molecule interaction is attributed to the hydrogen bond. The energy attributed to the hydrogen bonded dimer is mainly coulombic ( $-97.6 \text{ kJ mol}^{-1}$ ) with a total energy of  $-59.9 \text{ kJ mol}^{-1}$ . Pixel calculations were also used to analyse the top dispersive interactions in the low and high pressure polymorphs of saccharin to gain insight into how the structure has changed. The decrease in the distance in the dimer layer has caused a steep increase in the repulsive nature of the energy attributed to the  $\pi$ - $\pi$  interaction as observed in interaction 1 in Figure 4.11. The total energy of the interaction remains constant at  $-28.3 \text{ kJ mol}^{-1}$  in the low pressure polymorph ( $\alpha$ -form) to  $\beta$ -saccharin. This is remarkable in and of itself as the pressure applied to the sample has increased by 1.24 GPa. There is a substantial increase in the interaction energy of the  $\pi$ - $\pi$  interaction (2 in Figure 4.11) in  $\beta$ -saccharin (Int. 2 corresponding to graphs in Figure 4.12).

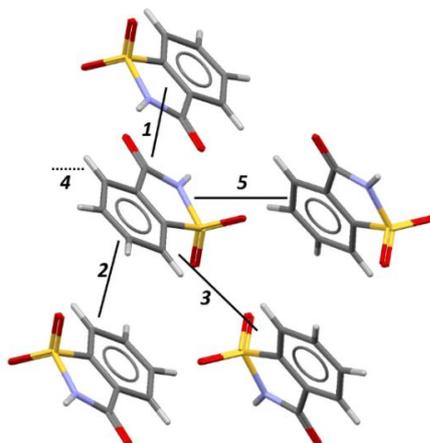
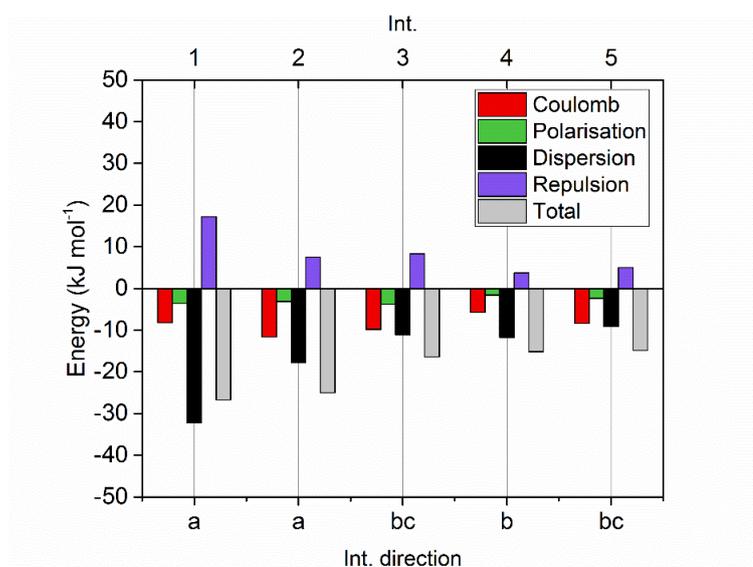


Figure 4.11 Dispersive molecule to molecule interactions are indicated by numbers which correspond to the numbers in Figure 4.12. The dashed line at interaction 4 indicates that the Pixel interaction is directly into the plane of the page, with another molecule of saccharin directly behind the centre molecule when viewed in this orientation.

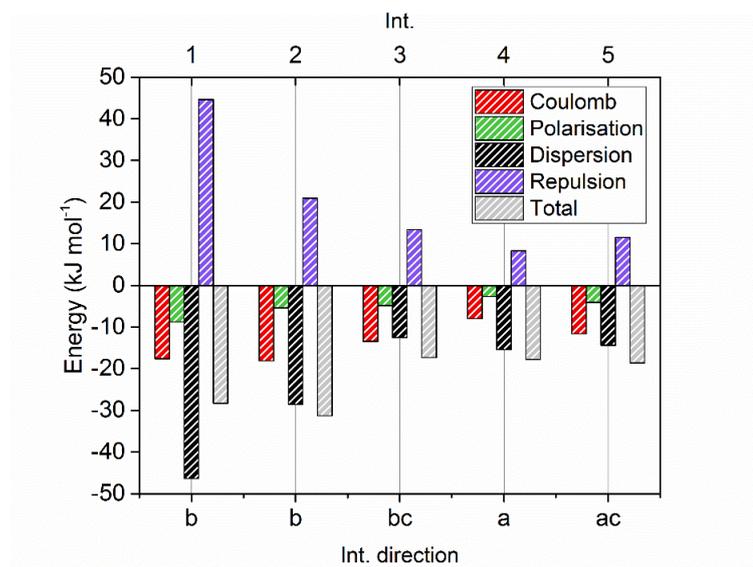
Table 4.2 Changes in energy between the high pressure (HP) and low pressure (LP) polymorph. Molecule to molecule interaction attributed to the  $\pi$ - $\pi$  interaction between saccharin dimer layers (Int. 2 corresponds to graphs in Figure 4.12).

Int. 2		Dist.	Coulomb	Polaris.	Disp.	Repul.	Total
		Å	kJ mol <sup>-1</sup>				
	LP	5.777	-11.6	-3.1	-17.8	7.5	-25
	HP	5.569	-18.1	-5.4	-28.6	20.9	-31.3

The energy attributed to the interaction has increased from -25 kJ mol<sup>-1</sup> in the low pressure polymorph to -31.3 kJ mol<sup>-1</sup> in  $\beta$ -saccharin. The energy associated with the  $\pi$ - $\pi$  interaction between the layers of saccharin dimers overtook the energy associated with the  $\pi$ - $\pi$  interaction in the hydrogen bonded layer (displayed in Figure 4.11 and Figure 4.12 as Int. 2). The most significant energetic difference observed in  $\beta$ -saccharin is the shortened distance between the, previously discussed, saccharin dimer layers. The numbering in Figure 4.11 corresponds to the numbering in Figure 4.12.



a)

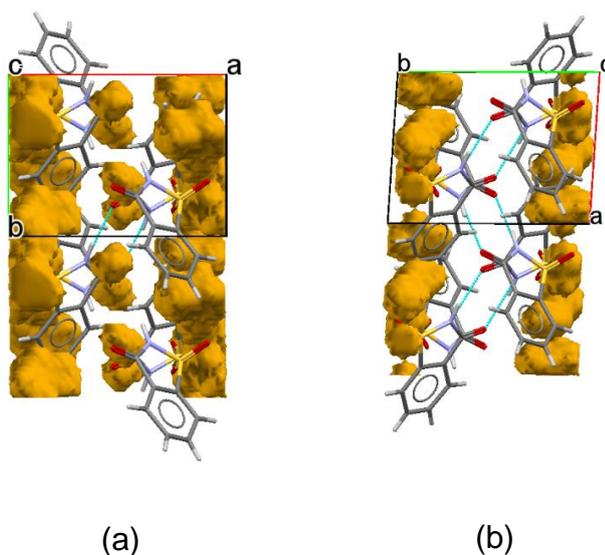


b)

Figure 4.12 Calculated Pixel energies of the five main interactions in  $\beta$ -saccharin. Coulombic, polarisation, dispersion, repulsion and total energies are plotted for each of the interactions 1-5, corresponding to Figure 4.11. The interaction directions are written as a function of the unit cell a) ambient pressure polymorph (at 0.08 GPa) in full colour; graph 1 b) high pressure polymorph (1.32 GPa) with pattern; graph 2.

#### 4.4.5. Void space analysis

The initial dataset at 0.08 GPa presented a void space of  $79.55 \text{ \AA}^3$ , 10.4 % of the unit cell. This void space reduced before the phase transition to  $55.46 \text{ \AA}^3$ , 7.7 % at 0.69 GPa. After the phase transition, at 1.32 GPa, the void space volume reduced to  $39.30 \text{ \AA}^3$ , 4.9 % of the total volume. The structure at 0.08 GPa reveals some void space in the saccharin dimer layer and significant void space at the edge of the unit cell between dimer layers, forming columns in the *b*-axis direction.  $\beta$ -saccharin presents no void spacing in the hydrogen bonded layer region direction. A significant reduction in void space at the unit cell edge is depicted in Figure 4.13 b) which is the location of the slip plane in the structure. The reduction in void space in the columns along *b*-axis of  $\beta$ -saccharin supports our hypothesis that the layered structure of saccharin with slip planes is less of a feature in the high pressure polymorph. In  $\beta$ -saccharin, the distance has decreased in this layer and the energetic interactions increased.



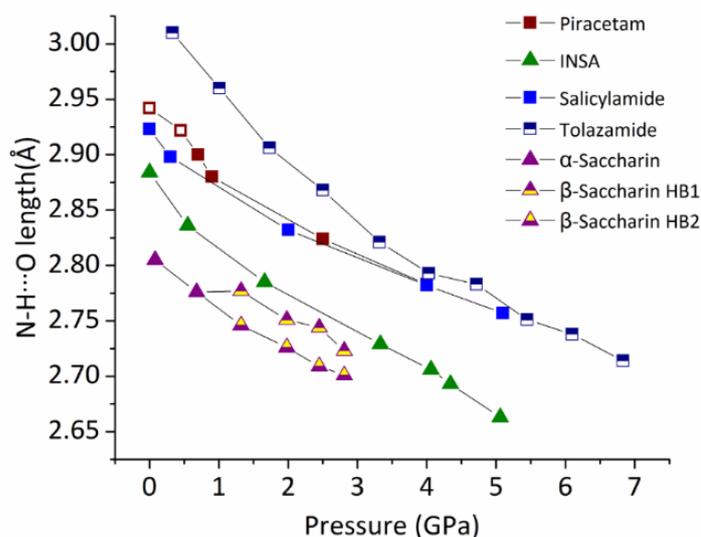
**Figure 4.13** Void space analysis with a probe of  $0.6 \text{ \AA}$  radius and  $0.2 \text{ \AA}$  grid spacing at 0.08 GPa before the phase transition and 1.32 GPa after the phase transition.

#### 4.4.6. Compression behaviour of amide dimer

The previous chapter in this thesis detailed a potential correlation between the rate of compression in the amide dimer  $\text{N-H}\cdots\text{O}$  contact and whether or not the dimer is involved in any secondary hydrogen bonding interactions. It is thought that any

## Lauren Connor

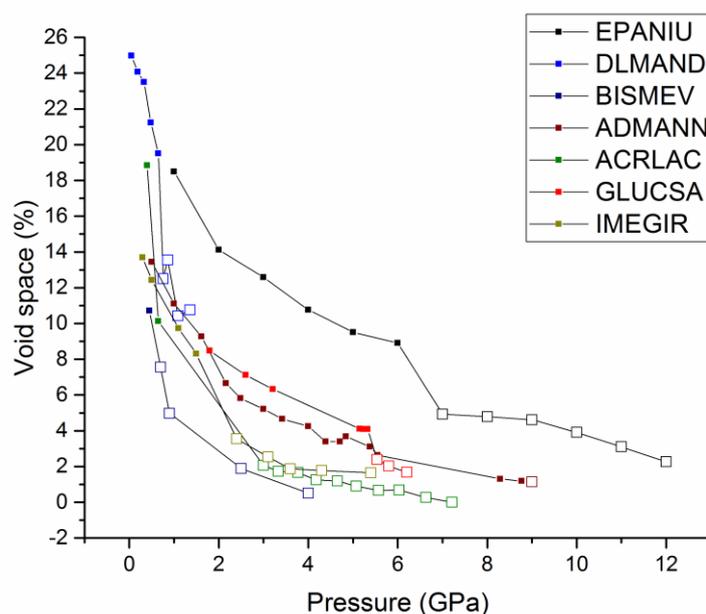
secondary hydrogen bonding will inherently be in an opposing direction and will stabilise the dimer interaction against shortening during compression. Of the 2454 structures in the CSD with pressure explicitly stated, there are only three compression studies on molecules containing an amide dimer. Piracetam<sup>37</sup> and Salicylamide<sup>67</sup> both contain a dimer that is involved in a secondary hydrogen bonding interaction and display a slower rate of compression. The dimer present in the structure of Tolazamide<sup>50</sup> consists of secondary amides and is not involved in any other hydrogen bonding, allowing for a faster rate of compression. The previous chapter in this thesis on a 1:1 co-crystal of indomethacin and saccharin (INSA) adds another compression study containing an amide dimer to this small sample size. The amide dimer in saccharin itself, like INSA and Tolazamide, is a secondary amide. The dimer is not stabilised against compression by another hydrogen bonding interaction and compresses at a similar rate. Shown in Figure 4.14, both  $\alpha$ -saccharin and  $\beta$ -saccharin display a similar rate of compression to Tolazamide and INSA. During the phase transition, the hydrogen bonding pattern of saccharin did not change. However, the loss of symmetry from a monoclinic  $P2_1/c$  cell to a triclinic  $P-1$  causes a difference in the lengths of the alternating (in and out of plane) amide dimer interactions. One dimer compression remains linear with the compression of the  $\alpha$ -saccharin (HB2 in Figure 4.14). The other hydrogen bond lengthens during the phase transition due to the strengthened interlayer  $\pi$ - $\pi$  interaction. Both hydrogen bonding contacts display a similar rate of compression from the first data point after the phase transition to 2.811 GPa.



**Figure 4.14** Plots showing how the length of piracetam (form II: half symbol & form V: filled symbol), INSA, tolazamide, and salicylamide’s N-H...O amide dimer respond to pressure. The compression of  $\alpha$  and  $\beta$  saccharin is plotted.  $\beta$ -saccharin contains amide dimers at two different lengths.

#### 4.4.7. Increasing density in phase changes at pressure

Crystal structures may undergo a phase transition to achieve a greater density in the crystalline lattice. A search of the CSD for polymorphs formed by high pressure means reveal a sharp decrease in void space at the point of transition with 6 of the 7 materials investigated. There does not seem to be any correlation between the size of the void space at ambient pressure and the relative void volume at the transition point. In terms of how the transformation occurs, i.e. through an entropic or enthalpic route, we do not have enough data points below the transition in order to be able to observe this using the Gibbs Free Energy equation:  $\Delta G = \Delta H - T\Delta S$  and using the enthalpy/pressure plot to identify the energy-driver.<sup>68</sup>



**Figure 4.15 Selected compression studies from the CSD showing a polymorphic phase transition in the compression series by an empty symbol<sup>37,38,43,69–72</sup>**

#### 4.5. Conclusion

Saccharin transforms to a new polymorph,  $\beta$ -saccharin, in the pressure range of 0.69 GPa and 1.32 GPa. The transformation fractures the crystal (monoclinic to triclinic system) but the domains remain large enough for the identification of the new phase.  $\beta$ -saccharin has a  $Z'=2$  structure where an increase in the interaction energies of the interplanar contacts is observed. The void space between the layers of the identified slip planes are reduced over the transition which validates our identification of these planes. Searches over the CSD reveal that saccharin's rate of compression in the amide dimer is similar to that of other dimers containing a pair of secondary amides. Primary amides seem to have a slower rate of compression as their nitrogen is involved in a secondary interaction directing away from the dimer bond and stabilising the dimer against compression.

## References

1. Cheng, R., Yang, H., Shao, M., Hu, T. & Zhou, X. Dental erosion and severe tooth decay related to soft drinks: a case report and literature review. *J. Zhejiang Univ. Sci. B* **10**, 395–399 (2009).
2. Malik, V., Schulze, M. & Hu, F. Sucrose-sweetened beverages increase fat storage in the liver, muscle, and visceral fat depot: a 6-mo randomized intervention study. *Am. J. Clin. Nutr.* **95**, 283–289 (2012).
3. Baker, P., Jones, A. & Thow, A. M. Accelerating the Worldwide Adoption of Sugar-Sweetened Beverage Taxes: Strengthening Commitment and Capacity Comment on 'The Untapped Power of Soda Taxes: Incentivizing Consumers, Generating Revenue, and Altering Corporate Behavior'. *Int. J. Heal. Policy Manag.* **7**, 474–478 (2017).
4. Evans, C. How successful will the sugar levy be in improving diet and reducing inequalities in health? *Perspect. Public Health* **138**, 85–86 (2018).
5. Huff, J. IARC Monographs, industry influence, and upgrading, downgrading, and under-grading chemicals: A personal point of view. *Int. J. Occup. Environ. Health* **8**, 249–270 (2002).
6. Banerjee, R., Bhatt, P. M., Ravindra, N. V. & Desiraju, G. R. Saccharin salts of active pharmaceutical ingredients, their crystal structures, and increased water solubilities. *Cryst. Growth Des.* **5**, 2299–2309 (2005).
7. Gao, Y. *et al.* Coformer selection based on degradation pathway of drugs: A case study of adefovir dipivoxil-saccharin and adefovir dipivoxil-nicotinamide cocrystals. *Int. J. Pharm.* **438**, 327–335 (2012).
8. Basavoju, S., Boström, D. & Velaga, S. P. Indomethacin-saccharin cocrystal: Design, synthesis and preliminary pharmaceutical characterization. *Pharm. Res.* **25**, 530–541 (2008).
9. Jung, M. *et al.* Bioavailability of indomethacin-saccharin cocrystals. *J. Pharm. Pharmacology* **62**, 1560–1568 (2010).
10. Kudo, S. & Takiyama, H. Production method of carbamazepine/saccharin cocrystal particles by using two solution mixing based on the ternary phase diagram. *J. Cryst. Growth* **392**, 87–91 (2014).
11. Balbani, A. P. S., Stelzer, L. B. & Montovani, J. C. Pharmaceutical excipients and the information on drug labels. *Braz. J. Otorhinolaryngol.* **72**, 400–406 (2006).
12. Pagire, S. K., Jadav, N., Vangala, V. R., Whiteside, B. & Paradkar, A. Thermodynamic Investigation of Carbamazepine-Saccharin Co-Crystal Polymorphs. *J. Pharm. Sci.* **106**, 2009–2014 (2017).
13. Aitipamula, S., Chow, P. S. & Tan, R. B. H. Dimorphs of a 1: 1 cocrystal of ethenzamide and saccharin: Solid-state grinding methods result in metastable polymorph. *CrystEngComm* **11**, 889–895 (2009).
14. Bart, J. C. J. The Crystal and Molecular Structure of Saccharin (o-Sulphobenzoic imide). *J. Chem. Soc. B Phys. Org.* 376 (1968).
15. Patyk, E., Skumiel, J., Podsiado, M. & Katrusiak, A. High-pressure (+)-sucrose polymorph. *Angew. Chemie - Int. Ed.* **51**, 2146–2150 (2012).
16. Cruz-Cabeza, A. J., Reutzel-Edens, S. M. & Bernstein, J. Facts and fictions about polymorphism. *Chem. Soc. Rev.* **44**, 8619–8635 (2015).
17. Bernstein, J. DISAPPEARING POLYMORPHS. WHERE DID THEY COME FROM? Joel Bernstein. 87–96

18. Thomas, L. H., Wales, C., Zhao, L. & Wilson, C. C. Paracetamol form II: An elusive polymorph through facile multicomponent crystallization routes. *Cryst. Growth Des.* **11**, 1450–1452 (2011).
19. Lou, B., Boström, D. & Velaga, S. P. Polymorph control of felodipine form II in an attempted cocrystallization. *Cryst. Growth Des.* **9**, 1254–1257 (2009).
20. Rafilovich, M. & Bernstein, J. Serendipity and four polymorphic structures of benzidine, C<sub>12</sub>H<sub>12</sub>N<sub>2</sub>. *J. Am. Chem. Soc.* **128**, 12185–12191 (2006).
21. Saha, B. K., Banerjee, R., Nangia, A. & Desiraju, G. R. A 1:1 molecular complex of 4-methylpyridine N-oxide and saccharin. *Acta Crystallogr. Sect. E Struct. Reports Online* **62**, (2006).
22. Eccles, K. S., Elcoate, C. J., Stokes, S. P., Maguire, A. R. & Lawrence, S. E. Sulfoxides: Potent co-crystal formers. *Cryst. Growth Des.* **10**, 4243–4245 (2010).
23. Shiraki, K., Takata, N., Takano, R., Hayashi, Y. & Terada, K. Dissolution improvement and the mechanism of the improvement from cocrystallization of poorly water-soluble compounds. *Pharm. Res.* **25**, 2581–2592 (2008).
24. Sanphui, P., Babu, N. J. & Nangia, A. Temozolomide cocrystals with carboxamide cofomers. *Cryst. Growth Des.* **13**, 2208–2219 (2013).
25. Caira, M. R. *et al.* Co-crystals of the antiretroviral nevirapine: Crystal structures, thermal analysis and dissolution behaviour. *CrystEngComm* **14**, 2541–2551 (2012).
26. Chadha, R., Bhandari, S., Haneef, J., Khullar, S. & Mandal, S. Cocrystals of telmisartan: Characterization, structure elucidation, in vivo and toxicity studies. *CrystEngComm* **16**, 8375–8389 (2014).
27. Oliveira, M. A., Peterson, M. L. & Davey, R. J. Relative enthalpy of formation for Co-crystals of small organic molecules. *Cryst. Growth Des.* **11**, 449–457 (2011).
28. Velaga, S. P., Basavoju, S. & Boström, D. Norfloxacin saccharinate-saccharin dihydrate cocrystal - A new pharmaceutical cocrystal with an organic counter ion. *J. Mol. Struct.* **889**, 150–153 (2008).
29. Lu, N., Tu, W. H., Shing, J. S., Wen, Y. S. & Liu, L. K. New type of saccharinate crystals: A possible supramolecular synthon. *J. Mol. Struct.* **980**, 101–107 (2010).
30. Fleischman, S. G. *et al.* Crystal Engineering of the Composition of Pharmaceutical Phases: Multiple-Component Crystalline Solids Involving Carbamazepine. *Cryst. Growth Des.* **3**, 909–919 (2003).
31. Babu, N. J., Reddy, L. S. & Nangia, A. Amide N-oxide heterosynthon and amide dimer homosynthon in cocrystals of carboxamide drugs and pyridine N-oxides. *Mol. Pharm.* **4**, 417–434 (2007).
32. Takata, N., Takano, R., Uekusa, H., Hayashi, Y. & Terada, K. A spironolactone-saccharin 1:1 cocrystal hemihydrate. *Cryst. Growth Des.* **10**, 2116–2122 (2010).
33. Lu, E., Rodríguez-Hornedo, N. & Suryanarayanan, R. A rapid thermal method for cocrystal screening. *CrystEngComm* **10**, 665–668 (2008).
34. Corpinot, M. K. *et al.* Are oxygen and sulfur atoms structurally equivalent in organic crystals? *Cryst. Growth Des.* **17**, 827–833 (2017).
35. Bhatt, P. M., Ravindra, N. V., Banerjee, R. & Desiraju, G. R. Saccharin as a salt former. Enhanced solubilities of saccharinates of active pharmaceutical ingredients. *Chem. Commun.* 1073–1075 (2005). doi:10.1039/b416137h

36. Porter, W. W., Elie, S. C. & Matzger, A. J. Polymorphism in carbamazepine cocrystals. *Cryst. Growth Des.* **8**, 14–16 (2008).
37. Fabbiani, F. P. A. *et al.* High-Pressure Studies of Pharmaceuticals: An Exploration of the Behavior of Piracetam. *Cryst. Growth Des.* **7**, 1115–1124 (2007).
38. Oswald, I. D. H. & Urquhart, A. J. Polymorphism and polymerisation of acrylic and methacrylic acid at high pressure. *CrystEngComm* **13**, 4503–4507 (2011).
39. Johnstone, R. D. L. *et al.* High-pressure polymorphism in salicylamide. *CrystEngComm* **12**, 983–1316 (2010).
40. Zakharov, B. a., Losev, E. a. & Boldyreva, E. V. Polymorphism of “glycine–glutaric acid” co-crystals: the same phase at low temperatures and high pressures. *CrystEngComm* **15**, 1693 (2013).
41. Boldyreva, E. V. High-pressure diffraction studies of molecular organic solids . A personal view. *Acta Crystallogr. A.* **64**, 218–231 (2008).
42. Podsiadło, M., Olejniczak, A. & Katrusiak, A. Structure-Property Relations and Polymorphism in Compressed Methylamines. *Cryst. Growth Des.* **17**, 2218–2222 (2017).
43. Patyk, E. & Katrusiak, A. Transformable H-bonds and conformation in compressed glucose. *Chem. Sci.* **6**, 1991–1995 (2015).
44. Dziubek, K., Podsiadło, M. & Katrusiak, A. Nearly isostructural polymorphs of ethynylbenzene: Resolution of :CH...(arene) and cooperative =CH...(C=C) interactions by pressure freezing. *J. Am. Chem. Soc.* **129**, 12620–12621 (2007).
45. Oswald, I. D. H. & Pulham, C. R. Co-crystallisation at high pressure—an additional tool for the preparation and study of co-crystals. *CrystEngComm* **10**, 1114 (2008).
46. Fabbiani, F. P. A. *et al.* Pressure-induced formation of a solvate of paracetamol. *Chem. Commun.* **3**, 3004–3005 (2003).
47. Connor, L. E. *et al.* The ecstasy and the agony; compression studies of 3,4-methylenedioxymethamphetamine (MDMA). *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **71**, 3–9 (2015).
48. Kolesov, B. a, Mikhailenko, M. a & Boldyreva, E. V. Dynamics of the intermolecular hydrogen bonds in the polymorphs of paracetamol in relation to crystal packing and conformational transitions: a variable-temperature polarized Raman spectroscopy study. *Phys. Chem. Chem. Phys.* **13**, 14243–14253 (2011).
49. Boldyreva, E. V. High-pressure-induced structural changes in molecular crystals preserving the space group symmetry: anisotropic distortion/isosymmetric polymorphism. *Cryst. Eng.* **6**, 235–254 (2003).
50. Fedorov, A. Y. *et al.* Effect of pressure on two polymorphs of tolazamide: Why no interconversion? *CrystEngComm* **19**, 2243–2252 (2017).
51. Merrill, L. & Bassett, W. A. Miniature diamond anvil pressure cell for single crystal x-ray diffraction studies. *Rev. Sci. Instrum.* **45**, 290–294 (1974).
52. Piermarini, G. J., Block, S., Barnett, J. D. & Forman, R. a. Calibration of the Pressure Dependence of the R//1 Ruby Fluorescence Line To 195 Kbar. *J. Appl. Phys.* **46**, 2774–2780 (1975).
53. Tateiwa, N. & Haga, Y. Appropriate pressure-transmitting media for cryogenic experiment in the diamond anvil cell up to 10 GPa. *J. Phys. Conf. Ser.* **215**, 012178 (2010).
54. Nowell, H., Barnett, S. A., Christensen, K. E., Teat, S. J. & Allan, D. R. I19, the small-molecule single-crystal diffraction beamline at Diamond Light Source. *J. Synchrotron Radiat.* **19**, 435–441 (2012).

55. Winter, G., Lobley, C. M. C. & Prince, S. M. Decision making in xia2. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1260–1273 (2013).
56. Rigaku OD. CrystalisPro 1.171.38.41k. (2015).
57. Dolomanov, O. V, Bourhis, L. J., Gildea, R. J., Howard, J. A. K. & Puschmann, H. OLEX2: a complete structure solution, refinement and analysis program. 339–341 (2009).
58. Thorn, A., Dittrich, B. & Sheldrick, G. M. Enhanced rigid-bond restraints. *Acta Crystallogr. Sect. A Found. Crystallogr.* **68**, 448–451 (2012).
59. Sheldrick, G. M. SHELXT - Integrated space-group and crystal-structure determination. *Acta Crystallogr. Sect. A Found. Crystallogr.* **71**, 3–8 (2015).
60. Sheldrick, G. M. Crystal structure refinement with SHELXL. *Acta Crystallogr. Sect. C Struct. Chem.* **71**, 3–8 (2015).
61. Okaya, Y. The crystal structure of saccharin o -sulfobenzoimide, C<sub>6</sub> H<sub>4</sub> CO.NH.SO<sub>2</sub>, an artificial sweetening. *Acta Crystallogr. Sect. B Struct. Crystallogr. Cryst. Chem.* **25**, 2257–2263 (1969).
62. Wardell, J. L., Low, J. N. & Glidewell, C. Saccharin, redetermined at 120 K: a threedimensional hydrogen-bonded framework. *Acta Crystallogr. Sect. C Cryst. Struct. Commun.* **62**, (2006).
63. Eccles, K. S. *et al.* Evaluation of the Bruker SMART X2S: Crystallography for the nonspecialist? *J. Appl. Crystallogr.* **44**, 213–215 (2011).
64. Gavezzotti, A. Calculation of Intermolecular Interaction Energies by Direct Numerical Integration over Electron Densities . 2 . An Improved Polarization Model and the Evaluation of Dispersion and Repulsion Energies. *J. Phys. Chem.* **107**, 2344–2353 (2003).
65. Maschio, L., Civalieri, B., Ugliengo, P. & Gavezzotti, A. Intermolecular Interaction Energies in Molecular Crystals : Comparison and Agreement of Localized Møller Å Plesset 2 , Dispersion-Corrected Density Functional , and Classical Empirical Two-Body Calculations. *J. Phys. Chem.* **115**, 11179–11186 (2011).
66. Kruszynski, R. & Sierański, T. Can Stacking Interactions Exist beyond the Commonly Accepted Limits? *Cryst. Growth Des.* **16**, 587–595 (2016).
67. Johnstone, R. D. L. *et al.* High-pressure polymorphism in salicylamide. *CrystEngComm* **12**, 1065–1078 (2010).
68. Moggach, S. A., Parsons, S. & Wood, P. A. High-pressure polymorphism in amino acids. *Crystallography Reviews* **14**, (2008).
69. Woodall, C. H. *et al.* High-pressure crystallographic and spectroscopic studies on two molecular dithienylethene switches. *CrystEngComm* **16**, 2119–2128 (2014).
70. Patyk-Kaźmierczak, E., Warren, M. R., Allan, D. R. & Katrusiak, A. Intermolecular Contacts in Compressed α- D- Mannose. *Cryst. Growth Des.* **16**, 6885–6890 (2016).
71. Cai, W., Marciniak, J., Andrzejewski, M. & Katrusiak, A. Pressure effect on d,l-Mandelic acid racemate crystallization. *J. Phys. Chem. C* **117**, 7279–7285 (2013).
72. Zakharov, B. A. & Boldyreva, E. V. A high-pressure single-crystal to single-crystal phase transition in dl-alaninium semi-oxalate mono-hydrate with switching-over hydrogen bonds. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **69**, 271–280 (2013).

***-Chapter Five-***

***Compression of 1:1 and 2:1 co-crystals of benzoic acid and  
isonicotinamide***

### 5. Compression of 1:1 and 2:1 co-crystals of benzoic acid and isonicotinamide

#### 5.1. Synopsis

This study attempted to evaluate and compare the compression behaviour of different stoichiometries of benzoic acid and isonicotinamide (1:1 & 2:1). Formation of the 1:1 co-crystal was straight forward and produced decent quality co-crystals for diffraction. Formation of 2:1 co-crystals suitable for diffraction was a challenge. Using the sealed sample chamber of a Diamond Anvil Cell (DAC) as a vessel, temperature cycling produced crystals of improved quality for diffraction. The 1:1 was solved in a monoclinic cell and the 2:1 was solved in a triclinic cell in line with previous literature on 1:1 and 2:1 benzoic acid and isonicotinamide.<sup>1</sup> The co-crystals remained in their respective space groups throughout the compression. It was possible to elucidate the compression behaviour in the form of a 'wine-rack-effect' for both stoichiometric systems. Each of the co-crystalline systems exhibited a stacked layer structure exhibiting slip planes between each layer. The slip planes lead to a large decrease in the axis corresponding to the stacked layers in each of the co-crystalline systems. The other two axes in each co-crystalline system reduced to a notably lesser extent upon compression. Long hydrogen bonds acted like a screw in a hinge between perpendicular chains of the molecules. The hinge mechanism lead to the wine-rack effect in the compressed systems. Each stoichiometric system displayed a plateau of the axes at higher pressures due to the wine-rack effect.

#### 5.2. Introduction

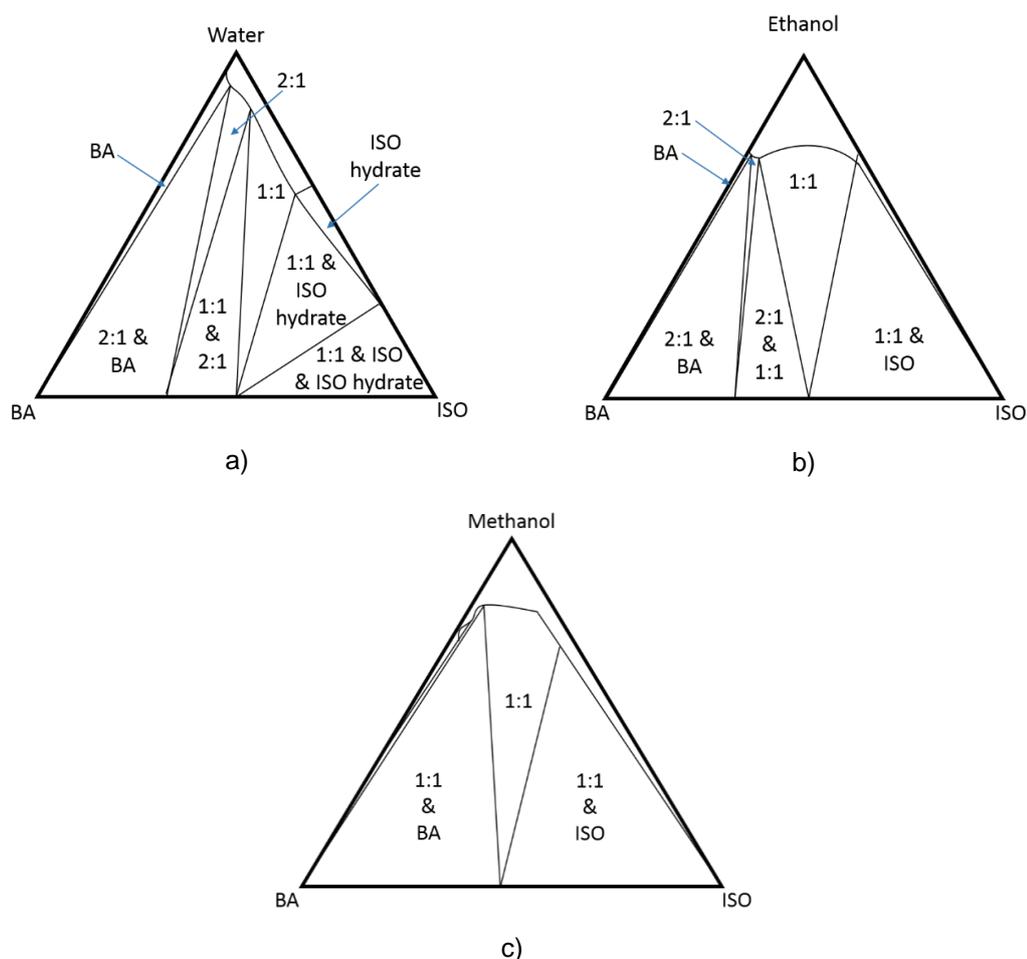
Multicomponent complexes are one route of many which can modify the physicochemical properties of a given compound. The design of these complexes is mostly based around the supramolecular synthon. The influence and importance of co-crystals has boomed over the past few years along with the literature on the topic.<sup>2-8</sup> Often, their improved physicochemical properties are analysed,<sup>8</sup> their comparison to the original single component carried out<sup>9</sup> and routes exploring the avenues to create co-crystals discussed.<sup>10</sup> The fundamental behaviour of co-crystals is of interest to crystal engineers as the co-crystal can open up new patterns of intermolecular interactions to explore. Knowledge of supramolecular synthons and how they behave are key to the 'supramolecular synthon' approach, applied by many with the aim of creating a new co-crystal.

## Lauren Connor

Studies on the influence of stoichiometry for a co-crystal have been conducted<sup>11–15</sup> but as of yet, there are no definitively outlined trends across all compounds on the influence of, for example, the addition of a more soluble co-former in larger stoichiometric ratios to improve solubility. In 2009, Bethune *et al.*<sup>16</sup> worked on co-crystals of carbamazepine: salicylic acid and carbamazepine: para-aminobenzoic acid in an attempt to show that co-crystals of carbamazepine can achieve pH-dependent solubility when co-crystallised with ionisable co-formers. However, there has been some recent study on one of the systems Bethune *et al.* worked on. In 2016, Li *et al.*<sup>13</sup> found that adding more para-aminobenzoic acid to carbamazepine in stoichiometric ratios of 1:1, 1:2 and 1:4 did not improve the dissolution in an expected pattern. Instead, the 1:4 co-crystal revealed the highest dissolution rate, followed by the 1:1 and then the 1:2.

In general, the synthesis of co-crystals is typically undertaken using a mix of the components at specific stoichiometric ratios to produce a co-crystal of consistent ratio. Saikia *et al.*<sup>17</sup> demonstrated the ability of theophylline and *o/m/p*-aminobenzoic acid to form different stoichiometric ratio co-crystals by liquid assisted grinding. The *m*- and *p*- co-former only allowed for formation of a 1:1 co-crystal via liquid assisted mechanical grinding whereas the *o*-aminobenzoic acid allowed for formation of four different stoichiometric ratios with solvent inclusion. The formation of co-crystals is dependent on many factors. In the search for co-crystals, the specific experimental region of formation can sometimes be difficult to hit.<sup>4,18</sup> In solution based co-crystallisation, it is evident that solubility of the starting components can influence the final stoichiometric ratio. Crystal engineers are making use of novel techniques to manipulate access to regions on the ternary phase diagram. Apshingekar *et al.*<sup>19</sup> demonstrated the use of ultrasound to navigate the ternary phase diagram when producing a co-crystal of caffeine and maleic acid by slurry methods in water.

The ternary phase diagrams for the formation of 2:1 and 1:1 co-crystals of benzoic acid and isonicotinamide have been studied.<sup>1,10,20</sup> Phase diagrams of the different stoichiometries in water, ethanol and methanol show a drastic difference as is shown in Figure 5.1. The formation of the 2:1 can be a difficult region to reach. The pure 2:1 region is only substantial in water (Figure 5.1). Svoboda *et al.* recently made use of a combined reactive and antisolvent approach to preferentially produce co-crystals of 2:1 and 1:1 depending on the conditions.<sup>10</sup> The study made use of the solubilities of the co-formers in different solvents (water and ethanol).



**Figure 5.1 Ternary phase diagrams for the formation of benzoic acid and isonicotinamide in a) water<sup>10</sup> b) ethanol<sup>10</sup> and c) methanol<sup>1</sup>**

As well as the analysis of the ternary phase and methods of formation, the benzoic acid and isonicotinamide co-crystal systems have been structurally characterised,<sup>1</sup> thermally analysed,<sup>18</sup> and their structural landscape explored.<sup>21</sup> The structural characterisation of the co-crystal systems revealed a difference in the bonding patterns. The 1:1 co-crystal forms an isonicotinamide dimer and is then capped at either side by an acid-pyridine interaction to form the 1:1. The 2:1 opened up a different supramolecular synthon, the acid: amide dimer. Some twinning was noted in the 2:1 co-crystal. In terms of thermal behaviour, Buanz *et al.* conducted thermal analysis which suggested that upon heating, the 2:1 co-crystal expands and then converts to the 1:1. Dubey *et al.* explored the structural landscape of the 1:1 co-crystal of benzoic acid and isonicotinamide. One route of exploration that has not been recorded in the literature is high pressure. High pressure techniques on multicomponent systems has been known to trigger phase transitions<sup>22–25</sup> and can be

## Lauren Connor

a useful tool to explore the structure-property relationship. The structure, molecular weight and molecular shape of benzoic acid and isonicotinamide as individual components are similar, and so, the compression behaviour of the two systems in comparison is a point of interest.

### 5.3. Experimental

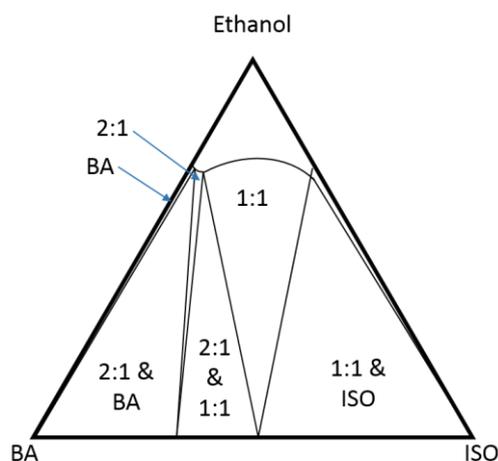
#### 5.3.1. The 1:1 co-crystal of benzoic acid and isonicotinamide

##### 5.3.1.1. General procedure

A Merrill-Bassett diamond-anvil cell (DAC) with a half-opening angle of  $40^\circ$  was used for the experiments. Diamond anvils with culet faces of  $600\mu\text{m}$  diameter sat in Bohler-Almax designed tungsten carbide backing seats. A  $300\mu\text{m}$  diameter hole was drilled (Almax EasyLab, Bohler  $\mu$ Driller) into a tungsten gasket to serve as the sample chamber. The sample chamber (flush with the backing anvil diamonds) was loaded with the sample and ruby sphere. A Horiba XplorA Raman spectrometer equipped with a 532nm laser was used to measure the ruby fluorescence. Throughout this study, 35:65 petroleum ether was used as the preferred pressure-transmitting medium (PTM). Petroleum ether has been known to maintain hydrostaticity for pressures of up to 6 GPa.<sup>26</sup>

##### 5.3.1.2. Preparation

A sample of 1:1 molar ratio benzoic acid and isonicotinamide co-crystals were formed according to the conditions outlined by Boyd et al.<sup>10,20</sup> Benzoic acid and isonicotinamide were added in a 1:1 molar ratio to a small amount of 99.9% ethanol, anhydrous. The vial was then closed with a pierced cap, permitting evaporation of solvent.



**Figure 5.2 Ternary phase diagram showing regions of the co-crystal formation in ethanol as produced by Svoboda *et al.***

#### 5.3.1.3. Ambient pressure single crystal X-ray diffraction

At ambient pressure, a crystal of ISBA11 was mounted onto a MiTiGen microloop of 100 $\mu\text{m}$  diameter. Silicon oil coated the co-crystal to aid in preparation and avoid crystal movement during SCXRD runs. X-ray diffraction intensities were collected from a sequence of four scans using a Bruker Apex II diffractometer with an Incotec 1 $\mu\text{S}$  microsource, Mo-K $\alpha$  radiation ( $\lambda = 0.71073\text{\AA}$ ), and equipped with an Oxford Cryosystems low temperature device operating at 123K. Data were reduced using SAINT as incorporated in Bruker Apex II Software.<sup>27</sup> An absorption correction was applied in two stages using both SHADE and SADABS, respectively.<sup>28,29</sup> The data were solved initially by intrinsic phasing and refined using *Olex2*<sup>30</sup> software ( $F^2$ ). All heavy atoms were anisotropically refined. All hydrogen atoms were placed in calculated positions and constrained to ride their parent atom.

#### 5.3.1.4. Diamond anvil cell preparation

For the compression study a needle shaped crystal of the 1:1 was loaded into the DAC with petroleum ether (35:65) to allow for hydrostatic compression before the sample chamber was sealed; the initial sealing of the cell introduced a pressure of 1.11 GPa.

### 5.3.1.5. High Pressure Studies

At each sequential increase in pressure, the cell was mounted onto the diffractometer for SCXRD. The high-pressure data were collected and treated in a similar manner to the ambient procedure described above. The X-ray diffraction intensities were collected from a sequence of twelve scans at 293K following the procedure described by Dawson *et al.* Data were indexed and integrated using SAINT in conjunction with dynamic masks.<sup>31</sup> All atoms were anisotropically refined. Hydrogen atoms were placed in calculated positions and constrained to ride their parent atom as per procedure in *Olex2*. RIGU vibrational restraints<sup>32</sup> were applied to all non-ambient pressure datasets.

The pressure of the cell was measured using fluorescence measurement of the ruby loaded into the DAC.<sup>33</sup> Data were collected at, 1.11 GPa, 2.29 GPa, 3.04 GPa, 3.77 GPa, 4.69 GPa and 5.32 GPa. A further dataset at 6.07 GPa was taken however the data were too poor for satisfactory refinement. RIGU restraints with a standard deviation of  $0.002 \text{ \AA}^2$  were applied to the data at 3.77 GPa, 4.69 GPa and 5.32 GPa.

**Table 5.1 Crystallographic tables for the 1:1 from ambient to 2.29 GPa**

For all structures:  $C_6H_6N_2O \cdot C_7H_6O_2$ ,  $M_r = 244.24$ , Monoclinic,  $C2/c$ ,  $Z = 16$ . Experiments were carried out with Mo  $K\alpha$  radiation using a Bruker SMART APEX2 area detector. H-atom parameters were constrained.

1:1 series	Ambient, 123 K	1.11 GPa, 297K	2.29 GPa, 297K
$a, b, c$ (Å)	22.764 (2), 5.1698 (4), 20.4939 (19)	21.494 (2), 5.0654 (2), 20.3705 (9)	20.502 (3), 4.9994 (3), 20.3196 (13)
$\beta$ (°)	96.703 (3)	97.742 (7)	98.353 (9)
$V$ (Å <sup>3</sup> )	2395.3 (4)	2197.6 (3)	2060.6 (4)
$\mu$ (mm <sup>-1</sup> )	0.10	0.11	0.11
Crystal size (mm)	0.2 × 0.07 × 0.01	0.25 × 0.04 × 0.03	0.25 × 0.04 × 0.03
Absorption correction	Multi-scan SADABS2014/3 (Bruker,2014/3) was used for absorption correction. $wR2(int)$ was 0.1098 before and 0.0399 after correction. The Ratio of minimum to maximum transmission is 0.9391. The $\lambda/2$ correction factor is Not present.	Multi-scan SADABS2014/3 (Bruker,2014/3) was used for absorption correction. $wR2(int)$ was 0.0719 before and 0.0459 after correction. The Ratio of minimum to maximum transmission is 0.9299. The $\lambda/2$ correction factor is Not present.	Multi-scan SADABS2014/3 (Bruker,2014/3) was used for absorption correction. $wR2(int)$ was 0.0688 before and 0.0499 after correction. The Ratio of minimum to maximum transmission is 0.8778. The $\lambda/2$ correction factor is Not present.
$T_{min}, T_{max}$	0.700, 0.746	0.693, 0.745	0.654, 0.745
No. of measured, independent and observed [ $I > 2.0\sigma(I)$ ] reflections	11741, 2658, 2016	4882, 632, 502	4463, 599, 497
$R_{int}$	0.024	0.043	0.044
$\theta_{max}$ (°)	26	23.3	23.3
$(\sin \theta/\lambda)_{max}$ (Å <sup>-1</sup> )	0.643	0.555	0.556
$R[F > 2\sigma(F)]$ , $wR(F)$ , $S$	0.040, 0.116, 1.01	0.034, 0.086, 1.07	0.034, 0.081, 1.05
No. of reflections	2658	632	599
No. of parameters	172	196	196
No. of restraints	5	120	120
$\Delta\rho_{max}, \Delta\rho_{min}$ (e Å <sup>-3</sup> )	0.21, -0.18	0.11, -0.10	0.10, -0.10

Computer programs: SAINTE v8.34A<sup>34</sup>, XT<sup>35</sup>, XL<sup>36</sup>, Olex2<sup>30</sup>

Table 5.1 *Contd.* Crystallographic tables for the 1:1 at 3.05 GPa and 3.77 GPa

1:1 series	3.05 GPa, 297K	3.77 GPa, 297K
$a, b, c$ (Å)	20.029 (2), 4.9652 (2), 20.2877 (9)	19.472 (5), 4.9289 (5), 20.2583 (18)
$\beta$ (°)	98.568 (7)	98.708 (15)
$V$ (Å <sup>3</sup> )	1995.1 (3)	1921.9 (6)
$\mu$ (mm <sup>-1</sup> )	0.12	0.12
Crystal size (mm)	0.25 × 0.04 × 0.03	0.25 × 0.04 × 0.03
Absorption correction	Multi-scan SADABS2014/3 (Bruker,2014/3) was used for absorption correction. $wR2(int)$ was 0.0616 before and 0.0446 after correction. The Ratio of minimum to maximum transmission is 0.9263. The $\lambda/2$ correction factor is Not present.	Multi-scan SADABS2014/3 (Bruker,2014/3) was used for absorption correction. $wR2(int)$ was 0.0694 before and 0.0509 after correction. The Ratio of minimum to maximum transmission is 0.8790. The $\lambda/2$ correction factor is Not present.
$T_{min}, T_{max}$	0.690, 0.745	0.655, 0.745
No. of measured, independent and observed [ $I > 2.0\sigma(I)$ ] reflections	4603, 581, 487	4215, 559, 452
$R_{int}$	0.041	0.050
$\theta_{max}$ (°)	23.3	23.3
$(\sin \theta/\lambda)_{max}$ (Å <sup>-1</sup> )	0.556	0.555
$R[F^2 > 2\sigma(F^2)], wR(F^2), S$	0.033, 0.081, 1.03	0.041, 0.088, 1.13
No. of reflections	581	559
No. of parameters	163	163
No. of restraints	120	120
$\Delta\rho_{max}, \Delta\rho_{min}$ (e Å <sup>-3</sup> )	0.13, -0.09	0.15, -0.11

Computer programs: SAINTE v8.34A<sup>34</sup>, XT<sup>35</sup>, XL<sup>36</sup>, Olex2<sup>30</sup>

Table 5.1 *Contd.* Crystallographic tables for the 1:1 at 4.32 GPa and 5.06 GPa

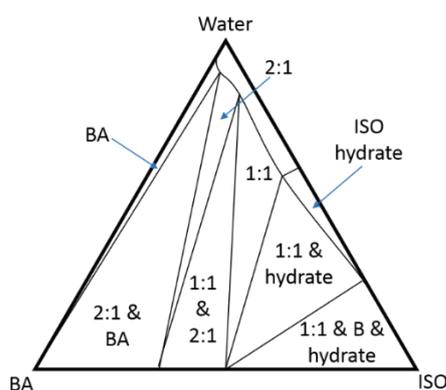
1:1 series	4.32 GPa, 297K	5.06 GPa, 297K
$a, b, c$ (Å)	19.215 (9), 4.9165 (7), 20.245 (3)	19.128 (8), 4.9124 (8), 20.236 (3)
$\beta$ (°)	98.78 (2)	98.78 (2)
$V$ (Å <sup>3</sup> )	1890.2 (10)	1879.2 (8)
$\mu$ (mm <sup>-1</sup> )	0.12	0.13
Crystal size (mm)	0.25 × 0.04 × 0.03	0.25 × 0.04 × 0.03
Absorption correction	Multi-scan SADABS2014/3 (Bruker,2014/3) was used for absorption correction. $wR2(int)$ was 0.1290 before and 0.0903 after correction. The Ratio of minimum to maximum transmission is 0.5560. The $\lambda/2$ correction factor is Not present.	Multi-scan SADABS2014/3 (Bruker,2014/3) was used for absorption correction. $wR2(int)$ was 0.1094 before and 0.0618 after correction. The Ratio of minimum to maximum transmission is 0.7060. The $\lambda/2$ correction factor is Not present.
$T_{min}, T_{max}$	0.414, 0.745	0.526, 0.745
No. of measured, independent and observed [ $I > 2.0\sigma(I)$ ] reflections	4256, 545, 413	2948, 499, 326
$R_{int}$	0.096	0.105
$\theta_{max}$ (°)	23.3	23.3
$(\sin \theta/\lambda)_{max}$ (Å <sup>-1</sup> )	0.556	0.556
$R[F^2 > 2\sigma(F^2)],$ $wR(F^2), S$	0.066, 0.179, 1.14	0.051, 0.126, 1.11
No. of reflections	545	499
No. of parameters	164	164
No. of restraints	120	120
$\Delta\rho_{max}, \Delta\rho_{min}$ (e Å <sup>-3</sup> )	0.24, -0.19	0.18, -0.15

Computer programs: SAINTE v8.34A<sup>34</sup>, XT<sup>35</sup>, XL<sup>36</sup>, Olex2<sup>30</sup>

### 5.3.2. The 2:1 co-crystal of benzoic acid and isonicotinamide

#### 5.3.2.1. Preparation of the 2:1 co-crystal

2:1 co-crystals of benzoic acid and isonicotinamide were supplied by Svoboda *et al.*<sup>10</sup> for the ambient pressure experiment. The co-crystals for the high pressure experiment were made utilising the phase diagram by Svoboda *et al.* 1:1 mole fraction of benzoic acid and isonicotinamide in water. Crystals grown were quickly harvested from the vial and checked by SCXRD.



**Figure 5.3 Ternary phase diagram showing regions of the co-crystal formation in water as produced by Svoboda *et al.***

#### 5.3.2.2. Ambient pressure single crystal X-ray diffraction

At ambient pressure, a dataset of the 2:1 co-crystal was collected and treated in a similar manner to the procedure described for the 1:1 co-crystal. A single crystal was loaded onto a MiTiGen microloop of 100 $\mu$ m diameter. X-ray diffraction intensities were collected from a sequence of eight scans using a Bruker D8 Venture diffractometer ( $\lambda = 1.54046 \text{ \AA}$ ) equipped with an Oxford Cryosystems low temperature device operating at 123K. Data were reduced using SAINT as incorporated in Bruker Apex II Software.<sup>27</sup> An absorption correction was applied using SADABS.<sup>28,29</sup> The data were solved initially by intrinsic phasing and refined using *Olex2*<sup>30</sup> software ( $F^2$ ). All non-H atoms were anisotropically refined. All hydrogen atoms were placed in calculated positions and constrained to ride their parent atom.

### 5.3.2.3. Diamond anvil cell preparation

The 2:1 co-crystals were weakly diffracting and problematic. To improve the aspect ratio and size of the crystals for high pressure experiments, the crystals were annealed in the sample chamber of the DAC. The DAC was loaded with a few crystals of 2:1 and water. An in-house heating apparatus was used to slowly temperature cycle the DAC (20 to 70°C) over the course of two days to anneal an improved quality crystal. An attempt to collect data in water after increasing the pressure yielded such poor quality data that the datasets were not included in the compression study. Regardless, datasets taken on the crystal were limited as water freezes out at ~1 GPa. Subsequent higher pressure datasets were taken from a second loading with petroleum ether (40:60). The sealing of the second loading was gradually taken up to 1.83 GPa as a means to freeze crystal movement within the DAC.

### 5.3.2.4. High Pressure Studies

At each pressure, a dataset of the 2:1 co-crystal was collected and treated in a similar manner to the 1:1 co-crystal high pressure datasets. The X-ray diffraction intensities were collected from a sequence of twelve scans at 297K following the procedure described by Dawson *et al.* Data were indexed and integrated using SAINT.<sup>31</sup> The initial dataset was solved by intrinsic phasing and the co-ordinates from the initial dataset were used for each subsequent pressure dataset. All data were refined isotropically due to poor data quality. Hydrogen atoms were placed in calculated positions and constrained to ride their parent atom as per procedure in *Olex2*. The pressure transmitting medium was petroleum ether (40:60).<sup>37</sup>

**Table 5.2 Crystallographic tables for the 2:1 from ambient to 2.73 GPa**

For all structures:  $2(C_7H_6O_2) \cdot C_6H_6N_2O$ ,  $M_r = 366.36$ , Triclinic,  $\bar{P}1$ ,  $Z'=2$ . All high pressure data were refined isotropically. H-atom parameters were constrained.

	Ambient P, 123 K	1.83 GPa, 297 K	2.73 GPa, 297 K
$a, b, c$ (Å)	10.0479 (7), 12.6692 (9), 14.2007 (9)	9.8417 (15), 12.226 (5), 13.269 (3)	9.7938 (10), 12.105 (3), 12.9481 (17)
$\alpha, \beta, \gamma$ (°)	80.475 (4), 79.784 (5), 90.001 (5)	80.34 (2), 79.323 (13), 90.02 (2)	80.588 (17), 79.120 (9), 90.044 (15)
$V$ (Å <sup>3</sup> )	1753.7 (2)	1545.9 (7)	1486.4 (5)
Radiation type	Cu $K\alpha$	Mo $K\alpha$	Mo $K\alpha$
$\mu$ (mm <sup>-1</sup> )	0.84	0.12	0.12
Crystal size (mm)	0.3 × 0.05 × 0.02	0.24 × 0.03 × 0.01	0.24 × 0.03 × 0.01
Diffractometer	Bruker D8-Venture	Bruker APEX-II CCD	Bruker APEX-II CCD
Absorption correction	Multi-scan SADABS2016/2 (Bruker,2016/2) was used for absorption correction. $wR2(int)$ was 0.1155 before and 0.0741 after correction. The Ratio of minimum to maximum transmission is 0.6823. The $\lambda/2$ correction factor is Not present.	Multi-scan SADABS2016/2 (Bruker,2016/2) was used for absorption correction. $wR2(int)$ was 0.1182 before and 0.0740 after correction. The Ratio of minimum to maximum transmission is 0.6823. The $\lambda/2$ correction factor is Not present.	Multi-scan SADABS2016/2 (Bruker,2016/2) was used for absorption correction. $wR2(int)$ was 0.1162 before and 0.0549 after correction. The Ratio of minimum to maximum transmission is 0.5864. The $\lambda/2$ correction factor is Not present.
No. of measured, independent and observed [ $I > 2\sigma(I)$ ] reflections	11408, 5658, 3664	6196, 1350, 837	7000, 1300, 842
$R_{int}$	0.062	0.077	0.074
$\theta_{max}$ (°)	66.5	23.3	23.3
$(\sin \theta/\lambda)_{max}$ (Å <sup>-1</sup> )	0.596	0.556	0.557
$R[F^2 > 2\sigma(F^2)], wR(F^2), S$	0.076, 0.212, 1.08	0.072, 0.202, 1.06	0.076, 0.204, 1.06
No. of reflections	5658	1350	1300
No. of parameters	491	173	173
No. of restraint	0	0	0
$\Delta\rho_{max}, \Delta\rho_{min}$ (e Å <sup>-3</sup> )	0.70, -0.49	0.22, -0.17	0.21, -0.19

Computer programs: SAINTE v8.37A,<sup>34</sup> XT,<sup>35</sup> XL,<sup>36</sup> Olex2<sup>30</sup>

Table 5.2 *Contd.* Crystallographic tables for the 2:1 from 3.29 GPa to 4.72 GPa

	3.29 GPa, 297 K	4.03 GPa, 297 K	4.72 GPa, 297 K
<i>a</i> , <i>b</i> , <i>c</i> (Å)	9.777 (2), 12.049 (7), 12.823 (3)	9.7452 (16), 11.991 (5), 12.596 (3)	9.738 (2), 11.999 (7), 12.465 (4)
$\alpha$ , $\beta$ , $\gamma$ (°)	80.56 (3), 79.003 (16), 90.06 (3)	80.89 (3), 78.883 (14), 90.09 (2)	81.14 (4), 78.729 (19), 90.14 (3)
<i>V</i> (Å <sup>3</sup> )	1462.0 (9)	1425.3 (7)	1410.6 (10)
Radiation type	Mo <i>K</i> $\alpha$	Mo <i>K</i> $\alpha$	Mo <i>K</i> $\alpha$
$\mu$ (mm <sup>-1</sup> )	0.12	0.12	0.13
Crystal size (mm)	0.24 × 0.03 × 0.01	0.24 × 0.03 × 0.01	0.24 × 0.03 × 0.01
Diffractometer	Bruker APEX-II CCD	Bruker APEX-II CCD	Bruker APEX-II CCD
Absorption correction	Multi-scan SADABS2016/2 (Bruker,2016/2) was used for absorption correction. <i>wR2</i> (int) was 0.0912 before and 0.0550 after correction. The Ratio of minimum to maximum transmission is 0.8382. The $\lambda/2$ correction factor is Not present.	Multi-scan SADABS2016/2 (Bruker,2016/2) was used for absorption correction. <i>wR2</i> (int) was 0.0841 before and 0.0518 after correction. The Ratio of minimum to maximum transmission is 0.8915. The $\lambda/2$ correction factor is Not present.	Multi-scan SADABS2016/2 (Bruker,2016/2) was used for absorption correction. <i>wR2</i> (int) was 0.0841 before and 0.0518 after correction. The Ratio of minimum to maximum transmission is 0.8915. The $\lambda/2$ correction factor is Not present.
No. of measured, independent and observed [ <i>I</i> > 2 $\sigma$ ( <i>I</i> )] reflections	6227, 1276, 815	6274, 1242, 825	6612, 1238, 798
<i>R</i> <sub>int</sub>	0.072	0.072	0.072
$\theta_{\max}$ (°)	23.3	23.3	23.4
( <i>sin</i> $\theta/\lambda$ ) <sub>max</sub> (Å <sup>-1</sup> )	0.557	0.557	0.558
<i>R</i> [ <i>F</i> <sup>2</sup> > 2 $\sigma$ ( <i>F</i> <sup>2</sup> )], <i>wR</i> ( <i>F</i> <sup>2</sup> ), <i>S</i>	0.070, 0.188, 1.07	0.556	0.073, 0.200, 1.07
No. of reflections	1276	1242	1238
No. of parameters	173	173	173
No. of restraint	0	0	0
$\Delta\rho_{\max}$ , $\Delta\rho_{\min}$ (e Å <sup>-3</sup> )	0.23, -0.16	0.24, -0.17	0.22, -0.18

Computer programs: S<sub>A</sub>I<sub>N</sub>T v8.37A,<sup>34</sup> X<sub>T</sub>,<sup>35</sup> X<sub>L</sub>,<sup>36</sup> Olex2<sup>30</sup>

Table 5.2 *Contd.* Crystallographic tables for the 2:1 from 5.25 GPa to 5.99 GPa

	5.25 GPa, 297 K	5.99 GPa, 297 K
<i>a</i> , <i>b</i> , <i>c</i> (Å)	9.737 (5), 11.938 (13), 12.400 (8)	9.706 (3), 11.882 (10), 12.314 (5)
$\alpha$ , $\beta$ , $\gamma$ (°)	80.96 (7), 78.67 (4), 90.20 (6)	81.19 (5), 78.58 (3), 90.05 (4)
<i>V</i> (Å <sup>3</sup> )	1394.9 (19)	1374.9 (13)
Radiation type	Mo <i>K</i> $\alpha$	Mo <i>K</i> $\alpha$
$\mu$ (mm <sup>-1</sup> )	0.13	0.13
Crystal size (mm)	0.24 × 0.03 × 0.01	0.24 × 0.03 × 0.01
Diffractometer	Bruker <i>APEX-II</i> CCD	Bruker <i>APEX-II</i> CCD
Absorption correction	Multi-scan <i>SADABS2016/2</i> (Bruker,2016/2) was used for absorption correction. <i>wR2(int)</i> was 0.1066 before and 0.0601 after correction. The Ratio of minimum to maximum transmission is 0.8754. The $\lambda/2$ correction factor is Not present.	Multi-scan <i>SADABS2016/2</i> (Bruker,2016/2) was used for absorption correction. <i>wR2(int)</i> was 0.0829 before and 0.0525 after correction. The Ratio of minimum to maximum transmission is 0.8757. The $\lambda/2$ correction factor is Not present.
No. of measured, independent and observed [ <i>I</i> > 2 $\sigma$ ( <i>I</i> )] reflections	6178, 1225, 796	5506, 1217, 788
<i>R</i> <sub>int</sub>	0.074	0.070
$\theta_{\max}$ (°)	23.4	23.4
( $\sin \theta/\lambda$ ) <sub>max</sub> (Å <sup>-1</sup> )	0.558	23.3
<i>R</i> [ <i>F</i> <sup>2</sup> > 2 $\sigma$ ( <i>F</i> <sup>2</sup> )], <i>wR</i> ( <i>F</i> <sup>2</sup> ), <i>S</i>	0.077, 0.234, 1.06	0.073, 0.213, 1.06
No. of reflections	1225	1217
No. of parameters	173	173
No. of restraint	0	0
$\Delta\rho_{\max}$ , $\Delta\rho_{\min}$ (e Å <sup>-3</sup> )	0.26, -0.19	0.22, -0.19

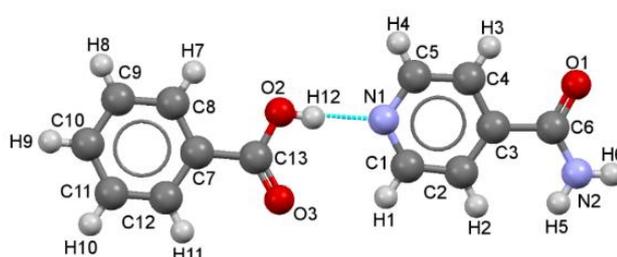
Computer programs: *SAINT* v8.37A,<sup>34</sup> *XT*,<sup>35</sup> *XL*,<sup>36</sup> *Olex2*<sup>30</sup>

## 5.4. Results and Discussion

### 5.4.1. Analysis of the 1:1 co-crystal of benzoic acid and isonicotinamide

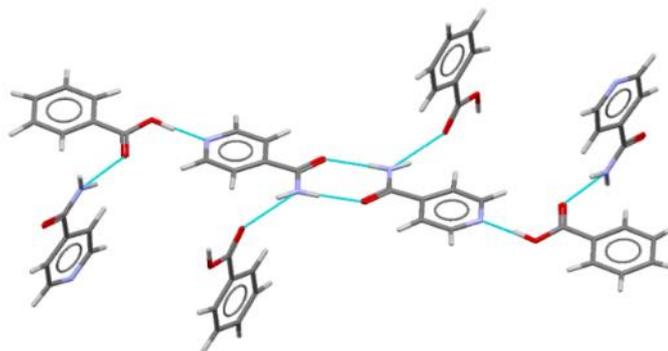
#### 5.4.1.1. Ambient pressure structural analysis

The 1:1 co-crystallises in  $C2/c$  with a  $Z'$  of 1. Benzoic acid and isonicotinamide are structurally similar in shape and size with molecular masses of  $122.123 \text{ g mol}^{-1}$  and  $122.127 \text{ g mol}^{-1}$ , respectively. The unit cell information was consistent with those found in the Cambridge Structural Database (CSD) for the 1:1 stoichiometric co-crystal of BAIS (CSD refcode: BUDWEC).<sup>38,39</sup>



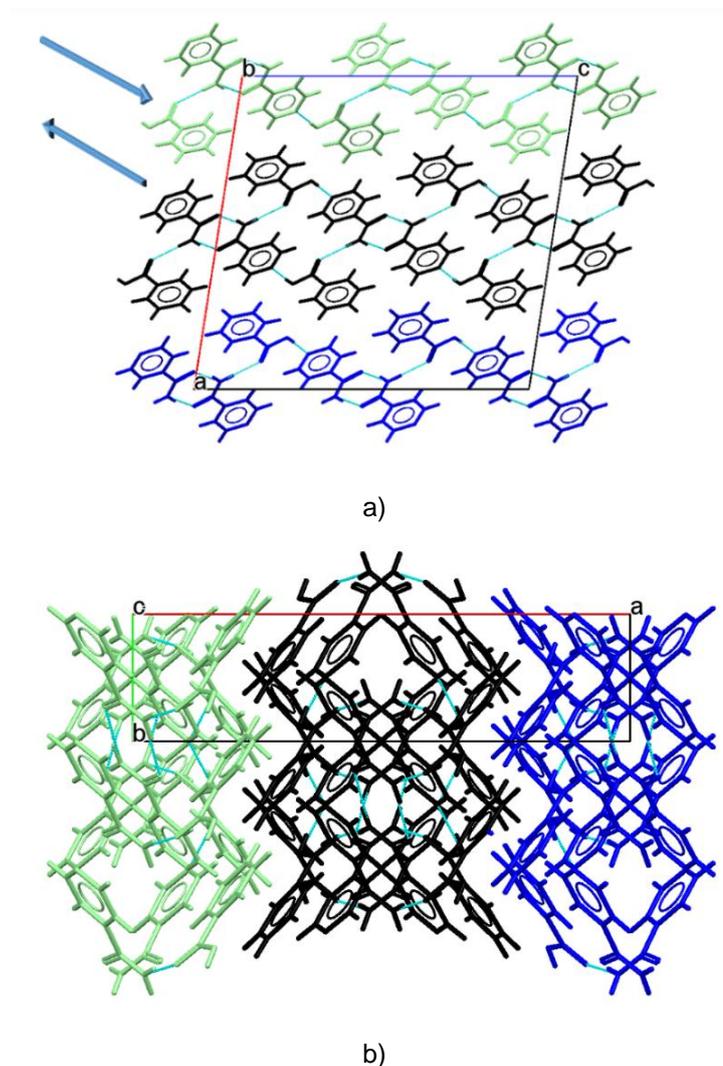
**Figure 5.4** The asymmetric unit of benzoic acid and isonicotinamide with numbering scheme, at ambient pressure. The colours of the elements are consistent throughout (carbon, grey; oxygen, red; nitrogen, blue; hydrogen, white).

Isonicotinamide components form an amide dimer. The mirroring isonicotinamide molecules have an amide dimer  $D \cdots A$  contact length of  $2.907 \text{ \AA}$ . Each isonicotinamide is hydrogen bonded via their pyridine moiety to the carboxylic acid moiety of a benzoic acid ( $2.633 \text{ \AA}$ ). The four molecules form the base tetramer of the 1:1. The network of hydrogen bonding expands from this tetramer. Each tetramer unit forms hydrogen bonds to four other tetramer units. The secondary amides of the isonicotinamide dimers donate the second hydrogen to the terminal benzoic acid of an almost perpendicular tetramer ( $2.961 \text{ \AA}$ ). The hydrogen bonding pattern of the tetramer unit is presented in Figure 5.5.



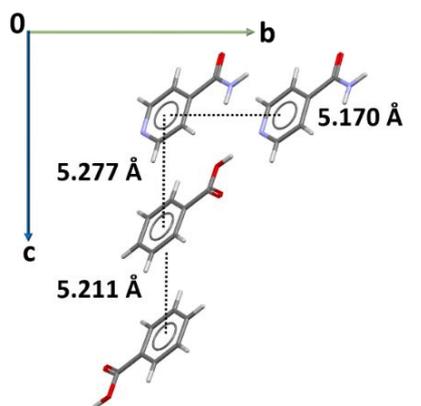
**Figure 5.5 Tetramer unit with the hydrogen bonding interactions of the 1:1**

The longest hydrogen bond formed is between the side bonded tetramer units in the *c*-axis direction (2.961 Å). Each planar tetramer is at an angle of 87.26° to its side bonded tetramer. The packing arrangement of the units is alternating in and out of plane when looking down the *b*-axis. A view down the *b* axis of the unit cell displays the network of hydrogen bonding. The terminal benzoic acid molecules of the tetramer unit form the borders of a layer of hydrogen bonding (Figure 5.5). Sheets of hydrogen bonded tetramer layers stack in the *a*-axis with no hydrogen bonding between them. Figure 5.6 b) reveals the layered structure in colour.



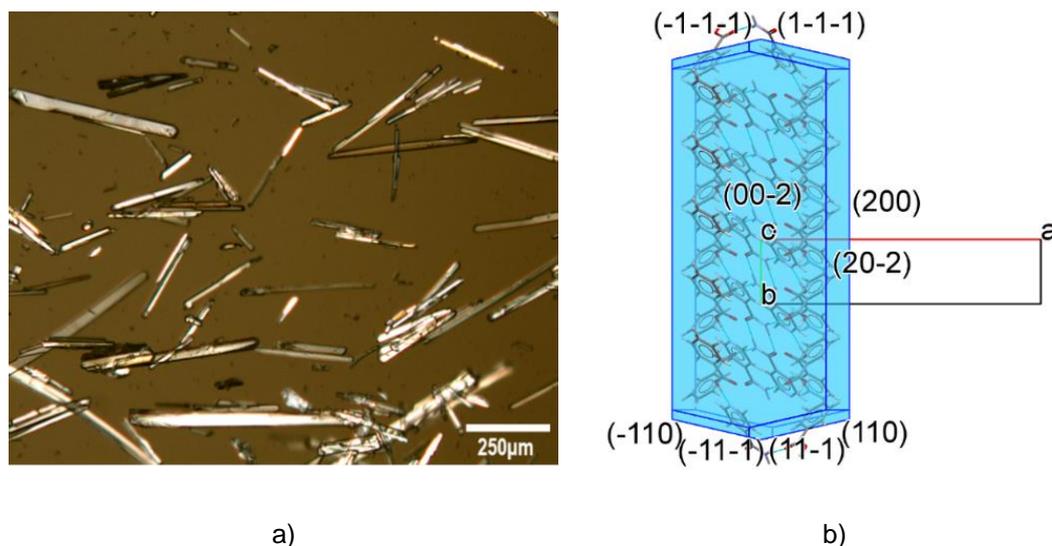
**Figure 5.6** Hydrogen bonded sheets of 1:1 benzoic acid and isonicotinamide in colour, made up of the base tetramer unit of the 1:1 a) complete tetramer shown in the middle layer in black, which is side bonded to another tetramer at an angle of  $87.26^\circ$  (arrows signify the lines of molecules which are in and out of plane) b) coloured layers of the structure in the *a*-axis showing a ‘wine-rack-type’ structure.

The aromatic nature of the 1:1 co-crystal compound presents askew  $\pi$ - $\pi$  stacking in two directions, the alternating tetramer directions present an edge to face type interaction. Directly down the *b*-axis at a ring centroid to ring centroid distance of 5.170 Å is the shorter centroid to centroid distance. In the *c*-axis direction a ring centroid to centroid distance of alternating, 5.277 Å and 5.211 Å was measured. The alternating chains that dominate the *a*-axis direction have an edge to face conformation.



**Figure 5.7** The aromatic stacking directions of the 1:1 relative to the unit cell axes, with ring to ring centroid distances noted.

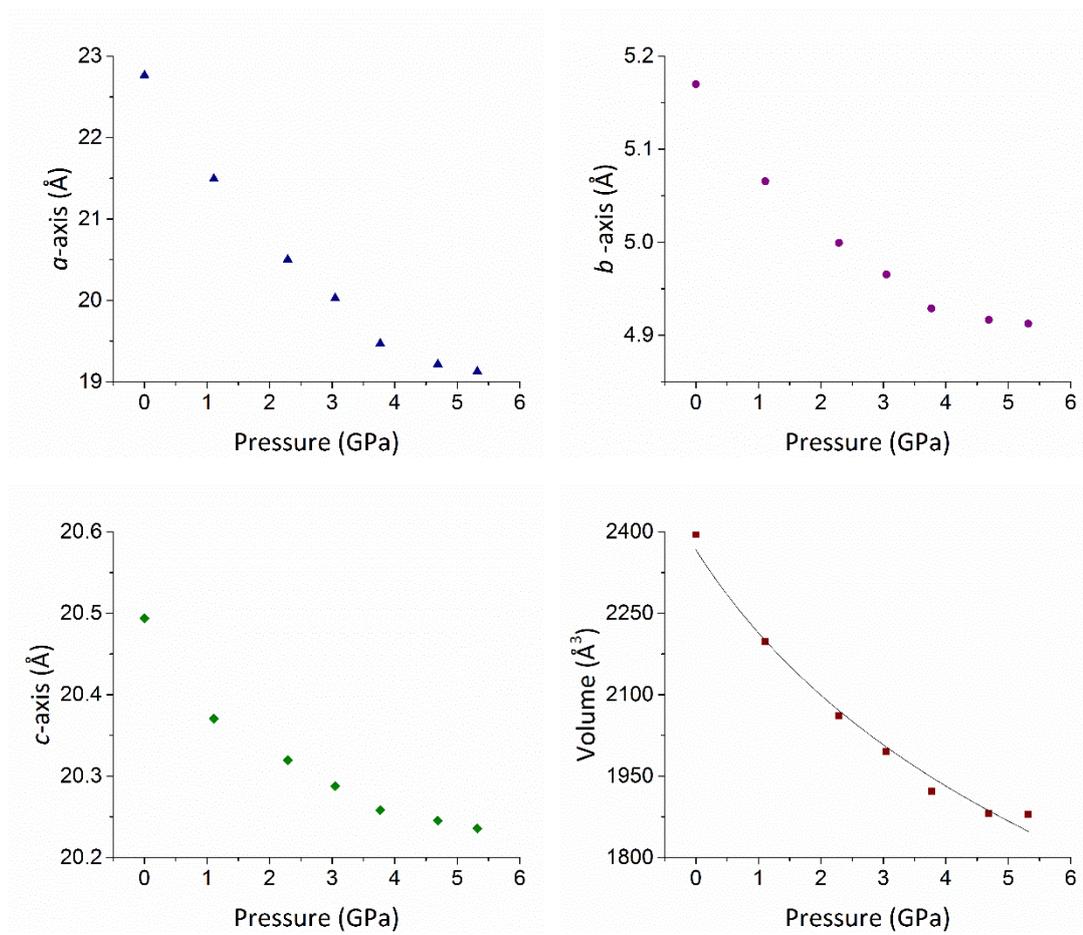
The Bravais Friedel Donnay Harker (BFDH) morphology prediction compares well with the observed morphology. BFDH predictions do not consider the growth medium or the energetics of the system, the prediction is based on geometric, crystallographic methods. Nonetheless, it is still useful to consider BFDH predictions when the observed morphology is a needle shaped crystal. In the case of the 1:1, the predicted and observed dominant faces match well. The BFDH prediction can be further complimented by considering the intermolecular interactions of the system. It is reasonable to assume that the edge to face conformation of the molecules in the *a*-axis is less energetic than the network of hydrogen bonds. The hydrogen bonded sheets of the 1:1 stack in the *a*-axis, the area between sheets has no strong hydrogen bonding interactions and would give rise to the largest face observed, the (2 0 0). The long needle shape presented by the 1:1 allows growth of the (2 0 0) face. The (2 0 0) plane is in the *a*-axis direction.



**Figure 5.8 a) BFDH predicted morphology of the 1:1 co-crystal of benzoic acid and isonicotinamide viewed down the c axis of the unit cell in agreement with the experimentally found morphology b).**

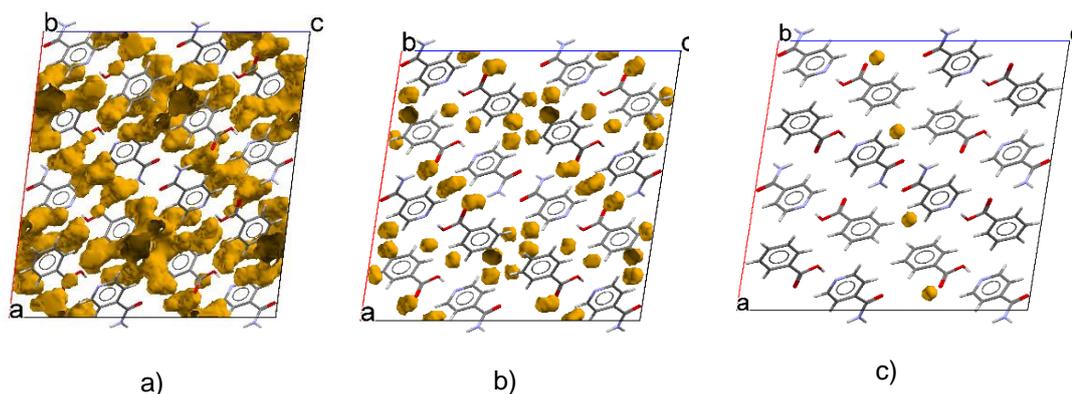
#### 5.4.1.2. Compression analysis

The compression of the 1:1 co-crystals axes is smooth and reveals no abrupt change in the unit cell parameters up to 5.32 GPa. The 2<sup>nd</sup> order Birch-Murnaghan Equation of State gives a bulk modulus ( $K_0$ ) of 13.1230 (14) GPa  $V_0 = 2366.4119$  (6) Å<sup>3</sup>  $K' = 4$ . The most prominent feature of the compression is the large decrease in unit cell length  $a$ . Unit cell length  $a$  decreased by 16 % from ambient to 5.32 GPa, the  $b$ -axis decreases by 5 % and the  $c$ -axis by 1 %. The higher pressure datasets begin to plateau.



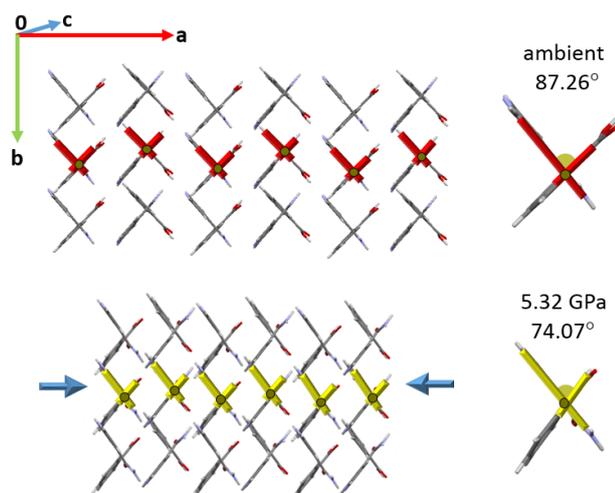
**Figure 5.9 Unit cell lengths as a function of pressure. The errors on each of the measurements is smaller than the symbol. Volume of the unit cell as a function of pressure (EoS fit 7).**

The hydrogen bonded network will stabilise against compression in the *c*-axis direction. There are two reasons for the extensive compression in the *a*-axis in particular. Firstly, the sheets of hydrogen bonding with no hydrogen bonding interaction between provide space to compress in the *a*-axis direction. Void space analysis was performed using a 0.5 Å probe radius with 0.2 Å approximate grid spacing. The initial ambient pressure structure comprised of 339.92 Å<sup>3</sup> void space (14.2 % of the unit cell). The void space decreased sharply upon compression dropping to 2.6 % of the unit cell at 3.05 GPa and dropping further to 0.1 % (2.18 Å<sup>3</sup>) at 5.32 GPa. Large areas of void space are present in the *a*-axis supporting the theory of the sheet like formation of the hydrogen bonding network.



**Figure 5.10 Unit cell and packing of the co-crystal viewed down the b-axis. Void spacing of the unit cell shown at (a) ambient, (b) 3.05 GPa and (c) 5.32 GPa.**

The second reason for the extensive compression in the *a*-axis is due to a herringbone arrangement of the alternating dimers. As previously noted, the molecules form an  $87.26^\circ$  angle with the side tetramer chains due to the secondary bonding of the amide dimer acting like a screw in a hinge mechanism. The hydrogen bond formed between the two almost perpendicular molecules appears to allow the molecules to rotate slightly around the bond. The hinge mechanism of the two molecules causes an overall effect similar to that of a wine-rack, in which, the herringbone network of molecules start to appear 'closed'. As depicted in Figure 5.11, this effect is predominantly in the *a*-axis direction. The plateau of all of the unit cell parameters at higher pressures may also be accounted for by the wine rack effect. Eventually the angle of the molecules across this bond becomes unfavourable, this, and the reduced void space for dissipation lead to a plateau at higher pressures. This effect has been observed in other compounds at high pressure, usually around an intramolecular torsion angle, and can lead to negative linear compressibility. The highly polymorphic and colourful system, ROY, has exhibited a piezochromic response to pressure as a result of the wine-rack effect whilst under hydrostatic pressure. Harty *et al.* determined that ROY has conformational flexibility and is able to change colour as a result of the compression and changing torsion angle around its CN bond.<sup>40</sup>



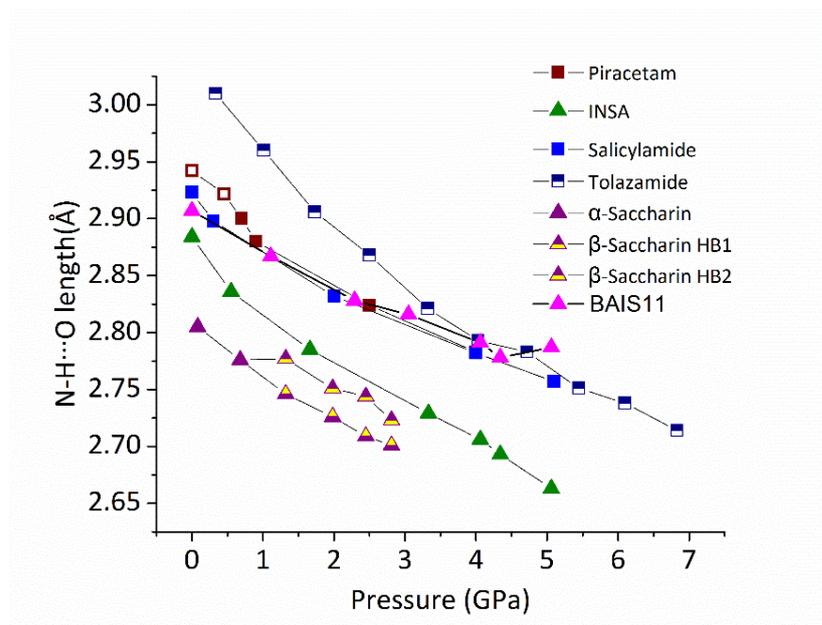
**Figure 5.11** Molecules of the 1:1 shown with stick images superimposed to enable easier visualisation of the ‘hinge’ mechanism. At ambient pressure the two molecules are ‘open’ with an angle of 87.26° and at 5.32 GPa the hinge mechanism ‘closes’ with an angle of 74.07° between the two molecules.

#### 5.4.2. Compression behaviour of the amide dimer

The two chapters precluding this have detailed the lack of compression studies in the CSD containing dimer interactions. Indomethacin and saccharin (INSA), saccharin as a singular component and, now, the 1:1 co-crystal of benzoic acid and isonicotinamide are added to the study. There appears to be a potential correlation between rate of compression of an amide dimer N-H···O contact and whether the dimer involves primary or secondary amides. The present study adds another amide dimer contact distance over the course of a compression to analyse. The rate of compression of the 1:1 co-crystals dimer agrees overall with the data presented in Figure 5.12. Previous work has detailed that structures such as Tolazamide, INSA and saccharin which contain secondary amides (not involved in any other hydrogen bonding) appear to compress at a faster rate. Structures such as piracetam, salicylamide and the present study on benzoic acid and isonicotinamide appear to compress at a slower rate due to secondary hydrogen bonding on the nitrogen of the amide. Overall the amide dimer compresses by 4%. The other two hydrogen bonding motifs in the structure (pyridine to carboxylic acid and the cross linking, secondary amide linker) decreased by 6% and 8%, respectively. The largest compression occurs in the longest hydrogen bond which is the weakest as both other interactions are linear. At the highest pressures in

## Lauren Connor

the present study, a sharp decrease is observed in the linker hydrogen bond. The sharp decrease coinciding with a slight increase in the amide dimer bond length. This is unexpected. The perpendicular bonding arrangement of the side chain is unfavourable for the hydrogen bond. Typically, the strongest hydrogen bonds are collinear.<sup>41</sup> The increase in pressure pushes this bond further from linearity.



**Figure 5.12** Plots showing how the length of piracetam (form II: half symbol & form V: filled symbol), INSA (1:1 co-crystal of indomethacin and saccharin), tolazamide, salicylamide, saccharin and the 1:1 co-crystal of benzoic acid and isonicotinamide's N-H...O amide dimer responds to pressure.

### 5.4.3. Analysis of the 2:1 co-crystal of benzoic acid and isonicotinamide

#### 5.4.3.1. Crystallisation and DAC loading

Formation of a 2:1 co-crystal of sufficient quality for SCXRD was challenging. The co-crystal investigated at ambient pressure was supplied by a fellow student as part of a collaborative project. The co-crystal supplied was formed by continuous reactive, anti-solvent co-crystallisation.<sup>10</sup> Co-crystals of 2:1 benzoic acid and isonicotinamide are brittle and it was difficult to cut the co-crystals to an appropriate size for DAC cell loading. Based on the ternary phase diagram produced by Svoboda *et al.*<sup>10</sup> in the collaborative project, a small scale experiment was performed to produce co-crystals for the high pressure experiments. From the ternary phase diagram in water and ethanol, it is evident that the region for formation in ethanol is narrow. Water was the

## Lauren Connor

chosen solvent for an evaporative crystallisation as the region for 2:1 formation is easier to reach. The co-crystals were indexed to ensure loading of the 2:1 as it is possible to crystallise 1:1 and benzoic acid from this method if solvent evaporates in excess. Co-crystals formed by the evaporative method were annealed in the DAC in water to produce better size and quality crystals for SCXRD. Datasets were taken in water at the lower pressures (<1 GPa) but were of such poor quality that they are not included in the compression study. The co-crystals were harvested after annealing and suitable co-crystal was reloaded with petroleum ether (40:60). Unfortunately, the loading of the annealed co-crystal in petroleum ether had to be taken up to 1.83 GPa to freeze the co-crystals movement within the cell.



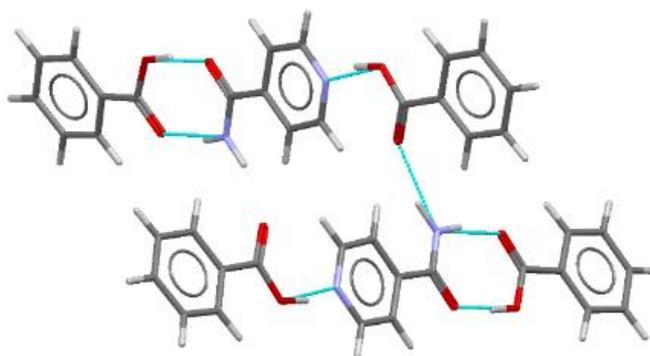
**Figure 5.13** Polarised light microscope image of the 2:1 co-crystals.

### 5.4.3.2. Ambient pressure structural analysis

The 2:1 co-crystallises in triclinic spacegroup P-1. The unit cell information was consistent with those found in the Cambridge Structural Database (CSD) for the 2:1 stoichiometric co-crystal of BAIS (CSD refcode: MOVTOH).<sup>38</sup> The crystal structure of the 2:1 was published in 2009 by Seaton *et al.*<sup>1</sup> The paper by Seaton *et al.* describes twinning along the (1 0 0) plane which corresponds to the twinning seen in our data. There are some similarities in the bonding of the 1:1 and the 2:1 co-crystal. The pyridine: carboxylic acid synthon present in the 1:1 co-crystal also features in the 2:1. The pyridine: carboxylic acid synthon has a D...A distance of 2.644 Å and is the shortest hydrogen bond present in the structure. On the other hand, there are some supramolecular synthon differences between the 1:1 and the 2:1 co-crystal. Contrary to the amide dimer in the 1:1 co-crystal, the isonicotinamide forms an amide: acid dimer (D...A distance of 2.836 Å) with benzoic acid in the 2:1 co-crystal. The new dimer sets the trimer base unit for the 2:1 stoichiometry co-crystal (benzoic acid: isonicotinamide: benzoic acid). Another notable difference in the 2:1 structure is the terminal benzoic acid that does not hydrogen bond to neighbouring trimers. Each

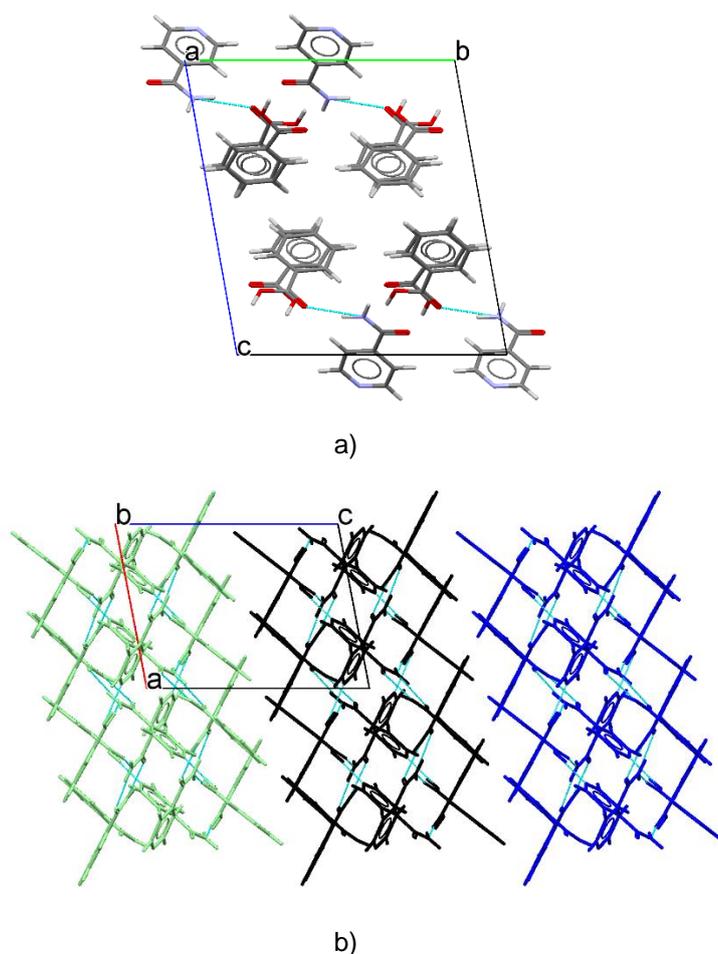
## Lauren Connor

planar trimer unit only donates in one direction, starting another trimer unit almost perpendicular ( $82.17^\circ$ ) to the original motif. The second hydrogen on the primary amide of the isonicotinamide forms a side chain hydrogen bond to the benzoic acid of a new perpendicular trimer unit. In comparison, the 1:1 co-crystal bonds in both directions due to the isonicotinamide dimer unit over the inversion center. The D...A distance of the side chain hydrogen bonds between perpendicular trimers is the longest hydrogen bond, with a length of 2.887 Å.



**Figure 5.14** Two trimers of the 2:1 co-crystal with hydrogen bonding motif shown.

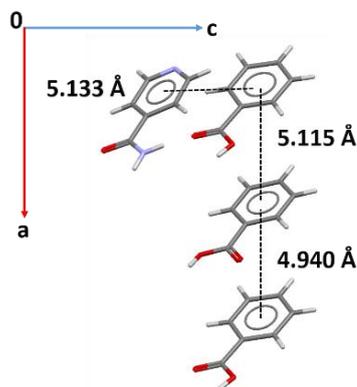
The unit cell is presented with hydrogen bonding in Figure 5.15 a), the trimer units bond primarily in the *a*-axis direction and the trimer bridging bonding is present in the *b*-axis direction. Despite the differences in hydrogen bonding patterns and space group, the overall crystal structures and spacing of the 1:1 and 2:1 co-crystals are relatively similar. The layered structure of the 1:1 is also present in the 2:1 co-crystal (**Error! Reference source not found.** b)). Sheets of hydrogen bonded tetramer layers stack in the *c*-axis with no hydrogen bonding between them. The lack of hydrogen bonding between the tetramer sheets in the *c*-axis highlights the layered structure which is depicted in colour in Figure 5.15 b).



**Figure 5.15 Hydrogen bonded layers of 2:1 benzoic acid and isonicotinamide in colour, the coloured layers of the structure in the *c*-axis display the same ‘wine-rack-type’ structure present in the 1:1 co-crystal.**

The aromaticity of the 2:1 co-crystal compound forms  $\pi$ - $\pi$  stacking in two directions. Benzoic acid molecules strengthen the overall structure in the *a*-axis direction by forming  $\pi$ - $\pi$  stacked channels. The benzoic acids stack in alternating ring centroid to centroid distances of 5.115 Å and 4.940 Å. The face to face conformation of the benzoic acids is the strongest type of  $\pi$ - $\pi$  conformation. A slightly longer askew  $\pi$ - $\pi$  interaction forms in the *c*-axis direction between benzoic acid and isonicotinamide molecules. The askew interaction is not stacked and has a ring centroid to ring centroid distance of 5.133 Å. Whilst the intercentroid distance is well within reasonable distance to qualify as a significant aromatic  $\pi$ - $\pi$  interaction, it is less influential to the overall structure than the stacked aromatic interactions in the *a*-axis. The aromatic interaction in the *c*-axis does not extend the length of the axis. Instead the  $\pi$ - $\pi$  interaction alternates with regions lacking any  $\pi$ - $\pi$  interaction, attributing to the layered structure which stacks in the *c*-axis. The lack of both hydrogen bonding and any  $\pi$ - $\pi$

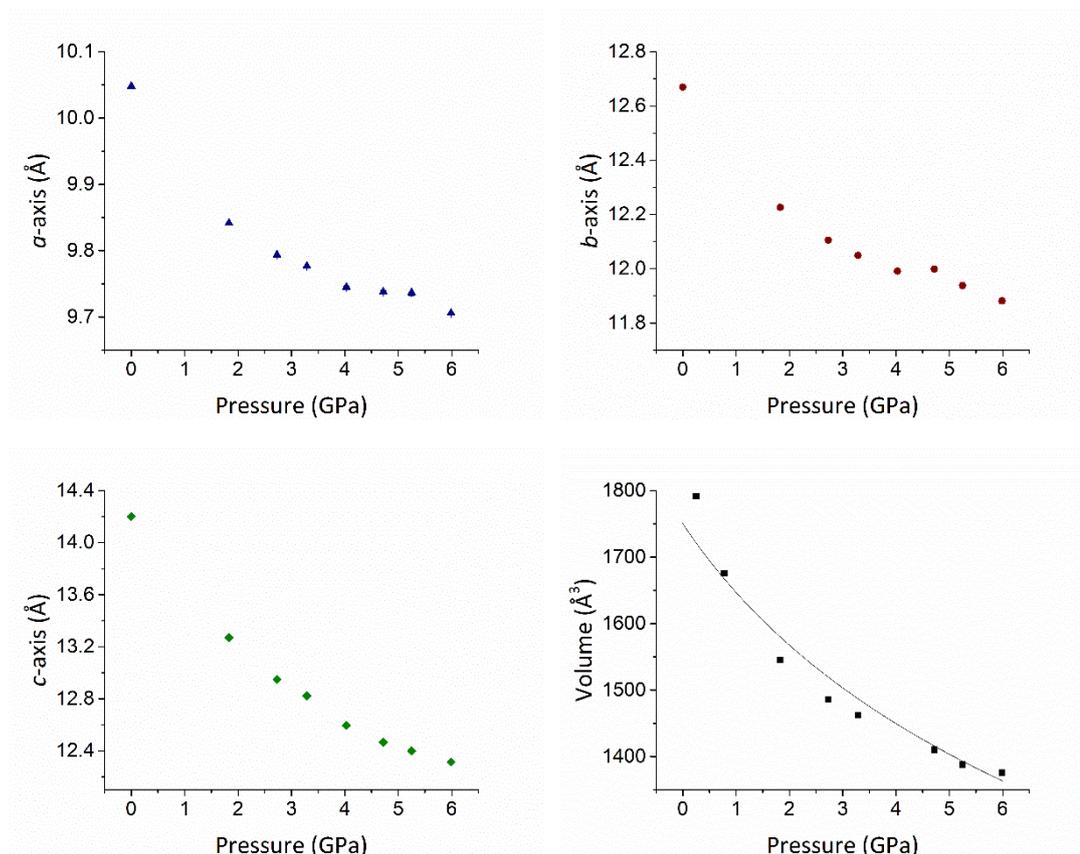
interaction between layers suggests that slip planes exist between the layers of the 2:1. Slip planes can be indicative of abnormal compression behaviour.



**Figure 5.16** The aromatic stacking directions of the 2:1 relative to the unit cell axes, with ring to ring centroid distances depicted.

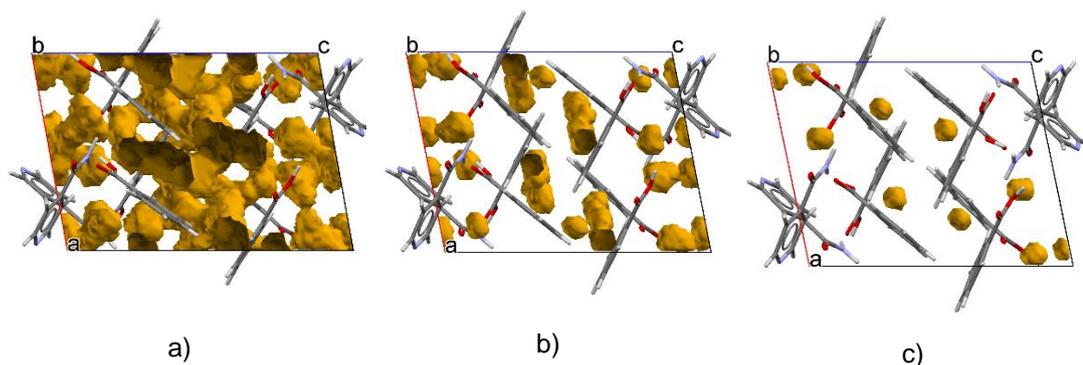
### 5.4.3.3. Compression analysis

In general, the 2:1 co-crystal of benzoic acid and isonicotinamide smoothly compressed up to 5.99 GPa. All of the axes begin to plateau at the higher pressures which emphasises the slight bump evident in Figure 5.17 between 4 and 6 GPa. The 2nd order Birch-Murnaghan Equation of State gives a bulk modulus ( $K_0$ ) of 14.55733 (12) GPa,  $V_0 = 1750.28833$  (14)  $\text{\AA}^3$   $K' = 4$ . The equation of state is reasonably comparable to the equation of state of the 1:1 co-crystal (bulk modulus ( $K_0$ ) of 13.1230 (14) GPa  $V_0 = 2366.4119$  (6)  $\text{\AA}^3$   $K' = 4$ ). Similar bulk modulus' confirm that the compressibility of the 1:1 and the 2:1 is comparable as would be expected from such a similar overall structure. As with the 1:1 system, the layers with slip planes between them facilitated the direction of greatest compressibility ( $c$ -axis - 13%). The  $b$ -axis followed with a decrease of 6% over the course of the compression. The  $a$ -axis decreased the least (3%) as the bulk of the hydrogen bonding and  $\pi$ - $\pi$  stacking extends in this direction.



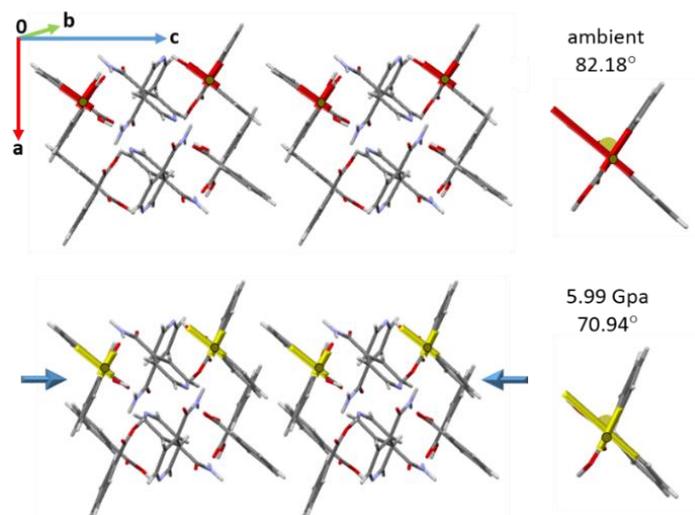
**Figure 5.17 Unit cell lengths as a function of pressure. The errors on each of the measurements is smaller than the symbol. Volume of the unit cell as a function of pressure (EoS fit 7).**

The bulk of the hydrogen bonding and  $\pi$ - $\pi$  stacking extends in the  $a$ -axis and helps to stabilise the structure against compression in this direction. The extensive compression in the  $c$ -axis is reasonable when considering the layered structure. The structure comprises of stacked layers with slip planes between them. The layers provide space for substantial compression in the  $c$ -axis direction. Void space analysis was performed using a 0.5 Å probe radius with 0.2 Å approximate grid spacing. The initial ambient pressure structure comprised of 227.13 Å<sup>3</sup> void space (13.0 % of the unit cell). The void space significantly decreased in the compressed structures, dropping to 2.7 % (40.35 Å<sup>3</sup>) of the unit cell at 2.73 GPa and dropping further to 0.5 % (6.22 Å<sup>3</sup>) at 5.99 GPa. The bulk of the structure extends down the  $a$ -axis and the bulk of the void space exists in the  $c$ -axis, depicted in Figure 5.18 **Error! Reference source not found.** Most of the void space is evident within the slip planes between the structural layers.



**Figure 5.18 Unit cell and packing of the co-crystal viewed down the b-axis. Void spacing of the unit cell shown at (a) ambient, (b) 2.73 GPa and (c) 5.99 GPa.**

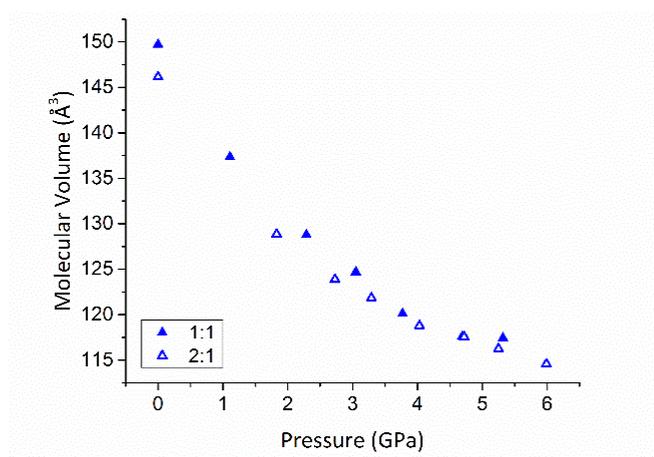
The 2:1 and 1:1 co-crystals both display a wine-rack-type structure which compresses in a similar manner. The wine-rack-type structure helps to explain the compression behaviour of the 2:1 in a similar manner to the 1:1 co-crystal. The planar trimers of the 2:1 are at an almost perpendicular angle of  $81.18^\circ$  with adjacent 2:1 trimers. The primary amide of the isonicotinamide facilitates this angle via the bridging bond formed to the adjacent, almost perpendicular benzoic acid. As previously noted with the 1:1 structure, the secondary bonding of the amide acts like a screw in a hinge mechanism. The rotation around this hydrogen bonding interaction gives rise to the wine-rack effect which is observed when the system is compressed. The hinge mechanism closes upon compression, this observed effect helps to explain the large decrease in the *c*-axis. The plateau of the *a*-unit cell length at higher pressure can be accounted for by the wine rack effect. Similar to the 1:1 co-crystal, the angle between the trimer units and the torsion around the hydrogen bond becomes resistant to the hinge mechanism. The torsional resistance and the reduced void space for dissipation leads to a plateau at higher pressures.



**Figure 5.19** The 2:1 co-crystal, viewed down the b-axis at ambient and 5.99 GPa. The benzoic acid molecules display a wine-rack-type compression behaviour similar to that of the 1:1 co-crystal.

#### 5.4.4. Compression comparison of the 1:1 and 2:1 co-crystal

The molecular masses of benzoic acid and isonicotinamide are similar as well as their size and shape. The similarity of the molecules allowed for easy calculation of their molecular volumes. The molecular volumes for each of the stoichiometric co-crystals was calculated and plotted to understand how the compression of the 2:1 and 1:1 compared on a molecular level. The stoichiometric difference did not lead to any compression differences and the decrease in volume remained consistent on a molecular level.



**Figure 5.20** Molecular volume comparison over the course of the compression study for the 1:1 and 2:1 co-crystal of benzoic acid and isonicotinamide.

### 5.5. Conclusion

The evaluation and comparison of the structures and compression behaviour of the 1:1 and 2:1 co-crystals of benzoic acid and isonicotinamide revealed subtle differences in hydrogen bonding. However, overall the structures of the 1:1 and 2:1 are similar and lend to a comparable mechanism of compression. The 1:1 structure is built up from an isonicotinamide dimer over an inversion centre which then bonds to a terminal benzoic acid at either end, forming a tetramer. The tetramer units are connected via a long hydrogen bond to another perpendicular tetramer. The 2:1 follows a similar overall structure of bonding, instead with trimer units and lacking the isonicotinamide dimer. The compression of both the 1:1 and 2:1 stoichiometry co-crystals revealed no abrupt change in crystal form and the crystal systems remained in their respective space groups throughout the compression study. The 1:1 co-crystal remained in space group  $C2/c$  and the 2:1 in  $P-1$ . The volume and each of the unit cell axes decreased throughout the compression as expected. A wine-rack effect is observed in the compression of the 1:1 and the 2:1 co-crystal. The wine-rack-effect arises due to the longer hydrogen bonds that interact with perpendicular units of molecules. The bond acts like a screw in a hinge and the molecules were observed to 'close' in both stoichiometry co-crystals. Further to the previous studies in chapter 3 and 4, the amide dimer present in the 1:1 co-crystal exactly matched the rate of compression of the other recorded amide dimers which were secondary hydrogen bonded. The 2:1 co-crystal did not present an amide dimer to enrich the secondary hydrogen bonded rate data.

## References

1. Seaton, C. C., Parkin, A., Wilson, C. C. & Blagden, N. Controlling the Formation of Benzoic Acid : Isonicotinamide Molecular Complexes Controlling the Formation of Benzoic Acid : Isonicotinamide. *Cryst. Growth Des.* **9**, 47–56 (2009).
2. Trask, A. V, Samuel, W. D. & Jones, W. Solvent-drop grinding : green polymorph control of cocrystallisation. *Chem. Commun.* 890–891 (2004).
3. Landenberger, K. B. & Matzger, A. J. Cocrystal engineering of a prototype energetic material: Supramolecular chemistry of 2,4,6-trinitrotoluene. *Cryst. Growth Des.* **10**, 5341–5347 (2010).
4. Nisar, M., Sung, H. H.-Y., Puschmann, H., Lakerveld, R., Haynes, R. & Williams, I. D. 11-Azaartemisinin cocrystals with preserved lactam: acid heterosynthons. *CrystEngComm* **20**, 1171–1314 (2018).
5. Issa, N. Towards more efficient screening of pharmaceutical cocrystals. *Thesis Submission, Univ. Coll. London* (2011).
6. Velaga, S. P., Basavoju, S. & Boström, D. Norfloxacin saccharinate-saccharin dihydrate cocrystal - A new pharmaceutical cocrystal with an organic counter ion. *J. Mol. Struct.* **889**, 150–153 (2008).
7. Trask, A. V. An overview of pharmaceutical cocrystals as intellectual property. *Mol. Pharm.* **4**, 301–309 (2007).
8. Elder, D. P., Holm, R. & De Diego, H. L. Use of pharmaceutical salts and cocrystals to address the issue of poor solubility. *Int. J. Pharm.* **453**, 88–100 (2013).
9. Chadha, R., Bhandari, S., Haneef, J., Khullar, S. & Mandal, S. Cocrystals of telmisartan: Characterization, structure elucidation, in vivo and toxicity studies. *CrystEngComm* **16**, 8375–8389 (2014).
10. Svoboda, V., Macfionnghaile, P., McGinty, J., Connor, L. E., Oswald, I. D. H. & Sefcik, J. Continuous Cocrystallization of Benzoic Acid and Isonicotinamide by Mixing-Induced Supersaturation: Exploring Opportunities between Reactive and Antisolvent Crystallization Concepts. *Cryst. Growth Des.* **17**, 1902–1909 (2017).
11. Zhou, Z., Chan, H. M., Sung, H. H. Y., Tong, H. H. Y. & Zheng, Y. Identification of New Cocrystal Systems with Stoichiometric Diversity of Salicylic Acid Using Thermal Methods. *Pharm. Res.* **33**, 1030–1039 (2016).
12. d'Agostino, S., Grepioni, F., Braga, D. & Ventura, B. Tipping the balance with the aid of stoichiometry: Room temperature phosphorescence versus fluorescence in organic cocrystals. *Cryst. Growth Des.* **15**, 2039–2045 (2015).
13. Li, Z. & Matzger, A. J. Influence of Coformer Stoichiometric Ratio on Pharmaceutical Cocrystal Dissolution: Three Cocrystals of Carbamazepine/4-Aminobenzoic Acid. *Mol. Pharm.* **13**, 990–995 (2016).
14. Chu, Q., Duncan, A. J. E., Papaefstathiou, G. S., Hamilton, T. D., Atkinson, M. B. J., Mariappan, S. V. S. & Macgillivray, L. R. Putting Cocrystal Stoichiometry to Work: A Reactive Hydrogen-Bonded 'superassembly' Enables Nanoscale Enlargement of a Metal-Organic Rhomboid via a Solid-State Photocycloaddition. *J. Am. Chem. Soc.* **140**, 4940–4944 (2018).
15. Evora, A. O. L., Castro, R. a E., Maria, T. M. R., Silva, M. R., Ter Horst, J. H., Canotilho, J. & Eusébio, M. E. S. A thermodynamic based approach on the investigation of a diflunisal pharmaceutical co-crystal with improved intrinsic dissolution rate. *Int. J. Pharm.* **466**, 68–75 (2014).

16. Bethune, S. J., Huang, N., Jayasankar, A. & Rodríguez-Hornedo, N. Understanding and predicting the effect of cocrystal components and pH on cocrystal solubility. *Cryst. Growth Des.* **9**, 3976–3988 (2009).
17. Saikia, B., Bora, P., Khatioda, R. & Sarma, B. Hydrogen Bond Synthons in the Interplay of Solubility and Membrane Permeability/Diffusion in Variable Stoichiometry Drug Cocrystals. *Cryst. Growth Des.* **15**, 5593–5603 (2015).
18. Buanz, A., Prior, T. J., Burley, J. C., Raimi-Abraham, B. T., Telford, R., Hart, M., Seaton, C. C., Davies, P. J., Scowen, I. J., Gaisford, S. & Williams, G. R. Thermal Behavior of Benzoic Acid/Isonicotinamide Binary Cocrystals. *Cryst. Growth Des.* **15**, 3249–3256 (2015).
19. Apshingekar, P. P., Aher, S., Kelly, A. L., Brown, E. C. & Paradkar, A. Synthesis of Caffeine/Maleic Acid Co-crystal by Ultrasound-assisted Slurry Co-crystallization. *J. Pharm. Sci.* **106**, 66–70 (2017).
20. Boyd, S., Back, K., Chadwick, K., Davey, R. J. & Seaton, C. Solubility, Metastable Zone Width Measurement and Crystal Growth of the 1:1 Benzoic Acid/Isonicotinamide Cocrystal in Solutions of Variable Stoichiometry. *J. Pharm. Sci.* **99**, 3779–3786 (2010).
21. Dubey, R. & Desiraju, G. R. Structural landscape of the 1 : 1 benzoic acid : Isonicotinamide cocrystal. *Chem. Commun.* **50**, 1181–1184 (2014).
22. Zakharov, B. a., Losev, E. a. & Boldyreva, E. V. Polymorphism of “glycine–glutaric acid” co-crystals: the same phase at low temperatures and high pressures. *CrystEngComm* **15**, 1693 (2013).
23. Woodall, C. H., Brayshaw, S. K., Schiffers, S., Allan, D. R., Parsons, S., Valiente, R. & Raithby, P. R. High-pressure crystallographic and spectroscopic studies on two molecular dithienylethene switches. *CrystEngComm* **16**, 2119–2128 (2014).
24. Losev, E. A., Zakharov, B. A. & Boldyreva, E. V. Polymorphic transformations in glycine co-crystals at low temperature and high pressure: Two new examples as a follow-up to a glycine–glutaric acid study. *CrystEngComm* **18**, 5869–5875 (2016).
25. Johnstone, R. D. L., Francis, D., Lennie, A. R., Marshall, W. G., Moggach, S. A., Parsons, S., Pidcock, E. & Warren, J. E. High-pressure polymorphism in L-serine monohydrate: Identification of driving forces in high pressure phase transitions and possible implications for pressure-induced protein denaturation. *CrystEngComm* **10**, 1758–1769 (2008).
26. Klotz, S., Chervin, J.-C., Munsch, P. & Le Marchand, G. Hydrostatic limits of 11 pressure transmitting media. *J. Phys. D. Appl. Phys.* **42**, 075413 (2009).
27. Bruker. APEX2. (2007).
28. Parsons, S. SHADE. Program for Empirical Absorption Corrections to High Pressure Data. (2004).
29. Sheldrick, G. M. SADABS, Programs for Scaling and Absorption Correction of Area Detector Data. *SADABS, Programs Scaling Absorpt. Correct. Area Detect. Data* University of Göttingen: Göttingen, Germany (2008).
30. Dolomanov, O. V, Bourhis, L. J., Gildea, R. J., Howard, J. A. K. & Puschmann, H. OLEX2: a complete structure solution, refinement and analysis program. 339–341 (2009).
31. Dawson, A., Allan, D. R., Parsons, S. & Ruf, M. Use of a CCD diffractometer in crystal structure determinations at high pressure. *J. Appl. Crystallogr.* **37**, 410–416 (2004).
32. Thorn, A., Dittrich, B. & Sheldrick, G. M. Enhanced rigid-bond restraints. *Acta Crystallogr. Sect. A Found. Crystallogr.* **68**, 448–451 (2012).

## Lauren Connor

33. Piermarini, G. J., Block, S., Barnett, J. D. & Forman, R. a. Calibration of the Pressure Dependence of the R//1 Ruby Fluorescence Line To 195 Kbar. *J. Appl. Phys.* **46**, 2774–2780 (1975).
34. Bruker. APEX3, SAINT and SADABS. (2015).
35. Sheldrick, G. M. SHELXT - Integrated space-group and crystal-structure determination. *Acta Crystallogr. Sect. A Found. Crystallogr.* **71**, 3–8 (2015).
36. Sheldrick, G. M. Crystal structure refinement with SHELXL. *Acta Crystallogr. Sect. C Struct. Chem.* **71**, 3–8 (2015).
37. Angel, R. J., Bujak, M., Zhao, J., Gatta, G. D. & Jacobsen, S. D. Effective hydrostatic limits of pressure media for high-pressure crystallographic studies. *J. Appl. Crystallogr.* **40**, 26–32 (2007).
38. Groom, C. R., Bruno, I. J., Lightfoot, M. P. & Ward, S. C. The Cambridge Structural Database. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **B72**, 171–179 (2016).
39. Aakery, C. B., Beatty, A. M. & Helfrich, B. A. 'Total synthesis' supramolecular style: Design and hydrogen-bond-directed assembly of ternary supermolecules. *Angew. Chemie - Int. Ed.* **40**, 3240–3242 (2001).
40. Harty, E. L., Ha, A. R., Warren, M. R., Thompson, A. L., Allan, D. R., Goodwin, A. L. & Funnell, N. P. Reversible piezochromism in a molecular wine-rack. *Chem. Commun.* **51**, 10608–10611 (2015).
41. Steiner, T. The hydrogen bond in the solid state. *Angew. Chem. Int. Ed.* **41**, 49–76 (2002).

***-Chapter six-***

*Data-driven approach to direct the co-crystallisation of  
anti-malarial Artemisinin*

## 6. Data-driven approach to direct the co-crystallisation of anti-malarial

### Artemisinin

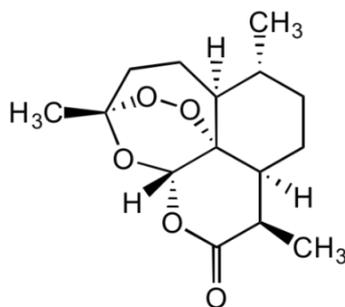
#### 6.1. Synopsis

A new solid form of antimalarial active pharmaceutical ingredient (API), Artemisinin, has been structurally determined by single crystal X-ray diffraction (SCXRD) methods. Random Forest (RF), a machine learning algorithm, was employed to direct the experimental screening, making use of open data on previous machine learning co-crystallisation approaches.<sup>1</sup>

#### 6.2. Introduction

##### 6.2.1. Artemisinin co-crystals

In the past twelve years, interest in co-crystallisation as a favourable method of solid state modification has risen in the pharmaceutical industry<sup>2-8</sup>. Artemisinin is an anti-malarial API and it is notoriously difficult to co-crystallise<sup>9</sup>.



**Figure 6.1 Artemisinin molecule.**

The molecule possesses oxygen H-bond acceptors, but no H-bond donor functionality. Therefore, and perhaps expectedly, a previous study focussed on the co-crystallisation of artemisinin yielded only two instances of co-crystal formation. The study comprised seventy-five liquid assisted grinding (LAG) experiments. Two successful co-crystals were formed with co-formers resorcinol (in a 2:1 ratio) and orcinol (in a 1:1 ratio). The supramolecular synthon present in both co-crystals is a phenol–lactone (O-H $\cdots$ O) moiety. The hydroxyl group of the phenol acts as the H-bond donor to artemisinin's carbonyl.<sup>10</sup> 11-azaartemisinin was the focus of a study by Nisar et al. aimed to co-crystallise this derivative with 25 different co-formers. Structurally, the sole difference is a nitrogen in place of the oxygen on the ether moiety of the ester group of artemisinin. Nitrogen creates a lactam moiety rather than the

## Lauren Connor

lactone which acted as an H-bond acceptor in the two known co-crystal examples of artemisinin. Atomic substitution in the place of one oxygen led to 13 new co-crystals of the 25 co-formers tested with 11-azaartemisinin.<sup>11</sup>

### 6.2.2. Computational screening

This study explores computationally-led approaches to the co-crystallisation of artemisinin as part of an academic placement in Nanyang Technological University, Singapore.

It is of interest to solid state scientists to minimise experimental co-crystal screening. Material waste, time spent in the laboratory, and risk of chemical exposure could all be reduced if computational methods were used to limit the number of screening experiments needed. Initial computational screening methods such as the synthon approach and hydrogen bond propensity approach were typical for hydrogen bonded entities.<sup>9</sup> The lack of hydrogen bonding in the Van der Waals solid, artemisinin, makes the application of synthon based approaches ineffective.

A statistical tool within the Cambridge Crystallographic Data Centre's (CCDC) Mercury programme was developed by Fábíán called Molecular complementarity tool.<sup>12</sup> The method is based on calculated descriptors of molecules within co-crystal structures of the Cambridge Structural Database (CSD). The molecular complementarity (MC) method suggests that molecules with similar molecular descriptors tend to co-crystallise. Fábíán calculated various molecular descriptors for components of co-crystals within the CSD and the most correlated descriptors were analysed. Polarity and shape of co-formers appear to be the strongest descriptor correlations. Fábíán determined threshold values for the differences in each molecular descriptor adjusted to the co-crystals in the CSD. The molecular complementarity tool used in Mercury is based on the threshold values. If the descriptors are within the threshold limit then a positive co-crystal result is predicted.

Other computational methods employ fluid phase thermodynamics to predict solid phase co-crystals. COSMO $therm$  software<sup>13</sup> uses COSMO-RS (COnductor like Screening MOdel for Real Solvents) theory to predict the thermodynamic equilibrium properties of liquids, which was originally developed by Klamt.<sup>14</sup> The co-crystal screening method breaks the solvent-accessible surface into sections and calculates the polarization charge density of the two surface sections (between cofomer 1) and 2). The pairwise surface interactions (polar and hydrogen bond) are quantified for each molecule and are then summed. The profile of the summed interactions is

## Lauren Connor

determined from density functional theory (DFT) calculations. The results are produced solely on excess enthalpy and do not consider entropy or sterics of the system. Excess enthalpy ( $H_{ex}$ ) can be taken as the tendency of the two components to associate in the mixture prior to co-crystallisation, and is defined as:

$$H_{ex} = H_{AB} - mH_{pureA} - nH_{pureB} \quad \text{Equation 6.1}$$

Where A and B denote co-formers 1 and 2, respectively; m and n denote the stoichiometric ratio of A and B, respectively;  $H_{pureA/B}$  denotes the enthalpies of the pure co-former reference states and  $H_{AB}$  denotes the enthalpy of the mixture of m:n stoichiometry co-formers.  $H_{ex}$  contains all enthalpic contributions.

### 6.2.3. Machine learning

Computational models and statistical approaches appear to have drawbacks in that they are specifically programmed on selected approaches. Co-crystallisation is still not fully understood and the ability to accurately model it is not possible yet, making predictions for challenging compounds such as artemisinin very difficult. A data-driven approach in the form of machine learning is favourable as such a method can train itself on any bank of known data and identify patterns without requiring any further context or fundamental understanding.<sup>15–18</sup> This is advantageous when applied to co-crystallisation, where it is not known precisely which variables are of significance. It is likely that mechanistic computational screening tools available do not take into account all variables significant to co-crystallisation. Moreover, the important variables may also change according to the co-crystallisation target molecules. On the other hand, machine learning improves with increasing data and diversity of cases considered.

Successful co-crystallisation using machine learning as a screening tool has already been explored by Wicker *et al.* using support vector machines (SVM).<sup>1</sup> The SVM algorithm enriched their co-crystal formation 2.6 fold. The group trained the SVM algorithm on a large acid and amide dataset which was produced in-house. Subsequently, the group tested an unseen bank of paracetamol co-crystallisation experiments. The dataset of 680 co-crystallisation results produced by this group could be the starting point of an open access bank of co-crystallisation data which can be utilised by other solid state groups.

Indeed, the training dataset by Wicker *et al.* was the starting point for this study. Machine learning requires sufficient data points representing the full range of

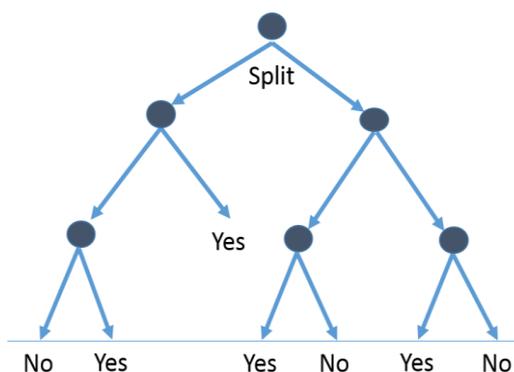
## Lauren Connor

outcomes it is to predict. In the context of co-crystallisation, this means that a similar quantity of negative experimental outcomes are required as positive; given the distinct lack of experimentally negative co-crystallisations presented in literature, this was the initial hurdle of a machine learning approach to co-crystallisation screening. The study by Wicker *et al.* is a rare example in literature not only providing a wealth of both positive and negative co-crystallisation outcomes, but also showcasing the equal value of negative results.

### 6.2.3.1. Random forest

There are many different machine learning algorithms available, each of which perform the same task in different ways. In this chapter, RF is used to generate a predictive model to direct the co-crystallisation of artemisinin towards more suitable co-formers. RF is a form of supervised machine learning, meaning it requires to be given the outcomes of each point of data in the set it is trained with (*c.f.* an unsupervised technique such as principle component analysis). It was used to classify co-former pairs as successfully formed or unsuccessful co-crystallisations (termed 'yes' and 'no'). When the outcome is a categorical variable such as this, the process of mapping it to the input variables is known as classification. Many algorithms, including RF, can also handle numerical outcomes, known as regression.

RF is a method based on decision trees. A decision tree makes a series of splits in the training dataset, each one increasing the so-called 'purity' (the ratio of one class to the other) at each node until only a single class remains at the terminals (Figure 6.2). The splits are based on a chosen input variable with an applied threshold. RF generates a series of decision trees (500 by default), introducing random elements in order to keep each tree diverse. The random elements of a RF include shortlists of the training set and shortlists of the variables. A random subset of the full training set is chosen (about 67 % by default), which leads to different biases throughout. At each node, the best variable to split with is chosen from a random shortlist (for classification this is defined as  $\sqrt{\text{total number of variables}}$ ). When it comes to making a prediction, the unknown is run through each tree, casting a vote for whichever class it reaches; the majority vote is then taken as the overall prediction for the unknown.



**Figure 6.2 Representation of a decision tree based on co-crystallisation with responses ‘yes’ for successful co-crystallisation and ‘no’ for unsuccessful co-crystallisation. A RF is composed of many decision trees.**

### 6.3. Computational screening methods

#### 6.3.1. Molecular complementarity

Co-formers chosen by considering the structure complementarity were selected in Materials Mercury. The molecular descriptors for each of the co-formers were calculated using the molecular complementarity tool. If descriptors for each molecule differ by less than the threshold values set in the software then co-crystal prediction is returned as ‘yes’.

#### 6.3.2. COSMOtherm based on excess enthalpy

COSMOtherm software was used to obtain the excess enthalpy of mixtures created from pure components. COSMOtherm<sup>13</sup> uses COSMO-RS theory.<sup>14</sup> COSMO-RS consists of using the screening charge density, computed from first principles calculations in combination with fast statistical thermodynamics to compute the chemical potential of a compound in solution. Each chosen co-former was tested in in a 1:1, 2:1 and 1:2 stoichiometric ratio. The results were ranked in terms of excess enthalpy.

#### 6.3.3. Machine learning

The RF algorithm was trained on a bank of 769 known co-crystallisation results. This style of classification machine learning is a probability approach with a 50% cut off between ‘successful’ and ‘unsuccessful’ co-crystallisation. Open data from a previous paper by Wicker *et al.* was used as the basis of the training set with additional data

## Lauren Connor

points added. Additional data points added include: negatives from experimental work carried out on artemisinin; known co-crystals of artemisinin from literature; known positive and negative results of 11-azaartemisinin (structurally similar to artemisinin); a variety of known co-crystals from the CSD.

Molecular SMILES (Simplified molecular-input line-entry system) codes for artemisinin and each co-former enabled calculation of molecular descriptors using Molecular Operating Environment (MOE) software.<sup>19</sup> Molecular descriptors with zero variance were eliminated leaving a remaining 435 molecular descriptors to use when training the dataset. 1000 trees were used in the training stage. The nature of the RF algorithm naturally handles correlated variables. A 10-fold cross validation was performed. The training data co-former cases were split tenfold and each tenth tested on the remaining 9/10 training data. This provided ten accuracy scores which were averaged. Once the model was sufficiently well trained, a test dataset was constructed and artemisinin was tested against each of the 118 co-formers from the training dataset. Top results from the machine learning testing dataset were tried experimentally.

### 6.4. Experimental

#### 6.4.1. Ball mill

1:1 molar ratio amounts of co-former with artemisinin were ball milled for 20 minutes at 30 Hz on a Retsch MM400 mechanical mill in 10 mL steel jars with two steel balls. On identification of a new phase, the ball mill experiment was repeated with a 50 $\mu$ L solvent drop of hexane.

#### 6.4.2. X-ray Powder Diffraction (XRPD)

Material from the ball mill experiments were analysed by XRPD. XRPD data were collected on a Bruker D8 Advance II diffractometer with Cu source radiation (1.540596 Å) at ambient temperature. The data were collected from 4-35°, 0.017° step in the 2 $\theta$ -range.

### 6.4.3. Differential Scanning Calorimetry

Aluminium DSC pans of 40 $\mu$ L volume, with pierced lids containing a 2-5mg of sample were heated from 25°C to 170°C at a rate of 5°C/min on a Netzsch DSC214 Polyma. Data were analysed using Netzsch Proteus Analysis Software (current version 7.1.0).

### 6.4.4. Single crystal growth

A new solid phase identified by XRPD led to single crystal experiments. A small solvent screen was performed using the 1:1 co-crystalline powder. Approximately 20mg of material was used per solvent. Single crystals were harvested from the vials and tested by single crystal X-ray diffraction (SCXRD).

Solvents tested: Acetone, acetonitrile, chloroform, ethanol, ethyl acetate, hexane, and methanol.

### 6.4.5. SCXRD

Crystals harvested from hexane were of sufficient size (0.3mm x 0.18mm x 1.12mm) and quality for SCXRD. X-ray diffraction intensities were collected on a Bruker D8 Venture diffractometer ( $\lambda = 1.54046 \text{ \AA}$ ) equipped with an Oxford Cryosystems low temperature device operating at 123K. Data were collected from a sequence of eleven scans. The data were indexed and integrated using SAINT<sup>20</sup>. An absorption correction was applied in using SADABS<sup>21</sup>. The crystal structure was solved by intrinsic phasing and refined in *Olex2* ( $F^2$ )<sup>22</sup>. All atoms were anisotropically refined. Hydrogen atoms were placed in calculated positions and constrained to ride their parent atom.

Table 6.1 Experimental table for Artemisinin and 1-Napthol co-crystal (1:1)

	art_nap_evap_hexane_3_0m
Chemical formula	C <sub>15</sub> H <sub>22</sub> O <sub>5</sub> ·C <sub>10</sub> H <sub>8</sub> O
$M_r$	426.49
Crystal system, space group	Orthorhombic, $P2_12_12_1$
Temperature (K)	123
$a, b, c$ (Å)	8.9041 (5), 9.5951 (6), 25.3104 (15)
$V$ (Å <sup>3</sup> )	2162.4 (2)
$Z$	4
Radiation type	Cu $K\alpha$
$\mu$ (mm <sup>-1</sup> )	0.76
Crystal size (mm)	0.1 × 0.05 × 0.04
Diffractometer	Bruker APEX-II CCD
Absorption correction	Multi-scan SADABS2016/2 (Bruker,2016/2) was used for absorption correction. $wR2(int)$ was 0.0873 before and 0.0434 after correction. The Ratio of minimum to maximum transmission is 0.9082. The $I/2$ correction factor is Not present.
$T_{min}, T_{max}$	0.685, 0.754
No. of measured, independent and observed [ $I > 2s(I)$ ] reflections	29744, 4506, 4431
$R_{int}$	0.028
$(\sin \theta/\lambda)_{max}$ (Å <sup>-1</sup> )	0.631
$R[F^2 > 2\sigma(F^2)], wR(F^2), S$	0.026, 0.070, 1.04
No. of reflections	4506
No. of parameters	284
H-atom treatment	H-atom parameters constrained
$\Delta\rho_{max}, \Delta\rho_{min}$ (e Å <sup>-3</sup> )	0.16, -0.18
Absolute structure	Flack $x$ determined using 1870 quotients $[(I+)-(I-)]/[(I+)+(I-)]$ (Parsons, Flack and Wagner, Acta Cryst. B69 (2013) 249-259).
Absolute structure parameter	0.02 (2)

### 6.5. Results and Discussion

#### 6.5.1. Negative experimental results from co-crystal computational methods

The results of computational tools such as the molecular complementarity tool (MC) and co-crystallisation tool in CosmoRS results are shown in Table 6.2. Co-crystallisation between artemisinin and each co-former was attempted by various experimental techniques such as neat grinding (NG), liquid assisted grinding (LAG), cooling and evaporative co-crystallisation. Various solvents were tested in each solvent based method, a total of 10 solvents were tried in solution methods. Only a 1:1 stoichiometric ratio was attempted in mechanical methods (neat grinding and liquid assisted grinding). For solution methods, the stoichiometric ratio was varied (1:1, 2:1 and 1:2). All experimental results listed in Table 6.2 were negative. During an industrial placement at AstraZeneca, access to a shortlist of 350 GRAS (Generally regarded as safe) co-formers was granted. The screen was run on *Cosmotherm* software and the results ranked by excess enthalpy. The larger the negative excess enthalpy the better the indication of co-crystal formation. MC and *Cosmotherm* were used to screen a wide range of co-formers before a selection of the results were experimentally tested. Notably, the 2009 paper on the development of MC by Fábíán *et al.*<sup>12</sup> suggests that a combination of computational methods should be employed to predict co-crystallisation. Artemisinin and selected co-formers were experimentally screened in stages based on separate types of virtual screening. If a combination of computational approaches were used from the outset only 7 of the 16 co-formers listed would have been tested. However, in challenging cases such as artemisinin, virtual false negatives are of more concern than false positives. It is experimentally difficult to validate a false positive and the costs in the case of a false positive are laboratory time and material. A false negative can cost the experimentalist a missed, novel, solid phase. Balance is required in the case of virtual and experimental screening, to this effect, a 2014 paper on co-crystal design by Wood *et al.*<sup>9</sup> recommends that the MC tool be used as a filter for co-crystallisation. The MC tool is tailored to predict 90% of the co-crystals present in the database bringing with it a host of false positives. Wood *et al.* also suggest a more statistical approach for systems where no hydrogen bonds or only weak hydrogen bonds are available for co-crystallisation (such as artemisinin).

Table 6.2 Computational and experimental screening methods tried prior to machine learning approach tried.

CF	Computational		Experimental	Co-crystal formation
	MC	CosmoRS-H <sub>ex</sub>		
	MC	CosmoRS-H <sub>ex</sub>		
<b>Lactide</b>	Y	-0.047655	NG, LAG, evap., cooling	N
<b>benzoic acid</b>	Y	-0.7569	NG, LAG, evap., cooling	N
<b>isonicotinamide</b>	Y	-0.1470175	NG, LAG, evap., cooling	N
<b>Saccharin</b>	Y	-0.65208	NG, LAG, evap., cooling	N
<b>paracetamol</b>	N	-0.763365	NG, LAG, evap., cooling	N
<b>acetic acid</b>	Y	-0.485725	NG, LAG	N
<b>Mannitol</b>	Y	0.08268	NG, LAG	N
<b>xylitol</b>	Y	0.20598	NG, LAG	N
<b>nicotinamide</b>	Y	0.20121	NG, LAG	N
<b>glycerol</b>	Y	0.264825	NG, LAG	N
<b>oxalic acid</b>	N	-1.87045	NG, LAG	N
<b>amidosulfonic acid</b>	N/A	-2.09134	NG, LAG	N
<b>para-aminobenzoic acid</b>	N	-1.37987	NG, LAG	N
<b>3,5 dihydroxybenzoic acid</b>	Y	-1.729405	NG, LAG	N
<b>cholesterol</b>	Y	N/A	NG, LAG	N
<b>3,5 dinitrobenzoic acid</b>	Y	-1.415515	NG, LAG	N

### 6.5.2. Machine learning success

Machine learning is a purely statistical approach which incorporates many descriptors without the user knowing which molecular descriptors are most influential to the result. The data used to train the model included more than 600 data points from open data published by Wicker *et al.* The main problem of machine learning for co-crystallisation is the volume of data required, as well as, in the case of co-crystallisation, the lack of negative co-crystallisation attempts published. A balance between positive and negative responses is vital to remove bias. The open data by Wicker *et al.* supplied 403 negatives. Wicker *et al.* ball milled components for 20 minutes at 30 Hz to produce the responses for the training dataset. The paper by Wicker *et al.* was tailored to suit the needs of a machine learning approach in that the experimental method was rigid. However, all of the training dataset were specifically acids and amides which was appropriate for their paracetamol target but did not suit the case of artemisinin. By introducing known co-crystals from the CSD database to improve structural diversity for the artemisinin target, the experimental diversity, inherently, increased. Results added to the dataset from the 11-azaartemisinin paper were tested by LAG<sup>11</sup> and some co-crystals from the CSD database included other solution based methods for co-crystal formation.<sup>5,23-31</sup> Varying the experimental techniques distorts the model. Nevertheless, the top results from the machine learning testing dataset were kept experimentally consistent with the method used by Wicker *et al.* The compounds were tested with artemisinin by neat grinding for 20 mins at 30 Hz.

**Table 6.3 Test dataset predictions. Highlighted in red are the co-crystals which failed to co-crystallise by ball milling. Highlighted in green is the successfully co-crystallised co-former.**

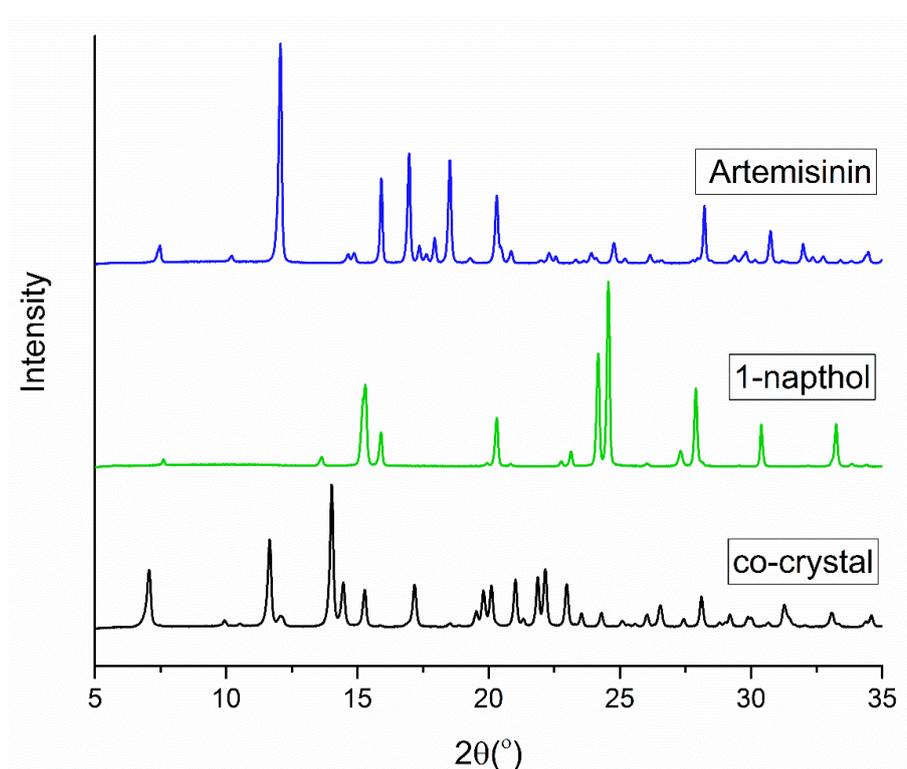
Co-former	Prediction	Probability (%)
Hydroquinone	Yes	68.2
1,4-diethynylbenzene	Yes	58.3
1-naphthol	Yes	57.6
Phenazine	Yes	54.1
1,4-di-4-pyridyl-ethylene	Yes	52.6
p-phenylenediamine	Yes	52.2
1,2-bis-4-pyridyl-ethane	Yes	51.8
Benzidine	Yes	51.8
4,4'-bipyridine	Yes	50.6
Piperazine	Yes	50.6
2,9-dimethylphenanthroline	Yes	50.4

The top tested predictions were co-formers orcinol and resorcinol, two known co-crystals of artemisinin with alcohol functionalities and were invalid for the purpose of this study and are not included in the table. Hydroquinone, 1,4-diethynylbenzene and 1-naphthol were tested with 1-naphthol readily forming a new co-crystal with artemisinin (ArtNap). The other co-formers tested are part of an ongoing study to validate the machine learning method.

### 6.5.3. Analysis of ArtNap

#### 6.5.3.1. Powder analysis of ArtNap

The ball milled 1:1 ratio mixture revealed almost complete conversion to a 1:1 co-crystal mix. Two samples of the material from the same ball mill batch revealed a new solid phase but only one sample indicated complete conversion to the pure sample. A subsequent grinding experiment with a solvent drop of hexane allowed full conversion to the 1:1 ArtNap co-crystal.



**Figure 6.3 XRPD patterns of Artemisinin, 1-naphthol and the ball milled sample revealing new solid phase ArtNap.**

### 6.5.3.2. Thermal analysis of ArtNap

DSC analysis confirmed the melting point of Artemisinin to be 152.3°C and the melting point of 1-Naphthol to be 96.5°C. The thermogram of ArtNap reveals a much lower thermal stability with a melting point of 68.7°C.

Table 6.4 Thermogram meta data

	Onset (°C)	Peak (°C)	Area (J/g)
<b>ArtNap</b>	68.7	71.5	-69.74
<b>Artemisinin</b>	152.3	153.3	-82.6
<b>1-Naphthol</b>	96.5	98.5	-172.2

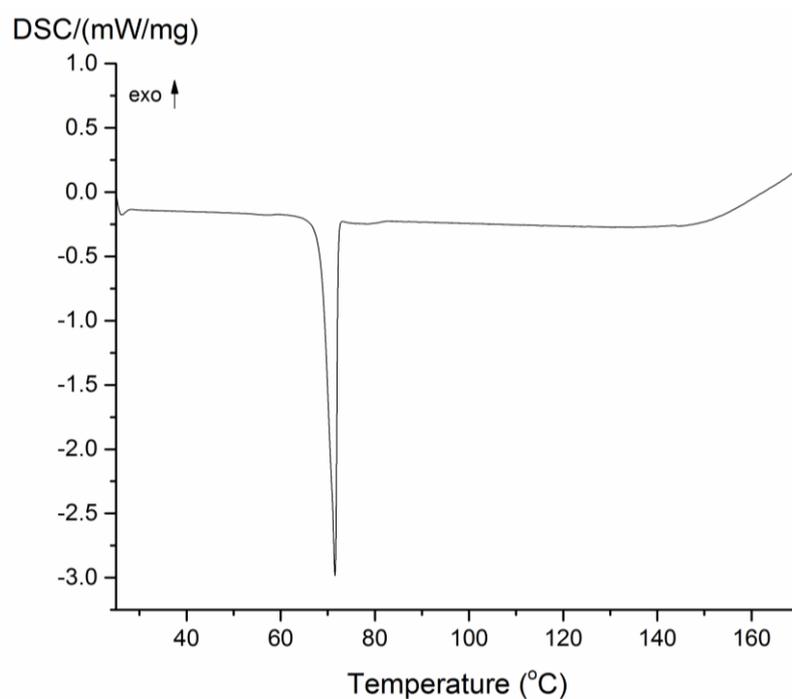


Figure 6.4 Thermogram of the ArtNap co-crystal.

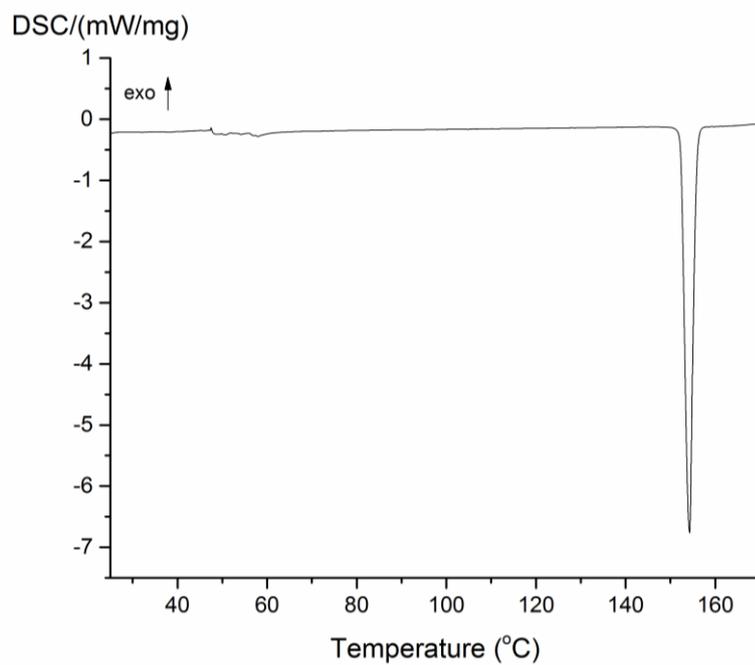


Figure 6.5 Thermogram of artemisinin.

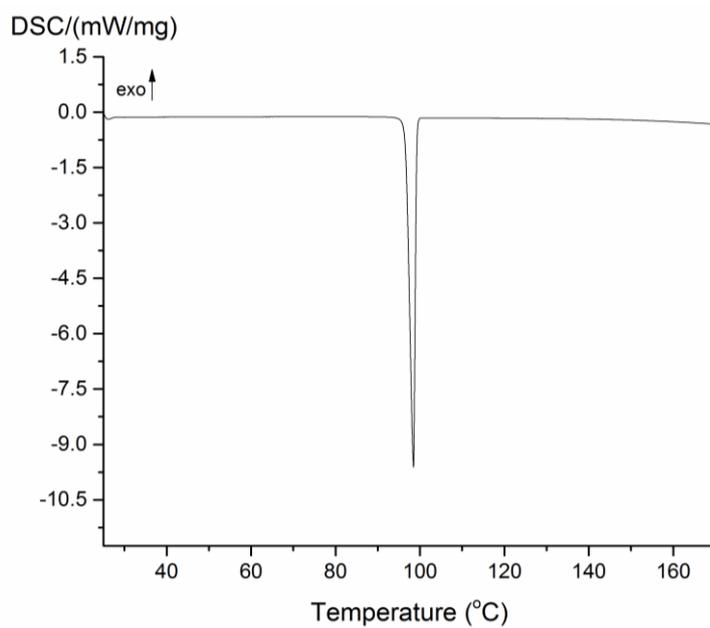
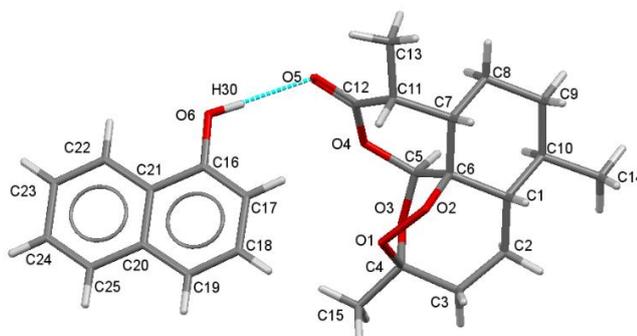


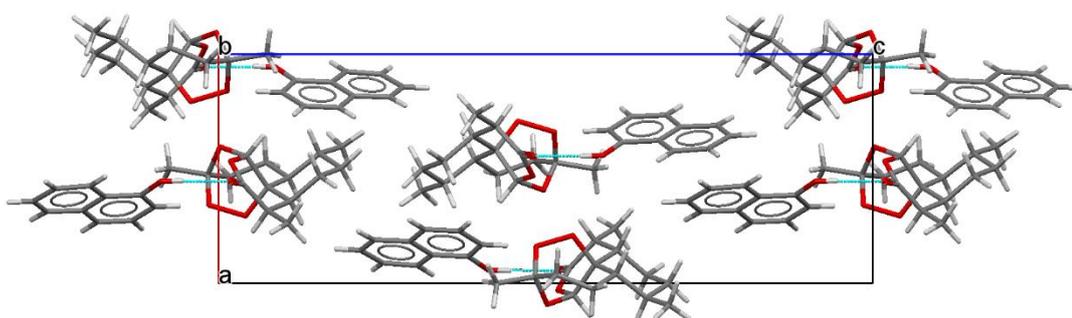
Figure 6.6 Thermogram of 1-naphthol.

### 6.5.3.3. Crystal Structure of ArtNap

The crystal structure of ArtNap was confirmed at low temperature with an R-factor of 2.64%. ArtNap forms as a 1:1 stoichiometric ratio co-crystal between artemisinin and 1-naphthol. Only one hydrogen bond is present in the ArtNap co-crystal. The hydroxyl of 1-naphthol's phenol ring donates to the lactone moiety of artemisinin. Both known co-crystals of artemisinin contain the same alcohol-lactone (O-H $\cdots$ O) supramolecular synthon. Other alcohols have been tested in this study and by Karki *et al.*<sup>10</sup> with no success. In the case of artemisinin, the machine learning approach did steer the experimental study in the same direction as a supramolecular synthon approach would. Nevertheless, machine learning may be a more comprehensive tool to predict co-crystal formation than the supramolecular synthon approach, based on known co-crystals.



**Figure 6.7** The labelled co-crystal of artemisinin and 1-naphthol, only the hydrogen atom involved in the supramolecular synthon is labelled for clarity. The colours of the elements are consistent throughout (carbon, grey; oxygen, red; hydrogen, white).



**Figure 6.8** View down the b axis showing 1:1 molecular units of the co-crystal and the hydrogen bonding between components.

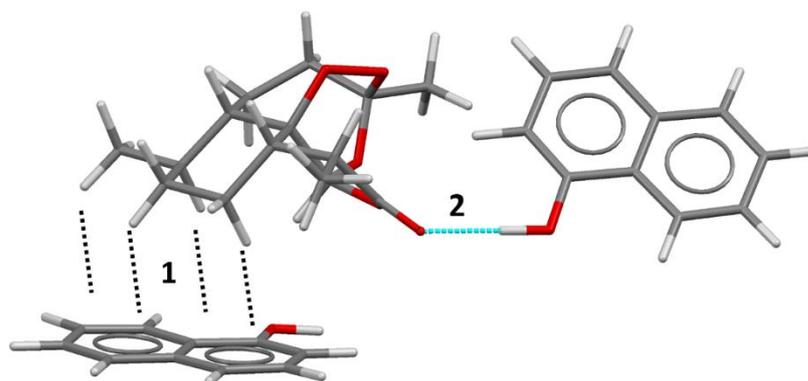
## Lauren Connor

ArtNap co-crystallises in orthorhombic  $P2_12_12_1$ , with  $Z'$  of 2. The hydrogen bonding interactions are parallel to the  $c$ -axis (Figure 6.8).

*In silico* calculation of molecule to molecule interactions in the structure were performed using Pixel.<sup>32,33</sup> The calculations reveal the strongest intermolecular force calculated between artemisinin and 1-naphthol co-former is not the hydrogen bonded dimer but is between the artemisinin and the naphthol closest to the alkyl backbone and as such is mainly dispersive in nature (Figure 6.9).

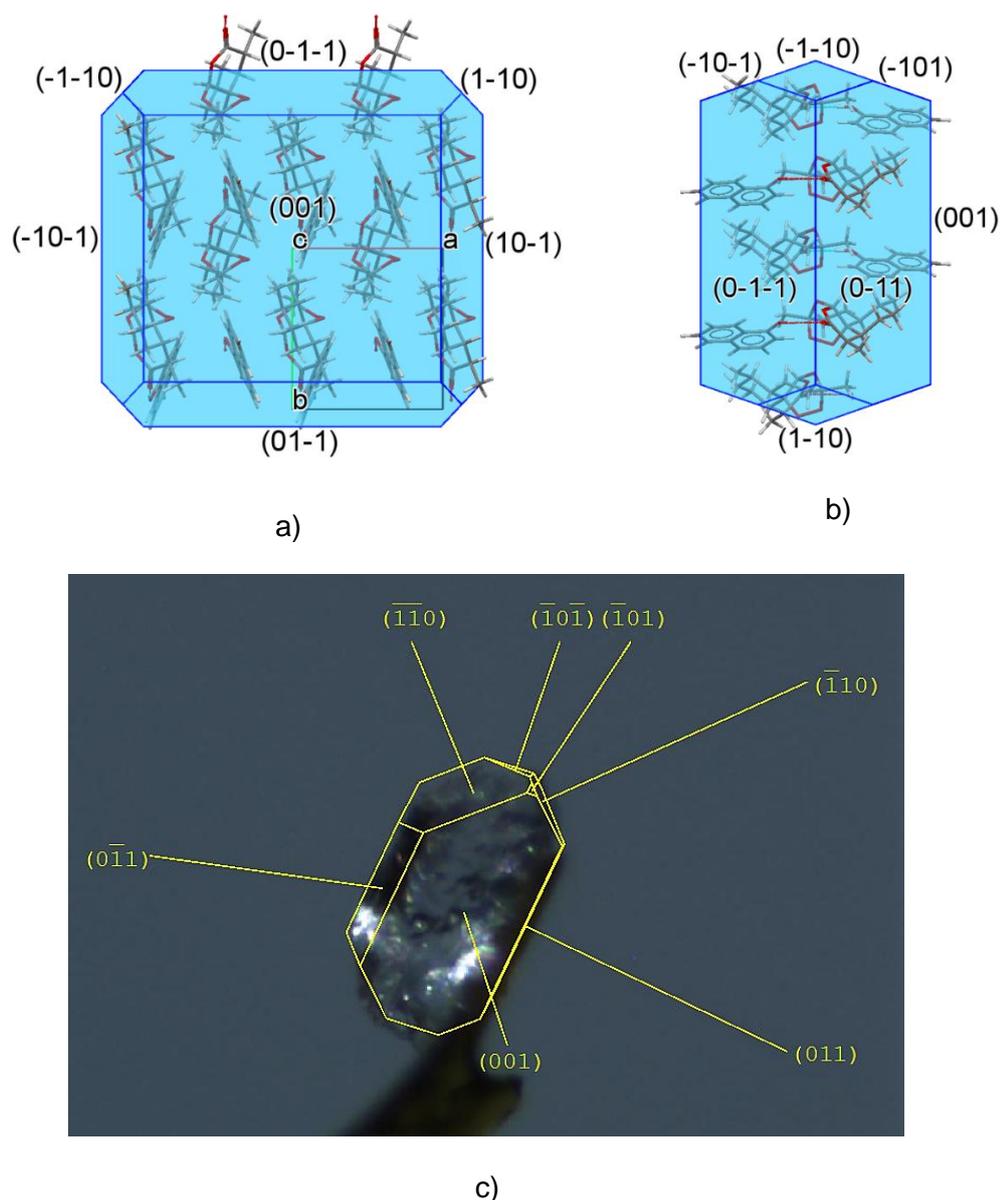
**Table 6.5 Top two molecule to molecule interaction energies from artemisinin to 1-naphthol. Distance given in Å. Columbic, polarisation, dispersion, repulsion and total energy terms given in  $\text{kJ mol}^{-1}$ .**

	Symmetry operator	Dist.(Å)	Cou.	Pol.	Disp.	Rep.	Total
1	1-x,0.5+y,0.5-z	4.904	-23.1	-9	-50.4	43.2	-39.3
2	1.5-x,-y,-0.5+z	6.978	-53	-24.9	-29.5	68.8	-38.6



**Figure 6.9 Artemisinin is hydrogen bonded to 1-naphthol in 1:1 stoichiometric ratio as well as interacting with a neighbouring 1-naphthol molecule via dispersive interactions. Two main interactions labelled to match Table 6.5.**

CH-  $\pi$  interactions dominate the ArtNap co-crystal energetic structure. 1-naphthol is a polycyclic aromatic system. Artemisinin has four C-H moieties directional with the two  $\pi$  systems of 1-Naphthol in an edge-face type interaction depicted in Figure 6.9 as interaction 1. This is known as the Gulliver effect, many smaller interactions amount to more than one stronger interaction.

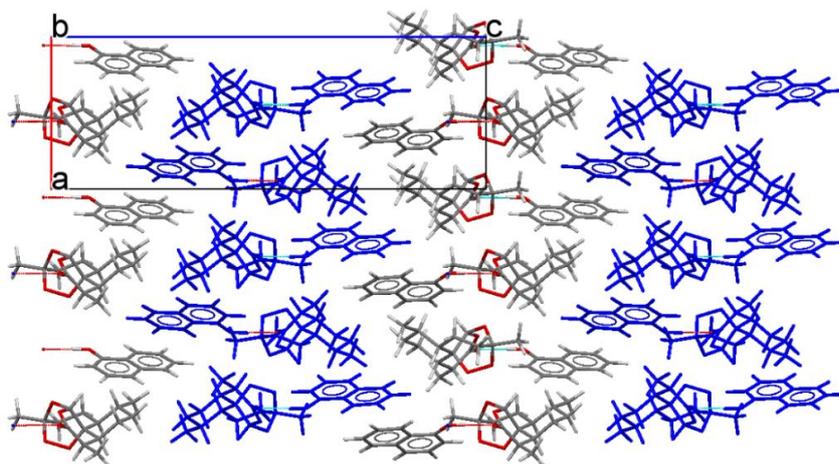


**Figure 6.10 BFDH predicted morphology viewed down the a) *c*-axis and b) *b*-axis. c) face indexing confirmed the largest face to be the (0 0 1) face, consistent with BFDH morphology predictions.**

Hydrogen bonding of the co-crystal is in the *c*-axis. Zig zagged chains of hydrogen bonded dimers that are stacked through the CH- $\pi$  dispersive forces exist along the *a*-axis. No strong interactions are present between these chains along the *c*-axis. The columbic force attributed to the hydrogen bonding in *c* (Figure 6.11) plays a secondary role, in terms of the overall lattice energy, to the dispersive interactions along the *a*-axis. The lack of interchain interaction in the *c*-axis direction leads to the largest observed and predicted crystal face (0 0 1). The observed crystal face (0 0 1) (Figure 6.10 a)) is elongated in the *a*-axis direction compared to the predicted crystal

## Lauren Connor

morphology (Figure 6.10 c)) which can be attributed to a quicker crystal growth rate along this direction; this is further experimental evidence to highlight the importance of these non-directional interactions. It would be more energetically favourable to add molecules to the bottom of the chain in the a-axis direction. Starting a new chain on the surface of (0 0 1) in the c axis would be less energetically favourable in comparison.



**Figure 6.11** View of the ArtNap down the b axis showing the layers of interactive components.

### 6.6. Conclusion

Machine learning proved a useful statistical tool to computationally direct experimental solid phase screening of artemisinin where previous computational methods had individually failed. The application of machine learning clearly has its uses for enriching the hit rate in a screen. The training dataset produced in this study is an extension of the dataset produced by Wicker *et al.* Introducing structural diversity to the dataset produced by Wicker *et al.* allowed for a more comprehensive approach in targeting the co-crystallisation of artemisinin. The cost of increased experimental diversity in the training set did not affect our aim but is an avenue which could be explored in further studies. The pure artemisinin and 1-naphthol co-crystal readily forms by LAG for 20 minutes with a solvent drop of hexane. Single crystals were made by evaporative co-crystallisation from hexane. The co-crystal was fully structurally analysed at 123K ( $R = 2.64\%$ ). The observed morphology and BFDH morphology are consistent but with an elongated (0 0 1) face due to zig zagged chains of coulombic and dispersive interactions in this direction as confirmed by Pixel methods

## References

1. Wicker, J. G. P., Crowley, L. M., Robshaw, O., Little, E. J., Stokes, S. P., Cooper, R. I. & Lawrence, S. E. Will they co-crystallize? *CrystEngComm* **19**, 5336–5340 (2017).
2. Buanz, A. B. M., Telford, R., Scowen, I. J. & Gaisford, S. Rapid preparation of pharmaceutical co-crystals with thermal ink-jet printing. *CrystEngComm* **15**, 1031–1035 (2013).
3. Steed, J. W. The role of co-crystals in pharmaceutical design. *Trends in Pharmacological sciences* **34**, (Durham University, 2013).
4. Maeno, Y., Fukami, T., Kawahata, M., Yamaguchi, K., Tagami, T., Ozeki, T., Suzuki, T. & Tomono, K. Novel pharmaceutical cocrystal consisting of paracetamol and trimethylglycine, a new promising cocrystal former. *Int. J. Pharm.* **473**, 179–186 (2014).
5. Basavoju, S., Boström, D. & Velaga, S. P. Indomethacin-saccharin cocrystal: Design, synthesis and preliminary pharmaceutical characterization. *Pharm. Res.* **25**, 530–541 (2008).
6. Jones, W., Motherwell, W. D. S. & Trask, A. V. Pharmaceutical Cocrystals : An Emerging Approach to Physical Property. **31**, 875–879 (2006).
7. Vishweshwar, P., A. McMahon, J., A.Bis, J. & J. Zaworotko, M. Pharmaceutical Cocrystals. *Journal of pharmaceutical sciences* **95**, 499–516 (2006).
8. McNamara, D. P., Childs, S. L., Giordano, J., Iarriccio, A., Cassidy, J., Shet, M. S., Mannion, R., O'Donnell, E. & Park, A. Use of a glutaric acid cocrystal to improve oral bioavailability of a low solubility API. *Pharm. Res.* **23**, 1888–1897 (2006).
9. Wood, P. A., Feeder, N., Furlow, M., Galek, P. T. A., Groom, C. R. & Pidcock, E. Knowledge-based approaches to co-crystal design. *CrystEngComm* **16**, 5839–5848 (2014).
10. Karki, S., Friscic, T., Fabian, L. & Jones, W. New solid forms of artemisinin obtained through cocrystallisation. *CrystEngComm* **12**, 4038–4041 (2010).
11. Nisar, M., Sung, H. H.-Y., Puschmann, H., Lakerveld, R., Haynes, R. & Williams, I. D. 11-Azaartemisinin cocrystals with preserved lactam: acid heterosynthons. *CrystEngComm* **20**, (2018).
12. Fabian, L. Cambridge Structural Database Analysis of Molecular Complementarity in Cocrystals. *Cryst. Growth Des.* **9**, 1436–1443 (2009).
13. COSMOlogic GmbH & Co KG. COSMOtherm.
14. Eckert, F. & Klamt, A. Fast Solvent Screening via Quantum Chemistry: COSMO-RS Approach. *AIChE J.* **48**, 369–385 (2002).
15. Johnston, A., Johnston, B. F., Kennedy, A. R. & Florence, A. J. Targeted crystallisation of novel carbamazepine solvates based on a retrospective Random Forest classification. *CrystEngComm* **10**, 23–25 (2008).
16. Raccuglia, P., Elbert, K. C., Adler, P. D. F., Falk, C., Wenny, M. B., Mollo, A., Zeller, M., Friedler, S. A., Schrier, J. & Norquist, A. J. Machine-learning-assisted materials discovery using failed experiments. *Nature* **533**, 73–76 (2016).
17. Bhardwaj, R. M., Johnston, A., Johnston, B. F. & Florence, A. J. A random forest model for predicting the crystallisability of organic molecules. *CrystEngComm* **17**, 4272–4275 (2015).
18. Lekamge, B. M. T., Sowmya, A. & Newman, J. Prediction of Protein X-ray Crystallisation Trial Image Time-courses. *Proc. 6th Int. Conf. Pattern Recognit. Appl. Methods* 663–668 (2017). doi:10.5220/0006246506630668

19. ULC, C. C. G. Molecular Operating Environment (MOE). (2013).
20. Bruker. APEX3, SAINT and SADABS. (2015).
21. Sheldrick, G. M. SADABS, Programs for Scaling and Absorption Correction of Area Detector Data. *SADABS, Programs Scaling Absorpt. Correct. Area Detect. Data* University of Göttingen: Göttingen, Germany (2008).
22. Dolomanov, O. V, Bourhis, L. J., Gildea, R. J., Howard, J. A. K. & Puschmann, H. OLEX2: a complete structure solution, refinement and analysis program. 339–341 (2009).
23. Childs, S. L., Wood, P. A., Rodriguez-Hornedo, N., Reddy, L. S. & Hardcastle, K. I. Analysis of 50 Crystal Structures Containing Carbamazepine Using the Materials Module of Mercury CSD Analysis of 50 Crystal Structures Containing Carbamazepine Using the Materials Module of Mercury CSD. *Cryst. Growth Des.* (2009). doi:10.1021/cg801056c
24. Buanz, A., Prior, T. J., Burley, J. C., Raimi-Abraham, B. T., Telford, R., Hart, M., Seaton, C. C., Davies, P. J., Scowen, I. J., Gaisford, S. & Williams, G. R. Thermal Behavior of Benzoic Acid/Isonicotinamide Binary Cocrystals. *Cryst. Growth Des.* **15**, 3249–3256 (2015).
25. Weyna, D. R., Shattock, T., Vishweshwar, P. & Zaworotko, M. J. Synthesis and Structural Characterization of Cocrystals and Pharmaceutical Cocrystals: Mechanochemistry vs Slow Evaporation from Solution. *Cryst. Growth Des.* **9**, 1106–1123 (2009).
26. Sarcevic, I., Orola, L., Veidis, M. V., Podjava, A. & Belyakov, S. Crystal and molecular structure and stability of isoniazid cocrystals with selected carboxylic acids. *Cryst. Growth Des.* **13**, 1082–1090 (2013).
27. Ghosh, S. & Malla Reddy, C. Co-crystals of caffeine with substituted nitroanilines and nitrobenzoic acids: Structure-mechanical property and thermal studies. *CrystEngComm* **14**, 2444–2453 (2012).
28. Yip, W., Territories, N. & Kong, H. Isolation and crystal structures of triethylenediamine-hydroquinone (1/1) and triethylenediamine-phenol (1/2). *J. Crystallogr. Spectrosc. Res.* **14**, 457–465 (1984).
29. Bosch, E. Cocrystals of 1,4-diethynylbenzene with 1,3-diacetylbenzene and benzene-1,4-dicarbaldehyde exhibiting strong nonconventional alkyne–carbonyl C—H...O hydrogen bonds between the components. *Acta Crystallogr. Sect. C Struct. Chem.* **72**, 748–752 (2016).
30. Perkin, J. C. S. I. Molecular complexes. Part XXIII. Crystal and molecular structures of  $\alpha$ -naphthol–1,2,4,5-tetracyanobenzene and  $\beta$ -naphthol–1,2,4,5-tetracyanobenzene. *J. Chem. Soc.* **0**, 160–164 (1978).
31. Ramon, G., Davies, K. & Nassimbeni, L. R. Structures of benzoic acids with substituted pyridines and quinolines: Salt versus co-crystal formation. *CrystEngComm* **16**, 5802–5810 (2014).
32. Gavezzotti, A. Calculation of Intermolecular Interaction Energies by Direct Numerical Integration over Electron Densities . 2 . An Improved Polarization Model and the Evaluation of Dispersion and Repulsion Energies. *J. Phys. Chem.* **107**, 2344–2353 (2003).
33. Maschio, L., Civalleri, B., Ugliengo, P. & Gavezzotti, A. Intermolecular Interaction Energies in Molecular Crystals: Comparison and Agreement of Localized Møller-Plesset 2, Dispersion-Corrected Density Functional, and Classical Empirical Two-Body Calculations. *J. Phys. Chem.* **115**, 11179–11186 (2011).

***-Chapter seven-***

*Concluding Remarks and Further Work*

### 7. Concluding remarks and further work

The overall aim of this research was to structurally investigate multicomponent constructs and small molecule systems under compression. Considering the main body of this thesis, the key conclusions drawn will be considered systematically.

Chapter three involved the continuous compression of a co-crystal of indomethacin and saccharin, this sparked interest in the rate of compression of the amide dimer. Data mining of the CSD revealed that there were few compression studies involving the amide dimer in the database for comparison. Nonetheless, a pattern was observed for the rate of compression in amides which are involved in a second hydrogen bonding interaction (primary amides) and those which are not (secondary amides). The secondary amides are not stabilised by a hydrogen bond in another direction, and so, compress at a faster rate. The hydrogen bonding observed in the ambient saccharin dimer of INSA is due to the dimer sitting in a 'pocket' with the indomethacin framework, at an ideal hydrogen bonding distance. A calculation is underway to determine the ideal hydrogen bonding distance of the saccharin dimer at pressure. It would be interesting to calculate the ideal bonding distances of a collection of primary and secondary amides by multiple point energy calculations.

Chapter Four followed directly on from the compression of the co-crystal to investigate the single component saccharin under compression as it had not been investigated in the literature. A new phase of saccharin occurs at 1.32 GPa and this is the first new polymorph of saccharin to be reported. The new polymorph possesses the same hydrogen bonding pattern as the ambient structure and we were able to compare the rate of compression of the saccharin dimer against the previous data. The rate was in good agreement with our hypothesis as saccharin's rate of compression was similar to that of other secondary amide dimers. The pressure at which saccharin undergoes a phase transition is a broad range due to a jump in pressure from 0.69 GPa to 1.32 GPa. For future work, it would be of interest to narrow the range of pressure at which the phase transition occurs. More data at lower pressures would also enable the understanding of whether the phase transition was driven by entropy or enthalpy. The phase transition fractures the crystal, however, it would be exciting to investigate whether the form would be recoverable to ambient pressure. If the phase transition is low enough after determining a smaller pressure range, it could potentially be reproduced in a large volume press to enable the production of larger volumes of the material for further analysis or experiments.

## Lauren Connor

The intention for chapter five was a structural comparison of a 2:1 and 1:1 co-crystal of benzoic acid and isonicotinamide to observe how pressure would affect the different stoichiometries of this material. In addition, the 1:1 co-crystal was the third in the study of the amide dimer under compression. The amide dimer present in the 1:1 co-crystal consisted of a pair of primary amides. On comparison of the rates of compression reported in previous studies, the rates and lengths of this dimer at pressure matched almost identically to those observed in other primary amide dimers. The 1:1 co-crystal displayed a wine-rack-effect upon compression, a large decrease down the longest length of the unit cell with minimal decrease in the other two directions. This phenomenon does appear to occur in the 2:1 co-crystal\*\* Further compression of the 1:1 co-crystal would be interesting to investigate as the wine rack effect has been known to cause negative linear compression. The last two data points of the study were taken at 5.06 GPa and the data in a further collection at 6.33 GPa were too poor. The cell axes begin to plateau at the higher pressures taken in the study, it would be interesting to elucidate the compression behaviour beyond 5.06 GPa. It is also apparent that it may undergo a reconstructive phase transition. The potential use of a pressure transmitting medium with a higher hydrostaticity limit would allow for collection at higher pressures. 4:1 methanol ethanol has a hydrostaticity limit of around 10 GPa. However, the co-crystals would be soluble in this medium. A saturated solution of the co-crystal in 4:1 methanol ethanol could be used and would make for an interesting study. Different pressure transmitting media may also lead to different compression behaviour. An extension of this suggestion could be proposed to all compression studies in this thesis in order to elucidate the effect of pressure transmitting medium on compression of the materials in this work.

The work in chapter six involved using a machine learning approach to co-crystallise artemisinin. Where previous computational approaches and statistical models had failed to lead to a successful co-former for artemisinin, machine learning proved a success. The co-crystallisation of artemisinin is difficult and as a pure compound, it exists as a van der Waals solid. The success of discovering a new co-crystal of artemisinin with 1-naphthol shows that the application of machine learning clearly has its uses for enriching the hit rate in a screen for difficult experiments where other mechanistic approaches may be unsuitable. This is a very exciting field and there is a lot of work still to be conducted in this area. Firstly, in terms of the work in chapter six, the aim was to co-crystallise artemisinin and this was successful. However, validation of the machine learning predictions is still to be carried out. An investigation

into the effects of diversifying the experimental methods should also be performed however this would require more negative data which is one of the biggest hurdles in machine learning approaches. The training data was an extended version of a bank of open data which was produced by another group. It would be advisable to continue collecting data and would be highly beneficial to set up a specific database. The purpose of the database would be to collect only co-crystallisation experiments to add to the training data for public use. However, it may be difficult to negate bias and keep the training data balanced between positive and negative results. The collection of data in this field is expected to naturally increase but should be honed to increase the likelihood of effective machine learning models.

The future of high pressure structural investigations could also be influenced by machine learning approaches. The combination of some of the elements contained in this thesis such as phase transitions, high pressure and machine learning approaches could lead to a fruitful investigation into the prediction of high pressure phase transitions. It may be worthwhile to consider a data based approach such as machine learning to predict whether a phase transition would occur in the ambient to 10 GPa pressure region. The intension of using a statistical based method would inevitably be to understand the scientific reasoning behind phase transitions at pressure. Often, crystallographers will use high pressure as a means to investigate a new area of chemical space in the search for new crystalline forms. At present, crystallographers do not understand the reason why certain molecules undergo a phase transition upon compression, whilst others do not. Therefore, it is not possible to employ a mechanistic, computational tool to predict the possibility of a phase transition for a given molecule. With a machine learning approach, a data driven method could offer predictions for molecules. Furthermore, it may be possible to understand which descriptors put into a training model are the most influential to the outcome. In this case, the molecular or crystalline descriptors used in the training model could be investigated to understand what aspects of the structure contribute to the instance of a phase transition. The study would not be without its hurdles. The main foreseen difficulty with this approach is a lack of consistent data. As aforementioned in previous chapters of this thesis, compression studies reported within the CSD are few and there is a great variation in experimental materials and methods used. The presently available data would be too distorted to build a machine learning training set likely to produce reliable results. For a clean bank of data, it would be wise to collect new compression studies with relatively consistent variables. In solving one issue (lack of

## Lauren Connor

data) another potential arises (time). Data collection is lengthy, however, synchrotron radiation would effectively reduce the collection time of the compression studies compared to standard laboratory collection. A fresh collection of data would allow the parameters of data collection and variables to be predefined. The pressure transmitting medium should be kept consistent and model compounds that are already recorded as polymorphic should be selected initially. Such a study may allow the relationship between the molecular structure and the likelihood of phase transition at pressure to be explored, advancing the field of high pressure crystallography.

In conclusion, this thesis has provided in depth structural studies on the compression of single and multicomponent constructs and in doing so has also produced a new polymorph for investigation. Furthermore, a new co-crystal of artemisinin has been structurally characterised and an extended set of positive and negative co-crystallisations will be published. The combined elements of this thesis contribute further ideas to explore the relationship between structure and high pressure phase transitions.

### Appendix

*All of the electronic appendices correspond to Chapter 6: Data-driven approach to direct the co-crystallisation of anti-malarial Artemisinin.*

*Each of the datasets are.txt files. Columns that are present in each of the dataset appendices include: cofomer 1 and cofomer 2, source (code to origin of datapoint) and SMILES codes for each of the cofomers.*

- **2018ConnorPhD\_Artemisinin\_training\_dataset.txt**

Random Forest training dataset. Output column contains a 'yes' or 'no' response. 'Yes' if cofomers in columns 1 and 2 form a co-crystal or 'No' if they do not (under the tested experimental conditions). Generated descriptors included.

- **2018ConnorPhD\_Artemisinin\_test\_dataset.txt**

Random Forest test dataset. Artemisinin tested against each of the cofomers present in the training dataset with more than 400 columns of generated descriptors.

- **2018ConnorPhD\_Artemisinin\_test\_dataset\_predictions.txt**

Predictions for Artemisinin with each of the cofomers in the training dataset. Numerical values in columns 7 and 8 combine to a value of 1. Any value above 0.5 in the 'yes' column is shown as likely to co-crystallise in column 6, the data.pred column. The ranking of the cofomers should be considered rather than the user defined cut-off value.

- **2018ConnorPhD\_SMILES.txt**

SMILES codes for each of the corresponding cofomers.

- **2018ConnorPhD\_RandomForestRscript.R**

Random Forest Script used for the datasets, showing user set parameters (for example the number of trees in the forest and the type (Classification)).