THE UNIVERSITY *of* EDINBURGH

It's Not All About U: The Role of Volume and Entropy in Weakly Bound Crystal Structures.

Cameron J G Wilson

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy to the School of Chemistry

The University of Edinburgh

2024

Declaration

I declare that this thesis has been composed by myself and that the work has not been submitted for any other degree or professional qualification. I confirm that the work submitted is my own, except where specifically noted or work which has formed part of co-authored publications has been included. My contribution and those of the other authors to this work have been explicitly indicated below. I confirm that appropriate credit has been given within this thesis where reference has been made to the work of others.

The work presented in Chapter 3 was previously published in reference 2 in the list of publications. This work was co-authored with Tomas Cervenka, Peter Wood and Simon Parsons. The conceptualisation of this work was by Simon Parsons. The CellVol code was written by myself and I completed all calculations with it. During a summer placement under my supervision, Tomas Cervenka completed much of the preliminary fitting of equations of state and calculations of bulk modulus presented in this chapter. I was responsible for writing the manuscript and edits were carried out by all authors.

The work presented in Chapter 4 was previously published in reference 5 in the list of publications. This work was co-authored with Peter Wood and Simon Parsons. I was responsible for the experimental work and data refinement with the support of Simon Parsons. I was responsible for writing the manuscript and edits were carried out by all authors.

The work presented in Chapter 5 was previously published in reference 6 in the list of publications. This work was co-authored with Jan Plesniar, Heike Kuhn, Jeff Armstrong, Peter Wood and Simon Parsons. The conceptualisation of this work was by Simon Parsons. As part of a BSc project under the supervision of Simon Parsons, Jan Plesniar completed preliminary

phonon calculations that were used to guide the design of phonon calculations completed in this work. As part of a MChem project under the supervision of Simon Parsons, Heike Kuhn completed force constant calculations that were used as the starting point for calculations presented in this chapter. Inelastic neutron scattering data were collected by Jeff Armstrong and Simon Parsons at ISIS neutron and muon source. All other calculations and analysis presented within this chapter were completed by myself. I was responsible for writing the manuscript and edits were carried out by all authors. In addition to the listed co-authors, the paper was additionally seen by Adam Michalchuk and Jonas Nyman who provided very useful advice on DFT methods and identified improvements that contributed towards its final form.

The work presented in Chapter 6 is currently unpublished. Measurement and processing of data collected in Edinburgh were completed by myself with the support of Simon Parsons. Preparation of the gas loaded cell used to recover phase II was completed by Bhaskar Tiwari who also assisted in data collection at DESY along with Simon Parsons and Nico Giordano. Data from ISIS neutron and muon source were collected by myself and Simon Parsons with the assistance of beamline scientists Nicholas Funnell and Christopher Ridley. Pawley and Rietveld refinements of neutron powder data were completed by Simon Parsons. Computational calculations for all structures and analysis of the component terms to Gibbs energy were completed by myself. We aim to publish this work in future and when published all contributors listed in this paragraph will feature as co-authors.

Peter Wood and Simon Parsons provided supervision throughout all projects.

Cameron Wilson 1st March 2024

List of Publications

1. Funnell, N. P., Allan, D. R., Maloney, A. G., Smith, R. I., Wilson, C. J. G., Parsons, S., Suppression of isotopic polymorphism. *CrystEngComm*. **2021**. 23(4), 769-776.
2. Wilson, C. J. G., Cervenka, T., Wood, P. A., Parsons, S., Behavior of Occupied and Void Space in Molecular Crystal Structures at High Pressure. *Cryst. Growth Des.* **2022**. 22(4), 2328-2341.
3. Broadhust, E. T., Wilson, C. J. G., Zissimou, G. A., Nudelman, F., Constantinides, C. P., Koutentis, P. K., Parsons, S., A first-order phase transition in Blatter's radical at high pressure. *Acta Crystallogr.* **2022**. B78(2) 107-116.
4. Broadhurst, E. T., Wilson, C. J. G., Zissimou, G. A., Padrón Gómez, M. A., Vasconcelos, D. M., Constantinides, C. P., Koutentis, P. A., Ayala, A. P. and Parsons, S., The Effect of High Pressure on Polymorphs of a Derivative of Blatter's Radical: Identification of the Structural Signatures of Subtle Phase Transitions. *Cryst. Growth Des.* **2023**. 23(3), 1915-1924.
5. Wilson, C. J. G., Wood, P. A. and Parsons, S., Discerning Subtle High-Pressure Phase Transitions in Glyphosate. *CrystEngComm*. **2023**. 25(6), 988-997.
6. Wilson, C. J. G., Plesniar, J., Kuhn, H., Armstrong, J., Wood, P. A. and Parsons, S., The Advantages of Flexibility: The Role of Entropy in Crystal Structures Containing C–H···F Interactions. *Cryst. Growth Des.* **2024**. DOI: 10.1021/acs.cgd.4c00042

This page has been left intentionally blank

Abstract

Crystals play a vital role in everyday life, from optical electronics and memory storage, to pharmaceuticals and flavourings. Many crystalline compounds can exist in a variety of interchangeable forms that can have remarkably different properties. To safely and reliably use crystalline compounds, it is vital to understand the relative stabilities of these forms and the conditions under which they can interconvert.

For a crystal form to exist it must represent a minimum in free energy, G, and relative phase stability can be rationalised using the Gibbs equation ($\Delta G = \Delta U + P\Delta V - T\Delta S$). Attention is often focused on calculations of relative internal energies (ΔU), for example through *ab initio* and force field methods, and descriptions of the enthalpy of individual bonds dominate descriptions of crystal packing. However, at high pressure the volume term ($P\Delta V$) becomes increasingly significant with the need to minimise volume driving almost all high-pressure phase transitions. The entropy term ($T\Delta S$) is often neglected from both high pressure and ambient pressure studies. Justified by a claim that its magnitude is too small to ever cause a re-ranking in polymorph stability, in the absence of strong structure directing interactions, this too can become significant.

Chapters 1 and 2 are the introductory chapters of this thesis. Chapter 1 starts by introducing the concepts of polymorphism and the classification of different types of phase transitions. The chapter then describes the origins of the terms of the Gibbs equation and their varying importance under different conditions. Chapter 2 describes the experimental and computational methods used in this thesis and outlines how different terms of the Gibbs equation were determined.

Chapter 3 aims to address the role of volume at pressure by introduction of the CellVol code, for calculation of occupied and void space in crystal structures. The code partitions crystal structures into occupied and unoccupied space based on van der Waals radii to define the limits of inter and intramolecular contacts around atom centres. High pressure crystallographic and spectroscopic studies often show misleading correlations but by studying the contribution of these volume elements individually we have uncovered a highly sensitive method for detecting subtle phase behaviour at high pressure. This method has been used to characterise phase transitions in polymorphs of *L*-histidine and to identify a crystallographic signature for a transition in naphthalene which has been debated for over 80 years. It is also shown to reveal premonitory behaviour before a reconstructive phase transition. The method was applied to the high-pressure entries in the Cambridge Structural Database (CSD) to reveal similarities in the way molecular compounds use void space to accommodate increased pressure, which explains their narrow range of bulk moduli.

Chapter 4 describes a high-pressure single crystal study on the herbicide glyphosate (*N*-(phosomethyl)glycine) to 5.17 GPa using a 4:1 mixture of methanol and ethanol as a pressure-transmitting medium. The CellVol code is used to show that two transitions previously suggested to be substantial first order structural reorganisations on the basis of high pressure Raman spectroscopy are both more subtle second order processes. The crystal structure of glyphosate consists of strongly hydrogen-bonded layers, and the first transition corresponds to the onset of a greater compressibility within these layers. The second is an intramolecular feature, corresponding to deformation of the molecular backbone. The previous Raman study and our own crystallographic data show that glyphosate undergoes a more substantial

transition above 5.6 GPa, which destroys the integrity of the single crystal. The nature of this transformation is explored further in Chapter 6.

Chapter 5 considers the role of entropy in stabilising weakly bound crystal structures through studies of the C—H···F intermolecular bond. Although C—H···F interactions are enthalpically weak, they occur much more commonly in crystal structures than would be predicted by chance. In the case of a series of five fluorobenzene derivatives, heat capacity data show that the entropic contribution to free energy at 150 K is of the order −10 to −15 kJ mol^{-1}, which is similar to the stabilisation afforded by N—H···N and O—H···O hydrogen bonds in ammonia and water, respectively. The origin of entropy in ordered, crystalline solids lies in access to low energy vibrational modes (phonons) which can be observed in inelastic neutron scattering spectra and simulated by periodic density functional theory (DFT) calculations. DFT calculations on the fluorobenzenes, which reproduce experimental entropies and inelastic neutron scattering data, demonstrate that the lowest energy phonons are dominated by motions of the F atoms. The low energies of these vibrations stem both from the mass of the F atom and the highly deformable nature of C—H···F interactions, revealed by relatively flat potential curves calculated using symmetry adapted perturbation theory. These data show that the stabilising influence of C—H···F interactions can be explained at least as much by the entropic contribution of their deformability as by the enthalpic contribution of the contact energies.

Chapter 6 describes formation of potentially four new high-pressure polymorphs of glyphosate. Although the integrity of a single crystal of glyphosate is lost on slow compression above 5.6 GPa, rapid compression to 5.9 GPa in a 4:1 mixture of methanol and ethanol and to 7.0 GPa in neon as pressure-transmitting media generates two new single-crystal phases. The

crystal structure of the first of these phases was determined from in-house single crystal diffraction data, that of the second from synchrotron data obtained on the P02.2 beamline at DESY. The phase at 7.0 GPa, named phase-II, has a volume 1.023 $Å^3$ per molecule higher than the phase obtained on rapid compression in methanol-ethanol at 5.9 GPa, named phase III. Formation of phase II from phase III with increasing pressure is therefore thermodynamically forbidden. The volume difference corresponds to an estimated $P\Delta V$-stabilisation of phase III relative to phase II of 17.3 kJ mol^{-1} at 7.0 GPa. In an attempt to recover the destructive transition seen previously and to potentially drive its transition to phase II or III, neutron powder diffraction data on glyphosate-d_4 was completed at ISIS neutron and muon source. This revealed a far more complex structural landscape and the formation of two further phases, named IV and V, which did not fit the structural parameters of phases I, II or III. So far, the structures of phases IV and V have not been characterised further and it is the aim of future work to attempt to recover structural parameters for these phases, possibly through crystal structure prediction. Free energy calculations involving internal energy, volume and entropy are used to establish the driving forces of the single crystal transitions. Extrapolation of the parameters of phase I to the pressures both phase II and III were observed reveals that both transitions are driven by volume as the only significant stabilising contribution.

Lay Summary

We encounter crystals in our daily life far more commonly than you may initially think. Crystals are everywhere, from the salt and sugar we use to flavour our food, to the pharmaceuticals many of us rely on daily.

Many compounds that form crystals can exist in a variety of interchangeable forms depending on the way their building blocks, molecules, are arranged together in space. Like a wall of bricks, the same building blocks can be put together in multiple different ways in what is referred to as polymorphism. This often has important impacts on the properties of the crystal. Strength, solubility, colour, shape and conductivity can all vary between polymorphs and this often affects their ability to perform their intended function. For example, paracetamol exists in multiple forms, some of which more easily form tablets than others. Ritonavir, a pharmaceutical used to treat HIV/Aids has one form which is effective and another that is not. Over time or as a result of the change in environmental conditions, such as temperature or pressure, crystals can covert between their forms. To safely and reliably use crystalline compounds, it is vital to understand the relative stabilities of these forms and the conditions under which they can interconvert.

Like the physical bonds that exist between the atoms in a molecule, the molecules in a crystal arrange themselves to form additional, longer range, favourable interactions. The relative stability of different crystal forms is often interpreted by the formation of high energy bonds such as these. There are however other factors which are often overlooked. The packing of molecules together in an efficient way to minimise volume is also an important factor. This is especially important at elevated pressures where the need to minimise volume becomes vital. Entropy, which is related to the statistical distribution

of molecules over different vibrational states, is also an important consideration and the ability to form deformable bonds is a favourable process.

This thesis addresses the role of entropy and volume in determining crystal stability and the driving forces of interconversion between forms. This work determines the limitations of considering the energies of bonding interactions between molecules. This study is completed with a particular focus on how the narrative changes under elevated pressures.

Acknowledgements

There are a great many people who have contributed towards the completion of this thesis that I would like to thank.

Firstly, my Supervisor Simon Parsons. Your knowledge and expertise on virtually every subject are boundless. Thank you for always finding the encouragement and enthusiasm needed when I get grumpy. Thousands of cups of tea, tens of thousands of biscuits and the same ten stories later, I do not think I would have gotten half as far without you as my supervisor.

To my industrial supervisor, Pete Wood. Thank you for your help and support over the years. Thank you for your contribution towards the development of both these projects and myself and giving me the opportunity to work on such interesting research topics.

To the Parsons group old and new; Giulia, Lewis, Alex, Bhaskar, Calum, Dominic, Gary and honorary member Ellie. Thank you for all your help, friendship and advice over the years. A special thanks to Edward Broadhurst for back in the dark days of lockdowns, face masks and distanced lunches in the office. Thank you for keeping me sane and showing me the way to the light at the end of the tunnel.

Other special mentions in the department go to Tom, Susanna and Andrew. I hope my endless shenanigans and visits to your office were not too distracting. Thank you for adopting me when my group went down to just one. Our coffee breaks and pub trips have come to be my favourite parts of the week. Joe, I hope my endless shenanigans and visits to your office WERE too distracting, you deserved it. I know you will go onto great things and I look forward to you telling me all about them… even when I don't ask. To Phoebe and Ailsa, thank you for being there start to finish. I will always cherish the early days of brunches and household bubbles.

Beth, you have been my rock. From the very start you have been there for me through everything. I simply would not have got to this point without you and I cannot wait for our next adventure together.

Grandma Barbara, Grandma Valerie, Grandad Don and Grandad George. I hope I made you proud. Ellie and Lottie, thank you for driving my competitive spirit that got me into this mess in the first place. Mum, your constant love and support, particularly during the tough times are endless. I will never be able to thank you enough for all you have done for me.

Finally, to my dad, who in many ways I am far too alike for either of our own good. At least all the propaganda fuelled trips to the science fair paid off for one of us. Your values have shaped so much of who I am today and what got me to this point.

Dad, I dedicate this thesis to you.

Abbreviations

1,2,4,5F	1,2,4,5-tetrafluorobenzene
1,2F	1,2-difluorobenzene
1,3,5F	1,3,5-trifluorobenzene
1,4F	1,4-difluorobenzene
1F	Monofluorobenzene
3D ED	3-dimensional electron diffraction
ADPs	Atomic displacement parameters
API	Application Programming Interfaces
BM2	Birch-Murnaghan 2nd order
BM3	Birch-Murnaghan 3rd order
BM4	Birch-Murnaghan 4th order
BSSE	Basis set superposition error
BZ	Brillouin Zone
CCSD(T)	Coupled-cluster perturbative triple excitation
CIF	Crystallographic information file
CSD	Cambridge structural database
CSP	Crystal structure prediction
DAC	Diamond anvil cell
DESY	Deutsches Elektronen-Synchrotron
DFPT	Density functional perturbation theory
DFT	Density functional theory
DFTB	Density functional tight binding
DoS	Density of states
DSC	Differential scanning calorimetry
EoS	Equation of state

EtOH	Ethanol
GUI	Graphical user interface
ILL	Institut Laue-Langevin
INS	Inelastic neutron scattering
LINAC	Linear accelerator
MeOH	Methanol
MOF	Metal organic framework
OTFG	On the fly generation
PBE	Perdew-Burke-Ernzerhof
PDoS	Partial density of states
PE	Paris-Edinburgh
PET	poly(ethylene terephthalate)
PV	Pressure-volume
PVT	Pressure-volume-temperature
PXRD	Powder X-ray diffraction
SAPT	Symmetry-adapted perturbation theory
SCDS	Semi-classical density sums
SCXRD	Single crystal X-ray diffraction
TOF	Time of flight
TV	Temperature-volume
V2	Vinet 2nd order
V3	Vinet 3rd order
VDP	Voronoi-Dirichlet Polyhedra
WC	Tungsten carbide
ZTA	Zirconia-toughened alumina

Table of Contents

CHAPTER 1: PHASE STABILITY IN CRYSTAL STRUCTURES 1

1.1 POLYMORPHISM AND PHASE TRANSITIONS 2

1.2 THE GIBBS EQUATION 9

1.3 THE DETERMINATION OF THE ENERGETIC CONTRIBUTIONS OF GIBBS FREE ENERGY 11

1.3.1 Intermolecular Interaction Energies, ΔU *11*

1.3.1 The Volume Term, $P\Delta V$ *20*

1.3.3 The Entropy Term, $T\Delta S$ *31*

1.4 THESIS AIM 35

1.5 REFERENCES 36

CHAPTER 2: GENERAL EXPERIMENTAL AND COMPUTATIONAL PROCEDURES. . 47

2.1 DIFFRACTION 48

2.2 DETERMINATION OF INTERMOLECULAR INTERACTION ENERGIES, ΔU 58

2.2.1 SAPT Calculations *58*

2.2.2 PIXEL Calculations *60*

2.3 DETERMINATION OF THE VOLUME CONTRIBUTION, $P\Delta V$ 66

2.3.1 High Pressure X-ray Crystallography *66*

2.3.2 High Pressure Neutron Powder Diffraction *70*

2.4 DETERMINATION OF THE ENTROPY CONTRIBUTION, $T\Delta S$ 72

2.4.1 Phonon Calculations *72*

2.5 REFERENCES 75

CHAPTER 3: BEHAVIOUR OF OCCUPIED AND VOID SPACE IN MOLECULAR CRYSTAL STRUCTURES AT HIGH PRESSURE. 83

OVERVIEW 84

3.1 INTRODUCTION 85

3.2 COMPUTATIONAL PROCEDURES 89

3.2.1 Network and Void Volume Calculations *89*

3.2.2 Calculation of Network and Void Bulk Moduli for Compression Studies in the CSD *93*

3.2.3 Periodic DFT Calculations *93*

3.2.4 Pixel Calculations *94*

3.3 RESULTS....95
3.3.1 Detection of First Order Transitions in L-histidine....95
3.3.2 Detection of Subtle High-Pressure Second Order Transitions....98
3.3.3 Premonitory Behaviour....107
3.4 GENERAL TRENDS IN PACKING COEFFICIENTS AND BULK MODULI....111
3.4.1 Variation of Packing Coefficients....111
3.4.2 Contributions of Networks and Voids to Overall Compressibility....112
3.4.3 The Range of the Bulk Moduli of Molecular Solids....117
3.5 CONCLUSIONS....119
3.6 REFERENCES....121

CHAPTER 4: DISCERNING SUBTLE HIGH-PRESSURE PHASE TRANSITIONS IN GLYPHOSATE....129

OVERVIEW....130
4.1 INTRODUCTION....131
4.2 MATERIALS AND METHODS....133
4.2.1 Single Crystal X-ray Diffraction....133
4.2.2 Periodic DFT Calculations....134
4.2.3 Occupied Volume ('CellVol') Calculations....134
4.2.4 Pixel Calculations....135
4.3 RESULTS AND DISCUSSION....136
4.3.1 Ambient Pressure Structure....136
4.3.2 Response of the Unit Cell Parameters to Pressure....140
4.3.3 Intermolecular Interactions....142
4.3.4 Volume Analysis....144
4.4 CONCLUSIONS....149
4.5 REFERENCES....150

CHAPTER 5: THE ADVANTAGES OF FLEXIBILITY: THE ROLE OF ENTROPY IN CRYSTAL STRUCTURES CONTAINING C—H···F INTERACTIONS. 157

OVERVIEW....158
5.1 INTRODUCTION....159
5.2 MATERIALS AND METHODS....162
5.2.1 Structures Studied....162

5.2.2 DFT Geometry Optimisation ... *162*
5.2.3 Phonon Calculations ... *163*
5.2.4 INS Measurements ... *164*
5.2.5 Lattice and Intermolecular Interaction Energies. ... *165*
5.2.6 Entropy Calculations ... *165*
5.2.7 Force Constants ... *166*
5.3 Results and Discussion ... 168
5.3.1 Intermolecular Interactions ... *168*
5.3.2 Phonon Calculations ... *171*
5.4 Conclusions ... 180
5.5 References ... 182

CHAPTER 6: DOMINANT DRIVING FORCES AT PRESSURE: A STUDY OF NEW POLYMORPHS OF GLYPHOSATE UNDER VARYING CONDITIONS. ... 191

Overview ... 192
6.1 Introduction ... 193
6.2 Materials and Methods ... 195
6.2.1 Single Crystal X-ray Diffraction ... *195*
6.2.2 Structure Solution and Refinement of X-ray Data ... *196*
6.2.3 Neutron Powder Diffraction ... *197*
6.2.4 Analysis of Neutron Data ... *199*
6.2.5 DFT Geometry Optimisation ... *200*
6.2.6 Phonon Calculations ... *200*
6.2.7 Pixel Calculations ... *201*
6.2.8 Volume Calculations ... *202*
6.3 Results and Discussions ... 203
6.3.1 Single Crystal X-ray Diffraction ... *203*
6.3.2 Neutron Powder Diffraction ... *216*
6.3.3 Thermodynamic Driving Forces of Transitions ... *226*
6.4 Conclusions ... 230
6.5 References ... 232

CHAPTER 7: CONCLUDING REMARKS ... 239

6.1 Concluding Remarks ... 240

APPENDICES **247**

APPENDIX A- CHAPTER 3 248

A.1 Optimising the Density of Points for CellVol Calculations *248*

A.2 Equation of State Fitting *250*

A.3 Comparison of Network and Void Bulk Moduli *252*

APPENDIX B- CHAPTER 4 254

B.1 Occupied Volume Calculations *254*

B.2 Crystallography Tables *257*

B.3 Strain Tensor Results *260*

APPENDIX C- CHAPTER 5 263

C.1 Cell Optimisation *263*

C.2 Comparison of Aromatic H···F Interactions *263*

C.3 Convergence of Phonon Calculations *264*

C.4 SAPT Calculation Convergence *267*

C.5 Diagrams and Energies of Dimers for All Structures *269*

C.6 Inelastic Neutron Scattering Plots *275*

C.7 Density of States Plots *278*

C.8 Partial Density of States Plots *281*

C.9 Example Calculation of the Contributions to Internal Energy and Entropy by Occupation of Energy Levels at 150 K in 1,2F *284*

C.10 Justification for the Neglect of U and Zero-Point Energy *289*

APPENDIX D- CHAPTER 6 291

D.1 Crystallographic Tables *291*

D.2 Fitting of Phase I to 5.93 and 7.03 GPa *292*

D.3 Morphology of Glyphosate Crystals *293*

D.4 Pixel Tables for Phase I and II *293*

Chapter 1: Phase Stability in Crystal Structures

1.1 Polymorphism and Phase Transitions

Molecular compounds crystallise in the solid state by packing molecules together to minimise free energy. The separation distances of the molecules are the result of a competition between the pushes and pulls of multiple interactions.[1] Packing can occur in several different ways for many compounds and this leads to the formation of multiple phases. Phases that exist under the same conditions are referred to as polymorphs, and interconversions between polymorphs, or more generally phases, is referred to as a phase transition. The compound with the largest number of polymorphs is 5-methyl-2-[(2-nitrophenylamino)]-3-thiophenecarbonitrile, nicknamed ROY due to the red, orange, and yellow colour of its different phases. ROY has 13 different known polymorphs in the Cambridge Structural Database (CSD).[2,3] Polymorphism, is very common for molecular compounds and has been estimated to apply to at least 50% of organic solids, rising to 74% for those extensively studied.[4,5] The number of polymorphs known for a compound is sometimes said simply to reflect the extent to which it has been studied.[6]

Polymorphs can have very different properties as a result of the way their constituents pack. A famous example is the difference in appearance, strength and electrical conductivity of the carbon allotropes diamond and graphite. Polymorphism is therefore of high relevance to the pharmaceutical industry where the therapeutic elements of a large proportion of drugs involve small organic molecules highly prone to polymorphism.[7] In these cases, polymorphism can affect the tabletability of a drug or its solubility, often directly affecting its bioavailability, the therapeutic dose due to the level of uptake.[8] An example of the importance of molecular polymorphism is Ritonavir, a pharmaceutical used in the treatment of HIV/AIDs. In 1998, a

second polymorph of Ritonavir was discovered, a year after it had been marketed, with a lower solubility and hence lower bioavailability. This new polymorph was more thermodynamically stable and was found to seed conversion of the original polymorph. Withdrawal of the drug from the market and design of a new suspension-based delivery cost an estimated $250 million to the manufacturer.[9]

Attempts to discover of all potential polymorphic forms of drug compounds, a process called polymorphic screening, is of vital importance. To ensure all accessible polymorphs of a molecular compound are known, exhaustive assays are undertaken under varying conditions. New phases can be formed using several different stimuli, some of the most common are time, temperature, seeding or, most relevantly for this thesis, pressure.[10]

According to the Ehrenfest classification, transitions between polymorphs can be broadly categorised as either first order or second order events, with the order defined by the derivative of free energy in which a discontinuity occurs at the transition.[11] First order transitions are often accompanied by significant structural rearrangements. The first derivative of free energy with respect to pressure at constant temperature is volume

$$\left(\partial G/\partial P\right)_T = V \tag{1.1}$$

where G is free energy, P is pressure, T is temperature and V is volume. For transitions driven by pressure at constant temperature a first order transition is often characterised by a discontinuous change in volume (Figure 1.1), allowing them to be detected crystallographically. The second derivative of free energy with respect to pressure at constant temperature is compressibility, β, the inverse of the bulk modulus, K

$$-\frac{1}{V}\left(\partial^2 G/\partial P^2\right)_T = -\frac{1}{V}\left(\partial V/\partial P\right)_T = \beta \tag{1.2}$$

$$\frac{1}{\beta} = -V\left(\partial P/\partial V\right)_T = K \tag{1.3}$$

Second order transitions driven by pressure feature a discontinuous change in the gradient of a pressure-volume curve (Figure 1.1) and hence also the bulk modulus.[12] Further details on the bulk modulus of compounds and its determination using equations of state are given in Section 1.3.2*b*.

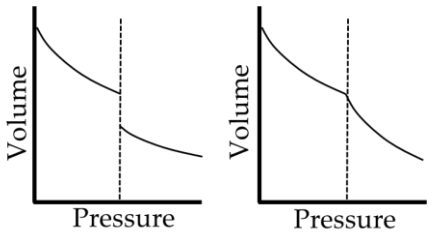


Figure 1.1: The Ehrenfest classification of transitions with pressure. Left, a 1st order transition is signalled by a discontinuous change in volume. Right, a second order transition features a discontinuous change in bulk modulus, seen here as a change in gradient. Redrawn based on diagrams in *Physical Chemistry* (Atkins, 10th Ed., 2014).[13]

The equations for the first and second derivatives of free energy with respect to temperature are analogous to those driven by pressure. First order transitions with respect to temperature are similarly accompanied by a discontinuous change in volume, also enabling them to be detected in crystallographic data. Since the first derivative of free energy with respect to temperature at constant pressure is entropy

$$\left(\partial G/\partial T\right)_p = -S \tag{1.4}$$

and at a transition

$$\Delta H = T\Delta S \tag{1.5}$$

where H is enthalpy. First order transitions with temperature are also often detected by enthalpic peaks in differential scanning colorimetry (DSC) traces. This can be seen in Figure 1.2 for the first order transition between the γ and α polymorphs of glycine.[14]

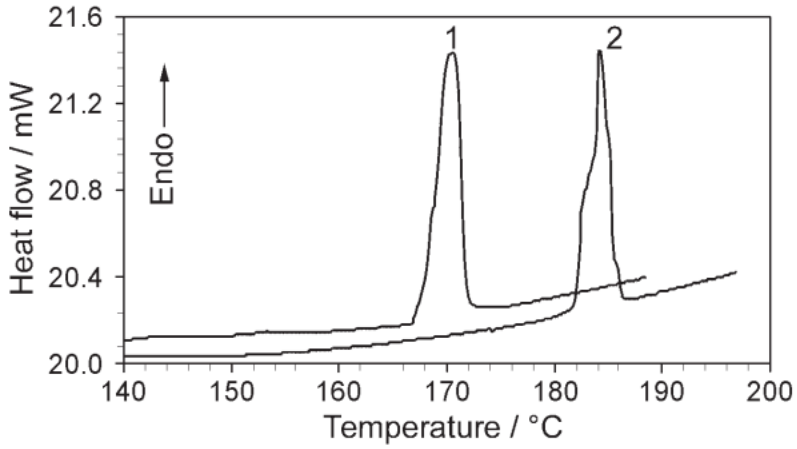


Figure 1.2: The DSC trace for the first order γ-α phase transition in glycine. Peaks 1 and 2 indicate different levels of pre-treatment of the samples. Reproduced from Ref 14 with permission from Springer Nature.

Second order transitions driven by temperature produce a change in the gradient of a temperature-volume curve, *i.e.* the thermal expansion coefficient $\alpha = (\partial V/\partial T)_P$, or the heat capacity

$$-T\left(\partial^2 G/\partial T^2\right)_P = -T\left(-\partial S/\partial T\right)_P = \left(\partial H/\partial T\right)_P = C_P \qquad (1.6)$$

leading to a change in baseline of a DSC trace, Figure 1.3. For example, in the case of a glass transition in a polymer, the heat capacity changes as new vibrational modes become accessible on increasing T, causing the polymer to transform from a rigid to a flexible state.

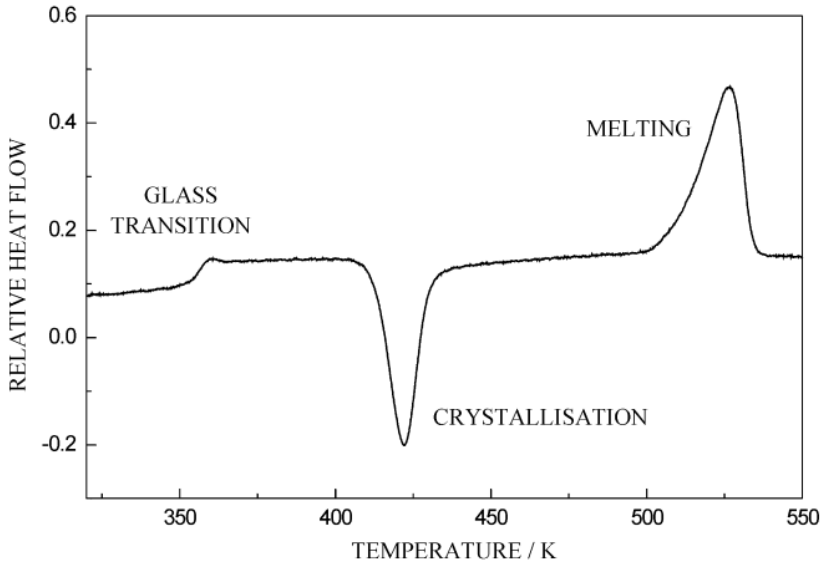


Figure 1.3: A DSC trace exemplifying a second order glass transition in amorphous poly(ethylene terephthalate) (PET). An exothermic crystallisation peak and endothermic melting peak are also seen in the trace. Reprinted from Thermochimica Acta, 367, Bailey, N. A., Hay, J. N., & Price, D. M., A study of enthalpic relaxation of poly(ethylene terephthalate) by conventional and modulated temperature DSC, 425-431, Copyright (2024), with permission from Elsevier.

Transitions can also be described as either displacive or reconstructive, depending on whether there are symmetry relationships between the two phases. Displacive transitions occur due to a loss of one or more symmetry elements, for example the transition between the high-pressure phases III ($P4_12_12$) and II ($P2_12_12_1$) of pyridine is the result of loss of 2-fold rotational symmetry along the *ab* face diagonals of phase III.[15] A transition without a formal symmetry relationship between its phases is a reconstructive transition and occurs via a nucleation and growth mechanism, for example the α ($P2_1/n$) and γ ($P3_1$) phases of glycine described above.[16] The reorganisations of crystal packing that occur in reconstructive transitions can result in loss of the integrity of a crystal.

A final classification for transitions and one limited in practice to those driven by temperature is as either enantiotropic or monotropic. The

classification of phase transitions into these categories depends on whether the transition point, the intersection of the extrapolated Gibbs energy surfaces, occurs before or after intersection with the liquid phase boundary, the melting point.[17] For enantiotropic systems, the temperature at which the free energy of the two phases is equal, the transition temperature, occurs below the melting point. One phase is stable below the transition temperature and the other phase is stable above the transition temperature up to the melting point. For a monotropic system, one phase remains the most stable phase for all temperatures below the melting point. Starting from the most stable phase under ambient conditions, it is impossible to obtain any other metastable phase by heating a monotropic system.[17] Both enantiotropic and monotropic phases relationships exist for the compound ROY (Figure 1.4). The phases ON (orange needles) and Y (yellow prisms) provide an example of an enantiotropic phase relationship, intersecting each other at the transition temperature, labelled T_t, before the melting points of the pure phases are reached, $T_{m\mathrm{ON}}$ and $T_{m\mathrm{Y}}$. The phases R (red prisms) and ON, on the other hand, provide an example of a monotropic relationship, any potential intersection of the Gibbs energy surfaces occurs after the melting points of the pure phases are reached, $T_{m\mathrm{R}}$ and $T_{m\mathrm{ON}}$.

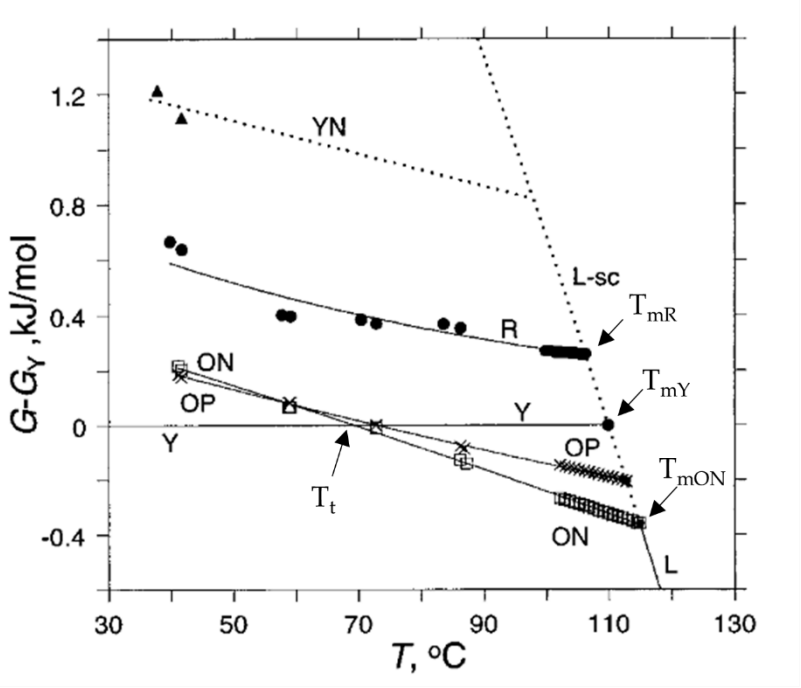


Figure 1.4: The stability relationships between phases of ROY with increasing temperature constructed from melting and eutectic melting data. T_{mR}, T_{mY} and T_{mON} indicate the melting points of the pure phases, T_t is the transition temperature. Adapted with permission from Yu *et al.* *J. Am. Chem. Soc.* 2000, 122, 4, 585–591. Copyright 2024 American Chemical Society.[18]

1.2 The Gibbs Equation

The second law of thermodynamics states '*The entropy of an isolated system increases in the course of a spontaneous change*'.[13] That is

$$\Delta S_{\text{tot}} > 0 \tag{1.7}$$

where S_{tot} is the total entropy of the isolated system of interest.[13] If there is a transfer of energy as heat between a system and its surroundings at thermal equilibrium at temperature, T, the Clausius inequality holds

$$dS - \frac{dq}{T} \geq 0 \tag{1.8}$$

where dS is the change in entropy of the system and dq is the change in heat (work) of the system. When the heat transfer occurs at constant pressure, P, all work is expansion work and

$$dq_p = dH \tag{1.9}$$

where dH is the change in enthalpy of the system. Hence from Equation 1.8

$$dS - \frac{dH}{T} \geq 0 \tag{1.10}$$

$$\therefore TdS \geq dH \tag{1.11}$$

Rearranging Equation 1.11 produces

$$dH - TdS \leq 0 \tag{1.12}$$

Which can be expressed alternatively by defining the Gibbs free energy, G, of the system

$$G = H - TS \tag{1.13}$$

An alternative treatment of heat transfer at constant volume provides the derivation of the Helmholtz function. For infinitesimal changes at constant temperature the Gibbs equation becomes the more familiar

$$\Delta G = \Delta H - T\Delta S \tag{1.14}$$

The inequality in Equation 1.12 defines the criterion for a spontaneous process as one that must reduce the Gibbs free energy of the system, ΔG must be lower

or equal to 0. The enthalpy of the system can be further decomposed into internal energy, U, pressure, P, and volume, V, contributions

$$\Delta G = \Delta U + P\Delta V - T\Delta S \quad (1.15)$$

As previously described, a stable phase comprises a structure with a packing arrangement that corresponds to a Gibbs energy minimum. Its formation is driven by the competing thermodynamic contributions of Equation 1.15 towards a geometry that provides a favourable compromise between the different terms. When considering the Gibbs energy difference, ΔG, between two polymorphs under ambient conditions, the ΔU term tends to dominate. Nyman and Day have shown through a study of 508 polymorphic molecular organic compounds that although the differences in polymorph lattice energies are typically modest, with over half under 2 kJ mol^{-1}, consideration of lattice enthalpies solely is sufficient to predict the ranking of polymorphic stability in 91% of compounds.[19]

At ambient pressure the volume contribution towards enthalpies is very small but at elevated pressures its contribution becomes increasingly significant. Virtually all phase transitions at high pressure are driven by a minimisation of volume.[20]

Entropies tend to be similar for polymorphs, and the entropy change of a solid-state phase transition is often small. However, the consideration of vibrational entropy in the remaining 9% of compounds from the Nyman and Day study led to a re-ranking of polymorph stability.[19] The ability of entropy to stabilise crystal structures is often overlooked and neglect of this term is potentially an oversimplification in some materials.

1.3 The Determination of the Energetic Contributions of Gibbs Free Energy

The relative contributions of the Gibbs equation are of fundamental interest for understanding the factors that drive phase transitions. The origins of the various terms and some common ways in which they are determined both experimentally and computationally are discussed in the remaining sections of this Chapter.

1.3.1 Intermolecular Interaction Energies, ΔU

Enthalpy, H, is the total heat content of a system. It has contributions from both the internal energy, U, determined by the formation of favourable intramolecular geometry and intermolecular contacts, and additional pressure, P, and volume, V, contributions. At constant volume, H reduces to U. Strongly stabilising interactions with numerically large but negative energies, such as hydrogen bonds, are often described as being structure directing, supramolecular synthons that can be used as motifs to design crystal structures with pre-determined architectures.[21] The relief of energy in strained contacts has also been rationalised as the driving force behind phase transitions.[22]

Intermolecular interaction energies can stabilise different conformational polymorphs displaying diverse molecular geometries.[23] The polymorph transition already noted for Ritonavir is an example of a transition between two conformational polymorphs.[24] In rare cases, intramolecular energy differences between polymorphs can be as high as 15-20 kJ mol^{-1}.[19] The polymorph containing the higher energy molecular conformation is stabilised

by intermolecular interactions, often as a result of an increase in accessible surface area of the conformer.[19,25]

The experimental measurement of individual internal energies can be complex. One method, that has been investigated by Cockroft and co-workers, is through the use of rotameric folding molecules, topoisomers, atropoisomers and tautomers as molecular balances. This allows the quantification of non-covalent interactions by probing the relative stability and position of equilibrium afforded to the different conformational states as a result of interactions present in one conformation but absent in others.[26]

Differential scanning calorimetry is commonly used for the measurement of enthalpy changes driven by temperature (see above). A differential scanning calorimeter measures the energy transferred as heat to or from a sample undergoing a physical or chemical change at constant pressure.[13] This is done by measuring heat transfer to the sample relative to a standard which does not undergo a change. When no change occurs, the heat transferred, q_P, is given by

$$q_P = C_P \Delta T \tag{1.16}$$

where it is assumed that the heat capacity C_P is also independent of temperature. When a chemical process or first order phase transition occurs, an additional transfer of heat is required, $q_{P,\mathrm{ex}}$, such that the total transfer of heat required to maintain the same change in temperature of the sample as the standard is given by $q_P + q_{P,\mathrm{ex}}$. This additional term is interpreted as a change in the heat capacity of the sample at constant pressure and can be written as

$$C_{P,\mathrm{ex}} = \frac{q_{P,\mathrm{ex}}}{\Delta T} \tag{1.17}$$

The enthalpy change associated with the process is

$$\Delta H = \int_{T_1}^{T_2} C_{P,\mathrm{ex}} dT \tag{1.18}$$

This is the area under the peak in a plot of $C_{P,\mathrm{ex}}$ against T in the limits T_1 to T_2 within which the change occurs. Use of DSC to study phase transitions is exemplified by Perlovich *et al.* for the transformation of γ to α glycine, Figure 1.2.[14]

The advent of modern computing and improvements in accessibility and cost has led to internal energies increasingly being calculated computationally. Intermolecular interaction energies between molecules can be portioned into individual energy components: electrostatic, polarisation, dispersion and repulsion.

$$E_{\mathrm{Tot}} = E_{\mathrm{Elec}} + E_{\mathrm{Pol}} + E_{\mathrm{Disp}} + E_{\mathrm{Rep}} \tag{1.19}$$

The principal binding energy in solids is electrostatic.[27] In principle all stabilising interactions, whether involving permanent, induced or instantaneous dipoles are electrostatic interactions. Within this work we adopt the nomenclature of describing electrostatic interactions as the attractive and repulsive interactions between permanent nuclear and electronic charges only. This electrostatic energy contribution is described by Coulomb's law

$$E_{\mathrm{Elec}}(r_{ij}) = \frac{Q_i Q_j}{4\pi\varepsilon_0 r_{ij}} \tag{1.20}$$

where Q_i and Q_j are charges separated by a distance r_{ij} and ε_0 is the vacuum permittivity (8.854×10^{-12} F m^{-1}). Modelling electrostatic energies by assigning point charges to atoms in a crystal structure is simple but it does not capture the shape of a charge distribution. More computationally expensive methods seek to model charge anisotropy, an important factor in reproducing the high directionality of hydrogen bonds.[28]

Polarisation forces, also called induction forces, describes the polarisation induced when a charge approaches a sphere of electron density. The result is an induced dipole that forms an attractive interaction with the permanent dipole. The contribution is given by

$$E_{\mathrm{Pol}} = -\frac{D_{ij}}{r_{ij}^6} \tag{1.21}$$

where

$$D_{ij} = \frac{\mu_i^2 \alpha_j}{4\pi\varepsilon_0} \tag{1.22}$$

and μ_i is the magnitude of the permanent dipole moment of molecule i and α_j is the polarizability volume of molecule j.

In materials where contributions from permanent dipoles are less pronounced, for example symmetric molecules such as CH_4 or N_2 or layered compounds such as graphite, the primary interaction is dispersion.[27] These interactions are commonly referred to as van der Waals or London dispersion forces and occur when fluctuations in electron density instantaneously forms a dipole, which in turn induces electron density deformation and hence a dipole in an adjacent molecule. Dispersion is best understood as a whole molecule interaction and is always attractive. The form of the interaction is very similar to that of polarisation energy with a similar $-1/r_{ij}^6$ dependence

$$E_{\mathrm{Disp}} = -\frac{C_{ij}}{r_{ij}^6} \tag{1.23}$$

where

$$C_{ij} = \frac{3}{2}\alpha_i \alpha_j \frac{I_i I_j}{(I_i + I_j)} \tag{1.24}$$

and I_i and I_j are the ionisation energies of the interacting molecules, and α_i and α_j are the polarizability volumes.

At shorter molecule separations, quantum mechanical repulsion effects become dominant. The repulsion is a consequence of the Pauli exclusion principle that forbids electrons of the same quantum number occupying an identical quantum state. In practice this is often modelled by a resistance to increasing overlap of electron clouds and the resulting distortion of electron density.[29] The electrons in the overlapping region experience a mutual

electrostatic repulsion and a splitting of energy levels which raises the overall energy of the system.[27] The repulsive energy is commonly parameterised using the Born-Mayer form with empirical constants $B_{ij}$and ρ_{ij}

$$E_{\mathrm{Rep}}(r_{ij}) = B_{ij} exp\left({}^{-r_{ij}}/_{\rho_{ij}}\right) \quad (1.25)$$

For a simple models of ionic crystals, electrostatic energies and Born-Mayer repulsive energies are sometimes considered sufficient.[27] When modelling molecular compounds, the repulsive energy term is almost always coupled with the dispersion term. Combining the dispersion interaction with the Born-Mayer repulsive equation above yields the Buckingham potential

$$E_{ij}(r_{ij}) = -\frac{C_{ij}}{r_{ij}^6} + B_{ij} exp\left({}^{-r_{ij}}/_{\rho_{ij}}\right) \quad (1.26)$$

Interactions in molecular crystal structures have contributions from all energetic terms. The relative contributions of each term can be used to characterise the type of interaction. For example, the π-stacking interaction in benzene and the C—H···N hydrogen bond in pyridine have similar interaction energies, −13.6 kJ mol^{-1} and −11.1 kJ mol^{-1}, but their contributions towards the total energy are very different.[15,30] The π-stacked interaction is dominated by a large stabilising dispersion contribution with much smaller stabilising electrostatic and polarisation contributions and a large repulsive contribution (Table 1.1). Dispersion and repulsion contributions are smaller in the C—H···N hydrogen bond, the polarisation contribution is similar while the electrostatic contribution is significantly more stabilising (Table 1.1). The large electrostatic contribution is due to the presence of permanent charges in the hydrogen bond.

Table 1.1: Comparison of the molecule-molecule energy contributions in kJ mol^{-1} for a π-stacking interaction in benzene and a C—H···N hydrogen bond in pyridine.[15,30] Component energies in both structures were calculated using the Pixel method at the MP2/6-31G** level of theory, see Chapter 2.

	E_{Elec}	E_{Pol}	E_{Disp}	E_{Rep}	E_{Tot}
Benzene, π-π	−0.8	−4.0	−32.3	23.5	−13.6
Pyridine, C—H···N	−7.9	−3.5	−12.2	12.5	−11.1

There are a number of ways the computation of internal energies can be approached. The standard used in most computational procedures is the supramolecular approach. This aims to calculate interaction energies using the following formalism

$$\Delta E_{\mathrm{int}} = E_{AB} - E_A - E_B \tag{1.27}$$

That is, the interaction energy ΔE_{int} is calculated as the difference between the dimer energy E_{AB} and the monomer energies E_A and E_B. The supramolecular definition is applicable to any quantum chemistry approach. It can be applied with any method and any basis set, increasing systematically the level of theory as a balance between computational expense and accuracy. Since the approach is fully quantum mechanical, it captures electron transfer in the calculated energies, which consideration of the individual terms described above does not.

The use of finite basis sets in supramolecular approach makes it vulnerable to basis set superposition error (BSSE). When molecules in the dimer approach each other, their basis sets will begin to overlap and each molecule will begin to 'borrow' basis set functions from the other molecule. This will effectively confer an improvement in the basis sets of the molecules in the dimer calculation over those used for the calculation of the isolated molecules. The result is an effectively varying basis set for the dimer

calculation with a dependence on the separation of the molecules, r_{AB}. This is the BSSE and can be described by

$$\Delta E_{\text{int}} = {E_{AB}}^{AB}(r_{AB}) - {E_A}^{A} - {E_B}^{B} \tag{1.28}$$

where the superscript denotes the basis sets used in the calculation of the terms on the right-hand side. To avoid the need for unphysical infinite basis sets, a counterpoise correction is commonly used.[31] This method works by improving the basis sets of the isolated molecules by introduction of empty basis sets. Referred to as ghost orbitals, these are introduced as basis functions of the other molecule but with no electrons in them. These ghost orbitals carry the same distance dependence as the dimer term. Overall, the interaction energy calculated via counterpoise corrected supramolecular approach is then described by

$${\Delta E_{\text{int}}}^{CP} = {E_{AB}}^{AB}(r_{AB}) - {E_A}^{AB}(r_{AB}) - {E_B}^{AB}(r_{AB}) \tag{1.29}$$

Counterpoise-corrected supramolecular calculations can provide very accurate calculations of overall energies, but they do not yield a break-down into the individual energy terms described above, which provide useful physical insight into the character of an intermolecular interaction. A selection of methods that do provide this breakdown and that are used commonly for the study of molecular crystals are outlined below.

Symmetry adapted perturbation theory (SAPT) calculations attempt to solve the total dimer Hamiltonian for an isolated dimer by first solving a Hartree-Fock description of the isolated molecules and then adding additional terms to account for intra and inter molecule correlation. The calculation is based on a double perturbation series of inter and intra molecule interaction terms from which individual electrostatic, polarisation, dispersion and repulsion energy terms can be directly extracted. SAPT is mathematically complex, but the level of perturbation theory applied can be systematically

improved from the simplest SAPT0 approach to the SAPT2+3 level. The basis sets used for the calculations can also be systematically improved and this combination allows the quality of SAPT calculations to be incrementally improved until convergence is reached. SAPT calculations are available in a number of quantum mechanical codes, and that implemented in the program Psi4 were used here.[32] Unlike the other methods presented below, SAPT calculations are applicable to the isolated dimer only and are not calculated in the context of the wider crystal structure. This method, along with appropriate choices of augmentation level and basis set will be discussed in further detail in Section 2.2.1.

The Pixel method introduced by Gavezzotti is a semi-classical density sum (SCDS) method, based on numerical integrations over molecular electron densities.[29,33,34] It is of relatively low computational expense when compared to the quantum mechanical approaches described above, but is applicable only to interactions where charge transfer is absent. Electron densities of the isolated molecules are first calculated (e.g. using Gaussian[35]) and then divided into pixels of defined size and uniform electron density. Interaction terms between individual pixels are calculated one by one to produce electrostatic, polarisation, dispersion and repulsion components. The equations used for calculation of these individual energy terms and further details on the inputs for the Pixel process are described in Section 2.2.2. Estimates of lattice energies are calculated in the Pixel process by summation of the total dimer energies within a defined radius in a crystal structure. Pixel calculations can be conveniently run through the Mercury interface by use of the MrPixel GUI.[36,37]

An alternative method to calculate intermolecular energies is available in the program CrystalExplorer.[38] CrystalExplorer computes pairwise intermolecular interaction energies in a method that was inspired by applying

a quantum mechanical approach to the Pixel process.[38] The method calculates the same individual physically meaningful terms with relatively low-level quantum mechanical calculations but with the addition of a scale factor calibrated using higher-level calculations, Equation 1.30. The scale factors are optimised against reference counterpoise-corrected supramolecular calculations for 1794 molecule/ion pairs.[38] Calculations of energies with CrystalExplorer require a complete crystal structure and unlike Pixel and SAPT, components of the structure cannot be calculated individually. Like Pixel, estimates of lattice energies can be calculated for a crystal by summing over individual pairwise contributions.

$$E_{\mathrm{Tot}} = k_{\mathrm{Elec}}E_{\mathrm{Elec}} + k_{\mathrm{Pol}}E_{\mathrm{Pol}} + k_{\mathrm{Disp}}E_{\mathrm{Disp}} + k_{\mathrm{Rep}}E_{\mathrm{Rep}} \qquad (1.30)$$

Finally, interaction terms can also be calculated using force field methods. These methods are widely used within the field of crystal structure prediction (CSP). Force field calculations are faster than quantum mechanical calculations and while simple force fields based on combining the Buckingham potential, Equation 1.26, and the electrostatic contribution, Equation 1.20, are common, they can become much more elaborate.[39] To achieve a level of accuracy that is competitive with the methods described above, this method requires the design of bespoke parameterisation models which can require substantial input and time.[40]

In this thesis both Pixel and SAPT methods were used exclusively. The ability to extract individual energy terms from both Pixel and SAPT calculations is very valuable for studies at pressure, which contribute a significant proportion of this thesis. Contributions towards individual contacts with increasing pressure can be studied to indicate when contacts start to become destabilising and can be used to rationalise high-pressure phase transitions. Higher levels of SAPT tend to be more computationally

expensive than Pixel calculations but produce more accurate results. This is demonstrated for a simple dimer of water, bonded through a single O—H···O hydrogen bond. The cohesive energy calculated via high level coupled cluster supramolecular calculations in the S22 database is −21.004 kJ mol^{-1}.[41] The Pixel method at the MP2/6-31G** level calculates the cohesive energy of this dimer as −24.237 kJ mol^{-1} whereas SAPT2+3 with an aug-cc-pVDZ basis set calculates the cohesive energy as −19.455 kJ mol^{-1}, slightly closer to the S22 result. The basis sets/levels used in these calculations are typical of those used in the majority of this work.

1.3.1 The Volume Term, $P\Delta V$

Under ambient conditions the pressure-volume term, or the work term, is usually very small and is often neglected entirely from the Gibbs equation. For example, the contribution to Gibbs energy from a change in molecular volume of 5 Å^3 at ambient pressure (~0.0001 GPa) is equivalent to 3×10^{-4} kJ mol^{-1}, which is negligible when compared to other molecular energy contributions. However, under elevated pressures, it becomes increasingly significant. At 5 GPa, a relatively common pressure in the study of small molecule organic solids, this same change in volume is equivalent to an energy penalty of 15 kJ mol^{-1}, larger than a N—H···N hydrogen bond in ammonia (9.3 kJ mol^{-1}) and much larger than the typical polymorph energy differences revealed in Nyman and Day's study (see below).[19,42] The application of pressure allows the study of the role of the pressure-volume term. Some of the main motivations of studies at pressure as well as its role in driving phase transitions will be outlined in Section 1.3.2*a*.

The change in volume of a compound with increasing pressure can reveal structural information on the hardness of a material through the

calculation of its bulk modulus, Equation 1.3. To calculate this, the pressure-volume curve must be fitted with an appropriate equation of state (EoS).[43] The thermodynamic grounding of EoSs and some common EoS used within this work will be outlined in Section 1.3.2*b*.

The assessment of the response of molecular contacts to pressure requires a definition of the extent of the molecular surface for calculation of component volumes such as void space. This defines the limit of influence (or domain) of individual atom centres and is often completed using van der Waals radii, which constitute one measure of atomic size.[44] Some background on the development of van der Waals radii and a description of the origins of those used in this work will be described in Section 1.3.2*c*

1.3.2*a* The Application of Pressure

High pressure has been applied in a broad range of different disciplines. A pressure of 0.1 GPa is the highest pressure at which living organisms have been observed in deep sea trenches, leading to investigations on the effect of pressure on biomelecules.[45] Pressure at the earth's core thought to be in excess of 360 GPa,[46] and geophysical and planetary science studies are commonly conducted at pressures of the of the order of 10s or 100s of GPa.[47]The highest pressure reached for which structural data is available in the CSD is 180 GPa, this belongs to a series of rhenium carbides.

Structural work on molecular materials has mostly been limited to pressures below 10 GPa, 86.57% of structural data in this field having been measured at 6 GPa or below.[48] Interest has been directed towards new phases of compounds of pharmaceutical relevance which appear to show no propensity for polymorphism under ambient conditions.[49] For example, the pharmaceutical material chlorothiazide crystallises as a single polymorph

under ambient conditions but readily undergoes an isosymmetric phase transition to a second phase at 4.4 GPa.[50] Phases formed at high pressure can sometimes be recovered to ambient conditions, as has been demonstrated for paracetamol form II.[8] This form has a packing arrangement which promotes improved tablettability, but is difficult to produce at ambient pressure. Its recovery enabled formation of the phase by seeding at ambient conditions. High pressure studies also enable investigation on the results of applying pressure during industrial processes, a pressure of 0.1 GPa being comparable to mechanical pressures exerted during tabletting.[49] Chlorpropamide, an anti-diabetic drug, has been shown to undergo a phase transition in this process.[51] Identifying the driving forces of phase transitions in terms of chemically meaningful models is a major motivation behind high pressure research.

Pressure has also been used to study structure-property relationships. In spin-crossover compounds, lower spin states generally have a smaller volume and increasing pressure commonly favours this state.[52] Molecular magnetic materials also display interesting results with pressure, enabling the properties of molecule-based magnets to be studied in varying states of distortion.[53-55] Research has also been conducted into the study of metal organic frameworks (MOFs) at high pressure.[56] MOFs are generally highly porous, with void space accounting for between 10-70% of the total unit cell volume.[49] Elevated pressures can enhance penetration of guests into these pores, revealing structural effects which control guest uptake even at much lower pressures.[49,57,58]

1.3.2*b* Equations of State

Equations of state (EoSs) are used to model the behaviour of a material when exposed to varying external conditions, and are applied extensively in Chapter 3.

The simplest EoS is the ideal gas law, which relates pressure on an ideal gas, P, to its volume, V, to the number of moles, n, the molar gas constant, R, and the temperature, T:

$$PV = nRT \tag{1.31}$$

The EoSs available for solids also aim to describe the relationship between pressure, temperature and volume. Both pressure-volume (*PV*) and temperature-volume (*TV*) EoSs are parametrised using a reference volume. *PV* EoSs are also parametrised using the first and second derivatives of pressure with respect to volume, whereas *TV* EoS are parametrised with the thermal expansion coefficient and derivatives of temperature with respect to volume. This makes both *PV* and *TV* EoS well suited to the field of crystallography where changes in volume are measured. Pressure-volume-temperature (*PVT*) EoS also exist and in these cases more complex 'cross derivatives' are required, such as $(\partial K/\partial T)_P$. *PV* EoS were used exclusively in this thesis for the study of high-pressure crystal structures.

The second and third order Birch-Murnaghan and Vinet EoSs are used most commonly in high-pressure crystallography. When fitted to volume versus pressure data sets, these are useful not only for characterising the mechanical properties of a material, but, in addition, deviations from extrapolated trends can also reveal first and second order phase transitions (see Chapter 3). The parameter of most interest is usually the bulk modulus, K, the inverse of the isothermal compressibility, β (Equation 1.3). A negative sign in Equation 1.3 ensures that compressibility is always a positive value and

a reciprocal volume term normalises the compressibility for sample size. Units are typically given in GPa with harder materials characterised by a larger bulk modulus. For example, benzene is a van der Waals solid and has a bulk modulus of 5.5 GPa.[59] Diamond, one of the hardest materials known, has a bulk modulus of 445 GPa.[60]

The widely-used 3rd order Birch-Murnaghan EoS (BM3) is derived from the thermodynamic definition of pressure under isothermal compression as the negative derivative of the Helmholtz free energy, F, with respect to volume at constant temperature,[61]

$$P = -\left({}^{\partial F(V)}/_{\partial V}\right)_T \tag{1.32}$$

BM3 makes the fundamental assumption that for homogeneous strain, free energy of a solid under compression can be described as a polynomial of finite strain, f. For BM3 this is truncated at the third order term such that

$$F = a_0 + a_1 f + a_2 f^2 + a_3 f^3 \tag{1.33}$$

where the finite strain is

$$f = \frac{1}{2}\left[\left({}^{V_0}/_{V}\right)^{\frac{2}{3}} - 1\right] \tag{1.34}$$

a_0 is the free energy of the un-strained material at $P = 0$. By establishing a reference state of zero strain, this term is set to zero and *excess* free energy as a result of compression is considered only. In addition, for a system at equilibrium at any pressure

$$\frac{\partial F(V)}{\partial f} = 0 = a_1 + 2a_2 f + 3a_3 f^2 \tag{1.35}$$

For this equation to be consistent a_1 must be equal to zero also. The remaining expansion is a polynomial description of strain which is consistent with liner elasticity for small perturbations. Inserting the remaining terms of the expansion into the definition of pressure in these boundary conditions, $P = 0$ and $V = V_0$, yields

$$P = -\frac{\partial F(V)}{\partial V} = -\frac{\partial F(V)}{\partial f} \times \frac{\partial f}{\partial V} = -\left[\frac{\partial f}{\partial V}(2a_2 f)\right]_T - \left[\frac{\partial f}{\partial V}(3a_3 f^2)\right]_T \quad (1.36)$$

Simplification and rearrangement in terms of the parameters V_0, K_0, K_0' and K_0'' gives the full BM3 expression

$$P(V) = \frac{3K_0}{2}\left[\left(\frac{V_0}{V}\right)^{\frac{7}{3}} - \left(\frac{V_0}{V}\right)^{\frac{5}{3}}\right]\left\{1 + \frac{3}{4}\left((K_0' - 4)\left[\left(\frac{V_0}{V}\right)^{\frac{2}{3}} - 1\right]\right)\right\} \quad (1.37)$$

A simpler 2nd order EoS (BM2) can be obtained by truncation of Equation 1.33 to include only the $a_2 f^2$ term. This assumption is equivalent to setting the value of K_0' to 4 in Equation 1.37. The resulting BM2 expression is given by

$$P(V) = \frac{3K_0}{2}\left[\left(\frac{V_0}{V}\right)^{-\frac{7}{3}} - \left(\frac{V_0}{V}\right)^{-\frac{5}{3}}\right] \quad (1.38)$$

While BM3 is the most commonly applied EoS in molecular systems, BM2 fits are occasionally seen for poorer data or more restricted data sets. BM4 fits are seen occasionally in detailed work.[62]

For highly compressible compounds where $V_P/V_0 < 0.6$ an improved fit may be achieved with a Vinet EoS.[63] This has been used to produce high quality fits at high levels of compression for $MgSiO_3$, perovskite and solid hydrogen.[64,65] Unlike BM EoS, it is derived from an empirical potential. This is an approximation of the binding energy of metals and has the form

$$F(a) - F_0(1 + a)\exp(-a) \quad (1.39)$$

where a, is a reduction of the atomic spacing and is given by

$$a = \frac{r - r_0}{l} \quad (1.40)$$

where r_0 and r are the interatomic distances at ambient pressure and high pressure respectively. l is a scaling length. Setting $V = (4/3)\pi r^3$ the definition for pressure in this case becomes

$$P = -\frac{\partial F(V)}{\partial V} = -\frac{F_0}{4\pi l r^2} a exp(-a) \quad (1.41)$$

Setting the bulk modulus and its first derivative to

$$K_0 = \frac{F_0}{12\pi l^2 r_0} \tag{1.42}$$

and

$$K_0' = 1 + \frac{2r_0}{3l} \tag{1.43}$$

yields the final form of the Vinet EoS

$$P = K_0 \frac{3f_v}{(1-f_v)^2} exp(\eta f_v) \tag{1.44}$$

by defining additional variables

$$f_v = 1 - \left(\frac{V_P}{V_0}\right)^{\frac{1}{3}} \tag{1.45}$$

and

$$\eta = \frac{3}{2}(K_0' - 1) \tag{1.46}$$

At ambient pressure, f_v is zero but increases with decreasing volume at higher pressures. The presence of K_0' in the Vinet EoS means it is often described as 3rd order (V3) by analogy to the BM set. In theory, also by analogy, setting $K_0' = 1$ would result in a 2nd order EoS (V2), however unlike BM2 where this corresponds to truncations in finite strain, this has no formal theoretical basis and is not used.[66] Vinet equations of state are often called universal equations of state due to their wide-ranging applicability.

1.3.2*c* van der Waals Radii

As will be discussed in Chapter 3, the volume of a material is the sum of the space occupied by the atoms and interstitial or void space that exists between them. Atomic electron density formally only drops to zero at infinite distance,[29] but it is useful to define a practical cut-off radius as a measure of atomic size. In principle this could be the distance at which the electron density falls below a certain threshold, but this is difficult to define because small changes in threshold choice will have significant impacts on the resulting

radii.[29] For these reasons, the defined radius is based on empirical observations, usually from experimental interatomic distances. The van der Waals radius is a single value defined for each element generally irrespective of oxidation state, environment or external conditions such as temperature and pressure. The radius is modelled as the same in all directions, resulting in a spherical van der Waals surface around an atom. It has been shown by Nyburg and Faerman that this is an oversimplification, illustrated in their work by the spheroidal shape adopted by the chalogens and halogens.[67] Modelling atoms with constant values for their van der Waals radii is also an approximation as the effective size is dependent on environment and hybridisation state.[68] The concept of the van der Waals radius has nevertheless proved to be an enduring and practically useful formalism.[68]

The most commonly used van der Waals radii are those produced by Bondi in his 1964 and 1966 papers.[69,70] These radii are embedded into many commonly used graphics programs such as Mercury and Platon.[36,71] As a result, the original paper still receives 600-700 citations annually (Figure 1.5). Bondi's radii were calculated based on X-ray crystallographic data, gas kinetic collision cross sections, critical densities and liquid state properties.[69] This was originally limited to 38 elements, including few transition metals and only uranium in the f-block.

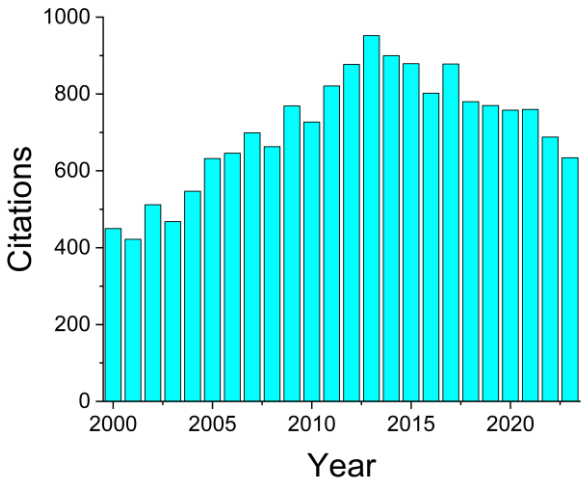


Figure 1.5: The rate of citations in Bondi, 1964.[69] These values are still used and cited widely.

There have been a number of attempts in recent years to extend and improve this dataset as the number of observed interatomic distances in the literature increases. In 2001, Batsanov provided a dataset for 68 elements, with better representation of the transition metals and the addition of thorium to the f-block.[72] These radii were calculated by assuming that the van der Waals radius of a given element is 0.8 Å longer than its covalent radius. Similar approaches of incrementing the sum of valence parameters by a constant were used by Nag *et al*., again for further transition metals and by Hu *et al*. to extend to the lanthanides and actinides.[44,73-75] The first fully computationally produced radii were calculated by Mantina *et al*. using noble gas atomic probes. These reproduced Bondi's set while adding the remaining 16 main group elements not included in Bondi's original study.[76]

Tables of van der Waals radii were, until 2013, therefore fragmentary and based on aged databases of crystallographic data. A more recent and comprehensive set of radii have since been deduced by Alvarez using large scale searches of the Cambridge structural database (CSD).[44,77] The distribution of experimental distances in the CSD between each element and oxygen (in most cases, with radius taken from Bondi's set) were plotted for all atom types.

For this process the van der Waals radius is still modelled as a sphere and the effects of oxidation state and coordination number are still neglected. A typical histogram of distances is shown schematically in Figure 1.6, next to a real distribution of Mg-O distances. The distribution partitions into regions at short distances characteristic of chemical bonds, those shown in grey for Mg-O between 1.8 and 2.5 Å, a 'van der Waals gap' where virtually no structures are observed, between 2.5 and 3.5 Å for Mg-O, and long intermolecular distances, those above 3.5 Å shown in blue for Mg-O. In this last region, a 'van der Waals peak' which characterises typical intermolecular contact distances, centred at 4.2 Å for Mg-O, gives way to a random distribution of very long, non-bonding interactions following the form of r^3, where r = interatomic distance.

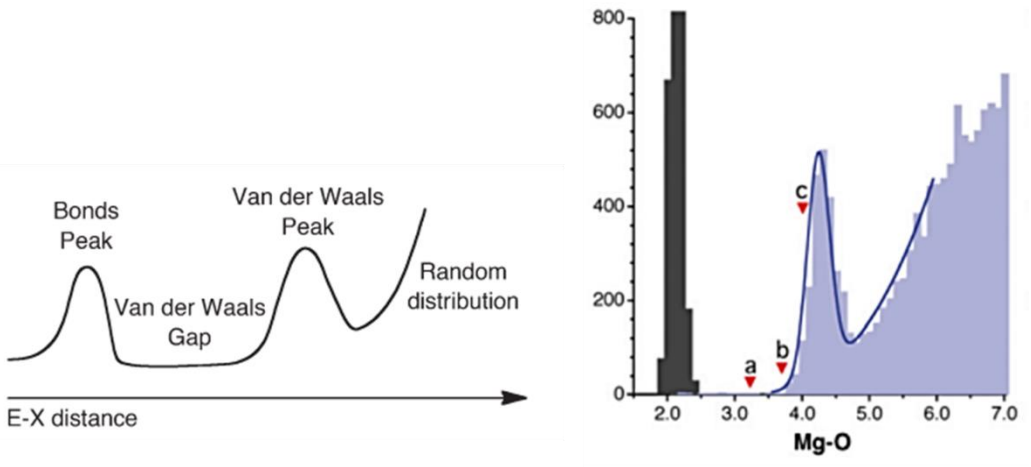


Figure 1.6: Left, a schematic representation of the relative bond distributions described by Alvarez. *E-X distance* is the distribution of distances between the studied element *E* and the probe element *R* (generally oxygen). Right, a real distribution of Mg-O contact distances, the grey bins indicate those attributed to covalent bonds and the blue represent intermolecular distances, *a* indicates the sum of Bondi's radii, *b* Batsanov's radii and *c* Alvarez's radii. Both figure parts are Reproduced from Ref 44 with permission from the Royal Society of Chemistry.

The van der Waals peak is the result of overlap between stabilising intermolecular interactions and the tail of the r^3 random distribution. The sum of the van der Waals radii is therefore not at the maximum of the van der Waals peak and must be extracted by removing the random distribution, non-bonding background.[44] The sum of the radii of element of interest, r_E, and probe element, r_X, are calculated from the Gaussian van der Waals peak centred at $d_{\max}$ at the full width at half maximum of the peak given by 2.3548σ, Figure 1.7. The radius of the studied atom is then extracted by assuming a constant radius of the probe atom, r_X (1.50 Å for O). In rare cases, low numbers of observations result in a poorly resolved van der Waals gap that restricts radii from being extracted. This is the case for some of the heaviest elements in the p-block, Tl and Pb. Attempts to use alternative atom probes for these atoms did not yield improved numbers of observations. The radii of the noble gases were later updated and extended to include Rn.[77] The Alvarez radii, now cover 94 elements and have been used exclusively in the present work.

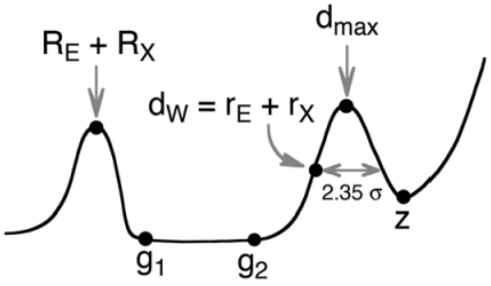


Figure 1.7: A schematic representation of the extraction of van der Waals radii from the van der Waals peak. The radius of element of interest is r_E, and the radius of the probe element is r_X. Reproduced from Ref 44 with permission from the Royal Society of Chemistry.

1.3.3 The Entropy Term, $T\Delta S$

The final term of the Gibbs equation is entropy. The Boltzmann equation for entropy is

$$S = k_B ln(W) \tag{1.47}$$

This equation, which was ultimately an assumption by Boltzmann, relates the entropy, S, to the number of microstates available, W, through the Boltzmann constant, k_B (1.38×10^{-23} JK^{-1}). The Boltzmann equation describes entropy in terms of the number of accessible states.[29] In a crystalline solid at absolute zero, 0 K, a system is restricted to just one microstate and the Boltzmann formula reduces to the third law of thermodynamics, entropy is zero. As temperature increases more microstates become accessible and the entropy of the system increases. This is referred to as configurational entropy and the availability of microstates leads to the common description of entropy as a measure of the order of a system.

Available microstates can be described in terms of degrees of freedom. For a monatomic, perfect gas, in the absence of any electronic effects the only degrees of freedom available are translations in all three dimensions. The molar entropy in this case is given by the Sackur-Tetrode equation

$$S_m = Rln\left(\frac{V_m e^{\frac{5}{2}}}{N_A \Lambda^3}\right) \tag{1.48}$$

where

$$\Lambda = \frac{h}{(2\pi m k_B T)^{\frac{1}{2}}} \tag{1.49}$$

V_m is the molar volume, N_A is Avogadro's number (6.022×10^{23} mol^{-1}), h is Planck's constant (6.626×10^{-34} J S), m is the molecular mass and all other symbols have the same meaning as above. Higher temperatures, higher masses and high molar volumes all contribute towards higher entropies. This is a very simple case, and for molecular gases additional rotational, vibrational

and electronic degrees of freedom exist. For most solids the origin of entropy lies in thermal vibrations and this has been the focus of the current work.[27]

Both experimental and computational determinations of entropies are generally more difficult than the corresponding analyses of enthalpies.[29] Entropy contributions are thus often neglected from crystallographic analysis. This simplification is analogous to reaction chemistry where the driving force of a reaction is largely considered in terms of bond breaking and bond making, ΔH, instead of a full analysis of the change in Gibbs free energy, ΔG.

The entropy of a material can be determined experimentally through heat capacity measurements. At constant pressure, the entropy, S, of a compound at temperature T_f is related to its entropy at T_i by measuring its heat capacity, C_p, at different temperatures and evaluating the integral

$$S(T_f) = S(T_i) + \int_{T_i}^{T_f} \frac{C_p(T)dT}{T} \tag{1.50}$$

Absolute entropies can be calculated by invoking the third law

$$S(T) = \int_0^T \frac{C_p(T)}{T} dT \tag{1.51}$$

The entropy of any phase transitions occurring over a temperature range of interest is accounted for using

$$\Delta S = \int_{T_1}^{T_2} \frac{C_{p,ex}}{T} dT = \frac{\Delta H}{T} \tag{1.52}$$

Commercial equipment for the measurement of heat capacities is generally limited to the study of solids. Although equipment capable of measuring heat capacities of liquids exists, it is far from standard apparatus and available data can be quite limited.

An alternative approach is to evaluate vibrational entropies using vibrational frequencies. These can, in principle, be measured using vibrational spectroscopy, but in solids dispersion of vibrational frequencies across the Brillouin zone needs to be taken into account. The most common experimental methods for measuring vibrational data, infra-red and Raman spectroscopies,

only sample vibrations at the Γ-point, where vibrations in all unit cells are in phase. Inelastic neutron scattering samples the entire Brillouin zone, but experiments require access to a neutron source, a very large sample (5-10 g is typical) and individual modes suffer considerable overlap.[78]

Due to these difficulties vibrational frequencies are often determined computationally. There are several methods commonly used, each attempting to find appropriate approximations to balance accuracy and very high computational expense. Nyman and Day have employed force field approaches to study entropy differences between polymorphs.[19] Their approach justified by the need to perform calculations for hundreds of different phases. Force fields of high enough quality for use in CSP studies were used for this work. Density functional tight-binding theory offers a semi-empirical, quantum mechanical approach which is also computationally inexpensive.[79] In this work, density functional theory based on density functional perturbation theory (DFPT) was used. DFPT is a first principles method that can be used to calculate $3N$ individual vibrational modes per **q**-point, where N is the number of atoms in the model. A deeper overview of the calculation procedure used in DFPT is given in Section 2.4.2.

Vibrational frequencies, v, contribute towards vibrational entropies according to the following equation

$$S_m^v = R\left\{\frac{\theta^v/_T}{e^{\theta^v/_T}-1} - ln\left(1 - e^{\frac{-\theta^v}{T}}\right)\right\} \tag{1.53}$$

where

$$\theta^v = \frac{hcv}{k_B} \tag{1.54}$$

c is the speed of light (2.99×10^8 m s^{-1}) and all other terms have the same meaning as above. The entropy contribution is highly sensitive to temperature and values are commonly quoted in terms of TS to provide values which can be compared to other energy terms at the temperature of interest. Low

frequency vibrations (< 200 cm^{-1}), corresponding in molecular crystals to the acoustic and external vibrations, contribute most to the entropy.

Frequencies from the output of DFPT and other calculations are often depicted in the form of density of states (DoS) plots. These present a histogramic representation of the density of vibrational frequencies, with larger peak heights representing increased numbers of vibrational modes.

The contributions of individual atom types to the DoS can be assessed using partial density of states (PDoS) plots. They are generated from vibrational frequencies in similar ways to DoS plots but with an additional scaling of each frequency by the contribution towards it of an individual atom type. The output of a DFPT calculation contains N eigenvector descriptions for x, y and z perturbations in atom positions for each of the $3N$ vibrational modes. For each eigenvector the fraction of the total motion attributed to the selected atom(s) is calculated and used to provide a total fractional contribution towards the mode.

1.4 Thesis Aim

The overall aim of this thesis is to investigate the role of volume and entropy on crystal structures, how they impact mechanical behaviour and phase transitions in molecular compounds. This will allow the assessment of the level to which a discussion of phase behaviour in terms of purely intermolecular interaction energies is an oversimplification. This will be completed with a particular focus on how this narrative adapts at elevated pressures.

To achieve this overall aim, this thesis will first investigate the behaviour of occupied and void space at elevated pressures and present a new method for calculating these component volumes in molecular compounds. Secondly, this method will be applied to a new high-pressure study on glyphosate to reinterpret phase transitions identified using spectroscopic data. Next the role of entropy will be assessed in a computational study of C—H···F containing compounds. Finally, new high-pressure phases of glyphosate formed under varying rates of compression will be presented. The contributions of volume, entropy and internal energy introduced in previous chapters will then be considered to deduce their importance in driving the transitions at elevated pressures.

1.5 References

(1) Dunitz, J. D.; Gavezzotti, A., How molecules stick together in organic crystals: weak intermolecular interactions. *Chem. Soc. Rev.* **2009**, 38 (9), 2622-2633.

(2) Warren, L. R.; McGowan, E.; Renton, M.; Morrison, C. A.; Funnell, N. P., Direct evidence for distinct colour origins in ROY polymorphs. *Chem. Sci.* **2021**, 12 (38), 12711-12718.

(3) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C., The Cambridge structural database. *Acta Crystallogr.* **2016**, B*72* (2), 171-179.

(4) Broadhurst, E. T.; Xu, H.; Clabbers, M. T. B.; Lightowler, M.; Nudelman, F.; Zou, X.; Parsons, S., Polymorph evolution during crystal growth studied by 3D electron diffraction. *IUCrJ.* **2020**, 7 (1), 5-9.

(5) Cruz-Cabeza, A. J.; Reutzel-Edens, S. M.; Bernstein, J., Facts and fictions about polymorphism. *Chem. Soc. Rev.* **2015**, 44 (23), 8619-8635.

(6) McCrone, W. C., Polymorphism. *Phys. Chem. Org. Solid State.* **1965**, 2, 725–767.

(7) Bernstein, J., *Polymorphism in Molecular Crystals*. Oxford University Press, **2002**.

(8) Oswald, I. D. H.; Chataigner, I.; Elphick, S.; Fabbiani, F. P. A.; Lennie, A. R.; Maddaluno, J.; Marshall, W. G.; Prior, T. J.; Pulham, C. R.; Smith, R. I., Putting pressure on elusive polymorphs and solvates. *CrystEngComm.* **2009**, 11 (2), 359-366.

(9) Galek, P. T. A.; Allen, F. H.; Fábián, L.; Feeder, N., Knowledge-based H-bond prediction to aid experimental polymorph screening. *CrystEngComm.* **2009**, 11 (12), 2634-2639.

(10) Moggach, S. A.; Parsons, S.; Wood, P. A., High-pressure polymorphism in amino acids. *Crystallogr. Rev.* **2008**, 14, 143-184.

(11) Ehrenfest, P. *Phasenumwandlungen im ueblichen und erweiterten Sinn, classifiziert nach den entsprechenden Singularitaeten des thermodynamischen Potentiales*; NV Noord-Hollandsche Uitgevers Maatschappij, **1933**.

(12) Jaeger, G., The Ehrenfest classification of phase transitions: introduction and evolution. *Archive for history of exact sciences* **1998**, 53, 51-81.

(13) Atkins, P.; De Paula, J.; Keeler, J., *Atkins' physical chemistry*. Oxford university press, **2023**.

(14) Perlovich, G. L.; Hansen, L. K.; Bauer-Brandl, A., The Polymorphism of Glycine. Thermochemical and structural aspects. *J. Therm. Anal. Calorim.* **2001**, 66, 699-715.

(15) Giordano, N.; Beavers, C. M.; Campbell, B. J.; Eigner, V.; Gregoryanz, E.; Marshall, W. G.; Peña-Álvarez, M.; Teat, S. J.; Vennari, C. E.; Parsons, S., High-pressure polymorphism in pyridine. *IUCrJ.* **2020**, 7, 58-70.

(16) Dawson, A.; Allan, D. R.; Belmonte, S. A.; Clark, S. J.; David, W. I. F.; McGregor, P. A.; Parsons, S.; Pulham, C. R.; Sawyer, L., Effect of high pressure on the crystal structures of polymorphs of glycine. *Cryst. Growth Des.* **2005**, 5 (4), 1415-1427.

(17) Lee, E. H., A practical guide to pharmaceutical polymorph screening & selection. *Asian J. Pharm. Sci.* **2014**, 9 (4), 163-175.

(18) Yu, L.; Stephenson, G. A.; Mitchell, C. A.; Bunnell, C. A.; Snorek, S. V.; Bowyer, J. J.; Borchardt, T. B.; Stowell, J. G.; Byrn, S. R., Thermochemistry and conformational polymorphism of a hexamorphic crystal system. *J. Am. Chem. Soc.* **2000**, 122 (4), 585-591.

(19) Nyman, J.; Day, G. M., Static and lattice vibrational energy differences between polymorphs. *CrystEngComm.* **2015**, 17 (28), 5154-5165.

(20) McMahon, M. I., *High-pressure crystallography. Advanced X-Ray Crystallography.* **2012**, 69-109.

(21) Desiraju, G. R., Supramolecular synthons in crystal engineering—a new organic synthesis. *Angew. Chem., Int. Ed. Engl.* **1995**, 34, 2311-2327.

(22) Wood, P. A.; Forgan, R. S.; Henderson, D.; Parsons, S.; Pidcock, E.; Tasker, P. A.; Warren, J. E., Effect of pressure on the crystal structure of salicylaldoxime-I, and the structure of salicylaldoxime-II at 5.93 GPa. *Acta Crystallogr.* **2006**, B62 (6), 1099-1111.

(23) Cruz-Cabeza, A. J.; Bernstein, J., Conformational polymorphism. *Chem. Rev.* **2014**, 114 (4), 2170-2191.

(24) Bauer, J.; Spanton, S.; Henry, R.; Quick, J.; Dziki, W.; Porter, W.; Morris, J., Ritonavir: an extraordinary example of conformational polymorphism. *Pharm. Res.* **2001**, 18, 859-866.

(25) Thompson, H. P. G.; Day, G. M., Which conformations make stable crystal structures? Mapping crystalline molecular geometries to the conformational energy landscape. *Chem. Sci.* **2014**, 5 (8), 3173-3182.

(26) Mati, I. K.; Cockroft, S. L., Molecular balances for quantifying non-covalent interactions. *Chem. Soc. Rev.* **2010**, 39 (11), 4195-4205.

(27) Dove, M. T., *Structure and dynamics: an atomic view of materials*. Oxford University Press, **2003**.

(28) Wood, P. A.; Allen, F. H.; Pidcock, E., Hydrogen-bond directionality at the donor H atom—analysis of interaction energies and database statistics. *CrystEngComm.* **2009**, 11 (8), 1563-1571.

(29) Gavezzotti, A., *Molecular aggregation: structure analysis and molecular simulation of crystals and liquids*. Oxford University Press, **2006**.

(30) Dunitz, J. D., Organic fluorine: odd man out. *ChemBioChem.* **2004**, 5 (5), 614-621.

(31) Boys, S. F.; Bernardi, F., The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **1970**, 19 (4), 553-566.

(32) Smith, D. G. A.; Burns, L. A.; Simmonett, A. C.; Parrish, R. M.; Schieber, M. C.; Galvelis, R.; Kraus, P.; Kruse, H.; Di Remigio, R.; Alenaizan, A., PSI4 1.4: Open-source software for high-throughput quantum chemistry. *J. Chem. Phys.* **2020**, 152.

(33) Gavezzotti, A., Calculation of lattice energies of organic crystals: the PIXEL integration method in comparison with more traditional methods. *Z. Kristallogr. Cryst. Mater.* **2005**, 220, 499-510.

(34) Gavezzotti, A., Efficient computer modeling of organic materials. The atom–atom, Coulomb–London–Pauli (AA-CLP) model for intermolecular electrostatic-polarization, dispersion and repulsion energies. *New J. Chem.* **2011**, 35, 1360-1368.

(35) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, 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.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; 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, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., Gaussian 09, Revision E. 01, Gaussian. *Inc:, Wallingford CT* **2009,** 201.

(36) Macrae, C. F.; Sovago, I.; Cottrell, S. J.; Galek, P. T. A.; McCabe, P.; Pidcock, E.; Platings, M.; Shields, G. P.; Stevens, J. S.; Towler, M., Mercury 4.0: From visualization to analysis, design and prediction. *J. Appl. Crystallogr.* **2020**, 53, 226-235.

(37) Reeves, M. G.; Wood, P. A.; Parsons, S., MrPIXEL: automated execution of Pixel calculations via the *Mercury* interface. *J. Appl. Crystallogr.* **2020**, 53, 1154-1162.

(38) Spackman, P. R.; Turner, M. J.; McKinnon, J. J.; Wolff, S. K.; Grimwood, D. J.; Jayatilaka, D.; Spackman, M. A., CrystalExplorer: A program for Hirshfeld surface analysis, visualization and quantitative analysis of molecular crystals. *J. Appl. Crystallogr.* **2021**, 54 (3), 1006-1011.

(39) Burrows, S. A.; Korotkin, I.; Smoukov, S. K.; Boek, E.; Karabasov, S., Benchmarking of molecular dynamics force fields for solid–liquid and solid–solid phase transitions in alkanes. *J. Phys. Chem.* **2021**, B125 (19), 5145-5159.

(40) Holden, D.; Jelfs, K. E.; Cooper, A. I.; Trewin, A.; Willock, D. J., Bespoke force field for simulating the molecular dynamics of porous organic cages. *J. Phys. Chem.* **2012**, C116 (31), 16639-16651.

(41) Takatani, T.; Hohenstein, E. G.; Malagoli, M.; Marshall, M. S.; Sherrill, C. D., Basis set consistent revision of the S22 test set of noncovalent interaction energies. *J. Chem. Phys.* **2010**, 132.

(42) Morrison, C. A.; Siddick, M. M., Determining the Strengths of Hydrogen Bonds in Solid-State Ammonia and Urea: Insight from Periodic DFT Calculations. *Chem. Eur. J.* **2003**, 9, 628-634.

(43) Angel, R. J., Equations of state. *Rev. Mineral. Geochem.* **2000**, 41 (1), 35-59.

(44) Alvarez, S. A., A cartography of the van der Waals territories. *Dalt. Trans.* **2013**, 42, 8617-8636.

(45) Kato, C.; Li, L.; Nogi, Y.; Nakamura, Y.; Tamaoka, J.; Horikoshi, K., Extremely barophilic bacteria isolated from the Mariana Trench, Challenger Deep, at a depth of 11,000 meters. *Appl. Environ. Microbiol.* **1998**, 64 (4), 1510-1513.

(46) Jeanloz, R., The nature of the Earth's core. *Annu. Rev. Earth Planet. Sci.* **1990**, 18, 357-386.

(47) Stan, C. V.; Beavers, C. M.; Kunz, M.; Tamura, N., X-ray diffraction under extreme conditions at the Advanced Light Source. *Quantum Beam Sci.* **2018**, 2 (1), 4.

(48) Kaźmierczak, M.; Patyk-Kaźmierczak, E., Kilobytes of kilopascals: high-pressure depositions of the Cambridge Structural Database. *Acta Crystallogr.* **2021**, B77 (6), 1012-1020.

(49) Moggach, S. A.; Oswald, I. D. H., *Crystallography under high pressures. 21st Century Challenges in Chemical Crystallography I: History and Technical Developments.* **2020**, 141-198.

(50) Oswald, I. D. H.; Lennie, A. R.; Pulham, C. R.; Shankland, K., High-pressure structural studies of the pharmaceutical, chlorothiazide. *CrystEngComm.* **2010**, 12 (9), 2533-2540.

(51) Otsuka, M.; Matsumoto, T.; Kaneniwa, N., Effects of the mechanical energy of multi-tableting compression on the polymorphic transformations of chlorpropamide. *J. Pharm. Pharmacol.* **1989**, 41 (10), 665-669.

(52) Gaspar, A. B.; Molnár, G.; Rotaru, A.; Shepherd, H. J., Pressure effect investigations on spin-crossover coordination compounds. *Comptes Rendus Chimie* **2018**, 21 (12), 1095-1120.

(53) Mingos, D. M. P., Raithby, P. R. eds., *21st Century Challenges in Chemical Crystallography I: History and Technical Developments*. Springer International Publishing, **2020**.

(54) Wang, X.; Kamenev, K. V., Review of modern instrumentation for magnetic measurements at high pressure and low temperature. *Low Temp. Phys.* **2014**, 40 (8), 735-746.

(55) Etcheverry-Berrios, A.; Parsons, S.; Kamenev, K. V.; Probert, M. R.; Moggach, S. A.; Murrie, M.; Brechin, E. K., Putting the Squeeze on Molecule-Based Magnets: Exploiting Pressure to Develop Magneto-Structural Correlations in Paramagnetic Coordination Compounds. *Magnetochemistry* **2020**, 6 (3), 32.

(56) Collings, I. E.; Goodwin, A. L., Metal–organic frameworks under pressure. *J. Appl. Phys.* **2019**, 126, 181101.

(57) Fairen-Jimenez, D.; Moggach, S. A.; Wharmby, M. T.; Wright, P. A.; Parsons, S.; Duren, T., Opening the gate: framework flexibility in ZIF-8 explored by experiments and simulations. *J. Am. Chem. Soc.* **2011**, 133 (23), 8900-8902.

(58) 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. Chem.* **2009**, 121 (38), 7221-7223.

(59) Moggach, S. A.; Parsons, S., High pressure crystallography of inorganic and organometallic complexes. *Spectrosc. Prop. Inorg. Organomet. Compd.* **2009**, 40, 324-354.

(60) Dewaele, A.; Datchi, F.; Loubeyre, P.; Mezouar, M., High pressure–high temperature equations of state of neon and diamond. *Phys. Rev.* **2008**, B77 (9), 094106.

(61) Birch, F., Finite elastic strain of cubic crystals. *Phys. Rev.* **1947**, 71 (11), 809.

(62) Pavese, A., Pressure–volume–temperature equations of state: a comparative study based on numerical simulations. *Phys. Chem. Miner.* **2002**, 29, 43-51.

(63) Poirier, J.-P., *Introduction to the Physics of the Earth's Interior*. Cambridge University Press, **2000**.

(64) Hama, J.; Suito, K., The search for a universal equation of state correct up to very high pressures. *J. Phys. Condens. Matter.* **1996**, 8 (1), 67.

(65) Hemley, R. J.; Mao, H. K.; Finger, L. W.; Jephcoat, A. P.; Hazen, R. M.; Zha, C. S., Equation of state of solid hydrogen and deuterium from single-crystal x-ray diffraction to 26.5 GPa. *Phys. Rev.* **1990**, B42 (10), 6458.

(66) Angel, R. J.; Alvaro, M.; Gonzalez-Platas, J., EosFit7c and a Fortran module (library) for equation of state calculations. *Z. Kristallogr. Cryst. Mater.* **2014**, 229 (5), 405-419.

(67) Nyburg, S. C.; Faerman, C. H., A revision of van der Waals atomic radii for molecular crystals: N, O, F, S, Cl, Se, Br and I bonded to carbon. *Acta Crystallogr.* **1985**, B41 (4), 274-279.

(68) Rowland, R. S.; Taylor, R., Intermolecular nonbonded contact distances in organic crystal structures: Comparison with distances expected from van der Waals radii. *J. Phys. Chem.* **1996**, 100 (18), 7384-7391.

(69) Bondi, A., van der Waals volumes and radii. *J. Phys. Chem.* **1964**, 68, 441-451.

(70) Bondi, A., van der Waals volumes and radii of metals in covalent compounds. *J. Phys. Chem.* **1966**, 70, 3006-3007.

(71) Spek, A. L., PLATON, an integrated tool for the analysis of the results of a single crystal structure determination. *Acta Crystallogr.* **1990**, A46 (s1), c34-c34.

(72) Batsanov, S. S., van der Waals radii of elements. *Inorg. Mater.* **2001**, 37 (9), 871-885.

(73) Nag, S.; Banerjee, K.; Datta, D., Estimation of the van der Waals radii of the d-block elements using the concept of bond valence. *New J. Chem.* **2007**, 31 (6), 832-834.

(74) Hu, S.-Z.; Zhou, Z.-H.; Xie, Z.-X.; Robertson, B. E., A comparative study of crystallographic van der Waals radii. *Z. Kristallogr. Cryst. Mater.* **2014**, 229 (7), 517-523.

(75) Hu, S.-Z.; Zhou, Z.-H.; Robertson, B. E., Consistent approaches to van der Waals radii for the metallic elements. *Z. Kristallogr.* **2009**, 224 (8), 375-383.

(76) Mantina, M.; Chamberlin, A. C.; Valero, R.; Cramer, C. J.; Truhlar, D. G., Consistent van der Waals radii for the whole main group. *J. Phys. Chem.* **2009**, A113 (19), 5806-5812.

(77) Vogt, J., Alvarez, S., van der Waals radii of noble gases. *Inorg. Chem.* **2014**, 53 (17), 9260-9266.

(78) Rivera, S. A.; Allis, D. G.; Hudson, B. S., Importance of vibrational zero-point energy contribution to the relative polymorph energies of hydrogen-bonded species. *Cryst. Growth Des.* **2008**, 8 (11), 3905-3907.

(79) Hourahine, B.; Aradi, B.; Blum, V.; Bonafé, F.; Buccheri, A.; Camacho, C.; Cevallos, C.; Deshaye, M. Y.; Dumitrică, T.; Dominguez, A., DFTB+, a software

package for efficient approximate density functional theory based atomistic simulations. *J. Chem. Phys.* **2020**, 152 (12).

This page has been left intentionally blank

Chapter 2: General Experimental and Computational Procedures.

2.1 Diffraction

The study of crystalline compounds via diffraction-based techniques is a staple of solid-state chemistry, it remains the only way to determine directly the structure of a compound at the atomic scale.[1] A crystal is *a material which exhibits an essentially sharp diffraction pattern*.[2] Diffraction occurs from crystals when the periodic atomic spacing is commensurate with the wavelength of the incident radiation. The periodicity of the crystal structure is characterised by the dimensions of the direct lattice, defined by the size of the simplest repeating unit, the unit cell, Figure 2.1.[1] Choice of the correct incident radiation is vital to the success of the diffraction experiment. There are currently three main sources of radiation routinely available for diffraction experiments; electrons, X-rays, and neutrons.[3]

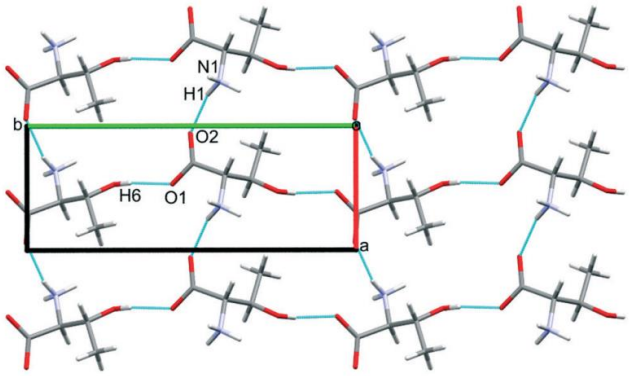


Figure 2.1: The simplest repeating unit for *L*-threonine, the unit cell, viewed down the *c* axis.[4] Molecules are connected by light blue contacts which indicate the direction and presence of hydrogen bonds. Reproduced from Ref 4 with permission from the Royal Society of Chemistry.

The use of electrons in diffraction experiments is at the very forefront of innovation in the field of crystallography. It is only within the last 15-20 years that 3D electron diffraction (3D ED) methods have made the reliable determination of crystal structure models from electron diffraction data possible.[5-8] Electron diffraction is particularly well suited to crystals on the

scale of 100s nm which are too small for conventional X-ray analysis. Dynamical diffraction due to multiple scattering events remains an unavoidable difficulty of electron diffraction and this can limit the precision of structural data determined by electron diffraction.[9]

Diffraction experiments are overwhelmingly carried out using X-ray sources. X-ray wavelengths used commonly are of the scale 0.5-1.5 Å, extending to 0.2-3.0 Å for specialised experiments (high pressure measurements at the shorter end and anomalous scattering measurements at the longer end). X-rays are most commonly generated in home labs by impact of an accelerated electron beam on a metal target.[1] The wavelength of the radiation emitted is characteristic of the metal target. Common X-ray sources, and their corresponding wavelengths, used within home labs are; Mo K_α, 0.7107 Å, Ag K_α, 0.5638 Å and Cu K_α, 1.5406 Å.[1,10] The choice of source (where available) should be made to complement the experimental conditions.[10] Shorter wavelengths scatter more weakly but can provide better resolution of the crystal structure by allowing access to lower atom spacings or a larger volume of reciprocal space.[11]

X-rays produced at synchrotron radiation sources are also common for less routine or extreme experimental conditions. Electrons, first generated in an electron gun are accelerated initially in a linear accelerator (LINAC) before being passed to a booster ring. In the booster ring the electrons are accelerated further to 99.9% of the speed of light.[12] Here the electrons are also bunched, with nanosecond separation between the bunches before being passed into the storage ring. Here bending magnets in second generation synchrotrons and wigglers and undulators in third generation synchrotrons are used to produce radiation as the bunches of electrons are bent around the ring. The change in momentum (direction) of the bunch of electrons results in the emission of an

intense beam of X-rays tangentially to the electron bunch path. Third generation synchrotrons are able to produce more focussed beams by additionally utilising the straight sections of the storage ring. Here wigglers or undulators bend the bunch back and forth in an oscillation to create a high emission rate of radiation. Interference between each bend results in a very intense and directional beam.[13] Radiation from bending magnets and insertion devices (wigglers and undulators) can be directed towards beamlines containing experimental hutches. The wavelength of the beam is broad and can be customised using optics such as mirrors and monochromators to select wavelengths of interest for the particular experiment. Synchrotron sources are capable of producing a significant flux of X-rays, up to 100 billion times brighter than standard lab X-ray tubes.[14] Synchrotron radiation was used in Chapter 6 of this thesis at the Deutsches Elektronen-Synchrotron (DESY) P02.2 extreme conditions beamline.

The process of diffraction can be described by the Thomson elastic scattering model.[11] The electrons within atoms oscillate in response to incoming X-rays, producing X-rays at the same wavelength, which are emitted in a spherical wave.[11] The interference between these waves when generated in a periodic lattice gives rise to the characteristic diffraction patterns obtained from crystalline materials, Figure 2.2. These patterns were first described mathematically using classical diffraction theory by von Laue,[15] but a simpler model based on reflections from sets of equally spaced planes constructed through the crystal structure was later proposed by Lawrence Bragg.[16-18] This model led to Bragg's law (Equation 2.1), where the wavelength, λ, of the diffracted X-rays at angle, θ, is related periodic spacing of planes, d, by

$$\lambda = 2d_{(hkl)} \sin\theta \tag{2.1}$$

Different sets of planes are identified using the Miller indices h, k and l in Equation 2.1. In single crystal diffraction each set of planes gives rise to a spot in a diffraction pattern which is observed when the angle, θ, between the incident beam and the miller planes satisfies Equation 2.1.

Rotation of the detector and the sample allows a complete set of spots to be collected, generating a three-dimensional diffraction pattern regularly distributed over a second lattice, referred to as the reciprocal lattice, Figure 2.2. Well-defined mathematical relationships enable the dimensions of the unit cell to be obtained from the dimensions of the reciprocal lattice.[11]

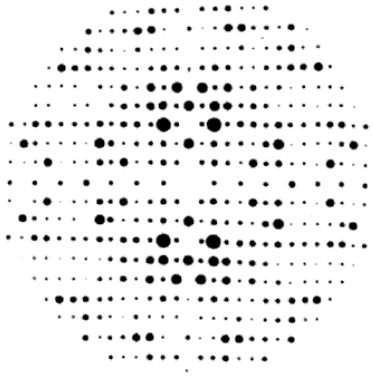


Figure 2.2: The distribution of diffraction spots in reciprocal space for oxalate monohydrate. The dimensions of the reciprocal lattice enable the dimensions of the unit cell to be determined. Crystal Structure Analysis- Principles and Practice. Reproduced with permission of Oxford University Press through PLSclear.

The variation of the intensities of the diffraction spots allows the crystal structure to be determined. The intensity of each Bragg spot is proportional to the square of the amplitude of the structure factor, $F_{(hkl)}$, the sum of the scattered waves by all j atoms in the unit cell with coordinates (x_j, y_j, z_j)

$$F_{(hkl)} = \sum_{j=1}^{N} f_j \exp\left[2\pi i\left(hx_j + ky_j + lz_j\right)\right] \exp\left(\frac{-8\pi^2 U_j \sin^2\theta}{\lambda^2}\right) \qquad (2.2)$$

where f_j is the scattering factor, characteristic of each atom type and a function of the number of electrons bound to the atoms.[11] Hence, heavier atoms will scatter X-rays much more strongly. Since electrons are scattered from diffuse

electron density, the value of *f* is resolution dependent, dropping-off rapidly at short *d*-spacings, Figure 2.3.

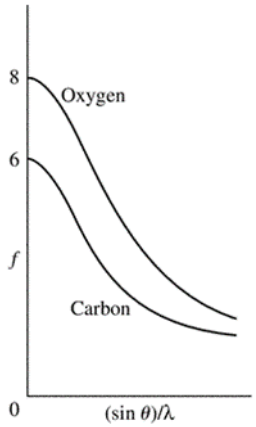


Figure 2.3: The dependence of scattering factor, f, on the reciprocal of the *d*-spacing ($sin(\theta)/\lambda$). At large spacings the scattering factor is proportional to the number of electrons in the atom but this drops off rapidly at shorter spacings.[11] Crystal Structure Analysis-Principles and Practice. Reproduced with permission of Oxford University Press through PLSclear.

The second exponential term in Equation 2.2 accounts for an isotropic treatment of atoms to account for thermal motion. Referred to as a displacement parameter, this is often expanded to account for anisotropic atomic motion in higher quality structural refinements where up to six independent parameters may be refined. The structure factor is subsequently related to the scattering electron density, $\rho_{(hkl)}$, via an inverse Fourier transform over all reflections

$$\rho_{(hkl)} = \frac{1}{V}\sum_{h,k,l} F_{(hkl)} \exp\left[-2\pi i(hx + ky + lz)\right] \qquad (2.3)$$

where V is the volume of the unit cell.[11] From this Fourier synthesis, an electron density map can be created, Figure 2.4, from which the positions of atoms can be identified. X-rays are diffracted from the whole crystal rather than a single unit cell. The finite time it takes to complete a diffraction experiment means we see a time-averaged picture of the electrons smeared out due to the thermal vibrations.[11]

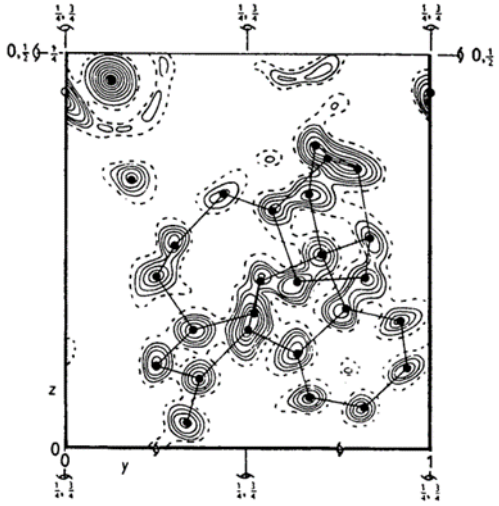


Figure 2.4: A Fourier projection down the x-axis for strychnine hydrogen bromide. Produced by Arnold Beevers in the same office as this thesis was written. Reproduced with permission of the International Union of Crystallography.[19]

Measured X-ray intensities provide only the structure factor amplitude and not their phases. Since phases cannot be measured experimentally, the calculation of electron density cannot be performed directly from experimental measurements. This is referred to as the phase problem.[20] There are several methods available that attempt to assign phases to reflections to solve a crystal structure. Some common methods are Patterson, direct or charge flipping methods.

The initial structure solution gives approximate values for the atomic coordinates, displacement parameters and other refinement parameters to produce a model. This model is then iteratively improved by adjusting the values of the parameters as to make the values of intensities calculated from the model match the measured intensities as closely as possible.[1] The R-factor provides a measure of the success of this process by comparing experimental and calculated structure factors.

$$R = \frac{\sum\left||F_i(\text{obs})|-|F_i(\text{calc})|\right|}{\sum|F_i(\text{obs})|} \qquad (2.4)$$

The third source used for structure determination is neutron radiation. Neutron diffraction is an established technique but the production of neutrons suitable for diffraction is limited to central facilities. Neutrons are generated by either fission or spallation type sources. Fission sources, such as that used at the Institut Laue-Langevin (ILL) in Grenoble, produce neutrons by thermal fission of nuclear isotopes, commonly ^{235}U.[3] This produces a continuous stream of neutrons which are slowed, or moderated, to thermal neutrons with energies (and de Broglie wavelengths) suitable for crystallographic diffraction experiments.[21] Spallation type sources, such as that at the ISIS neutron and muon source in Harwell, bombard a metal target, typically tungsten or tantalum, with high energy protons generated using a synchrotron.[3] This bombardment triggers emission of high energy neutrons which are then moderated to energies suitable for diffraction. As protons are produced from the synchrotron in pulses, spallation sources produce a pulsed neutron beam. This allows time of flight (TOF) set ups to be incorporated into spallation source facilities, discussed further in Section 2.3.3*a*.[3,22]

Neutron diffraction differs from X-ray diffraction in that there is no systematic trend scattering factor. Neutron diffraction involves interactions of a neutral beam and scattering can occur through magnetic or nuclear interactions. Nuclear scattering is defined by the elastic neutron scattering cross section[23]

$$\frac{d\sigma}{d\Omega} = b^2 = \frac{\sigma_t}{4\pi} \tag{2.5}$$

where b is the neutron scattering length and σ_t is the total scattering into a elementary scattering cone of solid angle $d\Omega$ per unit time. Scattering lengths can be positive or negative and are also unique to isotopes. This can make neutron diffraction attractive for studies where the exact positions of light atoms such as hydrogen are important.[24] Additionally, careful control of

isotopic substitution can be used to *hide* certain parts of an extended structure by matching scattering lengths.[25,26] The scattering length fulfils a similar role in modelling diffracted intensities to the scattering factor in X-ray diffraction. The nuclei themselves act as point scatters rather than diffuse electron clouds meaning neutron diffraction does not suffer the same attenuation in intensity at high resolution which characterises X-ray diffraction patterns.[23]

Samples in Chapter 6 of this thesis were studied using neutron powder diffraction. Powder diffraction removes the need to grow large, high-quality single crystals and can also give access to more diverse sample conditions. The theory of powder diffraction (PXRD) is the same as single crystal diffraction, except that polycrystalline powders are fine-grained mixtures of very small crystallites all with different orientations.[1] If sufficient orientations are present, *i.e.* the sample is large enough, all possible orientations of the reciprocal lattice can be present at once.[1] Increasing numbers of crystallites leads to the formation of multiple diffraction patterns which overlap and coalesce into a series of rings. The radii of these rings can then be used to recover the lattice spacings.

In single crystal X-ray diffraction (SCXRD) the Bragg equation is satisfied only at certain angles and rotation of the crystal and the detector is required to sample a sufficient amount of reciprocal space. In powder diffraction, the presence of many different orientations means that multiple Bragg conditions are satisfied all at once and the need for rotation is significantly reduced. The use of multiple detector banks can even remove the requirement for rotation entirely.

The refinement of PXRD data presents several challenges. Each reflection and its symmetry-equivalents,[1] will have the same scattering angle 2θ and will contribute towards the same ring. More seriously, the intensity of

the ring also contains contributions from any other reflections which happen to have the same, or approximately the same, *d*-spacing, making deconvolution of measured intensities into contributions from individual reflections a complex or even impossible process. The diffraction experiment is therefore limited to the measurement of several 10s or at best 100s of peaks, most of which contain contributions from several reflections. Despite this, a remarkable amount of structural information can be recovered from powder diffraction, it has even been used by Wright and co-workers to recover phase information of proteins.[27] Several different methods exist to recover structural information from the low number of peaks, including the Pawley and Rietveld refinement methods used in Chapter 6.

Because the exact peak width function is not known, Pawley methods attempt to perform pattern decomposition of the overlapping peaks as opposed to deconvolution.[1] The Pawley method generates peak positions by using an input space group and set of unit cell parameters which are refined to give the best fit to the experimental data. Pawley refinement is an iterative process which attempts to minimise *R*,[28]

$$R = \sum_{\theta}[y(\text{obs},\theta) - y(\text{calc},\theta) - Bg(\theta)]^2 \qquad (2.6)$$

where $y(\text{obs},\theta)$ at $y(\text{calc},\theta)$ are the observed and calculated intensities at the angle θ respectively. $y(\text{calc},\theta)$ has an additional dependence on the peak shape function used. $Bg(\theta)$ is the background and is modelled with a polynomial in θ.[28] The intensities of reflections, background parameters and peak shape parameters are all refined simultaneously along with unit cell parameters during the refinement process.[28] The Pawley method can be used to recover the unit cell parameters for phase identification but is 'structure-free' and is not suitable for modelling atom positions.

Rietveld refinement is the most common way to recover full structure refinement from PXRD data.[1] Unlike Pawley methods, intensities are reproduced using structure factors of a full crystal structure model. An initial structure model can be obtained using simulated annealing methods if a structure is unknown.[29] Atom positions, site occupancies, thermal parameters and functions to model the background and peak shapes are all refined in order to reproduce the whole diffraction pattern.[28] Observed and calculated diffraction patterns are compared to produce an R-factor similar to above with the calculated intensity at point θ_i given by[30]

$$y_{\mathrm{cal}}(\theta_i) = s\sum_k m_k (Lp)_k |F_k|^2 P_k \Psi(\theta_i - \theta_k) + b(\theta_i) \qquad (2.7)$$

where s is a scale factor, m_k is the reflection multiplicity, Lp is an angle dependent factor, P_k a function to deal with preferred orientation, Ψ a line profile function and b an isotropic atomic displacement factor. Although minimisation of terms such as the R-factor are useful in the iterative fitting process, the quality of the fit in powder diffraction is equally commonly assessed by visual inspection of a difference curve between observed and calculated patterns over the whole angular range.[30]

2.2 Determination of intermolecular interaction energies, ΔU

As discussed in Chapter 1, internal energies were determined exclusively with SAPT and Pixel methods in this thesis. These two methods were briefly introduced in Chapter 1 but will be discussed in greater detail below. Where calculations were performed on experimental structures directly, intramolecular hydrogen bond lengths were adjusted to their neutron values to account for artificial shortening. More commonly calculations were run on structures following geometry optimisation and the X—H bond lengths were not further adjusted in these cases.

2.2.1 SAPT Calculations

The input for SAPT calculations is an isolated dimer. The total dimer Hamiltonian is solved by first calculating the 'first order' Hamiltonian, a Hartree-Fock description of the isolated monomers.[31] This is then perturbed with intramonomer correlation terms, equivalent to Møller-Plesset perturbation theory on each monomer,[32] and an additional intermonomer correlation term for electron-nuclear interactions. This can be described by[32]

$$\hat{H} = \hat{F}_A + \hat{F}_B + \hat{W}_A + \hat{W}_B + \hat{V}_{AB} \quad (2.8)$$

where $\hat{H}$ is the total dimer Hamiltonian, $\hat{F}_A$ and $\hat{F}_B$ are the monomer Fock operators, $\hat{W}_A$ and $\hat{W}_B$ are the intramolecular correlation terms and $\hat{V}_{AB}$ is the intermolecular correlation term. In practice this forms a double perturbation series indexed by m and n, the order of the inter and intramonomer perturbation series, respectively. The individual terms from the perturbation series have physical meaning and are usually quoted in terms of electrostatics, induction, dispersion and exchange energies.[31] The induction term is

equivalent to polarisation energy and the exchange term is equivalent to repulsion energy. This alternate naming has been used throughout for consistency with other methods.

SAPT calculations can be truncated to different levels with increasing orders m and n also increasing both accuracy and computational cost.[31] For most dimers, this truncation converges rapidly meaning lower order truncations are often valid.[31] The simplest SAPT level, SAPT0, is simply a Hartree-Fock treatment of the monomers, the 0 indicating the absence of intramolecular correlation terms, but with second order intermolecular correlation terms included. The higher levels of SAPT, SAPT2 and SAPT2+ additionally include increasing levels of intramonomer correlation to the calculation. While the highest levels of SAPT, SAPT2+(3) and SAPT2+3 add increasingly higher levels of intermonomer correlation terms. The relative SAPT terms can be described briefly as below[31]

$$E_{\mathrm{SAPT}} = \sum_{m>0}\sum_{n\geq 0} E_{\mathrm{SAPT}}^{(mn)} = E_{\mathrm{Elec}} + E_{\mathrm{Pol}} + E_{\mathrm{Disp}} + E_{\mathrm{Rep}} \tag{2.9}$$

$$E_{\mathrm{SAPT0}} = E_{\mathrm{Elec}}^{(10)} + E_{\mathrm{Rep}}^{(10)} + E_{\mathrm{Pol,resp}}^{(20)} + E_{\mathrm{Rep-Pol,resp}}^{(20)} + \delta E_{\mathrm{HF}}^{(2)} + E_{\mathrm{Disp}}^{(20)} + E_{\mathrm{Rep-Disp}}^{(20)} \tag{2.10}$$

$$E_{\mathrm{SAPT2}} \mathrel{+}= E_{\mathrm{Elec,resp}}^{(12)} + E_{\mathrm{Rep}}^{(11)} + E_{\mathrm{Rep}}^{(12)} + {}^{t}E_{\mathrm{Pol}}^{(22)} + {}^{t}E_{\mathrm{Rep-Pol}}^{(22)} \tag{2.11}$$

$$E_{\mathrm{SAPT2+}} \mathrel{+}= E_{\mathrm{Disp}}^{(21)} + E_{\mathrm{Disp}}^{(22)} \tag{2.12}$$

$$E_{\mathrm{SAPT2+(3)}} \mathrel{+}= E_{\mathrm{Elec,resp}}^{(13)} + E_{\mathrm{Disp}}^{(30)} \tag{2.13}$$

$$E_{\mathrm{SAPT2+3}} \mathrel{+}= E_{\mathrm{Rep-Disp}}^{(30)} + E_{\mathrm{Pol-Disp}}^{(30)} + E_{\mathrm{Rep-Pol-Disp}}^{(30)} \tag{2.14}$$

The Hartree-Fock calculation requires basis sets with both polarisation and diffuse functionals to be defined.[31] These are available in augmented (aug) Dunning basis sets which are also correlation-consistent.[33] They are available in double, triple, quadruple, and quintuple-ζ forms and can be truncated to improve computational cost in what is known as the 'calendar' set, reducing

backwards in 'month' (aug, jul, jun…) for each diffuse shell removed from the calculation.[31] An excellent review on the construction of Dunning basis sets used within SAPT calculations can be found in Ref. 31. Within Chapter 5, SAPT2+3 was used with a double ζ basis set and an 'august' augmentation level for diffuse functions on all atoms as the best compromise between computational cost and accuracy, see Appendix C, Section C.4. This basis set was used for fluorobenzene compounds containing carbon, fluorine and hydrogen only and is labelled as aug-cc-pVDZ.[31] Functions are added in shells and for carbon and fluorine the double- ζ form consists of 3 *s*-functions, 2 *p*-functions and 1 *d*-function.[31] The augmentation adds a set of diffuse functions for every angular moment quantum number present and for carbon and fluorine this results in 4 *s*-functions, 3 *p*-functions and 2 *d*-functions, totalling 23 basis functions.[34] For hydrogen and helium the largest angular momenta in diffuse space are 1 smaller than for the elements Li-Kr.[34] Hence the functions for hydrogen at this level were 2 *s*-functions and 1 *p*-function which were augmented to 3 *s*-functions and 2 *p*-functions, totalling 9 basis functions.[31]

Bordner has found SAPT calculations can provide similar accuracy to experimental noble gas interaction potential measurements as coupled-cluster perturbative triple excitation (CCSD(T)) calculations.[35] SAPT calculations were run through the program Psi4.[36]

2.2.2 PIXEL Calculations

Pixel calculations were used throughout this thesis for the calculations of intermolecular and lattice energies based on crystal structure data. Pixel calculates interaction energies between pixels (or voxels) of electron density. Electron densities are determined through wavefunction calculations, typically using the programme Gaussian at the B3LYP or MP2 level of theory

with a 6-31G** basis set.[37] The electron density is then portioned into pixels, usually of dimensions of 0.06 or 0.08 Å in this work, where it is then averaged to form a pixel of homogeneous electron density in what is known as a 'cube' format electron density file.[38]

Even for quite small molecules the number of pixels can become very large, typically of the order 10^6.[39] In order to accelerate the calculations, the pixels are 'condensed' into super pixels. These super pixels contain $n \times n \times n$ original pixels, with n termed the condensation level, typically 3 or 4.[38] Super pixels are assigned to individual atoms on the basis of the smallest ratio of pixel-nuclear distance to atomic radius. If the pixel lies outside of any atomic radius, it is assigned to the atom with the closest atomic surface.[39] This association of pixels to individual atoms allows the electron density to be assigned different parameters based on its chemical environment.

A cluster of molecules is built up around the central reference molecule to a defined radius using the space group symmetry of the crystal structure. For materials of low polarity, the cluster radius is typically 14-20 Å, but for ionic or zwitterionic materials larger values are required, 40-50 Å.[40] Intermolecular energies between the central molecule and each of the additional molecules in this cluster are calculated through pair-wise evaluation of the pixel-pixel energies to produce electrostatic, polarisation, dispersion, repulsion, and total energies.

Intermolecular electrostatic energies are calculated using Coulomb's law by treating each pixel or nucleus as a point charge; pixel-pixel, pixel-nuclear and nuclear-nuclear interactions are evaluated.[39] If there are two molecules, A and B, where molecule A has nuclei of charge Z_j at positions $j = (x_j, y_j, z_j)$ and pixels of charge q_k at positions $k = (x_k, y_k, z_k)$, and molecule B similarly has nuclei of charge Z_m at positions $m = (x_m, y_m, z_m)$ and pixels of

charge q_i at positions $i = (x_i, y_i, z_i)$, the electrostatic potential at point i of the charge density of molecule B, generated by molecule A is[39]

$$\phi_i = \frac{1}{4\pi\varepsilon_0}\left[\sum_k \frac{q_k}{R_{ik}} + \sum_j \frac{Z_j}{R_{ij}}\right] \quad (2.15)$$

where charge pixels are separated by a distance R_{ik}, nuclei of A and pixels of B are separated by a distance R_{ij} and ε_0 is the vacuum permittivity (8.854×10^{12} F m^{-1}). A similar equation can be written for the electrostatic potential at nuclei m of molecule B, generated by molecule A[39]

$$\phi_m = \frac{1}{4\pi\varepsilon_0}\left[\sum_k \frac{q_k}{R_{km}} + \sum_j \frac{Z_j}{R_{jm}}\right] \quad (2.16)$$

where charge pixels of A and nuclei of B are separated by a distance R_{km} and nuclei of A and nuclei of B are separated by a distance R_{jm}. The sum of these electric potentials multiplied by their respective associated charges gives the total electrostatic energy[39]

$$E_{\text{Elec}} = \sum_i q_i\phi_i + \sum_m Z_m\phi_m \quad (2.17)$$

In calculations of the electrostatic term the inverse dependence on separation means any overlapping densities at very short separation distances will lead to singularities. To avoid this, any pixel-pixel distances that are less than half the dimension of a super pixel are reset to exactly half the super pixel dimension in a 'collision avoidance' procedure.[39]

Polarisation forces are calculated by considering atomic polarisability as opposed to molecular polarisability. The polarisability of a pixel i is taken to be

$$\alpha_i = \left(\frac{q_i}{Z_{atom}}\right)\alpha_{atom} \quad (2.18)$$

where q_i is the charge of the pixel, and Z_{atom} and α_{atom} are respectively the charge and polarisability of the atom with which it is associated. An electric field of strength ε induces a dipole moment μ_i[39]

$$\mu_i = \alpha_i\varepsilon \quad (2.19)$$

If an increasing electric field is applied, the change in energy associated with the pixel is given by

$$\Delta E_i = -\mu_i d\varepsilon \tag{2.20}$$

where $d\varepsilon$ is the increase in the electric field. In the range from 0 to ε this becomes[39]

$$\Delta E_i = -\int_0^{\varepsilon} \mu_i d\varepsilon \tag{2.21}$$

Combining this with Equation 2.19 yields[39]

$$\Delta E_i = -\int_0^{\varepsilon} \alpha_i \varepsilon d\varepsilon = -\frac{1}{2}\alpha_i \varepsilon^2 \tag{2.22}$$

As for the electrostatic term, very short pixel-pixel separation distances can be problematic for the polarisation calculation. These would produce very high, unphysical, electric fields on the scale 10^{13} V m^{-1}.[39] To avoid this, the same collision avoidance procedure is applied and then a damping term, d_i, added to Equation 2.22

$$\Delta E_i = -\frac{1}{2}\alpha_i [\varepsilon d_i]^2 \tag{2.23}$$

for $\varepsilon < \varepsilon_{max}$

$$d_i = \exp\left(\frac{-\varepsilon}{\varepsilon_{max} - \varepsilon}\right) \tag{2.24}$$

When $\varepsilon > \varepsilon_{max}$ the contribution is zero. ε_{max} is an empirical parameter and is normally set to 150×10^{10} V m^{-1}. The total polarisation energy of a molecule is then calculated as the sum of the polarisation energies of the associate pixels[39]

$$E_{\mathrm{Pol}} = \sum \Delta E_i \tag{2.25}$$

Dispersion energies are calculated for molecules A and B as the sum of individual pixel-pixel dispersion terms. Each pixel is considered as a separate oscillating charge. In the London approximation, the oscillation frequency is associated with an ionisation potential. In a Pixel calculation the effective ionisation potential of a pixel I_i is estimated from a function of the ionisation

potential of its corresponding atom, I^0, and its distance to the atomic nucleus, R_i[39]

$$I_i = I^0 exp(-\beta R_i) \tag{2.26}$$

where β is a dispersion energy coefficient, an empirical value almost always taken to be 0.4 Å^{-1}. This contributes towards an oscillator strength for a pair of pixels, E_{OS}[39]

$$E_{\mathrm{OS}} = \left(I_i I_j\right)^{\frac{1}{2}} \tag{2.27}$$

The individual dispersion terms are then approximated by the London dispersion equation as an extension of the Drude model[39,41]

$$E_{\mathrm{Disp}\ A,B} = -\frac{3}{4}\frac{\sum_{i,A}\sum_{j,B} E_{OS} f(R)\alpha_i\alpha_j}{(4\pi\varepsilon_0)^2\left(R_{ij}\right)^6} \tag{2.28}$$

$$f(R) = \exp\left[-\left(\frac{D}{R_{ij}-1}\right)^2\right] \text{ for } R_i < D \tag{2.29}$$

$$f(R) = 1 \text{ for } R_i > D \tag{2.30}$$

where $f(R)$ is a damping factor which again restricts the appearance of singularities at very short pixel-pixel separation distances. D is the damping threshold and is also an empirical parameter, normally this is set to 3.50 Å.[39]

Pauli, or short-range, repulsion is taken to be proportional to the overlap of the electron densities of two molecules. For molecules A and B, the overlap is given by[39]

$$S_{AB} = \sum_{i,A}\sum_{j,B}\left[\rho_i(A)\rho_j(B)\right]V \tag{2.31}$$

where V is the pixel volume and ρ_i and ρ_j are the respective charge densities of molecules A and B. By assigning pixels to specific atoms, the overlap is split into contributions from pairs of individual atoms. For atoms m and n this gives the contribution S_{mn}. For an electronegativity difference $\Delta\chi_{mn}$ between the atoms this gives the repulsion energy as[39]

$$E_{\mathrm{Rep},mn} = (K_1 - K_2\Delta\chi_{mn})S_{mn} \tag{2.32}$$

where $K_1$and K_2 are empirical parameters, normally 4800 and 1200 respectively.[42,43] Pixel uses undeformable electron densities and the overall term in the brackets is used to approximate the deformation of electron densities as molecules approach one another. A study by Wood *et al*. on comparative energies calculated by Pixel and DFT (SIESTA) calculations over the range 0-8 GPa has shown high correlation over the entire range, indicating that the deformation of electron densities is (i) small and (ii) well approximated by the term in brackets.[44]

Short-range repulsion effects are highly sensitive to molecule separation. The dimer energies that contain a non-zero repulsion energy are used to define the first coordination sphere of the reference molecule.[38]

The total intermolecular energy for pixel process is the sum of the electrostatic, polarisation, dispersion and repulsion terms.[39] In principle, lattice energies are calculated for molecular compounds as a summation of the total energy terms for all dimers present. This is with the exception of the contribution from polarisation which depends on the field induced by all molecules in the cluster surrounding the reference molecule, and is calculated in an additional step.[39]

Pixel lattice energies typically recover ~90-95% of experimental sublimation enthalpies.[39] Lattice energies also compare well to those calculated by intermolecular perturbation theory, as shown by Gavezzotti, and periodic DFT calculations, as shown by Wood *et al*.[44,45]

2.3 Determination of the Volume Contribution, $P\Delta V$

The volume of a unit cell is calculated from the lattice parameters determined from a diffraction pattern, and its variation with pressure established using high-pressure crystallographic data. In the present work, the role played by volume in determining the phase stability of molecular structures was studied in Chapters 3, 4 and 6. The variation of volume with pressure is modelled using the equations of state discussed in Chapter 1.

High pressure crystallography is a rapidly expanding sub-discipline of chemical crystallography and significant developments in equipment and data processing over the last 25-30 years have allowed high-pressure crystallographic studies to become a practically routine process.[46] The instrumentation used in this thesis to obtain high pressures is described below.

2.3.1 High Pressure X-ray Crystallography

High pressure crystallographic studies can use either X-ray of neutron sources. In this work, samples were investigated with neutron powder diffraction and single-crystal X-ray diffraction.

Diamond anvil cells (DACs) are the most versatile and readily available apparatus for the application of pressure.[47] DACs are capable of reaching megabar pressures[48] and are applicable to a wide range of experimental techniques including infrared, Raman, magnetisation measurements, Mössbauer and even pair distribution function measurements as well as structure determination by X-ray diffraction.[49]

DACs were first introduced by Bridgman, and generate pressure using a pair of diamond anvils opposed across a gasket hole forming the sample

chamber.[47] Measurements are made through the diamonds, utilising the highly crystalline nature and high strength of diamond, along with its transparency to a wide range of the electromagnetic spectrum.[50] The gasket is typically steel, tungsten or rhenium with a hole which has a diameter smaller than that of the culets of the diamond anvils.[46,51]

The sample is attached to the culet of one diamond in the cell. This is loaded alongside a pressure marker and pressure transmitting medium. Common pressure markers are quartz[52], with pressure measured through changes to its unit cell,[47] or ruby, which enables pressure to be measured via the fluorescence method.[53] The pressure transmitting medium ensures that pressure is applied hydrostatically to the sample and the gasket. Its choice depends on the pressure range of interest and it should not react with or dissolve the sample being studied.[47] Commonly used media are a 4:1 mixture of MeOH and EtOH, a 1:1 mixture of pentane and isopentane, and gases such as N_2, He, Ne or Ar. Pressure is applied by applying a force to the upper faces of the diamond anvils, using a clamp arrangement with screws, lever arms or inflatable membranes.[51]

The DAC restricts the pathways that incident and diffracted X-rays can take through the sample during diffraction experiments. In the case of single-crystal diffraction measurements, this limits the volume of reciprocal space that can be sampled during data collection ('shading'). Collection strategies have to be optimised to maximise the available data.[54] Use of high-energy X-ray radiation ($\lambda < 0.5$ Å) can also improve access to a larger fraction of the diffraction pattern. Multiple crystals can also be loaded into one DAC and their data sets combined. However, completeness is almost always compromised in the low-symmetry crystal systems typically adopted by molecular materials.

DACs also produce diffraction data contaminated by additional Bragg reflections due to the diamonds and sometimes the pressure marker or pressure transmitting medium, which must be taken into account.[55] DACs also suffer relatively coarse pressure control that can make it difficult to target the pressure range of interest, though this can be alleviated using membrane-control.[56]

2.3.1*a* The Merrill-Bassett Cell

The Merrill-Bassett DAC, Figure 2.5, is relatively small and lightweight, allowing single crystal diffraction experiments to be performed using a standard goniometer with minimal modifications.[46,57] These cells were used in Chapters 4 and 6 of this thesis. For use of the most common Merrill-Bassett cells the modification of a standard goniometer head, a longer beamstop arm, a shorter collimator and refocussing of the diffractometer video camera are all that is required. The standard Merrill-Bassett cell applies pressure by the operation of three screws and has an opening angle of 76°.

In this work, the diamonds used in Merrill-Bassett DACs were Boehler-Almax-cut[58] 0.6 mm culet, 1.95 mm high diamonds glued in place to tungsten carbide conical backing disks. The conical disk setup allows wider angular access to the sample chamber.[59] The two cell plates were steel, each approximately 7 mm thick and 37 mm across. Gaskets were made from tungsten sheets of thickness 0.3 mm indented by 0.2 mm. The total mass of the cell is typically around 120 g. Cells were centred iteratively using diffraction until the crystal offsets refined to within 10 μm of zero.

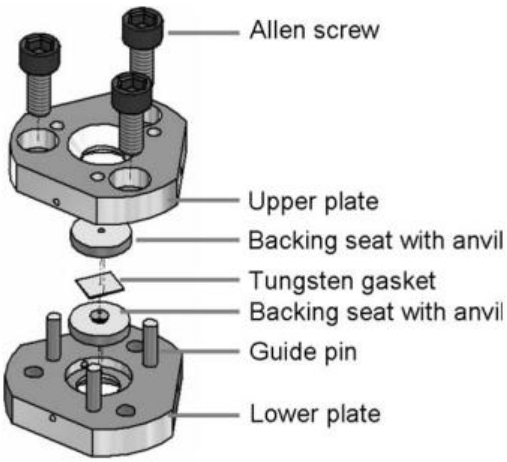


Figure 2.5: The Merrill-Bassett cell. Reproduced with permission of the International Union of Crystallography.[59]

2.3.1*b* The Short Symmetric Cell

The second DAC type used in this thesis were short symmetric cells (Figure 2.6). These cells were designed at DESY for use on beamline P02.2, but no publication describing them is yet available. Unlike the Merrill-Bassett cell, the short symmetric cell is based on a piston like design, similar to that of the commonly used BX90 cell.[60] This results in an improved control of pressure through four screws. These cells feature a slightly wider opening angle of 80˚, so fractionally more reciprocal space can be sampled in measurements. The short symmetric cell was chosen for its compatibility with gas loading equipment for loading of neon gas as the pressure transmitting medium for studies in Chapter 6.

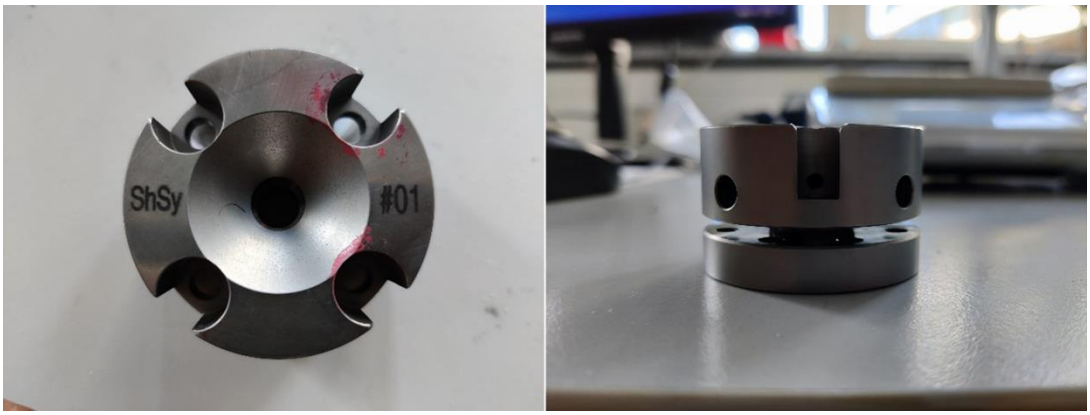


Figure 2.6: Images of the short symmetric DAC used for measurements at DESY in Chapter 6.

2.3.2 High Pressure Neutron Powder Diffraction

High pressure neutron powder diffraction was used to complement X-ray studies in Chapter 6. Neutron powder experiments were completed at ISIS neutron and muon source on the PEARL instrument.

2.3.2*a* PEARL

PEARL is a time-of-flight (TOF) neutron powder diffractometer optimised for use with a Paris-Edinburgh (PE) press (below) for the application of pressure.[61] Rather than vary the angle of diffraction while the wavelength is held fixed, TOF diffractometers fix the angle of the source, sample, and detector while irradiating with a pulsed beam of a continuous spectrum of wavelengths. The time between the pulse and the arrival of a neutron at the detector can be used to calculate its wavelength, λ, and this can then be used to calculate the d-spacings in the crystal at the fixed angle using[3]

$$\lambda = \frac{ht}{mL} = 2d \sin\theta \tag{2.33}$$

where h is Planck's constant, t is the time between the initial pulse and arrival of the neutron at the detector, m is the mass of a neutron (1.674927×10^{-27} kg), L is the source-detector distance, d is the d-spacing and θ is the source to

detector angle.[3] Detectors are situated at both transverse and longitudinal geometries and produce a usable *d*-spacing range of 0.4-4.1, 8.0-12.5 Å respectively.[61] PEARL has a total of 14 detectors, 1-9 are transverse detectors situated around θ = 90°, 10-12 are longitudinal back-scattering detectors at 107°, 127° and 147° respectively and 13 and 14 are longitudinal forward scattering detectors at 26° and 46° respectively.[62]

The PE press contains anvils that operate in a similar way to those in a DAC. There are several choices of anvil type available for the PEARL instrument, zirconia-toughened alumina (ZTA), tungsten carbide (WC) or sintered diamond,[61] which were used for the experiment described in Chapter 6. The choice of anvil is largely dictated by the pressure range of interest. The gasket consists of a null-scattering Ti-Zr capsule which contains the sample, pressure marker and pressure transmitting medium.[61] The PEARL diffractometer has an operating range of 0.5-28 GPa in pressure and 80-1400 K in temperature.[61] Pressure is achieved through hydraulic loads to the main press piston.[61]

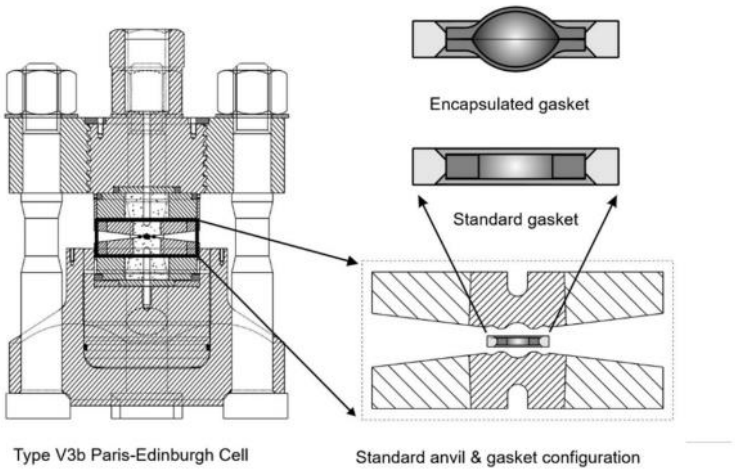


Figure 2.7: A schematic of the Paris-Edinburgh press used on the PEARL instrument at ISIS. Reproduced with permission of the International Union of Crystallography.[63]

2.4 Determination of the Entropy Contribution, $T\Delta S$

Vibrational contributions to entropy were calculated in this work using the DFPT methods introduced in Chapter 1. A more in-depth description of the phonon calculations completed this work are described below.

2.4.1 Phonon Calculations

The entropy of crystalline solids was calculated by periodic DFT in the CASTEP code using linear response (DFPT) calculations.[64] Phonon calculations can be described in terms of an expansion of a Taylor series. Consider *N* atoms in a unit cell of indices *l*, each with three degrees of freedom, collectively described by α, to describe displacement along *x*, *y* and *z* axes. The effect on energy *E* of small perturbations, *u*, of atoms in the direction α can be described by[65]

$$E = E_0 + \sum_{l,\alpha} \frac{\partial E}{\partial u_{l,\alpha}} \cdot u_{l,\alpha} + \frac{1}{2} \sum_{l,\alpha,l',\alpha'} \frac{\partial^2 E}{\partial u_{l,\alpha} \partial u_{l',\alpha'}} \cdot u_{l,\alpha} \cdot u_{l',\alpha'} + \ldots \quad (2.34)$$

where E_0 is the lattice energy of the equilibrium structure. The first derivative term describes the atomic forces of the atoms. By converging geometries to very tight levels prior to calculations this term becomes very close to zero and can be neglected.[65] In the harmonic approximation, all terms above the second order term are also neglected[66] and the resulting equation provides an approximation very similar to Hooke's law from classical physics, where the force constant *k* equates to the second derivative in Equation 2.34[67]

$$U = {}^1/_2 \, kx^2 \quad (2.35)$$

The double derivatives in Equation 2.34 can be calculated either directly using the finite displacement method or using the linear response method.

They form the elements of a Hessian or dynamical matrix, and frequencies are obtained from the eigenvalues and eigenvectors of this matrix.[65] The inclusion of anharmonicity into calculations requires the consideration of higher order terms of this expansion. These calculations however are very computationally expensive and were not applied in this work.

In crystalline materials phonon calculations are carried out at different points, referred to as **q**-points, in the Brillouin zone (the primitive cell in reciprocal space) corresponding to the manner in which the phases of vibrations in different unit cells are related to each other. The number and distribution of these points determines the quality of the calculation. For unbiased sampling, wave vectors are usually dictated using a Monkhorst-Pack grid.[65] This spaces **q**-points evenly throughout the Brillouin zone with a weighting contribution which totals to 1.

To improve computational time, calculations of reduced quality can be carried out at the Γ-point only, a single q-point at the centre of the BZ where the vibrations in all unit cells are in phase with each other.[65] These calculations represent an under sampling of the BZ and result in a reduced number of vibrational states and therefore an underestimated entropy.

A more complete sampling of the Brillouin zone can be extracted from a limited number of points by use of Fourier interpolation.[68] Although not applicable to Γ-point calculations, this method was used in Chapter 5 to recover comparable performance of lower-level calculations to those of much higher densities.

Density of states (DoS) and partial density of states (PDoS) plots described in Section 1.3.3 were used for the analysis of frequency distributions. Plots replace a list of frequencies by a distribution based on standard histogram methods.[65] The density of states, $g(\omega)$, defines $g(\omega)d\omega$ as the

number of vibrations with angular frequencies between ω and $\omega + d\omega$.[3] This distribution is valuable in identifying occupation of low energy bands responsible for contributing towards high entropies.

2.5 References

(1) Girolami, G. S., *X-ray Crystallography*. University Science Books, **2016**.

(2) Authier, A.; Chapuis, G., *A little dictionary of crystallography*. International Union of Crystallography, **2014**.

(3) Dove, M. T., *Structure and dynamics: an atomic view of materials*. Oxford University Press, **2003**.

(4) Giordano, N.; Beavers, C. M.; Kamenev, K. V.; Marshall, W. G.; Moggach, S. A.; Patterson, S. D.; Teat, S. J.; Warren, J. E.; Wood, P. A.; Parsons, S., High-pressure polymorphism in L-threonine between ambient pressure and 22 GPa. *CrystEngComm.* **2019**, 21 (30), 4444-4456.

(5) Palatinus, L.; Corrêa, C. A.; Steciuk, G.; Jacob, D.; Roussel, P.; Boullay, P.; Klementová, M.; Gemmi, M.; Kopeček, J.; Domeneghetti, M. C., Structure refinement using precession electron diffraction tomography and dynamical diffraction: tests on experimental data. *Acta Crystallogr.* **2015**, B71 (6), 740-751.

(6) Kolb, U.; Gorelik, T.; Kübel, C.; Otten, M. T.; Hubert, D., Towards automated diffraction tomography: Part I—Data acquisition. *Ultramicroscopy* **2007**, 107 (6-7), 507-513.

(7) Kolb, U.; Gorelik, T.; Otten, M. T., Towards automated diffraction tomography. Part II—Cell parameter determination. *Ultramicroscopy* **2008**, 108 (8), 763-772.

(8) Wan, W.; Sun, J.; Su, J.; Hovmöller, S.; Zou, X., Three-dimensional rotation electron diffraction: software RED for automated data collection and data processing. *J. Appl. Crystallogr.* **2013**, 46 (6), 1863-1873.

(9) Kolb, U.; Mugnaioli, E.; Gorelik, T. E., Automated electron diffraction tomography–a new tool for nano crystal structure analysis. *Cryst. Res. Technol.* **2011**, 46 (6), 542-554.

(10) Krause, L.; Herbst-Irmer, R.; Sheldrick, G. M.; Stalke, D., Comparison of silver and molybdenum microfocus X-ray sources for single-crystal structure determination. *J. Appl. Crystallogr* **2015**, 48 (1), 3-10.

(11) Blake, A. J.; Cole, J. M.; Evans, J. S. O.; Main, P.; Parsons, S.; Watkin, D. J., *Crystal structure analysis: principles and practice*. Oxford University Press, **2009**.

(12) Sorensen, T. L. M.; McAuley, K. E.; Flaig, R.; Duke, E. M. H., New light for science: synchrotron radiation in structural medicine. *Trends Biotechnolo.* **2006**, 24 (11), 500-508.

(13) Wille, K., Introduction to insertion devices. **1998**.

(14) McCormick, L. J.; Giordano, N.; Teat, S. J.; Beavers, C. M., Chemical crystallography at the advanced light source. *Crystals* **2017**, 7 (12), 382.

(15) Friedrich, W.; Knipping, P.; Laue, M., Interferenzerscheinungen bei roentgenstrahlen. *Ann. Phys.* **1913**, 346 (10), 971-988.

(16) Glazer, A. M., The first paper by WL Bragg–what and when? *Crystallogr. Rev.* **2013**, 19 (3), 117-124.

(17) Bragg, W. H., The Reflection of X-rays by Crystals.(II.). *Proc. R. Soc. Lond.* **1913**, A89 (610), 246-248.

(18) Bragg, W. L., The Specular Reflection of X-rays. *Nature.* **1912**, 90 (2250), 410-410.

(19) Robertson, J. H.; Beevers, C. A., The crystal structure of strychnine hydrogen bromide. *Acta Crystallogr.* **1951**, 4 (3), 270-275.

(20) Harrison, R. W., Phase problem in crystallography. *J. Opt. Soc. Am.* **1993**, A10 (5), 1046-1055.

(21) Willis, B. T. M.; Carlile, C. J., *Experimental neutron scattering*. Oxford University Press, **2017**.

(22) Colognesi, D.; Celli, M.; Cilloco, F.; Newport, R. J.; Parker, S. F.; Rossi-Albertini, V.; Sacchetti, F.; Tomkinson, J.; Zoppi, M., TOSCA neutron spectrometer: the final configuration. *Appl. Phys.* **2002**, A74, s64-s66.

(23) Boothroyd, A. T., *Principles of neutron scattering from condensed matter*. Oxford University Press, **2020**.

(24) Pinna, R. S.; Rudić, S.; Parker, S. F.; Armstrong, J.; Zanetti, M.; Škoro, G.; Waller, S. P.; Zacek, D.; Smith, C. A.; Capstick, M. J., The neutron guide upgrade of the TOSCA spectrometer. *Nucl. Instrum. Methods Phys. Res.* **2018**, A896, 68-74.

(25) Guthrie, M., Future directions in high-pressure neutron diffraction. *J. Phys. Condens. Matter.* **2015**, 27 (15), 153201.

(26) Enderby, J. E., Neutron diffraction, isotopic substitution and the structure of aqueous solutions. *Philos. Trans. R. Soc.* **1980**, B290 (1043), 553-566.

(27) Wright, J. P.; Besnard, C.; Margiolaki, I.; Basso, S.; Camus, F.; Fitch, A. N.; Fox, G. C.; Pattison, P.; Schiltz, M., Molecular envelopes derived from protein powder diffraction data. *J. Appl. Crystallogr.* **2008**, 41 (2), 329-339.

(28) David, W. I. F.; Shankland, K.; McCusker, L. B.; Baerlocher, C., *Structure determination from powder diffraction data*. Oxford University Press, **2006**.

(29) Coelho, A. A., Whole-profile structure solution from powder diffraction data using simulated annealing. *J. Appl. Crystallogr.* **2000**, 33 (3), 899-908.

(30) Louër, D., Powder X-ray diffraction, applications. *Encycl. Spectrosc. Spectrom.* **2017**, 723–731.

(31) Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D., Levels of symmetry adapted perturbation theory (SAPT). I. Efficiency and performance for interaction energies. *J. Chem. Phys.* **2014**, 140.

(32) Hohenstein, E. G.; Sherrill, C. D., Efficient evaluation of triple excitations in symmetry-adapted perturbation theory via second-order Møller–Plesset perturbation theory natural orbitals. *J. Chem. Phys.* **2010**, 133 (10).

(33) Dunning Jr, T. H., Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, 90 (2), 1007-1023.

(34) Kendall, R. A.; Dunning Jr, T. H.; Harrison, R. J., Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, 96 (9), 6796-6806.

(35) Bordner, A. J., Assessing the accuracy of SAPT (DFT) interaction energies by comparison with experimentally derived noble gas potentials and molecular crystal lattice energies. *ChemPhysChem.* **2012**, 13 (17), 3981-3988.

(36) Smith, D. G. A.; Burns, L. A.; Simmonett, A. C.; Parrish, R. M.; Schieber, M. C.; Galvelis, R.; Kraus, P.; Kruse, H.; Di Remigio, R.; Alenaizan, A. PSI4 1.4: Open-source software for high-throughput quantum chemistry. *The Journal of chemical physics* **2020**, *152*.

(37) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, 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.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven,

T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; 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, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., Gaussian 09, Revision E. 01, Gaussian. *Inc:, Wallingford CT* **2009,** 201.

(38) Reeves, M. G.; Wood, P. A.; Parsons, S., MrPIXEL: automated execution of Pixel calculations via the *Mercury* interface. *J. Appl. Crystallogr.* **2020**, 53, 1154-1162.

(39) Gavezzotti, A., *Molecular Aggregation*. Oxford University Press, **2007**.

(40) Wood, P. A.; Francis, D.; Marshall, W. G.; Moggach, S. A.; Parsons, S.; Pidcock, E.; Rohl, A. L., A study of the high-pressure polymorphs of L-serine using ab initio structures and PIXEL calculations. *CrystEngComm.* **2008**, 10 (9), 1154-1166.

(41) Rigby, M., *The Forces Between Molecules*. Clarendon Press, **1986**.

(42) Dunitz, J. D.; Gavezzotti, A., Toward a quantitative description of crystal packing in terms of molecular pairs: Application to the hexamorphic crystal system, 5-methyl-2-[(2-nitrophenyl) amino]-3-thiophenecarbonitrile. *Cryst. Growth Des.* **2005**, 5 (6), 2180-2189.

(43) Gavezzotti, A., Quantitative ranking of crystal packing modes by systematic calculations on potential energies and vibrational amplitudes of molecular dimers. *J. Chem. Theory Comput.* **2005**, 1 (5), 834-840.

(44) Wood, P. A.; Francis, D.; Marshall, W. G.; Moggach, S. A.; Parsons, S.; Pidcock, E.; Rohl, A. L., A study of the high-pressure polymorphs of L-serine using ab initio structures and PIXEL calculations. *CrystEngComm.* **2008**, 10, 1154-1166.

(45) Gavezzotti, A., Towards a realistic model for the quantitative evaluation of intermolecular potentials and for the rationalization of organic crystal structures. Part I. Philosophy. *CrystEngComm.* **2003**, 5 (76), 429-438.

(46) Moggach, S. A.; Oswald, I. D. H., *Crystallography under high pressures. 21st Century Challenges in Chemical Crystallography I: History and Technical Developments.* **2020**, 141-198.

(47) Collings, I. E.; Goodwin, A. L., Metal–organic frameworks under pressure. *J. Appl. Phys.* **2019**, 126 (18).

(48) Jayaraman, A., Diamond anvil cell and high-pressure physical investigations. *Rev. Mod. Phys.* **1983**, 55 (1), 65.

(49) Stan, C. V.; Beavers, C. M.; Kunz, M.; Tamura, N., X-ray diffraction under extreme conditions at the Advanced Light Source. *Quantum Beam Sci.* **2018**, 2 (1), 4.

(50) Weir, C. E.; Lippincott, E. R.; Van Valkenburg, A.; Bunting, E. N., Infrared studies in the 1-to 15-micron region to 30,000 atmospheres. *J. Res. Natl. Bur. Stand.* **1959**, A63 (1), 55.

(51) Miletich, R.; Allan, D. R.; Kuhs, W. F., High-pressure single-crystal techniques. *Rev. Mineral. Geochem.* **2000**, 41 (1), 445-519.

(52) Angel, R. J.; Allan, D. R.; Miletich, R.; Finger, L. W., The use of quartz as an internal pressure standard in high-pressure crystallography. *J. Appl. Crystallogr.* **1997**, 30 (4), 461-466.

(53) 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.* **1975**, 46 (6), 2774-2780.

(54) Dawson, A.; Allan, D. R.; Parsons, S.; Ruf, M., Use of a CCD diffractometer in crystal structure determinations at high pressure. *J. Appl. Crystallogr.* **2004**, 37 (3), 410-416.

(55) Gaspar, A. B.; Molnár, G.; Rotaru, A.; Shepherd, H. J., Pressure effect investigations on spin-crossover coordination compounds. *Comptes Rendus Chimie.* **2018**, 21 (12), 1095-1120.

(56) McMonagle, C. J.; Allan, D. R.; Warren, M. R.; Kamenev, K. V.; Turner, G. F.; Moggach, S. A., High-pressure sapphire capillary cell for synchrotron single-crystal X-ray diffraction measurements to 1500 bar. *J. Appl. Crystallogr.* **2020**, 53 (6), 1519-1523.

(57) Merrill, L.; Bassett, W. A., Miniature diamond anvil pressure cell for single crystal x-ray diffraction studies. *Rev. Sci. Instrum.* **1974**, 45 (2), 290-294.

(58) Boehler, R.; De Hantsetters, K., New anvil designs in diamond-cells. *High Press. Res.* **2004**, 24 (3), 391-396.

(59) Moggach, S. A.; Allan, D. R.; Parsons, S.; Warren, J. E., Incorporation of a new design of backing seat and anvil in a Merrill–Bassett diamond anvil cell. *J. Appl. Crystallogr.* **2008**, 41 (2), 249-251.

(60) Kantor, I.; Prakapenka, V.; Kantor, A.; Dera, P.; Kurnosov, A.; Sinogeikin, S.; Dubrovinskaia, N.; Dubrovinsky, L., BX90: A new diamond anvil cell design for X-ray diffraction and optical measurements. *Rev. Sci. Instrum.* **2012**, 83, 125102.

(61) Bull, C. L.; Funnell, N. P.; Tucker, M. G.; Hull, S.; Francis, D. J.; Marshall, W. G., PEARL: the high pressure neutron powder diffractometer at ISIS. *High Press. Res.* **2016**, 36 (4), 493-511.

(62) Ridley, C., *PEARL technical information*. 2022. https://www.isis.stfc.ac.uk/Pages/PEARL-technical-information.aspx (accessed 03/02/2024).

(63) Marshall, W. G.; Francis, D. J., Attainment of near-hydrostatic compression conditions using the Paris–Edinburgh cell. *J. Appl. Crystallogr.* **2002**, 35 (1), 122-125.

(64) Nyman, J.; Day, G. M., Static and lattice vibrational energy differences between polymorphs. *CrystEngComm.* **2015**, 17 (28), 5154-5165.

(65) Sholl, D. S.; Steckel, J. A., *Density functional theory: a practical introduction*. John Wiley & Sons, **2022**.

(66) Rivera, S. A.; Allis, D. G.; Hudson, B. S., Importance of vibrational zero-point energy contribution to the relative polymorph energies of hydrogen-bonded species. *Cryst. Growth Des.* **2008**, 8 (11), 3905-3907.

(67) Giustino, F., *Materials modelling using density functional theory: properties and predictions*. Oxford University Press, **2014**.

(68) Michalchuk, A. A. L.; Rudić, S.; Pulham, C. R.; Morrison, C. A., Vibrationally induced metallisation of the energetic azide α-NaN_3. *Phys. Chem. Chem. Phys.* **2018**, 20, 29061-29069.

Chapter 3: Behaviour of Occupied and Void Space in Molecular Crystal Structures at High Pressure.

Overview

This Chapter describes a Monte Carlo algorithm for calculation of occupied ('network') and unoccupied ('void') space in crystal structures. The variation of the volumes of the voids and the network of intermolecular contacts with pressure sensitively reveals discontinuities associated with first and second order phase transitions, providing insight into the effect of compression (and, in principle, other external stimuli) at a level between those observed in individual contact distances and the overall unit cell dimensions. Within this Chapter this method has been applied to pre-existent high-pressure data to demonstrate its utility. The method is shown to be especially useful for the correlation of high-pressure crystallographic and spectroscopic data, illustrated for naphthalene, where a phase transition previously detected by vibrational spectroscopy, and debated in the literature for over 80 years, has been revealed unambiguously in crystallographic data for the first time. Premonitory behavior before a phase transition and crystal collapse at the end of a compression series has also been detected. The network and void volumes for 129 high-pressure studies taken from the Cambridge Structural Database (CSD) were fitted to equations of state to show that networks typically have bulk moduli of between 40 and 150 GPa, while those of voids fall into a much smaller range, 2-5 GPa. These figures are shown to reproduce the narrow range of overall bulk moduli of molecular solids (*ca.* 5-20 GPa). The program, called CellVol, has been written in Python using the CSD Python API, and can be run through the command line or through the Cambridge Crystallographic Data Centre's Mercury interface.

3.1 Introduction

Use of high pressure to investigate polymorphism and the mechanical properties of different classes of intermolecular interaction is increasingly popular; Figure 3.1 shows the number of depositions on the Cambridge Structural Database (CSD) of structures determined at above 0.1 GPa.[1] As discussed in Chapter 1, an observed crystal structure represents a minimum in free energy. Under ambient conditions this corresponds to optimisation of the balance between internal energy (determined by formation of favourable intramolecular geometry and intermolecular contacts such as H-bonds) and entropy, but at high pressure the need to minimise volume becomes increasingly important. For a system of constant composition, the volume always decreases with increasing pressure and first order phase transitions occurring at high pressure are always accompanied with a decrease in volume. Volume minimisation is the dominant driving force in almost all high-pressure phase transitions, although the relief of unfavourably compressed contacts can also play a role.[2] For example, over the course of the phase transition between *L*-serine-I to *L*-serine-II that occurs above 4.8 GPa, a decrease in volume overcomes a destabilising lattice energy change.[3,4] Analysis of volume changes are therefore critical in the interpretation of phase transitions at high pressure.

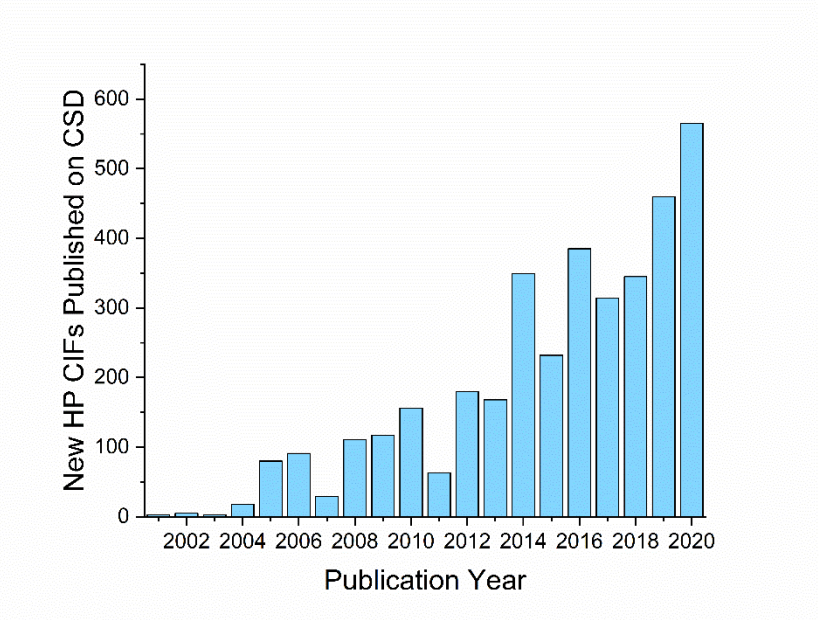


Figure 3.1: The number of annual depositions of high-pressure crystal structures in the Cambridge Structural Database (version 5.42 with updates up to and including February 2021).

The volume of a crystal structure is made up of contributions from the atoms, ions or molecules, their network of interactions and the interstitial voids. In close-packed hard-sphere structures the fraction of the volume occupied by atoms is 0.72 (the packing coefficient), with unoccupied interstitial voids comprising the remainder.[5] Kitaigorodsky recognised that the underlying topology of molecular crystal structures is also often found to correspond to close packing or body centred cubic arrangements,[6] a conclusion which has been extensively explored by Peresypkina and Blatov using formal topological analysis based on Voronoi-Dirichlet partitioning.[7] Accordingly the packing coefficients of molecular structures are broadly similar to those of hard sphere structures and usually fall into the range 0.6-0.8 at ambient pressure.[5]

Some crystal structures, such as those of metal organic frameworks, contain large void spaces capable of accommodating guest species, including gases, common solvents and even relatively large molecules such Keggin

anions in porous chromium terephthalate MIL-101 (OCUNAC, unit cell volume = 701 860 Å^3) with pore volumes up to ~20 600 Å^3.[8] At least 20% of the total unit cell volume is unoccupied at ambient pressure even in non-porous molecular crystals. For these compounds space is distributed over small interstitial sites around and between the molecules and is not accessible to guest species.

There are a number of algorithms available to evaluate occupied and void space in both classes of molecular crystal structures. Some of these use a spherical probe which is rolled over the van der Waals surface of a molecule, mapping where the surfaces of the probe and molecule meet. An alternative 'accessible surface' can be defined by mapping the position of the centre of the probe.[9,10] A topological approach has been described by Blatov[11] where voids are constructed by first partitioning the crystal structure into Voronoi-Dirichlet polyhedra (VDPs) using the atomic positions, and then reapplying the same partitioning algorithm to both the atoms and the vertices of the original VDPs. The new VDPs centred on the original vertices then define the voids. Gavezzotti introduced a method where the unit cell volume is divided into pixels of typical volume of 0.001 Å^3; the void volume is the fraction of pixels not within the van der Waals radius of an atom multiplied by the total unit cell volume.[5] Spackman and co-workers have partitioned structures into occupied and void regions using a value (typically 0.002 au) of the promolecule electron density to define molecular surfaces.[12] The methods of Blatov, Gavezzotti and Spackman sample the volume within the nooks and crannies on a molecular surface that might be missed by the rolling probe method. The methods described have been implemented in numerous crystallographic and molecular graphics programs including Mercury,

PLATON, OLEX-2, X-seed, CrystalExplorer, ATOMS, Oscail and TOPOS.[9,10,12-17]

Molecular structures often respond to high pressure initially by compression of void space, as shown by high-pressure analysis of the voids in α-D-mannose using the rolling probe method.[18] The inference is that the void space is softer than the network of inter- and intramolecular contacts and one of the aims of this chapter is to quantify this difference. A simple algorithm for partitioning the volume of a unit cell (V) into contributions from the occupied network of inter- and intramolecular interactions, V_{net}, and unoccupied void volume, V_{void}, will be described that has the advantage that it is easily combined with tools available in the CSD Python API.[1] The variations in V_{net} and V_{void} with pressure show how the intermolecular interactions and the void space combine to determine the overall compressibility of a material. Discontinuities in this behaviour can be used to detect subtle structural phase transitions and even premonitory behaviour.

3.2 Computational Procedures

3.2.1 Network and Void Volume Calculations

The aim of the computational procedure described below was to partition the volume of a unit cell (V) into regions occupied by atoms and their network of intermolecular contacts (V_{net}), and regions that consist of unoccupied void space (V_{void}). The values of the network and void volumes were evaluated in a Monte Carlo procedure, which can be illustrated by calculation of the volume of a sphere of radius r. A cube of known volume, V_{cube}, which encloses the sphere, is populated at random with a large number of n points. A point will lie within the sphere if its distance from the centre of the sphere is less than r. Hence, if there are n_s points within the sphere, the quantity $(n_s/n)V_{cube}$ converges to the volume of the sphere as n increases.

The same procedure was used to evaluate the void and network volumes in a crystal structure. Points at fractional coordinates [x, y, z] were generated from a uniform probability distribution between 0 and 1. The number of points generated, n, was of order 10^6. The distance from each point to atoms within and just beyond (see below) the edges of the unit cell was calculated. Points which were beyond the van der Waals radii of the atoms were defined as belonging to the voids, while those within the radius of any atom belong to the network of molecules and contacts. The van der Waals radii used were those given by Alvarez as described in Chapter 1.[19] If the number of points in the network volume is n_{net} then

$$V_{net} = \left(\frac{n_{net}}{n}\right) V \tag{3.1}$$

$$V_{void} = V - V_{net} \tag{3.2}$$

The input to the procedure for the calculation is a Crystallographic Information File (CIF). The first step of the calculation is to normalise X—H

distances to neutron values if the structure had been determined using X-rays. A list of atoms is generated consisting of the contents of the reference unit cell and any atoms within the van der Waals radius of the largest atom (R_{max}) of the unit cell edges. Inclusion of these atoms is necessary because a point inside the unit cell may sit within the van der Waals radius of an atom just outside. An atom was included if $|x| < R_{max}/a$, $|y| < R_{max}/b$ and $|z| < R_{max}/c$, where a, b and c are unit cell dimensions. The calculation then proceeds as described above.

The precision of the volume estimations obtained from the Monte Carlo procedure vary as $\sqrt{n}$,[20] and because the calculation is based on random numbers the values of the void and network volumes differ between individual runs. The standard deviation, $\sigma(V_{net})$, of the network volume can be estimated by performing multiple runs. If after three such runs the value of $\sigma(V_{net})/V_{net}$ is less than a target value, the calculation finishes, otherwise the calculation will continue to a user-defined maximum. The precision is defined by the 'population' standard deviation of the network volumes, rather than the standard deviation of the mean network volume. It does not decrease with the number of repeated runs but does become better defined. If the required precision has not been achieved a warning is printed and the calculation should be repeated with a larger value of n. The target value of $\sigma(V_{net})/V_{net}$ used for this work was 0.1%, and was achieved with a million points in three runs for almost all structures (see Figure A.1 in Appendix A).

The total precision of the volume estimates has two sources. One is that described above, arising from the reproducibility of the Monte Carlo calculations. The other arises from the precision of the structural parameters themselves. These can be propagated into the final volume estimates in a second Monte Carlo procedure in which multiple structural models are

generated by perturbing each atomic coordinate with independent Gaussian random deviates taken from a distribution with a mean of zero and a standard deviation equal to the coordinate standard uncertainty.[21] The network and void volume can be calculated for each perturbed structure generating a distribution. The mean and standard deviation of this distribution, which reflects the scatter obtained from both Monte Carlo procedures, are then taken as the network volume and its standard deviation.

In this work the standard deviations of the volumes quoted for the structures in Section 3.3.3 were generated using 100 perturbed sets of coordinates at each pressure point. The structures analysed in Sections 3.3.1 and 3.3.2 are the results of DFT optimisations, for which no coordinate standard uncertainties are available, and in these cases the precision is based solely on the spread of the volume estimates from a single set of coordinates. When the number of points used in the Monte Carlo volume calculations is chosen to give $\sigma(V_{net})/V_{net} = 0.1\%$, the spread of values obtained with coordinate error propagation is not very different from that obtained from a single structure, suggesting that the spread is dominated by the reproducibility within the individual Monte Carlo volume calculations. If the number of points is chosen to reduce $\sigma(V_{net})/V_{net}$ significantly below 0.1%, or if low precision structures are analysed, then the propagated error would be expected to become more significant.

Our approach for volume calculations is similar to the algorithm outlined by Gavezzotti for calculation of molecular volumes. In Gavezzotti's approach the unit cell volume would have been divided into pixels with a volume in the region of 0.001 Å^3 and each one classified according to whether it was or was not within the van der Waals radius of an atom. This calculation is more efficient than our Monte Carlo method for small unit cells, but becomes

slower for large unit cells. For example, a unit cell with a volume of 5000 Å^3 would require calculations for 5 million pixels, as opposed to multiple Monte Carlo runs of 1 million pixels, which are often found to complete in three runs. The Monte Carlo method also enables iteration enabling a target precision to be defined. The use of dimensionless points instead of pixels (or a rolling probe) may also sample small features on molecular surfaces.

The algorithm has been implemented in a program called CellVol, written in Python using the NumPy library and functions available in the CSD Python API. Investigations were made into speeding up the calculation by basing them on the asymmetric unit rather than the unit cell. The CSD Python API does not currently include a function for identifying the coordinate limits of the asymmetric unit of a space group, but testing the algorithm using the CrysFML Fortran library,[22] which does have this feature, did not lead to a sufficient increase in speed or precision to make consideration of space group symmetry worthwhile. The majority of structures studied at high pressure have unit cell volumes of less than 1000 Å^3, and the calculation time shows no appreciable dependence of volume of the unit cell in this range. At higher volumes there is a roughly linear dependence. Calculations were shown to complete within 5 minutes for 97% of structures in the CSD using a modest desktop personal computer with an Intel® Core™ i7-9700 CPU with a base speed of 3.00 GHz (see Appendix A, Figure A.2).

The program can be invoked from within Mercury or the command line for both CSD entries and user-supplied CIFs and may be used for single or multiple structures. The values of n, the maximum number of runs and the required precision can be set by the user via a GUI or as command line arguments. The code is available and maintained in open-source format from https://github.com/CwilsonEd/CellVol.

3.2.2 Calculation of Network and Void Bulk Moduli for Compression Studies in the CSD

Network and void bulk moduli were calculated using the EoSFit program by equation of state (EoS) fitting of the variation of void or network volumes with pressure[23] for compounds in the CSD for which variable pressure data are available, the list being taken from the study by Giordano *et al*. 2019.[24] The form of Birch-Murnaghan and Vinet EoS used in this fitting are described in Chapter 1. Ambient pressure reference structures were added to the high-pressure refcode sets where necessary. Structures which contained no coordinates, had unhandled errors, or contained disorder were removed. Refcode sets with less than five members were also removed. The final dataset contained 1472 separate refcodes in 129 refcode families. Systematic differences were often seen when several separate studies were combined, and where possible data were taken from contiguous sets of measurements from a single investigation. Network and void volumes were scaled to *Z* (the number of formula units per unit cell) to allow a comparison across phase transitions where *Z* changes. The results of these calculations are given in Table 1 in the electronic supplementary material.

3.2.3 Periodic DFT Calculations

High-pressure crystal structures of naphthalene were optimised using periodic Density Functional Theory (DFT) using the CASTEP program[25] in order to remove geometric variation related to instabilities in Rietveld refinements. The published structures were used for starting coordinates with the unit cell parameters and space group held fixed to those that were determined experimentally. A basis set cut-off of 950 eV was used with a

Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional with 'on the fly generated' (OTFG) pseudopotentials embedded in the program. The **k**-point spacing was 0.08 $Å^{-1}$. These parameters converged the total energy to less than 0.1 meV per atom. For the geometry optimisation, a tolerance of 5 $\times 10^{-6}$ eV $atom^{-1}$ was used for energy convergence with a maximum force tolerance of 0.01 eV $Å^{-1}$ and a maximum displacement tolerance of 5×10^{-4} Å.

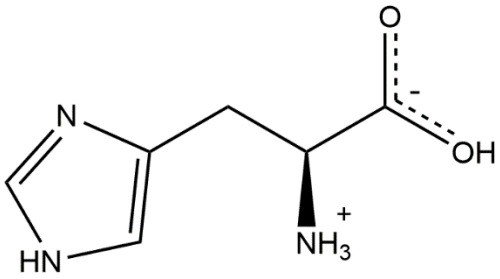
3.2.4 Pixel Calculations

Intermolecular interaction energies in naphthalene were calculated with the Pixel method[5,26,27] using the MrPixel interface.[28] Gaussian-09[29] was used to calculate the electron density at the MP2 level of theory with the 6-31G** basis set. The molecular electron density was calculated on a grid of 0.08 × 0.08 × 0.08 $Å^3$ and a condensation level equal to 4 was used for the Pixel calculations out to a cluster radius of 14 Å. The total energy of each of the contacts was taken as the sum of the electrostatic, polarisation, dispersion, and repulsion interactions.

3.3 Results

3.3.1 Detection of First Order Transitions in L-histidine

The amino acid *L*-histidine (Scheme 1) has two stable ambient pressure polymorphs, an orthorhombic ($P2_12_12_1$) and a monoclinic ($P2_1$) form, both of which were recently studied to between 6 and 7 GPa.[30] High-pressure phase transitions were identified for both, at 4.5 GPa for the orthorhombic form and at 3.1 GPa for the monoclinic form. The space group symmetry was preserved in both transitions, which were driven by a reduction in volume. Despite being shown to have remarkably similar ambient pressure interaction energies, crystal packing and molecular conformations, the two polymorphs were shown to be different in their response to compression. Prior to the phase transition in the monoclinic form chains formed by N—H···O interactions shear relative to one another, but the same compression mechanism cannot occur in the orthorhombic form without a change in space group symmetry.



Scheme 1: The molecular structure of L-histidine.

Volume and packing energy analyses in the original study were based on experimentally determined coordinates which had been optimised by periodic DFT. The same coordinates were used in the present work, with the results shown in Figures 3.2*i-iv*. For the orthorhombic form, the network shows a small increase of 0.05 Å^3 per molecule at the phase transition at 4.5

GPa (Figure 3.2*i*) which is compensated by the reduction of the void volume of 3.84 Å^3 per molecule (Figure 3.2*ii*). This increase in the network volume signals a rearrangement at the transition which enables the molecules to make more efficient use of the void space. Fitting of the network volume using second and third order Birch-Murnaghan EoSs before and after the phase transition reveals a reduction in bulk modulus following the transition from 121(5) to 83(5) GPa. The 22.2% change in void volume at the transition compares to a unit cell reduction of only 2.6%, showing that discontinuities can be considerably larger in the component (*i.e.* network or void) volumes than in the total volume.

For the monoclinic form the change is more substantial. At the phase transition (3.1 GPa) the network volume undergoes a discontinuous drop (Figure 3.2*iii*). This reduction in the network accounts for 0.78 Å^3 per molecule, accompanied by a 3.14 Å^3 per molecule reduction in the void space. The reduction of volume in the void (Figure 3.2*iv*) appears to be the driving force for this transition with a reduction of 15.4% of the void on transition compared to the 2.6% reduction in the total unit cell volume. A Vinet EoS fit for the void volume before and after the transition reveals an increase in void bulk modulus from 2.9(2) to 9(3) GPa. The change in the network bulk modulus is harder to discern because the trend shown in Figure 3.2*iii* after the transition does not follow the typical functional form of a Birch-Murnaghan or Vinet EoS, but we estimate that it increases marginally from 101(9) to 139 GPa by using a 3rd order Birch-Murnaghan EoS before the transition and Equation 3.3 afterwards.[30]

$$K = -\tilde{V} \times \frac{\Delta P}{\Delta V} \tag{3.3}$$

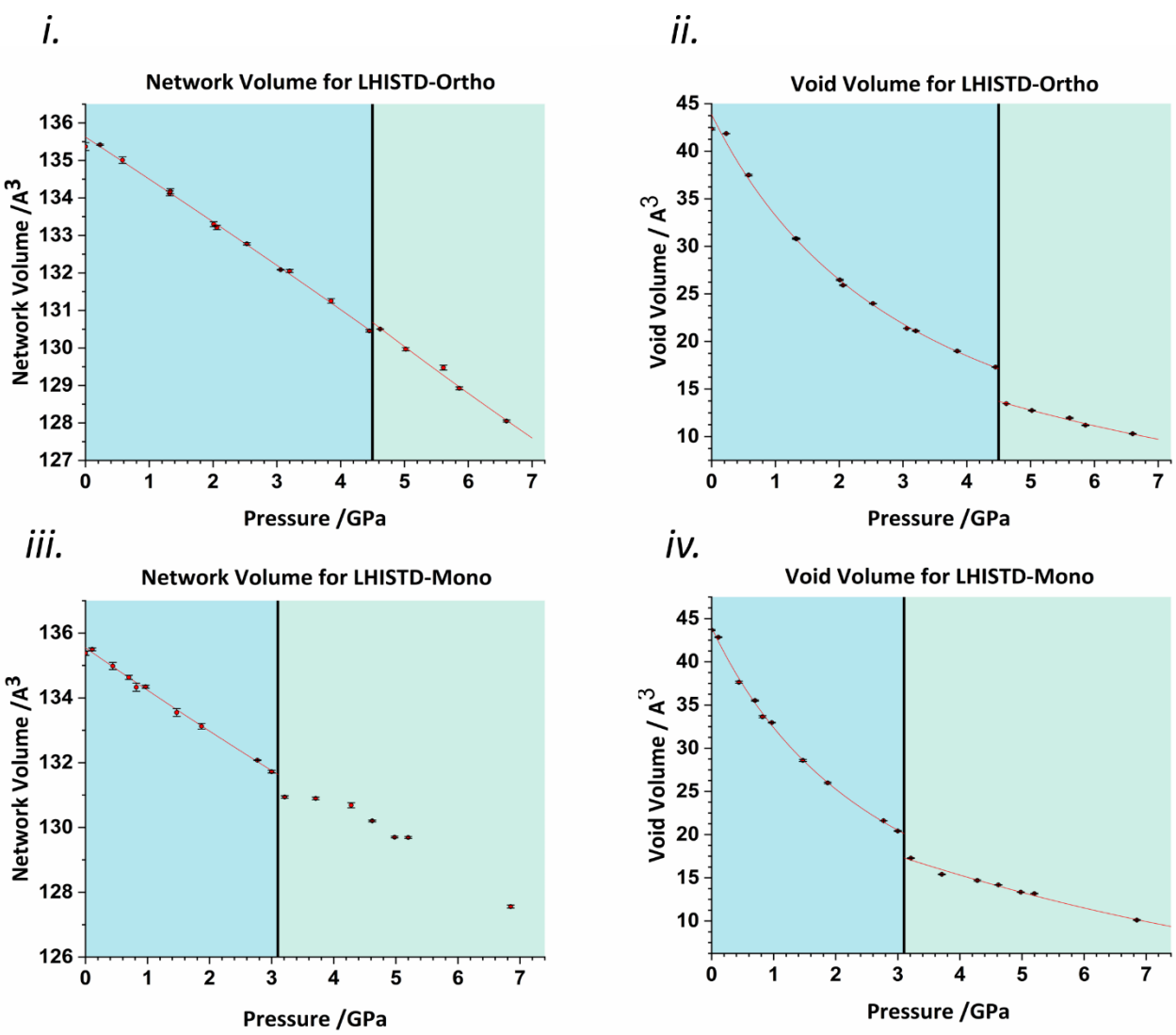


Figure 3.2: Void analysis for L-histidine i. Form I network volume against pressure, a 3rd order Birch-Murnaghan EoS is fitted before the transition and a 2nd order Birch-Murnaghan EoS is fitted afterwards. ii. Form I void volume against pressure, 3rd order Vinet EoSs are fitted before and after the transition. iii. Form II network volume against pressure, a 3rd order Birch-Murnaghan EoS is fitted before the transition. iv. Form II void volume against pressure , 3rd order Vinet EoSs are fitted before and after the transition. Black vertical lines mark phase transitions.

Despite the similarity of the geometries and energies of the intermolecular interactions in the two polymorphs, the way in which the networks and voids interact to minimise volume is different. These results

correct the analysis presented by Novelli *et al*. 2020,[30] where both transitions were described as occurring with an increase in network volume and premonitory behaviour detected in the monoclinic form. The trends reported previously were the result of a programming error.

3.3.2 Detection of Subtle High-Pressure Second Order Transitions

Compression studies of molecular materials using Raman and other forms of spectroscopy can reveal structural rearrangements which are difficult to detect in crystallographic data. The existence of a phase transition between 2 and 4 GPa in naphthalene has been the subject of debate in the literature for over eighty years owing to the importance of polycyclic aromatic hydrocarbons in oil, tar and coal deposits in the Earth's crust and the relevance of their mechanical behaviour in modelling the geophysics of planets and their moons in the Solar System.[31] An 'unmistakable' but 'sluggish' transition was first identified by Bridgman in 1938 using volumetric measurements to 3.0 GPa.[32] A study by Block *et al*. 1970 detected this transition in the range 2-3.5 GPa through optical observation of a sample in a diamond anvil cell (DAC).[33] However, this was followed by a volumetric study by Vaidya and Kennedy in 1971 and a Raman study by Nicol *et al*. in 1975 which both found no evidence of a structural phase transition up to 4 and 3.6 GPa, respectively.[34,35] An infrared study by Hamann *et al*. 1978 provided support for the hypothesis of a sluggish transition by demonstrating a discontinuous shift in all bands in the range 2-4 GPa.[36] A Raman study over the same pressure range by Meletov *et al*. 2013, proposed that discontinuous changes in relative phonon frequencies signalled a phase transition near 3.5 GPa.[37] Most recently, a study by O'Bannon

and Williams using both Raman and infrared spectroscopy confirmed a subtle transition at 2-3 GPa,[31] the greater sensitivity of the infrared data when compared to the Raman data in this paper being suggested as an explanation for the lack of transition observed in Nicol's study. The same study identifies a further phase transition above 15 GPa which involves dimerisation or polymerisation of the molecules.

The CSD contains two crystallographic studies on naphthalene.[38,39] The first of these studies (NAPHTA19-22) was a single-crystal X-ray study by Fabbiani *et al*. 2006 which features four high-pressure points to 2.1 GPa;[38] no transitions were observed, as would be expected from the spectroscopic data. The second was a synchrotron powder study (NAPHTA39-48) by Likhacheva *et al*. in 2014 which detected anisotropic compression but no transitions up to 5.60 GPa.[39] These data were also included in a later study by the same authors, extending the pressure limit to 13 GPa with all data fitted to a single EoS.[40]

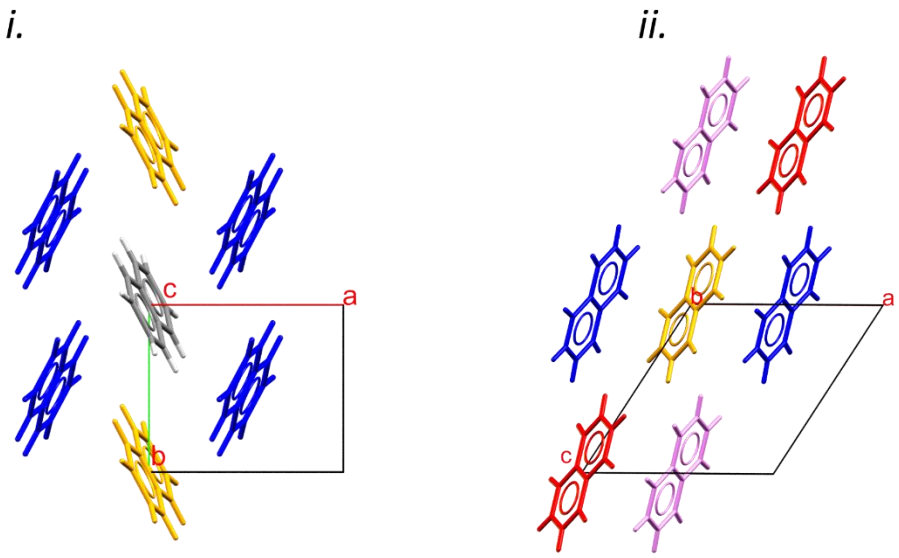


Figure 3.3: The crystal structure of naphthalene. i. Viewed down the c axis. ii. Viewed down the b axis. Molecules are colour coded to match Pixel contact graphs below with respect to the central grey molecule.

The structure of naphthalene at ambient pressure (NAPHTA47) can be viewed as consisting of layers formed in the *ab* planes (Figure 3.3*i*). Within the layers each molecule interacts with six nearest neighbours. Four of these interactions are through a set of equivalent C—H···π contacts (Figure 3.3*i*), contact #1; shortest H···C distance = 2.84 Å, 'herringbone' angle between the planes of the naphthalene molecules = 51.51°, centroid-centroid distance = 5.099 Å, energy −17.2 kJ mol^{-1} according to the Pixel method). Two longer contacts (#2 in Figure 3.3*i*) form in which the molecules are related by lattice translations along *b*. In the two longer contacts, the H-atoms of one molecule lie above the C-atoms in the other, but the displacement is too large to describe it as a 'stacking interaction', and it is probably best regarded as a non-specific dispersion interaction.[41] It is, nevertheless, almost as strong as contact #1 (centroid-centroid = 5.984 Å, energy = −15.6 kJ mol^{-1}). The layers stack along *c*, each molecule interacting with three molecules in the layers above and below, featuring close H···H contacts (Figure 3.3*ii*, shortest H···H distances = 2.40 Å, two equivalent contacts labelled #3, and one unique contact labelled #4, centroid-centroid = 7.913 and 8.675 Å, energies −7.6 and −6.4 kJ mol^{-1}, respectively), to give an overall molecular coordination number of 12. The stacking follows an ABC sequence, and the structure topologically resembles cubic close packing.

Analysis of effect of pressure on the crystal structure of naphthalene revealed that after initial compression of the network volume by 2.5 Å^3 between 0 and 1 GPa, the trend levels off between 1 and 2 GPa, followed by a collapse above 2 GPa (Figure 3.4*i*). This trend was eventually traced to a parallel decrease in the volume of the naphthalene molecules themselves rather than in the intermolecular contacts, for example C–C bond distances span a range of 1.385(4) – 1.448(3) Å at ambient pressure and 1.339(6) – 1.400(7)

at 5.60 GPa. While changes in covalent distances are possible with pressure, in other structures[30,42] they tend to be of the order of 0.02 Å at around 7 GPa and it appears that change in the bond distances and molecular volume are related to instabilities in the Rietveld refinements used in structure analysis. The crystal structures were therefore optimised using periodic DFT. Application of volume analysis to the optimised structures showed a more subtle, but nevertheless clear change in the gradient of the network volume (Figure 3.4*ii*) with the points between 0 and 2 GPa lying on one line, and those between 2.65 and 5.60 GPa lying on another, steeper, line. Fitting of the network volume using a 2nd order Birch-Murnaghan EoS before and after 1.90 GPa revealed a reduction in the network bulk modulus from 125(4) to 82(4) GPa. The signs of the transition are less obvious in the void and total volumes (Figure 3.4*iii* and *iv*). The void volume can be fairly well fitted with a single 3rd order Vinet EoS (K_0 = 2.1(1) GPa K' = 0.88), though there is some scatter in the fit. The trends indicate that the phase transition involves the network assuming a greater role in accommodating pressure. The onset of additional modes of compression is comparable to a second order thermal phase transition such as a glass transition in a polymer described in Chapter 1, in which the heat capacity increases as new vibrational modes become available on heating.

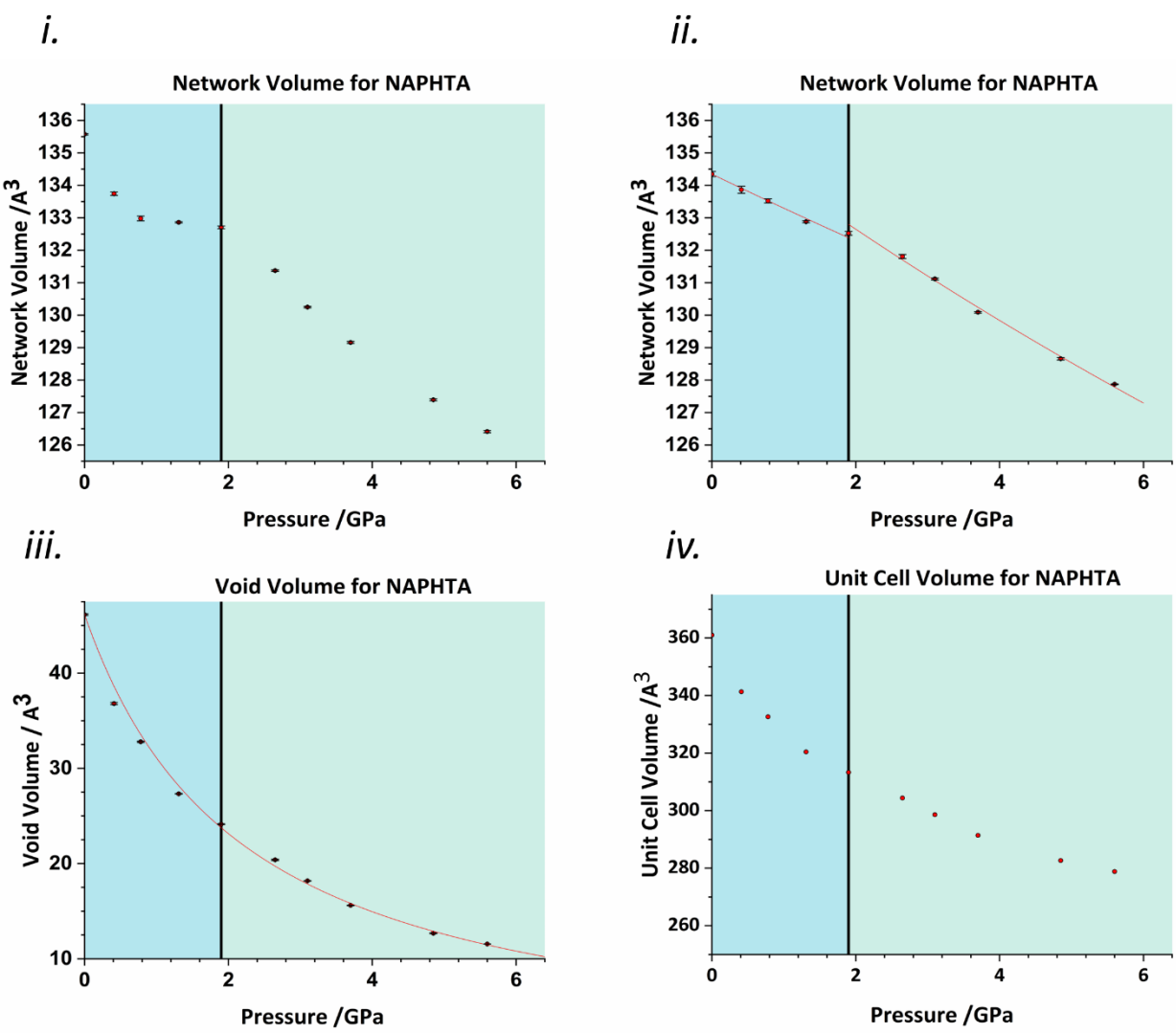


Figure 3.4: i. Network volume plot derived using data from the high-pressure powder diffraction study of naphthalene by Likhacheva et al.[39] ii. Network volume plot derived using DFT-optimised data. 2nd order Birch-Murnaghan EoSs have been fitted before and after 1.90 GPa. iii. Void volume plot using DFT-optimised data. A 3rd order Vinet EoS has been fitted across the full pressure range. iv. Unit-cell volume as a function of pressure. Black vertical lines mark the discontinuity discussed in the text.

It is possible to analyse the origin of the network volume changes by calculating the volume of isolated structural fragments as a function of pressure (Figure 3.5*i*, volumes have been offset vertically for the sake of clarity). The volume of an individual molecule (calculated as described in ref[30]) decreases linearly by only ~0.5 Å^3 up to 5.60 GPa, with the discontinuity seen

in the experimental volume no longer present. When the calculation is applied to a fragment of a layer consisting of a central reference molecule and the six molecules connected to it by contacts #1 and #2, a discontinuity in the gradient is observed at the same pressure as in the overall network volume. The volume of the set of five molecules connected by C—H···π contacts is approximately linear, while that of the three molecules connected by contact #2 shows a change in gradient above 2 GPa. A calculation applied to a fragment consisting of two layers shows no new features. Since the van der Waals surfaces of the different fragments overlap, the sum of the fragment volumes shown in Figure 3.5*i* is not equal to the network volume, but the results suggest that the phase transition has its origin in the structure of the layers, particularly in contact #2. This conclusion is in contrast to the original analysis by Likhacheva *et al*, who suggested that the anisotropic compression was associated with suppression of compression of interlayer C···C contacts.

Contact #2 is a non-specific dispersion contact, and the lack of any characteristic geometrical constraint (such as an optimal herringbone angle in a C—H···π contact) makes the interaction very flexible. This feature can be seen in the interaction potentials calculated for the four contacts as a function of pressure using the Pixel method (Figure 3.6). All four contacts are pushed into progressively less stabilising regions of their potentials between ambient pressure and 5.60 GPa. Destabilisation occurs for contacts #1 and #3 first, suggesting that optimisation of these directs the crystal structure of naphthalene. The effect is smallest for contact #2, showing that compression of this interaction represents a low energy pathway for accommodating increased pressure. The change in energy in the contacts is still clearer in an energy difference framework (Figure 3.5*ii* and *iii*),[43] where the struts represent contact energy differences after compression to 5.60 GPa: contact #2 not only

has the thinnest strut, it is also the only contact which is more stable at 5.60 GPa than it was at ambient pressure.

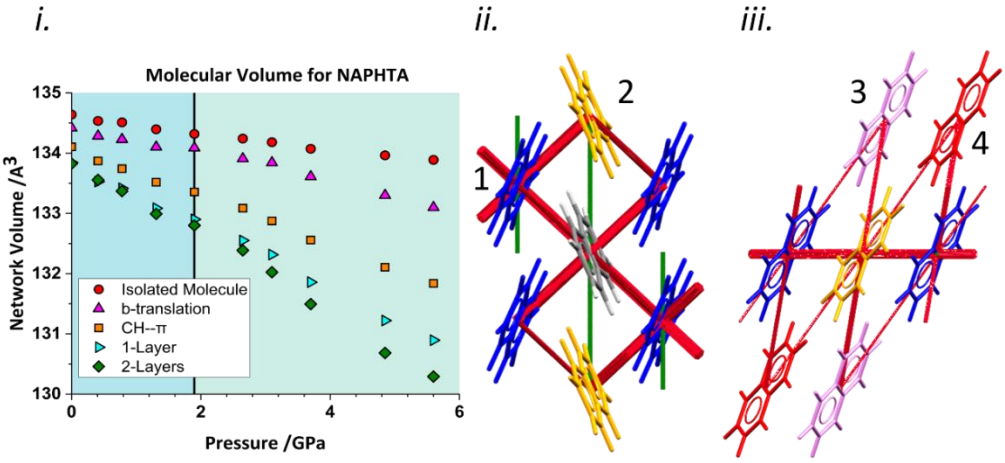


Figure 3.5: i. Applications of the CellVol code to molecular subsets in the crystal structure of naphthalene. Data have been scaled to the number of molecules in the subset. For clarity, the isolated molecule is offset by +0.25 Å^3, the CH---π plot is offset by −0.25 Å^3 and both the single and double layer are offset by −0.5 Å^3. Black vertical line marks the discontinuity. ii. and iii. Energy difference frameworks for the DFT-optimised crystal structures viewed down **c** and **b**, respectively. Strut thickness represents energy differences between ambient pressure and 5.6 GPa. Green and red respectively represent stabilisation and destabilisation on increasing pressure.

The points in the curves in Figure 3.6 are shown as diamonds and circles below and above 2 GPa, respectively, and all four curves show an upward trajectory in energy after the transition, which correlates with the loss of void space as atoms are pushed within the sums of the van der Waals radii (Figure 3.7[44]). The upward trend appears at slightly lower pressure in the case of contact #1, reflecting the hardening of the herringbone angle identified by Likhacheva *et al.* and the step in the trend of the angle made between the long axis of the molecule and the *a* axis immediately prior to the transition (Figure 3.8). The destabilisation in contact #1 is the result of a marked increase in Pauli repulsion, which increases from 12.5 to 68.4 kJ mol^{-1} over the pressure range.

It seems that the transition occurs as the result of the hardening of contact #1, which signals the need to change the mechanism of compression.

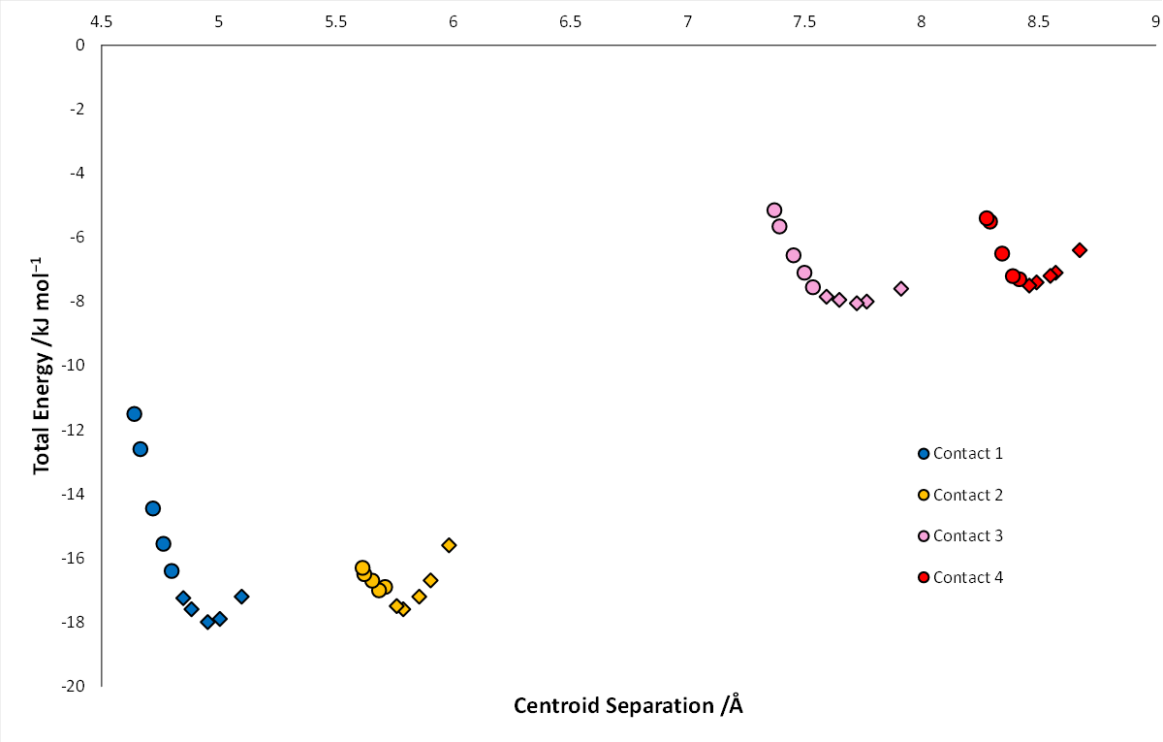


Figure 3.6: Contact energy versus centroid-centroid distance for contacts within the first coordination sphere for the DFT-optimised crystal structure of naphthalene. The colours match those in Figure 3.3. Points before 2 GPa are represented by diamonds, those after are circles.

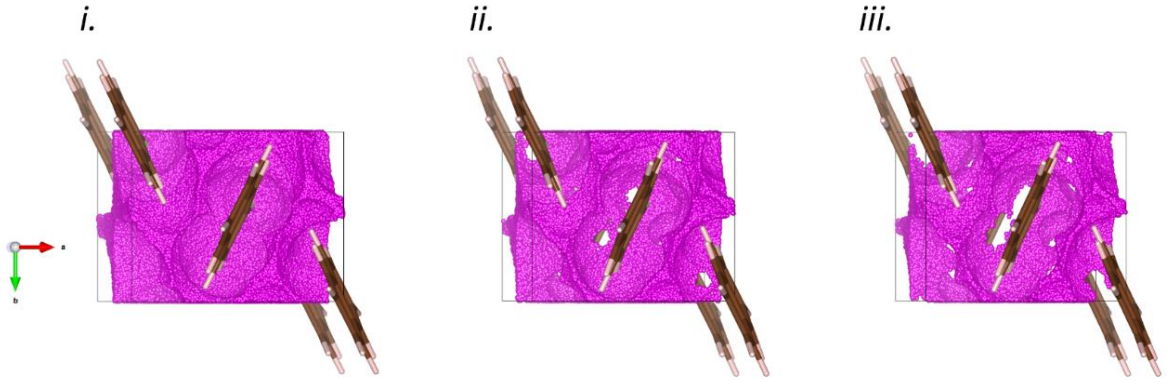


Figure 3.7: Visualisation of void volume calculated with the CellVol code using the VESTA program at i. 1.31, ii. 1.9 and iii. 2.65 GPa. The voids are shown in pink, and white regions correspond to areas where van der Waals surfaces of atoms overlap. Interlayer void volume decreases across the pressure range, disappearing at around 2 GPa.

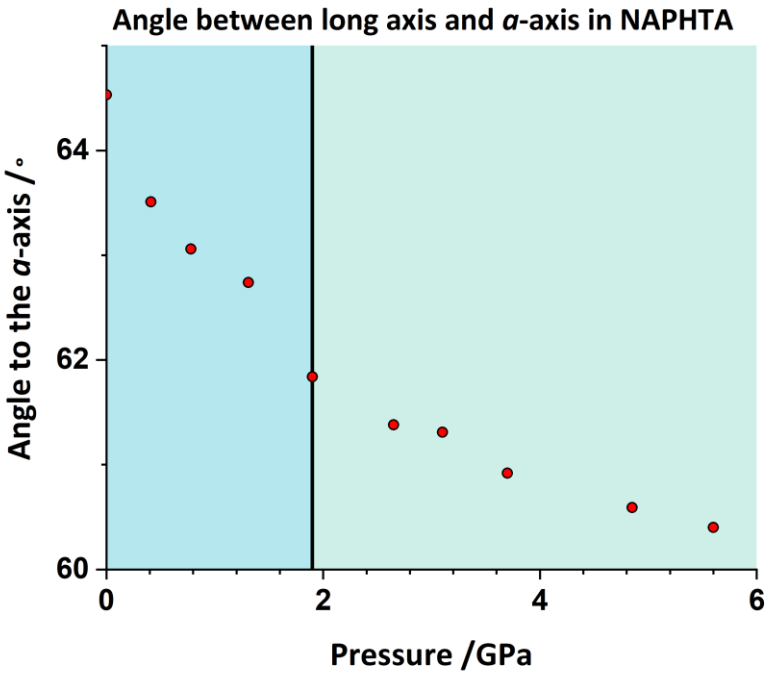


Figure 3.8: The angle between the long molecular principal axis and the *a* unit-cell axis in naphthalene. Values calculated using PLATON. The black vertical line marks the discontinuity.

3.3.3 Premonitory Behaviour

In [Cu(L-Asp)(H_2O)$_2$] (LEFJAH, Figure 3.9, Asp = aspartate) aspartate ligands bridge Cu(II) ions, binding at one end in bidentate fashion through the amine and one oxygen of the α-carboxylate group, and at the other through one oxygen atom of the β-carboxylate group to form a one dimensional coordination polymer.[45] The square based pyramidal geometry of the Cu centres is completed by two water ligands, one of which forms in the axial direction of the pyramid. The Cu-N and Cu-O bond distances in the equatorial plane are all below 2 Å, but the distance to the axial water at ambient pressure (LEFJAH01) is 2.311(2) Å. In addition, the second oxygen atom of the β-carboxylate forms a distant interaction [2.925(2) Å] with the Cu, making an angle of 143.16(7)° with the axial water. The Cu···O distances along the axial direction begin to compress above 0.3 GPa, the longer reaching 2.883(6) Å at 0.9 GPa. This was interpreted in terms of a conversion of a long interaction into a primary coordination bond, changing the Cu coordination from [4+1] to [4+2].[45] At 6.8 GPa the shorter axial distance is 2.234(10) Å and the longer 2.662(13) Å. On increasing the pressure 7.9 GPa, a phase transition occurs in which the two distances shorten discontinuously to 2.02(2) Å and 2.57(2) Å, respectively.

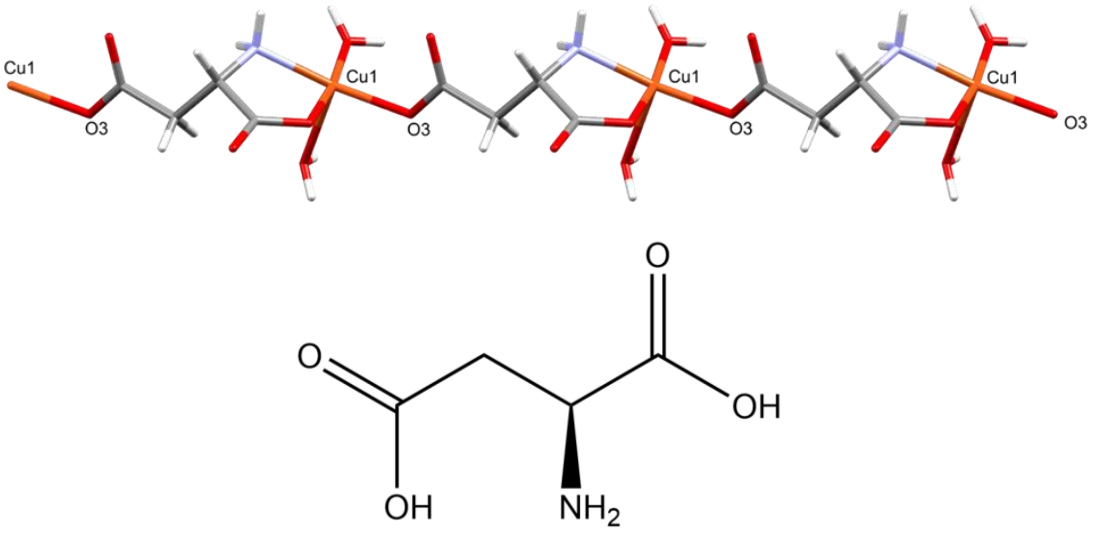


Figure 3.9: Top, the one dimensional coordination polymer structure of Cu(L-Asp)(H_2O)$_2$] (Asp = aspartate, LEFJAH). Blue: nitrogen, grey: carbon, white: hydrogen, red: oxygen and orange: copper. Bottom, L-aspartic acid that forms the aspartate ligand through loss of hydrogens from both OH groups.

The phase transition produces discontinuities in the unit cell dimensions, but not very obviously in the unit cell volume (Figure 3.10*i*), as an increase in the *a* and *b* dimensions are compensated for by a compression of *c*. The transition is clearly visible in the network and void plots (Figure 3.10 *ii* and *iii*). The point at 7.9 GPa in the void volume is at *higher* volume than expected on the basis of a Vinet EoS applied to the points between 0 and 6.8 GPa. At the same time, the network volume decreases by 1.7 Å^3 per molecule. The transition can therefore be understood in terms of more compact bonding in the network. The bulk moduli of the network and voids between 0 and 6.80 GPa are 38(11) and 6.7(4) GPa, respectively, and the abrupt halt in the reduction in void volume at the transition implies that, as in the case of naphthalene there is a point at which the mechanism of compression switches from the voids to the network. In addition to stepwise compression in the axial Cu-O bonds, Gould *et al.*[45] identify shortening of O—H···O H-bonds which

connect the polymer chains and a conformational change in the chelate ring which allows compression along the polymer chain as mechanisms through which the network compresses after the phase transition.

The network volume calculations also reveal some structural instabilities before the phase transition, suggesting premonitory structural effects in the bonding. This is shown in Figure 3.10*iii*, in which a 3rd order Birch-Murnaghan EoS is seen to fit the points between 0 and 4.92 GPa, but not those at 5.90 and 6.77 GPa, which fall below the line. In other structures, for example 3-fluorosalicylaldoxime (MIGPAU), similar features in the network volume (Figure 3.10*iv*) occur at the end of the pressure series before a destructive phase transition or loss of long-range order which prevented collection of further diffraction data.[46] A recent paper on the study of 4-methylpyridine pentachlorophenol (GADGUN) associated a similar feature with a destructive transition which occurred on further increase in pressure.[47] Detection of premonitory behaviour will aid the interpretation of why transitions take place in some compounds but not others, and for rationalising the limit of compression while maintaining long-range order.

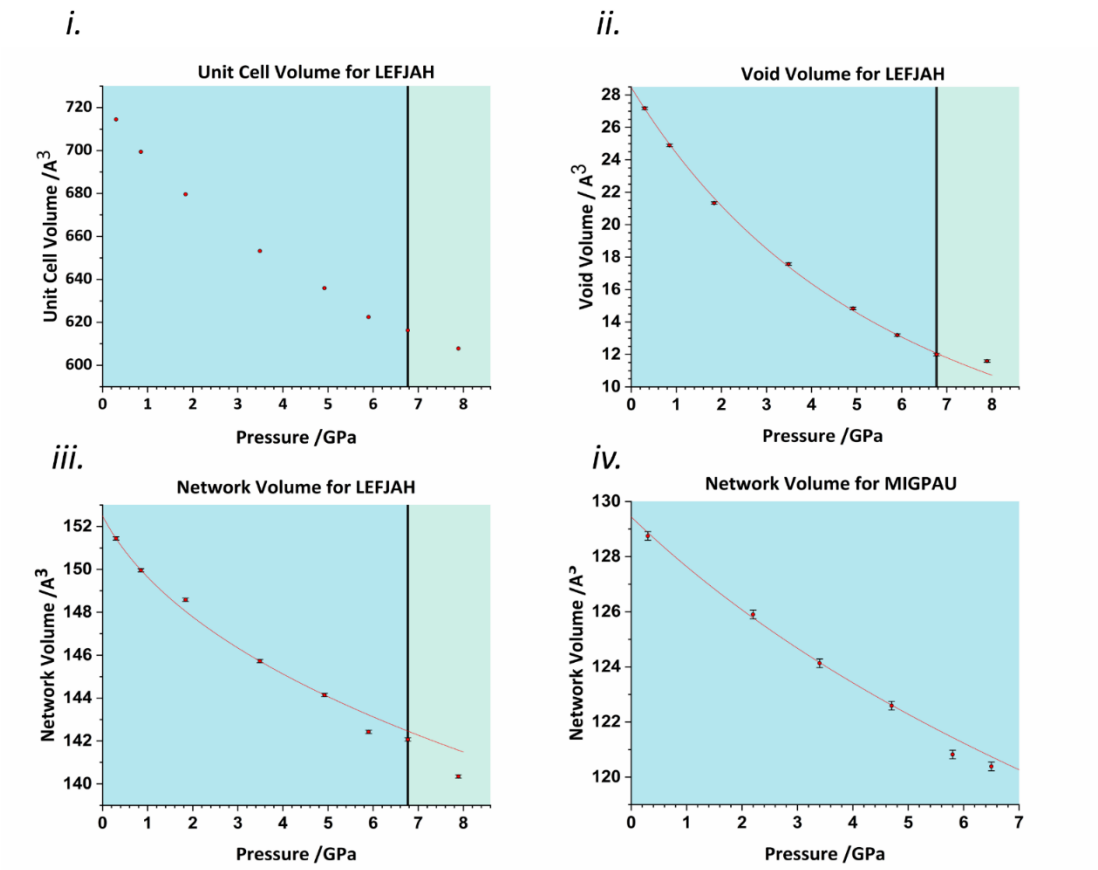


Figure 3.10: i. Unit cell volume of LEFJAH against pressure. ii. Void volume of LEFJAH against pressure with a 3rd order Vinet EoS fitted to all but the last point. iii. Network volume of LEFJAH against pressure with a 3rd order Birch-Murnaghan EoS fitted to the first five points. iv. Network volume against pressure in MIGPAU with a 3rd order Vinet EoS fitted to the first four points. This provided a slightly improved statistical fit over 2nd or 3rd order Birch-Murnaghan EoS, the more common choice for network plots. Black vertical lines mark the phase transition.

3.4 General Trends in Packing Coefficients and Bulk Moduli

3.4.1 Variation of Packing Coefficients

The packing coefficient x of a crystal structure measures the fraction of the unit cell which is occupied by atoms, and is readily obtained from the network volume,

$$x = \frac{V_{net}}{V} \tag{3.4}$$

The packing coefficients of all 1472 structures identified in this study are plotted as a function of pressure in Figure 3.11. At ambient pressure, the distribution of x mostly falls between 0.6 and 0.8, which is typical for molecular crystals at ambient pressure (see Section 3.1).[5] The values increase with pressure, tending towards 1 at above 10 GPa as void space is compressed. This is expected, but the implied total loss of void space reflects the use of constant values of van der Waals radii determined from ambient pressure crystal structures and may obscure a high degree of overlap between molecular van der Waals surfaces in some regions of a structure, but less in others. The narrowing with increasing pressure is also a consequence of the importance of the pressure-volume contribution to the free energy which places a premium on efficient packing as pressure increases.

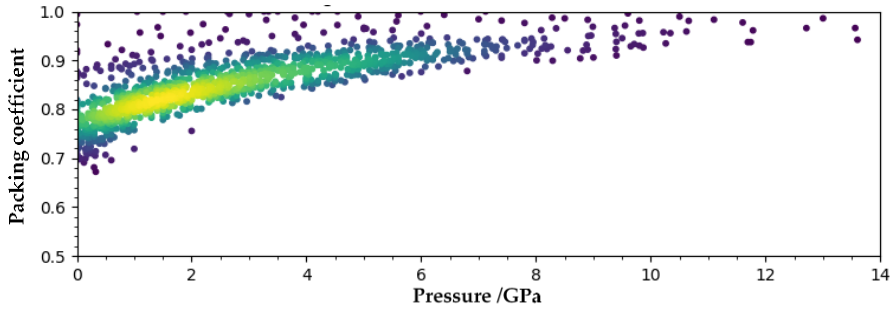


Figure 3.11: Packing coefficients as a function of pressure for molecular crystal structures.

3.4.2 Contributions of Networks and Voids to Overall Compressibility

The bulk modulus (see Chapter 1) is a measure of the compressibility or hardness of a material. It has units of pressure and some typical values for H-bonded molecular solids are 14.0(5) and 11.6(6) GPa for *L*-histidine I and II respectively[30] and 13.1(6) GPa for *L*-alanine.[48] Values for van der Waals solids usually fall below 10 GPa, e.g. benzene, bianthrone and $Ru_3(CO)_{12}$ with bulk moduli 5.5(7), 8.1(5) and 6.6 GPa respectively.[49-51] Ionic salts such as NaCl and CaF_2 have values of 25 and 82.0(7) GPa,[49,52] and those for moderately hard metals such as scandium and titanium are 57 and 110 GPa, respectively.[53] The bulk modulus of diamond is 445 GPa.[54] The pressure derivative, *K*′, is dimensionless and has typical values between 4 and 10 for molecular solids.

The examples discussed above show that the magnitudes of compression exhibited by voids and networks are very different. The contributions made by the voids and network to the compression of the overall volume can be analysed by recognising that since the total volume $V = V_{net} + V_{void}$, where V_{net} and V_{void} are the network and void volumes, then

$$dV = dV_{net} + dV_{void} \tag{3.5}$$

Substituting Equation 3.5 into the inverted form of the bulk modulus (Equation 1.3) yields Equation 3.6.

$$\frac{V}{K} = -\frac{dV_{net}}{dP} - \frac{dV_{void}}{dP} \tag{3.6}$$

Multiplication of the first and second terms on the right-hand side respectively by V_{net}/V_{net} and V_{void}/V_{void} gives Equation 3.7:

$$\frac{V}{K} = V_{net}\left(-\frac{1}{V_{net}}\frac{dV_{net}}{dP}\right) + V_{void}\left(-\frac{1}{V_{void}}\frac{dV_{void}}{dP}\right) \tag{3.7}$$

By analogy with the bulk modulus equation, the first term in brackets on the right-hand side of Equation 3.7 can be described as the reciprocal of the

network bulk modulus K_{net}, and the second, the reciprocal of the void bulk modulus K_{void}, so that the overall bulk modulus (K) is related to K_{net} and K_{void} by

$$\frac{V}{K} = \frac{V_{net}}{K_{net}} + \frac{V_{void}}{K_{void}} \tag{3.8}$$

Although the concept of the 'bulk modulus of a void' seems unphysical, it exists within a model in which the total volume of a unit cell is described in terms of occupied and unoccupied space. It arises as a term in Equation 3.7, which expresses the different responses of these spaces to pressure. Whereas the overall bulk modulus is a precisely defined quantity expressed in terms of pressure and volume, its partitioning into networks and voids is a matter of definition (here, using van der Waals radii), and so K_{net} and K_{void} do not have the same formal thermodynamic status as K. Nevertheless, values can be calculated by fitting V_{net} and V_{void} to an EoS and are useful for comparative purposes.

Some representative values of K_{net} and K_{void} are given in Table 3.1, a more extensive listing for 129 different compounds, including values of the ambient pressure volumes, bulk moduli and their pressure derivatives, is available in the electronic supplementary material, Table 1. A discussion of the fitting and the quality of the fits can be found in Appendix A, Section A2. After some experimentation, we found that network curves were generally fitted well with a second or third order Birch-Murnaghan EoS. The volumes of the much softer voids were usually better modelled with Vinet EoSs, as is found generally for the total volumes of soft solids and materials under very high compression.[55] The forms of these EoS are given in Chapter 1. For some structures it proved very difficult to fit V_{net}, and in these cases K_{net} was estimated using Equation 3.3.

Table 3.1: Bulk moduli and packing coefficients for selected compounds. Full fitting parameters available in Table 1 of the electronic supplementary material. * = V_0 was not refined. † = Values obtain using Equation 3.3. Packing coefficients were calculated for the lowest pressure structure available from the study. For NIBSOG the network volume was better fitted to a Vinet EoS rather than a Birch-Murnaghan EoS.

Compound	Network Bulk Modulus /GPa	Void Bulk Modulus /GPa	Overall Bulk Modulus /GPa	Packing Coefficient
D,*L*-serine	134(4)	4.9(5)	19(2)	0.7832
bromo-substituted bisdiselenazolyl radical (R_1 = ethane, R_2 = Br)	130(3)*	4.5(1)*	15.0(4)*	0.7828
L-alanine	121(2)*	4.1(1)*	13.1(6)[48]	0.7585
L-histidine orthorhombic	121(5)	3.2(2)	14.0(5)[30]	0.7617
L-cysteic acid monohydrate	102(8)	5(1)	17(5)	0.7817
L-histidine monoclinic	101(9)	2.9(2)	11.6(6)[30]	0.7561
bianthrone	95(9)	3.5(7)	8.1(5)[51]	0.8195
ROY OP	87†	2.2(1)	4.3(3)[56]	0.7496
ROY Y	84(4)	2.2(2)	6.0(7)[57]	0.7527
bis(3-Methoxysalicylaldoximato)-nickel(ii)	80(4)*	2.5(1)*	9.7(10)[58]	0.7961
$Ag_2Cu_2L_4$ (L = 2-diphenylphosphino-3-methylindole)	67(2)	2.2(5)	6(1)[59]	0.7362
bis(3-Fluorosalicylaldoximato)-nickel(ii)	48(8)	0.9(6)	9.1(17)[58]	0.7836
$Na_5[Mn(L\text{-}tart)_2]\cdot 12H_2O$ (1, L-tart = L-tartrate)	45(4)	5.0(6)	23.9(6)[60]	0.8898
Iron trifluoride	40.7(8)*	2.3(1)*	14(1)[61]	0.9153
KCp (Cp = cyclopentadienyl)	8(2)	1.9(3)	4.9(3)[62]	0.8729

The effect of error propagation in Equation 3.8 can be exemplified by taking typical values and errors of the bulk modulus of the network and void of 100(5) and 3.2(4) GPa respectively with an average cell volume of 1000 Å^3 with a typical packing coefficient of 0.75. The error in the overall calculated bulk modulus [11.7(13) GPa] is 11.1% of the total value. This is very close to the error in the void bulk modulus (12.5%) which provides the most significant contribution to the overall error.

The void bulk moduli fall into narrow range, typically 2-5 GPa. By contrast, network bulk moduli generally lie within the range 40-150 GPa, varying with the class of intermolecular interaction (see below), and are comparable to moderately hard metals.[53] Although the compressibility of the voids is determined by the distortion of the surrounding network, there appears to be no correlation between void and network bulk moduli (Figure A.3 in Appendix A, Section A3). The response of a given crystal structure to pressure depends more on the specific relationship between the network and void space than the strength of interactions that are present. However, the narrow range of K_{void} implies that no matter how compressible a network is, the way in which the void space adapts to elevated pressure to minimise volume is consistent across a very broad range of molecular solids.

Amongst molecular materials, H-bonded materials tend to have higher network bulk moduli, *e.g. L*-histidine 121(5) and 101(9) GPa for orthorhombic and monoclinic respectively, *D,L*-serine 134(4) GPa and *L*-alanine 121(2) GPa in Table 3.1. Histograms separating the network bulk moduli of hydrogen bonded and non-hydrogen bonded compounds are shown in Figure 3.12. Hydrogen bonded compounds produce a tighter distribution centred around 105 GPa while non-hydrogen bonded compounds produce a flatter distribution and centred around 85 GPa. Hydrogen bonds are the strongest

class of intermolecular interaction and are sensitive to geometry for example the D—H···A angle,[63] making them less deformable than other classes of intermolecular interaction.

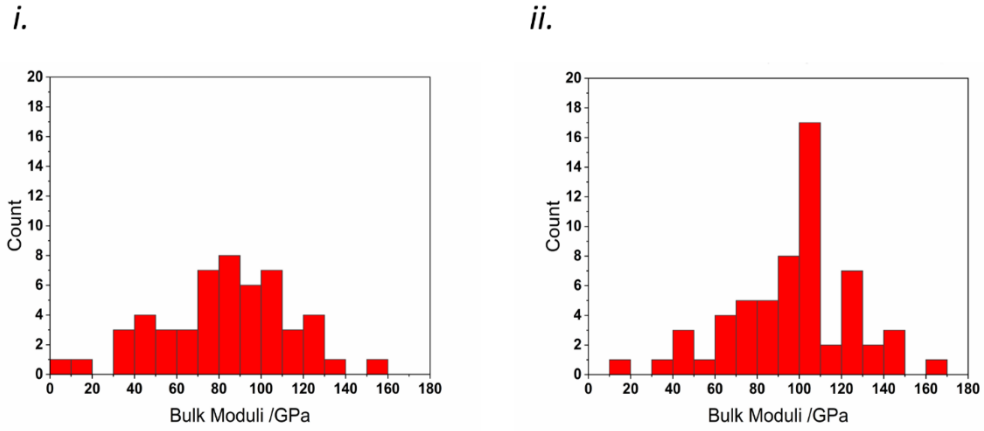


Figure 3.12: Network bulk moduli histograms for i non-hydrogen bonded compounds and ii hydrogen bonded compounds. Bulk moduli values calculated using Equation 3.3 have been excluded from the plots.

Some non-hydrogen bonded compounds also have high network bulk moduli. The compound IZOXOL (bromo-substituted bisdiselenazolyl radical (R_1 = ethane, R_2 = Br)) has a network bulk modulus of 130(3) GPa. It contains numerous short Se···Se intermolecular interactions and has quite a high packing coefficient (0.783). The typically low bulk modulus of the voids offsets the high network bulk modulus and the material has an overall bulk modulus of 15.0(4) GPa, which is relatively high for a non-H-bonded material.

Van der Waals crystals tend to have lower network bulk moduli, this reflects the lack of a characteristic or sharply defined geometrical signature for dispersion interactions which can therefore deform with pressure without incurring a high energy penalty, as seen for interaction #2 in naphthalene in Figure 3.5. An extreme example of this is the compound KCp (Cp = cyclopentadienyl) which features a zigzag polymeric chain with two cyclopentadienyl rings coordinated to each potassium centre. Despite the high

packing coefficient (0.8729), highly compressible networks and voids, 8(2) and 1.9(3) GPa respectively, result in low overall bulk moduli (4.9(3) GPa).

Data for the polymorphic compounds *L*-histidine and ROY are also listed in Table 3.1. Polymorphs tend to have similar overall bulk moduli, and while there may be some variation in the network bulk moduli and packing coefficients, the void bulk moduli are similar. Orders of the overall bulk moduli of these polymorphs follow packing coefficients, less densely packed polymorphs tending to be the more compressible.

3.4.3 The Range of the Bulk Moduli of Molecular Solids

The bulk moduli of minerals, metals and ceramics span ranges of many 10s or even 100s of GPa. Those of molecular materials fall into a much narrower range, typically falling between 5 and 20 GPa. The analysis described above can be used to provide some insight as to why this should be the case.

By substituting Equation 3.4 into Equation 3.8 we obtain

$$\frac{1}{K} = \frac{x}{K_{net}} + \frac{(1-x)}{K_{void}} \tag{3.9}$$

Overall bulk moduli for different combinations of the extreme values of K_{net} and K_{void} in the range x = 0.6-0.8 are shown in Table 3.2. The results reproduce the typical range of bulk moduli for molecular compounds. Since it has a numerically much smaller value than the network bulk modulus, the most significant factor in this calculation is the void bulk modulus. As this lies in a narrow range, so to do the bulk moduli of molecular solids.

Table 3.2: Bulk moduli calculated using Equation 3.9 from the range of network and void bulk moduli using typical packing coefficients.

K_{net}/GPa	K_{void}/GPa	Range of bulk moduli for x = 0.6-0.8
40	2	4.65 - 8.33
40	5	10.53 - 16.67
150	2	4.90 - 9.49
150	5	11.90 - 22.06

3.5 Conclusions

The methods that we have described enable the changes in overall unit cell volume that occur at high pressure to be decomposed into contributions from the interstitial voids and the network of intra- and intermolecular bonds. The partitioning is based on whether or not random points lie within or outside the van der Waals surfaces of the atoms which compose the structure. In the examples studied, relatively low pressure was seen to be taken up by interstitial voids, but at some point the network assumed a greater role, and phase transitions could be associated with the onset of an increase in the compressibility of the network.

There is a 'compensation' that can occur between the network and void volumes, which means that they are individually more sensitive to phase transitions than the overall volume. Large unit cell volume discontinuities are of course seen in the network and void volumes too, but more subtle effects are also revealed. For example, the literature contains many examples of discontinuities seen in vibrational spectra that seem not to be reproduced by conventional crystallographic analysis, but which are seen in the partitioned volumes. This feature was exemplified for naphthalene, where spectroscopic and crystallographic data could be reconciled, and unambiguous evidence of a phase transition seen in structural data for the first time since it was discovered in 1938. Detection of effects premonitory to phase transitions and loss of long-range order was also possible.

Network bulk moduli are usually between 40 and 150 GPa, the value reflecting the types of intermolecular interactions present, H-bonded networks being the least compressible. The higher end of these values is comparable to moderately hard metals. Void bulk moduli are over an order of magnitude smaller than network bulk moduli, and surprisingly perhaps, fall into quite a

narrow range (usually about 2-5 GPa). Because $K_{net} >> K_{void}$, the value of the overall bulk modulus, which depends on the reciprocals of the component moduli (Equation 3.8), is more strongly influenced by the value of K_{void} than that of K_{net}, and combination of the typical ranges for these quantities with typical packing coefficients recovers the range of bulk moduli seen for molecular solids, which shows much less variation (5-20 GPa) than metals (30-160 GPa), ionic salts (20-90 GPa) or ceramics (50-300 GPa).

The network and void bulk moduli have been shown to be useful parameters for the purposes of comparison of different structures and the same structure in different pressure ranges. The numerical values obtained depend on the method of partitioning and so they do not have the same fundamental thermodynamic significance as the overall bulk modulus. It is important to remember that the results of calculations on fragments of networks are not additive because the van der Waals surfaces of atoms and molecules within a network can overlap, so that volume of a network is usually less than the sum of the volumes of the component molecules.

Nevertheless, analysis of the partitioned volume allows the effects of external stimuli to be identified at a 'mesoscopic' level between the microscopic level of individual atom-atom distances and the macroscopic level of overall parameters such as the unit cell dimensions. Although the stimulus studied here has been pressure, the approach is equally applicable to crystal structures studied at variable temperature and with variable composition, such as in the up-take of guest species by framework materials. The method described should therefore prove extremely useful in the interpretation of crystallographic data collected under varying conditions, including the correlation of effects seen by spectroscopy and other measurements and in characterising the driving forces of phase transitions.

3.6 References

(1) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C., The Cambridge structural database. *Acta Crystallogr.* **2016,** B72, 171-179.

(2) Wood, P. A.; Forgan, R. S.; Henderson, D.; Parsons, S.; Pidcock, E.; Tasker, P. A.; Warren, J. E., Effect of pressure on the crystal structure of salicylaldoxime-I, and the structure of salicylaldoxime-II at 5.93 GPa. *Acta Crystallogr.* **2006,** B62, 1099-1111.

(3) Moggach, S. A.; Allan, D. R.; Morrison, C. A.; Parsons, S.; Sawyer, L., Effect of pressure on the crystal structure of L-serine-I and the crystal structure of L-serine-II at 5.4 GPa. *Acta Crystallogr.* **2005,** B61, 58-68.

(4) Rychkov, D. A.; Stare, J.; Boldyreva, E. V., Pressure-driven phase transition mechanisms revealed by quantum chemistry: L-serine polymorphs. *Phys. Chem. Chem. Phys.* **2017,** 19, 6671-6676.

(5) Gavezzotti, A., *Molecular Aggregation*. Oxford University Press, **2007**.

(6) Kitaigorodsky, A. L., *Molecular Crystals and Molecules*. Academic Press, **1973**.

(7) Blatov, V. A., Voronoi–dirichlet polyhedra in crystal chemistry: theory and applications. *Cryst. Rev.* **2004,** 10, 249-318.

(8) Férey, G.; Mellot-Draznieks, C.; Serre, C.; Millange, F.; Dutour, J.; Surblé, S.; Margiolaki, I., A chromium terephthalate-based solid with unusually large pore volumes and surface area. *Science* **2005,** 309, 2040-2042.

(9) Macrae, C. F.; Sovago, I.; Cottrell, S. J.; Galek, P. T. A.; McCabe, P.; Pidcock, E.; Platings, M.; Shields, G. P.; Stevens, J. S.; Towler, M.; Wood, P. A., Mercury 4.0: From visualization to analysis, design and prediction. *J. Appl. Crystallogr.* **2020,** 53, 226-235.

(10) Spek, A. L., A multipurpose crystallographic tool-PLATON. *Utrecht University, Netherlands.* **2007,** 61, 70-77.

(11) Blatov, V. A.; Shevchenko, A. P., Analysis of voids in crystal structures: the methods of 'dual' crystal chemistry. *Acta Crystallogr.* **2003,** A59, 34-44.

(12) Turner, M. J.; McKinnon, J. J.; Jayatilaka, D.; Spackman, M. A., Visualisation and characterisation of voids in crystalline materials. *CrystEngComm.* **2011,** 13, (6), 1804-1813.

(13) Dolomanov, O. V.; Bourhis, L. J.; Gildea, R. J.; Howard, J. A. K.; Puschmann, H., OLEX2: a complete structure solution, refinement and analysis program. *J. Appl. Crystallogr.* **2009,** 42, 339-341.

(14) Barbour, L. J., X-Seed—A software tool for supramolecular crystallography. *J. Supramol. Chem.* **2001,** 1, 189-191.

(15) Dowty, E., ATOMS for Windows and MacIntosh Shape Software. **2002**.

(16) McArdle, P., Oscail, a program package for small-molecule single-crystal crystallography with crystal morphology prediction and molecular modelling. *J. Appl. Crystallogr.* **2017,** 50, 320-326.

(17) Blatov, V. A., Multipurpose crystallochemical analysis with the program package TOPOS. *IUCr CompComm Newslett.* **2006,** 7, 4-38.

(18) Patyk-Kaźmierczak, E.; Warren, M. R.; Allan, D. R.; Katrusiak, A., Intermolecular contacts in compressed α-D-mannose. *Cryst. Growth Des.* **2016,** 16, 6885-6890.

(19) Alvarez, S. A., A cartography of the van der Waals territories. *Dalton Trans.* **2013,** 42, 8617-8636.

(20) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T., *Numerical Recipes in Fortran: The Art of Scientific Computing*. 2 ed. Cambridge University Press. **1992**.

(21) Anderson, G. M., Error propagation by the Monte Carlo method in geochemical calculations. *Geochim. Cosmochim. Acta.* **1976,** 40, 1533-1538.

(22) Rodriguez-Carvajal, J.; González-Platas, J., Crystallographic Fortran 90 Modules Library (CrysFML): a simple toolbox for crystallographic computing programs. *IUCr CompComm Newslett.* **2003,** 1, 50-58.

(23) Gonzalez-Platas, J.; Alvaro, M.; Nestola, F.; Angel, R., EosFit7-GUI: a new graphical user interface for equation of state calculations, analyses and teaching. *J. Appl. Crystallogr.* **2016,** 49, 1377-1382.

(24) Giordano, N.; Beavers, C. M.; Kamenev, K. V.; Marshall, W. G.; Moggach, S. A.; Patterson, S. D.; Teat, S. J.; Warren, J. E.; Wood, P. A.; Parsons, S., High-pressure polymorphism in L-threonine between ambient pressure and 22 GPa. *CrystEngComm.* **2019,** 21, 4444-4456.

(25) Clark, S. J.; Segall, M. D.; Pickard, C. J.; Hasnip, P. J.; Probert, M. I. J.; Refson, K.; Payne, M. C., First principles methods using CASTEP. *Z. Kristallogr. Cryst. Mater.* **2005,** 220, 567-570.

(26) Gavezzotti, A., Calculation of lattice energies of organic crystals: the PIXEL integration method in comparison with more traditional methods. *Z. Kristallogr. Cryst. Mater.* **2005,** 220, 499-510.

(27) Gavezzotti, A., Efficient computer modeling of organic materials. The atom–atom, Coulomb–London–Pauli (AA-CLP) model for intermolecular electrostatic-polarization, dispersion and repulsion energies. *New J. Chem.* **2011,** 35, 1360-1368.

(28) Reeves, M. G.; Wood, P. A.; Parsons, S., MrPIXEL: automated execution of Pixel calculations via the *Mercury* interface. *J. Appl. Crystallogr.* **2020,** 53, 1154-1162.

(29) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, 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.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; 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, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., Gaussian 09, Revision E. 01, Gaussian. *Inc:, Wallingford CT* **2009,** 201.

(30) Novelli, G.; Maynard-Casely, H. E.; McIntyre, G. J.; Warren, M. R.; Parsons, S., Effect of High Pressure on the Crystal Structures of Polymorphs of L-Histidine. *Cryst. Growth Des.* **2020,** 20, 7788-7804.

(31) O'Bannon, E.; Williams, Q., Vibrational spectra of four polycyclic aromatic hydrocarbons under high pressure: implications for stabilities of PAHs during accretion. *Phys. Chem. Miner.* **2016,** 43, 181-208.

(32) Bridgman, P. W., Polymorphic Transitions up to 50,000 kg/cm^3 of Several Organic Substances. *Proc. Am. Acad. Arts. Sci.* **1938,** 72, 227-268.

(33) Block, S.; Weir, C. E.; Piermarini, G. J., Polymorphism in benzene, naphthalene, and anthracene at high pressure. *Science.* **1970,** 169, 586-587.

(34) Vaidya, S. N.; Kennedy, G. C., Compressibility of 18 molecular organic solids to 45 kbar. *J. Chem. Phys.* **1971,** 55, 987-992.

(35) Nicol, M.; Vernon, M.; Woo, J. T., Raman spectra and defect fluorescence of anthracene and naphthalene crystals at high pressures and low temperatures. *J. Chem. Phys.* **1975,** 63, 1992-1999.

(36) Hamann, S. D., Infrared spectra and phase transitions of solids under pressure. *High Temp. High Press.* **1978,** 10, 503-510.

(37) Meletov, K. P., Phonon spectrum of a naphthalene crystal at a high pressure: Influence of shortened distances on the lattice and intramolecular vibrations. *Phys. Solid State.* **2013,** 55, 581-588.

(38) Fabbiani, F. P. A.; Allan, D. R.; Parsons, S.; Pulham, C. R., Exploration of the high-pressure behaviour of polycyclic aromatic hydrocarbons: naphthalene, phenanthrene and pyrene. *Acta Crystallogr.* **2006,** B62, 826-842.

(39) Likhacheva, A. Y.; Rashchenko, S. V.; Litasov, K. D., High-pressure structural properties of naphthalene up to 6 GPa. *J. Appl. Crystallogr.* **2014,** 47, 984-991.

(40) Likhacheva, A. Y.; Rashchenko, S. V.; Chanyshev, A. D.; Inerbaev, T. M.; Litasov, K. D.; Kilin, D. S., Thermal equation of state of solid naphthalene to 13 GPa and 773 K: In situ X-ray diffraction study and first principles calculations. *J. Chem. Phys.* **2014,** 140, 164508.

(41) Gavezzotti, A., The lines-of-force landscape of interactions between molecules in crystals; cohesive *versus* tolerant and 'collateral damage' contact. *Acta Crystallogr.* **2010,** B66, 396-406.

(42) Casati, N.; Kleppe, A.; Jephcoat, A. P.; Macchi, P., Putting pressure on aromaticity along with in situ experimental electron density of a molecular crystal. *Nat. Commun.* **2016,** 7, 10901.

(43) Reeves, M. G.; Tailleur, E.; Wood, P. A.; Marchivie, M.; Chastanet, G.; Guionneau, P.; Parsons, S., Mapping the cooperativity pathways in spin crossover complexes. *Chem. Sci.* **2021,** 12, 1007-1015.

(44) Momma, K.; Izumi, F., VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data. *J. Appl. crystallogr.* **2011,** 44, 1272-1276.

(45) Gould, J. A.; Rosseinsky, M. J.; Moggach, S. A., Tuning the coordination chemistry of a Cu (II) complex at high-pressure. *Dalton Trans.* **2012,** 41, 5464-5467.

(46) Wood, P. A.; Forgan, R. S.; Parsons, S.; Pidcock, E.; Tasker, P. A., 3-Fluorosalicylaldoxime at 6.5 GPa. *Acta Crystallogr.* **2009,** E65, 2001.

(47) Funnell, N. P.; Allan, D. R.; Maloney, A. G. P.; Smith, R. I.; Wilson, C. J. G.; Parsons, S., Suppression of isotopic polymorphism. *CrystEngComm.* **2021,** 23, 769-776.

(48) Funnell, N. P.; Dawson, A.; Francis, D.; Lennie, A. R.; Marshall, W. G.; Moggach, S. A.; Warren, J. E.; Parsons, S., The effect of pressure on the crystal structure of L-alanine. *CrystEngComm.* **2010,** 12, 2573-2583.

(49) Moggach, S. A.; Parsons, S., High pressure crystallography of inorganic and organometallic complexes. *Spectrosc. Prop. Inorg. Organomet. Compd.* **2009,** 40, 324-354.

(50) Ciabini, L.; Gorelli, F. A.; Santoro, M.; Bini, R.; Schettino, V.; Mezouar, M., High-pressure and high-temperature equation of state and phase diagram of solid benzene. *Phys. Rev.* **2005,** B72, 094108.

(51) Johnstone, R. D. L.; Allan, D.; Lennie, A.; Pidcock, E.; Valiente, R.; Rodríguez, F.; Gonzalez, J.; Warren, J.; Parsons, S., The effect of pressure on the crystal structure of bianthrone. *Acta Crystallogr.* **2011,** B67, 226-237.

(52) Speziale, S.; Duffy, T. S., Single-crystal elastic constants of fluorite (CaF_2) to 9.3 GPa. *Phys. Chem. Miner.* **2002,** 29, 465-472.

(53) Gray, T.; Whitby, M.; Mann, N., Bulk modulus for all the elements in the periodic table. https://periodictable.com/Properties/A/BulkModulus.html (accessed 05/11/2021)

(54) Dewaele, A.; Datchi, F.; Loubeyre, P.; Mezouar, M., High pressure–high temperature equations of state of neon and diamond. *Phys. Rev.* **2008,** B77.

(55) Poirier, J.-P., *Introduction to the Physics of the Earth's Interior*. Cambridge University Press, **2000**.

(56) Funnell, N. P.; Bull, C. L.; Ridley, C. J.; Capelli, S., Structural behaviour of OP-ROY at extreme conditions. *CrystEngComm.* **2019,** 21, 4473-4483.

(57) 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.* **2015,** 51, 10608-10611.

(58) Byrne, P. J.; Richardson, P. J.; Chang, J.; Kusmartseva, A. F.; Allan, D. R.; Jones, A. C.; Kamenev, K. V.; Tasker, P. A.; Parsons, S., Piezochromism in nickel salicylaldoximato complexes: tuning crystal-field splitting with high pressure. *Chem. Eur. J.* **2012,** 18, 7738-7748.

(59) Jarzembska, K. N.; Kamiński, R.; Dziubek, K. F.; Citroni, M.; Paliwoda, D.; Durka, K.; Fanetti, S.; Bini, R., Impact of high pressure on metallophilic interactions and its consequences for spectroscopic properties of a model

tetranuclear silver(I)–copper(I) complex in the solid state. *Inorg. Chem.* **2018,** 57, 8509-8520.

(60) Craig, G. A.; Woodall, C. H.; McKellar, S. C.; Probert, M. R.; Kamenev, K. V.; Moggach, S. A.; Brechin, E. K.; Parsons, S.; Murrie, M., A high-pressure crystallographic and magnetic study of $Na_5[Mn(L\text{-tart})_2]\cdot 12H_2O$ (L-tart = L-tartrate). *Dalton Trans.* **2015,** 44, 18324-18328.

(61) Jørgensen, J.-E.; Smith, R. I., On the compression mechanism of FeF_3. *Acta Crystallogr.* **2006,** B62, 987-992.

(62) Dinnebier, R. E.; van Smaalen, S.; Olbrich, F.; Carlson, S., Effect of crystal packing on the structures of polymeric metallocenes. *Inorg. Chem.* **2005,** 44, 964-968.

(63) Wood, P. A.; Allen, F. H.; Pidcock, E., Hydrogen-bond directionality at the donor H atom—analysis of interaction energies and database statistics. *CrystEngComm.* **2009,** 11, 1563-1571.

Chapter 4: Discerning Subtle High-Pressure Phase Transitions in Glyphosate

Overview

Within this Chapter the common garden herbicide glyphosate, N-(phosphonomethyl)glycine (glyphosate), has been studied between ambient pressure and 5.17 GPa using single crystal x-ray diffraction. Glyphosate forms a structure composed of layers parallel to the (1 0 -2) planes. Hydrogen bonds form along the stacking direction, which is very incompressible so that the effects of pressure are accommodated mostly within the layers. This study has confirmed two high pressure phase transitions previously observed by Raman spectroscopy, enabling the structural signatures of the transitions to be identified. Both transitions are very subtle and second order, involving changes to the way the structure responds to pressure rather than changes to structure. The first transition occurs between 0.93-1.21 GPa and corresponds to the onset of greater compressibility within the layers. The second transition between 3.78-4.23 GPa is an intramolecular feature signalling a deformation of the molecular backbone. In the absence of a first order phase transition, the packing remains in a compressed form of its ambient pressure form up until the highest pressure measured. A reconstructive phase transition is believed to occur at 5.98 GPa to a polycrystalline high-pressure phase.

4.1 Introduction

Vibrational spectroscopy provides a sensitive method for studying subtle transitions, but these sometimes do not appear to have an obvious structural effect on, for example, cell dimensions, volumes, or atomic positions. Identification of these transitions crystallographically can be complex and involve sifting through large amounts of numerical data, with the risk that subtle features could be missed.[1] For example, the presence or absence of a phase transition in naphthalene discussed in Chapter 3.[2-11] Clear discontinuities seen in high-pressure infra-red spectra of naphthalene are hardly discernable in crystal structure data.

Glyphosate, N-(phosphonomethyl)glycine, is a common garden herbicide, used widely in domestic products such as Weedol® and Roundup®. The Cambridge Structural Database (CSD)[12] features five entries for glyphosate under the refcode family PHOGLY. The first four (PHOGLY-PHOGLY03) are studies on crystals of the same phase (space group $P2_1/c$) obtained directly from commercial weed-killer samples at ambient temperature and pressure.[13-15] The final entry (PHOGLY04) is again an ambient temperature and pressure structure but features a unique polymorph in the space group $P2_1$, obtained on using MnO_2 in the presence of H_2SO_4 to oxidise N-(phosphonomethyl)iminodiacetic acid.[16]

Glyphosate was recently studied to 6.2 GPa using high pressure Raman spectroscopy in the spectral range 45 to 3700 cm^{-1} by Holanda *et al.*[17] Two high pressure phase transitions were identified. The first transition was located at between 0.97 and 1.5 GPa and the second between 4.29 and 4.63 GPa. The transitions were identified based on discontinuities in plots of vibrational frequency versus pressure, splitting and changes in the intensities of bands as well as the appearance of new modes. For example, the appearance at 0.97 GPa

of a new mode at 61 cm^{-1} was taken to indicate a change in space group symmetry. The second transition was identified above 3.86 GPa, where a band at 82 cm^{-1} merges to form a single wide band with another previously at 90 cm^{-1}. Glyphosate has not previously been studied using high pressure X-ray diffraction and the motivation for this work was to identify the structural character of the transitions with the aid of the component volume analysis described in Chapter 3.

4.2 Materials and Methods

4.2.1 Single Crystal X-ray Diffraction

Crystals of glyphosate (Sigma Aldrich) were grown by slow evaporation of an aqueous solution of concentration 10 mg mL^{-1}. Diffraction data were collected on a Bruker AXS D8 Venture diffractometer using Ag Kα radiation (λ = 0.56083 Å), first at ambient conditions and then at high pressure in two separate studies. In each case, a crystal was loaded into a Merrill–Bassett diamond-anvil cell (DAC) with half opening angle of 38°, 600 μm Boehler–Almax-cut diamonds and tungsten carbide backing plates.[18,19] A tungsten gasket of thickness 300 μm indented to 150 μm and hole diameter of 300 μm was used, with a 4:1 mixture of methanol and ethanol as a pressure transmitting medium.[20] A small ruby chip was also included in the sample loading and the ruby fluorescence method was used to measure the pressure; the values quoted have an uncertainty of 0.05 GPa.[21] The data at ambient pressure were collected in the DAC prior to addition of the pressure transmitting medium. The maximum pressure reached was 5.17 GPa. The data quality deteriorated at 5.60 GPa, and increasing pressure further to 5.98 GPa led to a proliferation of diffraction spots which pointed to the sample undergoing a destructive phase transition forming a new polycrystalline phase. Attempts to determine the unit cell dimensions of the new phase were unsuccessful, see Chapter 6.

The diffraction patterns were indexed using APEX4 and integrated using SAINT.[22] Corrections for the gasket shading, absorption and other systematic errors were applied using the multiscan procedures in SADABS.[23] The structure at ambient pressure was solved using dual-space methods (SHELXT),[24] and refinement of all structures was against $|F|^2$ (SHELXL),[25] from within the OLEX-2 interface.[26] The starting model at each pressure point

was taken from the previous point in the series and all intramolecular bond lengths and angles restrained to their values at ambient pressure. All non-hydrogen atoms were refined with anisotropic displacement parameters (adps) and hydrogen atoms attached to carbon were placed at calculated positions and allowed to ride their parent atoms; those attached to oxygen were treated as rotating rigid groups. Adps were subject to enhanced rigid bond restraints.[27] Listings of crystal and refinement data are available in Appendix B Section B2, Table A.1. Structures were visualised using Mercury[28] and Diamond.[29]

4.2.2 Periodic DFT Calculations

The high-pressure crystal structures were optimised using periodic density functional theory (DFT) using CASTEP with the Perdew-Burke-Ernzerhof (PBE)[30] exchange-correlation functional and 'on the fly generated' (OTFG) pseudopotentials.[31] The basis set cut-off and **k**-point spacing were 920 eV and 0.07 $Å^{-1}$, giving a convergence in the total energy of 0.1 meV/atom. The convergence criteria for the geometry optimisations were: 5×10^{-6} $eVatom^{-1}$ (energy), 0.01 eV $Å^{-1}$ (force) and 5×10^{-4} Å (displacement). The unit cell parameters were fixed at observed values throughout. DFT optimised structures were used in all calculations and figures unless stated otherwise.

4.2.3 Occupied Volume ('CellVol') Calculations

Occupied ('network') and unoccupied ('void') volumes in crystal structures were evaluated in a Monte Carlo procedure using the CellVol code[32] described in Chapter 3.

4.2.4 Pixel Calculations

Intermolecular interaction energies were calculated via the Pixel method[33-35] using the MrPixel interface.[36] Gaussian-09[37] was used to calculate the electron density at the MP2 level of theory with the 6-31G** basis set. The molecular electron density was calculated on a grid of 0.08 × 0.08 × 0.08 $Å^3$ and a condensation level of 4 was used for the Pixel calculations with a cluster radius of 14 Å. The positions of H-atoms, which strongly influence intermolecular interaction energies, are systematically in error when determined by X-ray diffraction as the result of the asymmetry of their local electron distributions. In addition, the position of H3 in the refined crystal structures exhibited some instability as a function of pressure, with an oscillation in the torsion at successive pressure points. This is likely to be an artefact, perhaps the result of the relatively low (~40%) completeness of high-pressure data sets, a result of shading by the diamond anvil cell. Pixel calculations were therefore based on coordinates obtained from geometry-optimisation of the crystal structures using periodic DFT.

4.3 Results and Discussion

4.3.1 Ambient Pressure Structure

Glyphosate, N-(phosphonomethyl)glycine, crystallizes in space group $P2_1/c$ with one molecule in the asymmetric unit.[38] The molecules adopt the zwitterionic tautomeric form (Figure 4.1), with formal positive and negative charges on the ammonium group based on N3 and phosphate groups based on P5, respectively. The bond distances, angles and conformation have been discussed in detail in ref [13].

Intermolecular interactions are dominated by H-bonding. Intermolecular interaction energies, evaluated using the Pixel method (Table 4.1) show that the first coordination sphere, consisting of those molecules forming interactions to a central reference molecule with non-zero Pauli repulsion terms, contains 13 molecules (Figure 4.2). The prediction of mechanical properties based on slip planes described by Bryant, Maloney and Sykes[39] indicates that the most probable slip plane in the structure of glyphosate is (1 0 −2), and the crystal structure can be described in terms of layers parallel to these planes (Figure 4.2i).

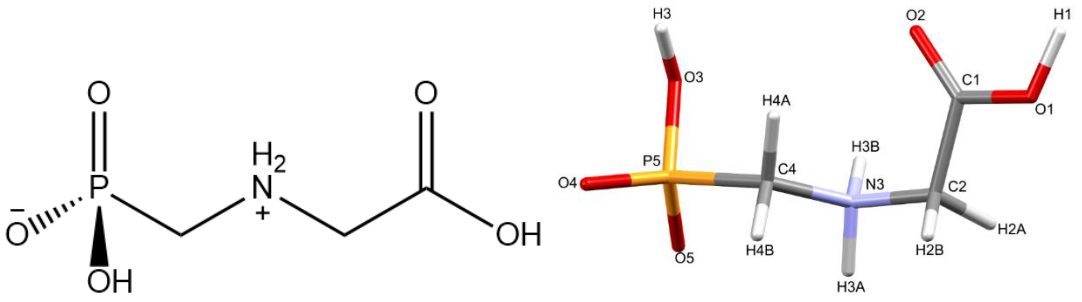


Figure 4.1: The zwitterionic molecular structure of glyphosate and the numbering scheme used in this work.

Within the layer (Figure 4.2ii) the phosphate moieties form pairs of O3—H3···O4 H-bonds across inversion centres in a ring motif (contact A,

centroid separation 6.868 Å, total energy −114.7 kJ mol^{-1}). These units are linked into a chain which runs along **b** by pairs of N3—H3B···O5 H-bonds, also forming ring motifs across inversion centres, between the ammonium and phosphate groups (contact B, centroid separation 5.033 Å, total energy −267.3 kJ mol^{-1}). This interaction between oppositely charged groups is the most stabilising contact in the structure. The chains are connected into a layer which is parallel to the (1 0 −2) planes on one side by weaker interactions C/C′ and D and on the other by a symmetry-equivalent interaction D′. Neither C/C′ nor D/D′ forms H-bonds with the central molecule. C is barely stabilising at all with a total energy equal to −0.6 kJ mol^{-1}, composed principally of destabilising electrostatic and stabilising dispersion components (+10.4 and −14.8 kJ mol^{-1}, respectively). The closest atom-atom distances, formed between C2H2A and carboxyl, measure 2.7 – 3.0 Å, and the centroid separation is 5.995 Å. Contact D is a longer-range electrostatic interaction (centroid separation 8.854 Å and total energy −28.4 kJ mol^{-1}).

Table 4.1: Interactions in the first coordination sphere. All energies are in kJ mol^{-1}. Classifications are relative to the interaction with the central molecule.

Label	Symmetry	Centroid distance (Å)	Electrostatic	Polarisation	Dispersion	Repulsion	Total	Contact description
				H-bonded				
Contact A	$1-x, 2-y, 1-z$	6.868	−244.8	−115.4	−32.3	277.9	−114.7	Pairs of O3H3···O4 H-bond ring motifs across inversion centres.
Contact B	$1-x, 1-y, 1-z$	5.033	−273.3	−95.9	−40.2	142.1	−267.3	Two N3H3B···O5 H-bond ring motifs across inversion centres.
Contact E/E′	$1-x, y \mp \frac{1}{2}, \frac{1}{2}-z$	6.532	−147.7	−50.6	−18.5	72.3	−144.5	Pairs of N3H3A···O4 H-bonds.
Contact F/F′	$x \pm 1, y, z$	8.669	−136.7	−67.3	−13.0	126.4	−90.6	Lattice translations of the central molecule in **a**. O1H1···O5 H-bond through a head to tail arrangement.
				Electrostatic Interactions				
Contact C/C′	$-x, y \mp \frac{1}{2}, \frac{1}{2}-z$	5.995	10.4	−8.5	−14.8	12.3	−0.6	Lattice translation of contact E/E′ along **a**. Composed principally of destabilising electrostatic and stabilising dispersion components.
Contact D/D′	$x \mp 1, \frac{3}{2}-y, z \mp \frac{1}{2}$	8.854	−24.2	−3.3	−4.1	3.1	−28.4	Lattice translation of contact F/F′ along **a**. Longer range electrostatic interaction.
Contact G	$-x, 1-y, 1-z$	6.460	16.8	−7.4	−11.9	8.4	5.8	Destabilising electrostatic interaction as a result of the juxtaposition of like charged moieties.
Contact H/H′	$x, \frac{3}{2}-y, z \mp \frac{1}{2}$	5.135	46.4	−12.7	−17.0	15.5	32.2	Destabilising electrostatic interaction as a result of the juxtaposition of like charged moieties.

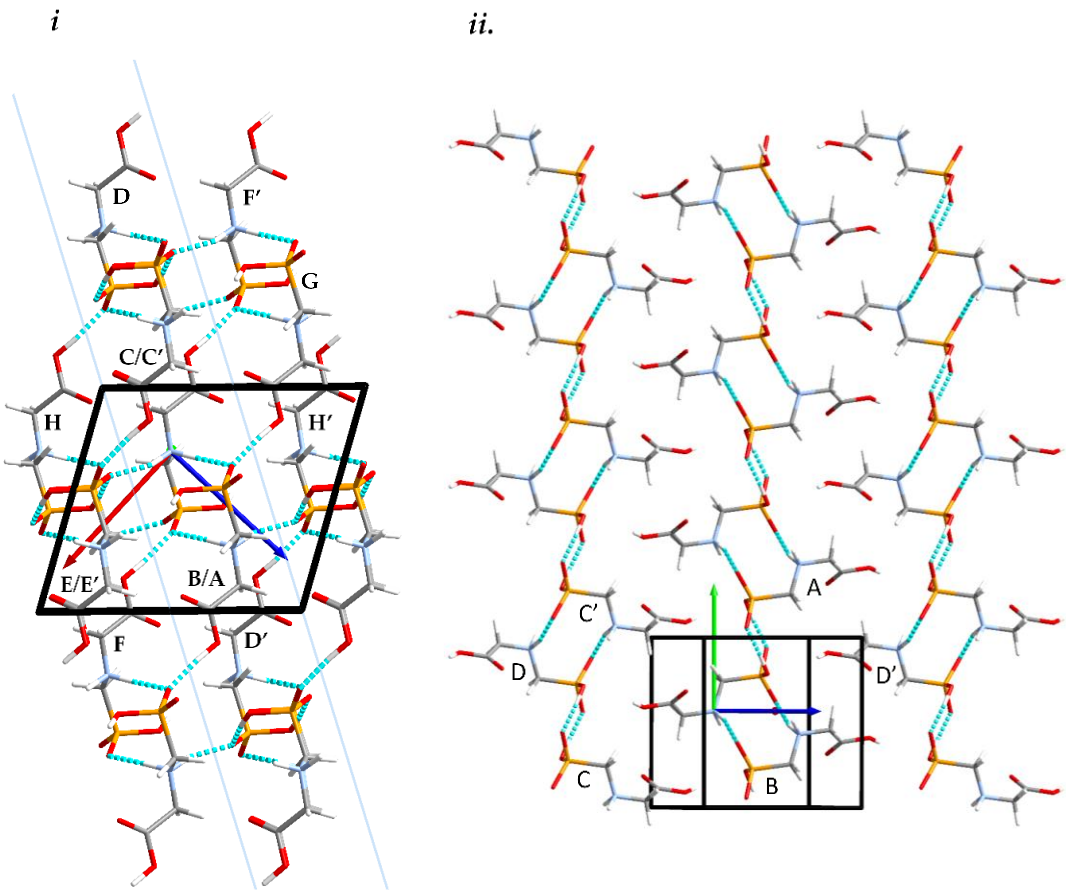


Figure 4.2: Packing diagrams for glyphosate. The interactions within the first coordination sphere are labelled. Slashes indicate those molecules hidden by symmetry from the view. Vectors drawn from the central reference molecule on the diagram represent the principal axes of the strain tensor; red, green and blue are used for the numerically smallest, middle and largest eigenvalues. i) A view down the *b*-direction reveals the layered structure parallel to (1 0 −2) slip planes, indicated by blue lines. ii) A view of a single layer. Note that labels refer to molecule - molecule interactions, as listed in Table 4.1, rather than individual contacts such as H-bonds.

The stacking of the layers involves four further H-bonded contacts. The most stabilising of these forms N3—H3A···O4 H-bonds with the central molecule (contact E/E′, centroid separation 6.532 Å, total energy −144.5 kJ mol^{-1}). Further H-bonds (contact F/F′) are formed through lattice translations along **a**, forming a head-to-tail arrangement involving O1—H1···O5 (centroid

separation 8.669 Å, total energy −90.6 kJ mol^{-1}). Destabilising electrostatic interactions G (centroid separation 6.460 Å, total energy 5.8 kJ mol^{-1}) and H/H′ (centroid separation 5.135 Å, total energy 32.2 kJ mol^{-1}) are also formed between the layers as the result of the juxtaposition of like-charged moieties. The formation of destabilising contacts has been noted in other zwitterionic structures such as amino acids.[40,41]

4.3.2 Response of the Unit Cell Parameters to Pressure

The total unit cell volume decreases monotonically (Figure 4.3i), with no apparent discontinuities between 0.97 and 1.5 GPa or between 4.29 and 4.63 GPa associated with the phase transitions reported in Holanda *et al.*'s study.[17] The volume versus pressure data can be fitted ($\chi^2 = 0.93$) to a single third-order Birch-Murnaghan equation of state (EoS) to yield a bulk modulus of 14.8 (4) GPa with a pressure derivative of 7.5 (4).[42] These values are typical of H-bonded molecular crystals, and are comparable to other amino acids such as serine [19 (2) GPa][32], alanine [13.1 (6) GPa][43] and histidine [14.0 (5) GPa].[40] No changes in the symmetry or space group of the crystals were observed over the pressure range studied.

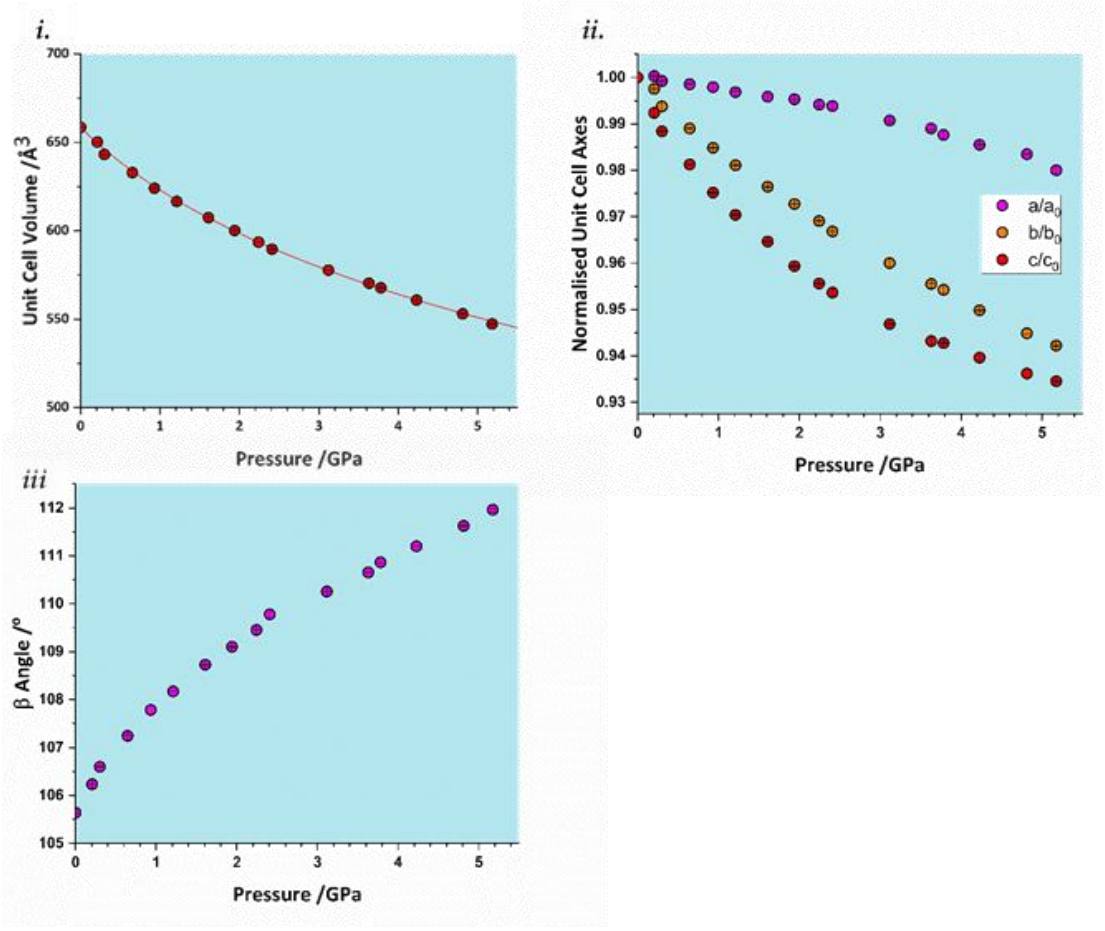


Figure 4.3: i) The unit cell volume as a function of pressure. A Birch-Murnaghan 3rd order equation of state has been fitted to yield a bulk modulus of 14.8 (4) GPa. ii) Normalised unit cell axes lengths as a function of pressure. iii) β angle as a function of pressure. Error bars are plotted for all parameters but are mostly smaller than the data symbols.

The compressibility of the unit cell parameters is anisotropic, following the order *c* (6.5% between ambient pressure and 5.17 GPa) > *b* (5.8%) > *a* (2.0%). (Figure 4.3ii); the value of β increases monotonically (Figure 4.3iii). None of the unit cell parameters show any significant features at the first reported transition pressure, though the **a** and **c** axis lengths show small discontinuities at 3.8 GPa, marking of the onset of the second transition.

The anisotropy of compression is much clearer in the strain tensor.[44,45] The smallest eigenvalue determined between ambient pressure and 5.17 GPa (Table 4.2) is over an order of magnitude smaller than the other two [−0.0012

(4) versus −0.0587 (3) and −0.1204 (4)]. The eigenvectors (Figure 4.2 and Table 4.2) show that the numerically largest eigenvalue is located in the planes described above in the direction of the very weak contacts C and D. By symmetry, one eigenvector must lie along **b**, and in glyphosate this is the middle axis, the smaller compression being consistent with the occurrence of the chain-forming H-bonds A and B described above. The axis of minimum compression lies perpendicular to the first two axes, approximately along the strongly H-bonded stacking direction, and has an eigenvalue which is barely significantly different from zero. The essentially two-dimensional character of the compression in glyphosate thus reflects the relative incompressibility of strong H-bonds.

Table 4.2: The principal values of the strain tensor calculated between ambient pressure and 5.17 GPa. <*a etc*. are the angles made by each eigenvector with the unit cell axes. The principal values of the strain tensor evaluated at each individual pressure are available in Appendix B, Section B3 Table A.2 and A.3.

Axis	Strain	<*a*/°	<*b*/°	<*c*/°
1	−0.0012 (4)	27.1	90.0	132.8
2	−0.0578 (3)	90.0	0.0	90.0
3	−0.1204 (4)	62.9	90.0	42.8

4.3.3 Intermolecular Interactions

The centroid-centroid separation of most contacts reduces with increasing pressure, except for contact G, which remains relatively constant at ~6.5 Å. Contacts which do not involve hydrogen bonds (C, D and H) generally show a greater reduction in centroid separation with the greatest reductions seen for contacts C and D in the direction of the most compressible strain axis (0.538, 0.761, 0.37 Å for contacts C, D and H respectively). Amongst the H-bonded

contacts (A, B, E and F), A shows a significantly higher change in centroid separation (0.46 Å) than the others, for which compression is more consistent (0.155, 0.148 and 0.173 Å for contacts B, E and F respectively). Both contact A and B align with the middle strain axis while contacts E and F align with the least compressible strain axis.

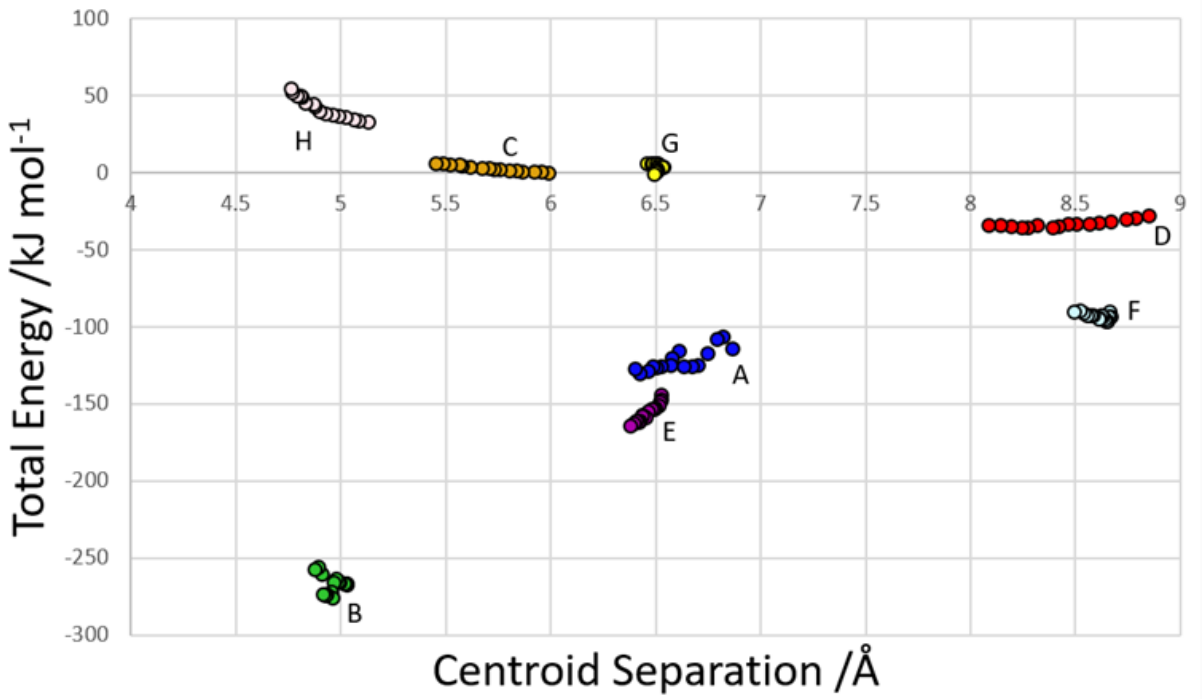


Figure 4.4: Total dimer energies versus centroid separation for contacts within the first coordination sphere.

The energies of each of the 13 intermolecular contacts are shown as a function of the centroid separation in Figure 4.4. The trends are monotonic for most of the contacts, but there is some scatter for the chain-forming H-bonded interactions A and B. These energies are highly sensitive to random errors in the atomic (particularly the H) positions. Even though the coordinates used in the Pixel calculations were optimised by periodic DFT, the cell dimensions were fixed at their experimental values during these optimisations, and so some random error persists. Since the scatter does not align with the phase transitions, it probably reflects random error.

4.3.4 Volume Analysis

As was introduced in Chapter 3, the volume of a crystal structure can be partitioned into regions which are occupied by molecules and their network of intra and intermolecular interactions and unoccupied interstitial voids. Trends in V_{net} and V_{void} with pressure can be more sensitive to subtle structural transitions than the overall unit cell dimensions. This has already been demonstrated to identify structural behaviour in naphthalene, 4-methylpyridine pentachlorophenol co-crystals and, in an additional study not reproduced here, the Blatter radical.[32,46,47]

The network and void volumes in glyphosate are shown in Figure 4.5 as a function of pressure. While there are no discontinuities in either the overall unit cell volume or intermolecular energies associated with the phase transitions revealed by Raman spectroscopy, small discontinuities do occur in the gradient of the network volume trends at 0.93 and 3.78 GPa. These are less clear in the void volume trends as they are overwhelmed by the greater compressibility of the voids. The trends in V_{net} between the transitions are quite linear, and fitting to second order Birch-Murnaghan equations of state yields bulk moduli of 90 (12), 105 (2) and 96 (7) GPa for the three phases with increasing pressure; there is no significant difference between these figures. The average network bulk modulus can be obtained by fitting a second-order Birch-Murnaghan EoS across the entire data range giving K = 104.2 (6) GPa (χ^2 = 0.86), a value typical of hydrogen bonded networks seen in Chapter 3. A third-order Vinet equation of state was applied to the void giving a bulk modulus of 4.4 (1) GPa, K'= 1.17 (9) and a χ^2 of 1.29, again typical of values seen in Chapter 3.

To identify which intermolecular interactions undergo the largest changes at the transitions, the volumes of the glyphosate molecule, dimers and

larger clusters were calculated by placing a box around each and evaluating the occupied volume using a similar Monte Carlo procedure to that described above for unit cells. Further details are available in the Appendix B, Section B1. The calculation was first applied to the interlayer volume, featuring contacts E, F, G, H and the central molecule (Figure A.4 in Appendix B, Section B1). This plot contains no clear discontinuity at the first transition, Figure 4.6i. The layer stacking is in the direction of the numerically smallest eigenvalue of the strain tensor and appears to remain unchanged through the first transition. A very slight discontinuity is seen at the second transition. The calculation was then applied to a single layer containing the central molecule and contacts A, B, C and D (also Figure A.4 in Appendix B, Section B1). This plot features a discontinuity at both transition pressures, indicating that both transitions have a structural signature within the layers, Figure 4.6ii. In particular, the layer compressibility appears to level-off immediately before the first transition but becomes re-established afterwards. The first transition can thus be associated with the re-emergence of compressibility within the layers.

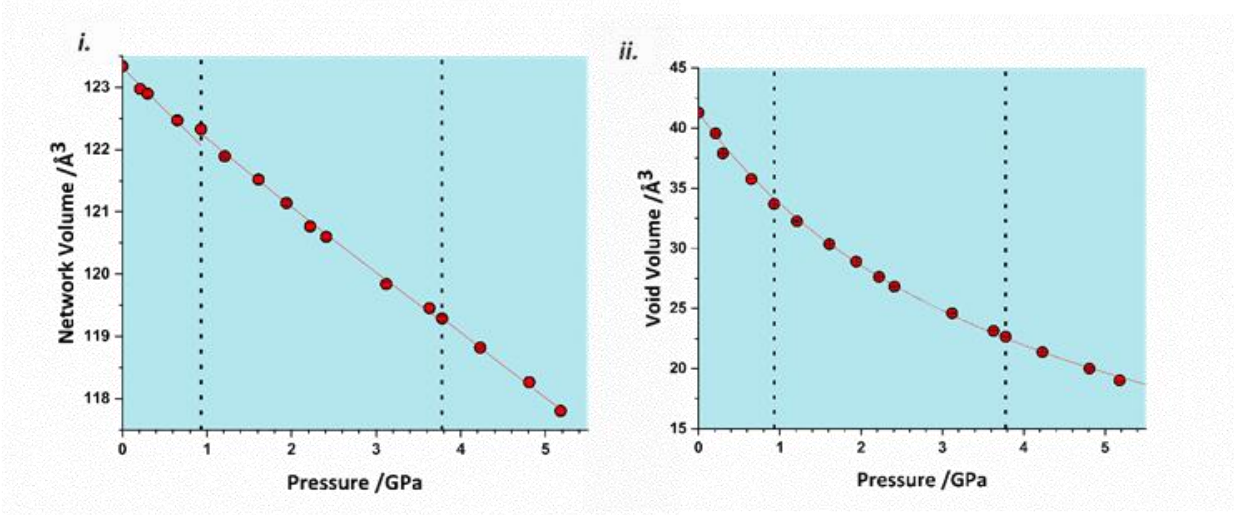


Figure 4.5: i) Network volumes at increasing pressure. ii) Void volumes at increasing pressure. Dashed lines in all plots represent phase transitions based on this work and that of Holanda *et al*.

To identify the interactions driving these transitions the Monte Carlo method was then applied to the individual dimers A-H listed in Table 4.1. Plots for all contacts can be found in Appendix B Section B1with selected contacts presented in Figures 4.6iii and iv. Volumes at this scale are very sensitive to structural instabilities and there is some scatter. Nevertheless, contact A and D/D′ show fairly clear discontinuities at the first transition. Contact A shows a reduction in dimer volume and contact D/D′ a smaller increase, so that the re-emergence of layer compressibility can be associated with contact A.

Both interlayer and intralayer plots and virtually all contacts can be interpreted as having a discontinuous change in dimer volume at the second transition. This suggests that the transition may have an intramolecular contribution, and application of the volume calculation to individual molecules reveals a small discontinuity in molecular volume at the second transition, Figure 4.7i. As pressure is increased, the long axis of the molecule hunches to a more compressed form, with most of this compression taken up at the carbonyl end of the molecule, Figure 4.7ii. As pressure increases, the longest axis of the inertial tensor (see Appendix B, Section B1) compresses, Figure 4.7iii, with the trend becoming steeper after the second transition. The alignment of the long axes with the layer explains why the second transition is more obvious in Figure 4.6ii than in 4.6i.

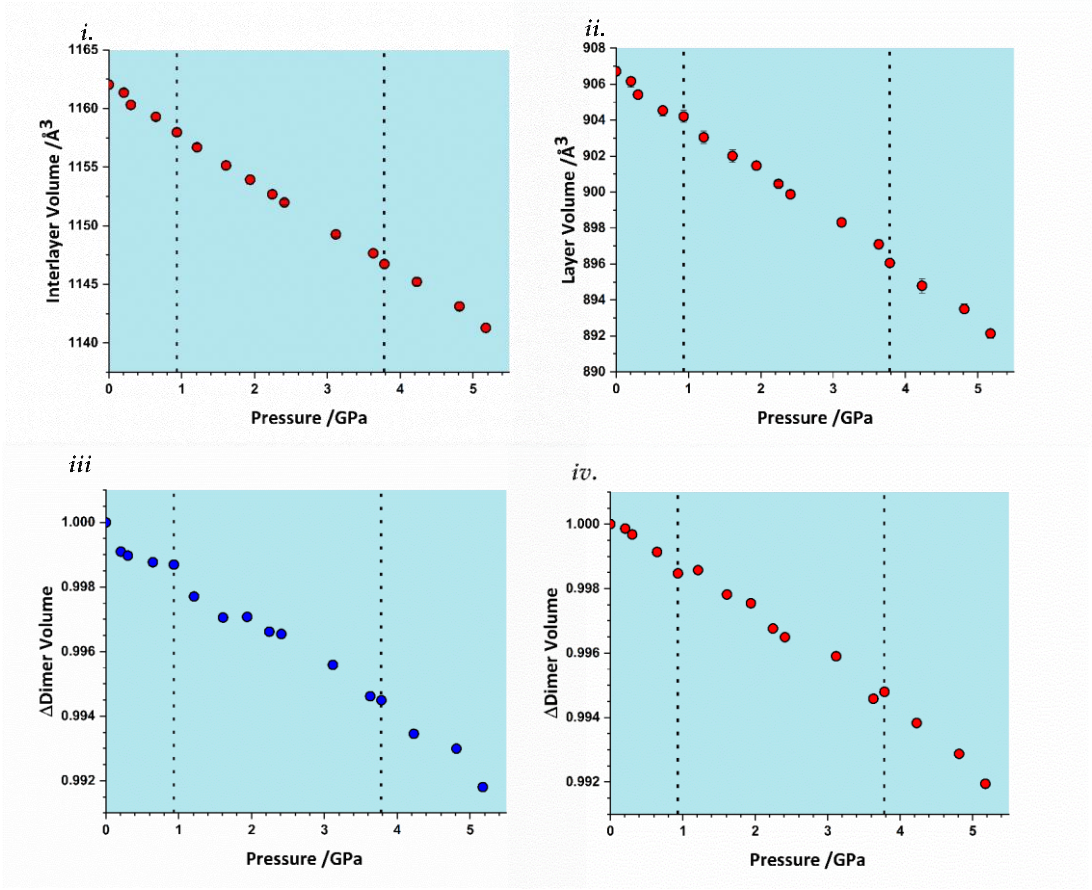


Figure 4.6: Variation in the volume of subsets of the glyphosate crystal structure. i) Interlayer volume, consisting of the central molecule and those in adjacent layers within the first coordination sphere. ii) Layer volume, consisting of a single layer containing the central molecule. iii) The volume of contact A (Figure 4.2). iv) The volume of contact D/D'.

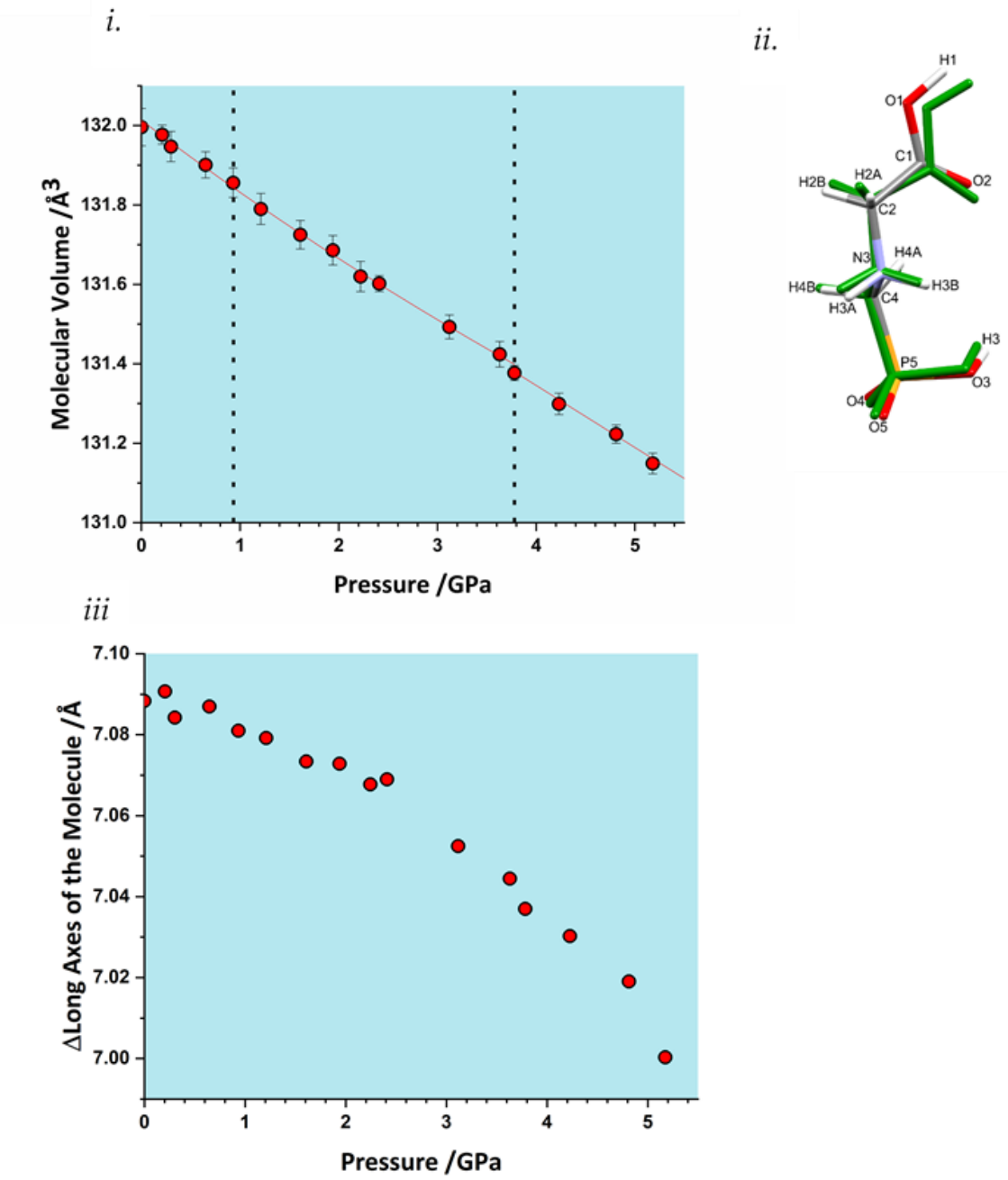


Figure 4.7: i) Molecular volume the glyphosate molecule as a function of pressure. ii) Hunching of the molecular backbone. The elementally coloured structure is from ambient pressure measurements, the overlaid green structure is from 5.17 GPa. iii) The change in length of the long axis of the molecule (the longest axis of the inertial tensor) with pressure.

4.4 Conclusions

The literature contains many examples of vibrational discontinuities from spectroscopic experiments at high pressure that often seem to not be reproduced by conventional crystallographic analysis. This seems to be especially relevant to second order transitions which instead of accounting for physical changes to structural packing, represent the onset of new mechanisms for accommodating pressure. The network and void volumes appear to be more sensitive than other structural features to these subtle changes.

In this Chapter we have applied these methods to two high pressure phase transitions in glyphosate which had previously been identified by Raman spectroscopy. In Chapter 3 it was noted that the onset of subtle transitions marks a shift from compression of the void space to compression of the more rigid network, and the same in true in glyphosate. The first transition, between 0.93-1.21 GPa corresponds to a discontinuous change in the overlap of van der Waals surfaces within the layers to accommodate further compression. The second transition between 3.78-4.23 GPa is an intramolecular effect, signalling an increased rate of contraction of the molecular backbone to allow further compression. Both transitions occur without a change to space group symmetry or a change to the structural packing. The lack of a discontinuity in either the total volume or energy for either transition, suggests that they are both second order.

4.5 References

(1) Moggach, S. A.; Oswald, I. D. H., *Crystallography under high pressures. 21st Century Challenges in Chemical Crystallography I: History and Technical Developments.* **2020**, 141-198.

(2) Bridgman, P. W., Polymorphic Transitions up to 50,000 kg/cm^3 of Several Organic Substances. *PNAS.* **1938**, 72, 227-268.

(3) Block, S.; Weir, C. E.; Piermarini, G. J., Polymorphism in benzene, naphthalene, and anthracene at high pressure. *Science.* **1970**, 169, 586-587.

(4) Vaidya, S. N.; Kennedy, G. C., Compressibility of 18 molecular organic solids to 45 kbar. *J. Chem. Phys.* **1971**, 55, 987-992.

(5) Nicol, M.; Vernon, M.; Woo, J. T., Raman spectra and defect fluorescence of anthracene and naphthalene crystals at high pressures and low temperatures. *J. Chem. Phys.* **1975**, 63, 1992-1999.

(6) Hamann, S. D., Infrared spectra and phase transitions of solids under pressure. *High Temp. High Press.* **1978**, 10, 503-510.

(7) Meletov, K. P., Phonon spectrum of a naphthalene crystal at a high pressure: Influence of shortened distances on the lattice and intramolecular vibrations. *Phys. Solid State.* **2013**, 55, 581-588.

(8) O'Bannon, E.; Williams, Q., Vibrational spectra of four polycyclic aromatic hydrocarbons under high pressure: implications for stabilities of PAHs during accretion. *Phys. Chem. Miner.* **2016**, 43, 181-208.

(9) Fabbiani, F. P. A.; Allan, D. R.; Parsons, S.; Pulham, C. R., Exploration of the high-pressure behaviour of polycyclic aromatic hydrocarbons: naphthalene, phenanthrene and pyrene. *Acta Crystallogr.* **2006**, B62, 826-842.

(10) Likhacheva, A. Y.; Rashchenko, S. V.; Litasov, K. D., High-pressure structural properties of naphthalene up to 6 GPa. *J. Appl. Crystallogr.* **2014**, 47, 984-991.

(11) Likhacheva, A. Y.; Rashchenko, S. V.; Chanyshev, A. D.; Inerbaev, T. M.; Litasov, K. D.; Kilin, D. S., Thermal equation of state of solid naphthalene to 13 GPa and 773 K: In situ X-ray diffraction study and first principles calculations. *J. Chem. Phys.* **2014**, 140, 164508.

(12) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C., The Cambridge Structural Database. *Acta Crystallogr.* **2016**, B72, 171-179.

(13) Knuuttila, P.; Knuuttila, H., The Crystal and Molecular Structure of N-(Phosphonomethyl)glycine (Glyphosate). *Acta Chem. Scand.* **1979**, B33, 623-626.

(14) Sheldrick, W. S.; Morr, M., N-Phosphonomethylglycin. *Acta Crystallogr.* **1981**, B37 (3), 733-734.

(15) Shkol'nikova, L. M.; Porai-Koshits, M. A.; Dyatlova, N. M.; Yaroshenko, G. F.; Rudomino, M. V.; Kolova, E. K., X-ray diffraction study of complexon-type organic ligands. III. Crystal and molecular structure of phosphonomethylglycine and iminodiacetylmethylphosphonic acid. *Zh. Strukt. Khim.* **1982**, 23, 98-107.

(16) Krawczyk, H.; Bartczak, T. J., New crystalline polymorphic form of glyphosate: synthesis, crystal and molecular structures of N-(phosphonomethyl) glycine. *Phosphorus, Sulfur, Silicon Relat. Elem.* **1993**, 82 (1-4), 117-125.

(17) Holanda, R. O.; da Silva, C. B.; Vasconcelos, D. L. M.; Freire, P. T. C., High pressure Raman spectra and DFT calculation of glyphosate. *Spectrochim. Acta.* **2020**, A242, 118745.

(18) Merrill, L.; Bassett, W. A., Miniature diamond anvil pressure cell for single crystal x-ray diffraction studies. *Rev. Sci. Instrum.* **1974**, 45 (2), 290-294.

(19) Moggach, S. A.; Allan, D. R.; Parsons, S.; Warren, J. E., Incorporation of a new design of backing seat and anvil in a Merrill-Bassett diamond anvil cell. *J. Appl. Crystallogr.* **2008**, 41, 249-251.

(20) Klotz, S.; Chervin, J. C.; Munsch, P.; Le Marchand, G., Hydrostatic limits of 11 pressure transmitting media. *J. Phys.* **2009**, D42 (7), 075413.

(21) Mao, H. K.; Bell, P. M.; Shaner, J. W. t.; Steinberg, D. J., Specific volume measurements of Cu, Mo, Pd, and Ag and calibration of the ruby R_1 fluorescence pressure gauge from 0.06 to 1 Mbar. *J. Appl. Phys.* **1978**, 49 (6), 3276-3283.

(22) *APEX4*; Bruker AXS Inc.: Madison, Wisconsin, USA, **2021**.

(23) Krause, L.; Herbst-Irmer, R.; Sheldrick, G. M.; Stalke, D., Comparison of silver and molybdenum microfocus X-ray sources for single-crystal structure determination. *J. Appl. Crystallogr.* **2015**, 48 (1), 3-10.

(24) Sheldrick, G. M., SHELXT–Integrated space-group and crystal-structure determination. *Acta Crystallogr.* **2015**, A71 (1), 3-8.

(25) Sheldrick, G. M., Crystal structure refinement with SHELXL. *Acta Crystallogr.* **2015**, C71 (1), 3-8.

(26) Dolomanov, O. V.; Bourhis, L. J.; Gildea, R. J.; Howard, J. A. K.; Puschmann, H., OLEX2: a complete structure solution, refinement and analysis program. *J. Appl. Crystallogr.* **2009**, 42, 339-341.

(27) Thorn, A.; Dittrich, B.; Sheldrick, G. M., Enhanced rigid-bond restraints. *Acta Crystallogr.* **2012**, A68 (4), 448-451.

(28) Macrae, C. F.; Sovago, I.; Cottrell, S. J.; Galek, P. T. A.; McCabe, P.; Pidcock, E.; Platings, M.; Shields, G. P.; Stevens, J. S.; Towler, M.; Wood, P. A., Mercury 4.0: From Visualization to Analysis, Design and Prediction. *J. Appl. Crystallogr.* **2020,** 53, 226-235.

(29) *Diamond – Crystal and Molecular Structure Visualization*. Crystal Impact, **2020**.

(30) Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, 77 (18), 3865-3868.

(31) Clark, S. J.; Segall, M. D.; Pickard, C. J.; Hasnip, P. J.; Probert, M. I. J.; Refson, K.; Payne, M. C., First principles methods using CASTEP. *Z. Kristallogr. Cryst. Mater.* **2005**, 220, 567-570.

(32) Wilson, C. J. G.; Cervenka, T.; Wood, P. A.; Parsons, S., Behavior of Occupied and Void Space in Molecular Crystal Structures at High Pressure. *Cryst. Growth Des.* **2022**, 22, 2328-2341.

(33) Gavezzotti, A., *Molecular Aggregation*. Oxford University Press, **2007**.

(34) Gavezzotti, A., Calculation of lattice energies of organic crystals: the PIXEL integration method in comparison with more traditional methods. *Z. Kristallogr. Cryst. Mater.* **2005**, 220, 499-510.

(35) Gavezzotti, A., Efficient computer modeling of organic materials. The atom–atom, Coulomb–London–Pauli (AA-CLP) model for intermolecular electrostatic-polarization, dispersion and repulsion energies. *New J. Chem.* **2011**, 35, 1360-1368.

(36) Reeves, M. G.; Wood, P. A.; Parsons, S., MrPIXEL: automated execution of Pixel calculations via the *Mercury* interface. *J. Appl. Crystallogr.* **2020**, 53, 1154-1162.

(37) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, 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.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; 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, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., Gaussian 09, Revision E. 01, Gaussian. *Inc:, Wallingford CT* **2009,** 201.

(38) Moggach, S. A.; Parsons, S.; Wood, P. A., High-pressure polymorphism in amino acids. *Crystallogr. Rev.* **2008**, 14 (2), 143-184.

(39) Bryant, M. J.; Maloney, A. G. P.; Sykes, R. A., Predicting mechanical properties of crystalline materials through topological analysis. *CrystEngComm.* **2018**, 20 (19), 2698-2704.

(40) Novelli, G.; Maynard-Casely, H. E.; McIntyre, G. J.; Warren, M. R.; Parsons, S., Effect of High Pressure on the Crystal Structures of Polymorphs of L-Histidine. *Cryst. Growth Des.* **2020**, 20, 7788-7804.

(41) Moggach, S. A.; Marshall, W. G.; Rogers, D. M.; Parsons, S., How focussing on hydrogen bonding interactions in amino acids can miss the bigger picture: a high-pressure neutron powder diffraction study of ε-glycine. *CrystEngComm.* **2015**, 17 (28), 5315-5328.

(42) Gonzalez-Platas, J.; Alvaro, M.; Nestola, F.; Angel, R., EosFit7-GUI: a new graphical user interface for equation of state calculations, analyses and teaching. *J. Appl. Crystallogr.* **2016**, 49, 1377-1382.

(43) Funnell, N. P.; Dawson, A.; Francis, D.; Lennie, A. R.; Marshall, W. G.; Moggach, S. A.; Warren, J. E.; Parsons, S., The effect of pressure on the crystal structure of L-alanine. *CrystEngComm.* **2010**, 12, 2573-2583.

(44) Ohashi, Y.; Burnham, C. W., Clinopyroxene lattice deformations. Roles of chemical substitution and temperature. *Am. Mineral.* **1973**, 58 (9-10), 843-849.

(45) Hazen, R. M.; Finger, L. W., *Comparative Crystal Chemistry*. John Wiley & Sons, **1982**.

(46) Broadhurst, E. T.; Wilson, C. J. G.; Zissimou, G. A.; Nudelman, F.; Constantinides, C. P.; Koutentis, P. A.; Parsons, S., A first-order phase transition in Blatter's radical at high pressure. *Acta Crystallogr.* **2022**, B78 (2), 107-116.

(47) Funnell, N. P.; Allan, D. R.; Maloney, A. G. P.; Smith, R. I.; Wilson, C. J. G.; Parsons, S., Suppression of isotopic polymorphism. *CrystEngComm.* **2021**, 23, 769-776.

This page has been left intentionally blank

Chapter 5: The Advantages of Flexibility: The Role of Entropy in Crystal Structures Containing C—H···F Interactions.

Overview

Molecular crystal structures are often interpreted in terms of strong, structure directing, intermolecular interactions, especially those with distinct geometric signatures such as H-bonds or π-stacking interactions. Other interactions, whether they are weak or lack a characteristic geometry, can be overlooked. Within this Chapter it is demonstrated that although the cumulative effect of weak interactions is significant, their deformability also leads to occupation of low energy vibrational energy levels which provides an additional stabilising entropic contribution. The entropies of five fluorobenzene derivatives have been calculated by periodic DFT calculations to assess the entropic influence of C—H···F interactions in stabilising their crystal structures. Calculations reproduce inelastic neutron scattering data and experimental entropies from heat capacity measurements. C—H···F contacts are shown to have force constants which are around half of those of more familiar interactions such as hydrogen bonds, halogen bonds and C—H···π interactions. This feature, in combination with the relatively high mass of F, means that the lowest energy vibrations in crystalline fluorobenzenes are dominated by C—H···F contributions. C—H···F contacts occur much more frequently than would be expected from their enthalpic contributions alone, but at 150 K the stabilising contribution of entropy provides, at −10 to −15 kJ mol^{-1}, a similar level of stabilisation to the N—H···N hydrogen bond in ammonia and O—H···O hydrogen bond in water.

5.1 Introduction

As described in Chapter 1, the structures formed by materials under different conditions are the consequence of minimisation of free energy, which can be interpreted as the result of a competition between enthalpic and entropic contributions. For example, sublimation, a change in phase from solid to gas, occurs, to a first approximation, when the temperature is high enough for the entropy of the gaseous state to overcome the enthalpic contribution made by the intermolecular interactions in the solid. Likewise, the crystal structures adopted by a compound are the result of a competition between the enthalpic and entropic characteristics of the intermolecular interactions.

The directing influence of strong intermolecular interactions such as hydrogen bonds in the determination of stable packing configurations in crystal structures have already been described. Motifs, such as the familiar dimers formed between carboxylic acid groups, occur so frequently and consistently that they are considered to be supramolecular synthons that can be used in strategies to engineer crystal structures with pre-determined architectures.[1] Such interactions typically have energies in the range of 10s kJ mol^{-1}.[2] The role of weak interactions is more subtle and open to different interpretations. Amongst these are weak C—H···halogen interactions, sometimes referred to as a class of weak hydrogen bond,[3] which have molecule-molecule energies well below 10 kJ mol^{-1}.

A study by Taylor has investigated the importance of weak C—H···X (X = O, N, F, Cl) interactions in crystal packing.[4] The frequency of formation of specific interactions was evaluated in the context of molecular surface area analysis to assess the extent to which interactions are formed because of their favourable characteristics as opposed to being formed randomly. A metric R_F for primary (shortest) interactions was defined and used to measure the

favourability of different classes of interaction, with values greater than unity indicating greater than random occurrence.

Charge density analysis has shown that there is an intrinsic polarisation of the electron density on the fluorine atoms[5] which promotes its participation in C—H···F, C—F···F—C, and C—F···π interactions in the solid state,[6,7] though overall, these interactions have weak, largely dispersion based, rather than electrostatic, enthalpic contributions. Nevertheless, the value of R_F for C—H···F intermolecular interactions in Taylor's study was, at 3.5, strikingly higher than either C—H···O (2.7) or C—H···N (3.0) interactions, which are occasionally also suggested in lists of supramolecular synthons.[3] The aim of this Chapter is to offer an explanation of why C—H···F interactions should occur so frequently.

It has been argued[8-11] that although individual C—H···F interactions are energetically weak, the presence of many of them becomes significant through a co-operative effect. Although these interaction energies are clearly significant, the free energy of a crystal structure is also influenced by other contributions which can exert a decisive influence on phase stability. For example, as has been seen in Chapter 3, volume minimisation is the most important driving force of high-pressure phase transitions, in which the change in lattice energy is often found to be positive.[12] A computational study on the role of entropy by Nyman and Day on 1061 experimentally determined structures of 508 polymorphic organic molecules has shown that vibrational contributions dominated by entropy differences contribute to significant differences in free energy. These were shown to be large enough in 9% of their polymorph pairs to cause a re-ranking of polymorph stability.[13] Kieslich *et al.* have also demonstrated the importance of entropy in hybrid organic-inorganic perovskites as a balance and interplay with hydrogen bonding interactions.[14]

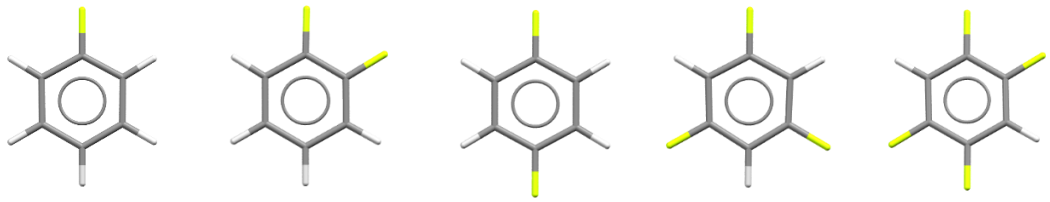
The ability of entropy to support negative volume changes in the presence of positive lattice energy changes has also been noted in a high-pressure polymorph of salicylamide.[15]

In this Chapter it is suggested on the basis of calculations on a series of fluorobenzene derivatives (Figure 5.1) that the shape and energetic characteristics of the potentials of C—H···F interactions leads to a high deformability and low vibrational frequencies, so that crystal structures in which they occur are favoured not only by the enthalpic influence of their interaction energies but also by entropy.

5.2 Materials and Methods

5.2.1 Structures Studied

Calculations were carried out on five fluorobenzene derivatives: monofluorobenzene (Cambridge Structural Database[16] refcode FACFAQ, **1F** hereafter), 1,2-difluorobenzene (FACFOE, **1,2F**), 1,4-difluorobenzene (FACGEV, **1,4F**), 1,3,5-trifluorobenzene (PVVAWA01, **1,3,5F**) and 1,2,4,5-tetrafluorobenzene (FACJAU, **1,2,4,5F**). These structures, which were all published in a landmark study by Thalladi *et al.*, were used to define the initial starting geometries for the DFT optimisations described below.[10] The five structures selected from Thalladi *et al*'s study sample different classes of interaction geometry, while minimising computational expense by choosing in all cases except **1,2F**, molecules that occupy special positions enabling symmetry to be exploited in the calculations.



FACFAQ	FACFOE	FACGEV	PVVAWA01	FACJAU
1F	**1,2F**	**1,4F**	**1,3,5F**	**1,2,4,5F**

Figure 5.1: The fluorobenzene derivatives investigated in this study and the definitions of the compound codes used in the text.

5.2.2 DFT Geometry Optimisation

The crystal structures of the compounds shown in Figure 5.1 were geometry-optimised using periodic density functional theory (DFT) with CASTEP

through the Materials Studio interface.[17,18] The Perdew-Burke-Ernzerhof (PBE)[19] exchange-correlation functional and norm-conserving pseudopotentials[17] were used with a basis set cut-off energy of 990 eV and a **k**-point spacing of 0.04 $Å^{-1}$, which converged the total energy to <0.1 meV per atom. The unit cell dimensions were fixed in the initial cycles of optimisation, but then allowed to vary along with the coordinates. The convergence criteria were: 1×10^{-8} eV $atom^{-1}$ (energy), 0.002 eV $Å^{-1}$ (force) and 1×10^{-4} Å (displacement). The cell optimisation was conducted in a soft compressibility scheme with fixed basis quality and a maximum stress of 0.01 GPa. Structures were visualised using both Mercury [20] and Avogadro.[21]

5.2.3 Phonon Calculations

Following geometry optimisation, vibrational frequencies were calculated in the harmonic approximation at the Γ-point and across the full Brillouin zone using the linear response method (*i.e.* density functional perturbation theory rather than the finite displacement method based on supercells).[22] Anharmonicity and thermal expansion were neglected. For calculations across the Brillouin zone the **q**-vector separation was set to 0.04 $Å^{-1}$. The number of **q**-vectors (*N*) at this density was 3 for **1F**, 8 for **1,2F**, 14 for **1,4F**, 24 for **1,3,5F** and 16 for **1,2,4,5F**. The frequencies of the acoustic modes at the Γ-point were within 0.06 cm^{-1} of 0 cm^{-1}; all other frequencies were positive. Density of states (DoS) plots were calculated for all structures with the contributions of different atom types assessed at the Γ-point by use of partial density of states (PDoS) plots, this follows the method used by Kieslich *et al*.[14]

The effect of interpolation of the **q**-vector grid onto a finer grid spacing was investigated.[23] For **1F** the phonon interpolation grid was increased stepwise from $1 \times 1 \times 1$ to $16 \times 16 \times 16$, converging the calculated value of

$-TS$ at 150 K to within 0.01 kJ mol^{-1} (Appendix C, Section C.3, Table A.6). A 16 × 16 × 16 interpolation grid was used for all further calculations in this work.

Phonon calculations of increased computational expense were also applied to **1F**. The **q**-vector separation in this calculation was decreased from 0.04 to 0.03 Å^{-1}, increasing N from 3 to 12. This led to a modestly increased occupation of the low energy vibrations, but not such as to improve substantially the simulation of inelastic neutron scattering (INS) data (Appendix C, Section C.3, Figures A.7 and A.8). The value of $-TS$ at 150 K also changed only modestly, from −10.31 to −10.42 kJ mol^{-1}. The results of the calculations at the coarser spacing with interpolation were therefore accepted for all other derivatives in this study. Further information on convergence analysis of the phonon calculations can be found in Section C.3 of the Appendix.

Note: INS measurements were completed by Professor Simon Parsons of the University of Edinburgh and Jeff Armstrong of ISIS in 2017. Experimental details are included here for completeness only.

5.2.4 INS Measurements

Comparison of the experimental and simulated inelastic neutron scattering (INS) spectra was used to assess the accuracy of phonon calculations. Experimental INS data were measured at ISIS Neutron and Muon spallation source on the TOSCA indirect geometry spectrometer. Liquid samples for all compounds were obtained from Fluorochem. The samples were cold-ground under nitrogen to produce solid powders before being loaded into cold aluminium cans and sealed.[24] Spectra were recorded at ~20 K, the base

temperature of the instrument. Spectra for all samples were compared to simulated spectra obtained from phonon calculations by use of the program AbINS in Mantid.[25] The temperatures used in the simulations were set to match the thermocouple readings taken from the experimental data files, and data from the front and the back detectors of TOSCA were merged.

5.2.5 Lattice and Intermolecular Interaction Energies.

The lattice and intermolecular interaction energies for each optimised structure were calculated via the semi-empirical Pixel method[26-28] using the MrPixel interface.[29] Gaussian-09[30] was used for electron density calculations at the MP2 level of theory using the 6-31G** basis set. Molecular electron density was calculated on a grid of 0.08 × 0.08 × 0.08 $Å^3$ with a condensation level of 4 and a cluster radius of 14 Å. Component energies were calculated for electrostatic, polarisation, dispersion and repulsion interactions with total energy taken as the sum of all contributions. Contacts (labelled A, B…) within the first coordination sphere of **1F** and **1,2F** are listed along with their energies in Table 5.1. Data for all dimers are available in Appendix Table A.8, with individual dimers shown in Figures A.10-14 in the Appendix, Section C.5.

5.2.6 Entropy Calculations

Vibrational entropies were calculated for all compounds from the frequency output of the phonon calculations across the Brillouin zone using routines available in CASTEP. Estimates of the entropy determined at the Γ-point, which are used below for illustrative purposes, were calculated as in Section C.9 of the Appendix. Although the experimental structures were determined

at between 123 and 215 K, all entropies were calculated at 150 K to facilitate direct comparison.

5.2.7 Force Constants

Force constants were calculated for selected dimers within the first coordination sphere of **1,2F** in order to assess the relative deformability of C—H···F interactions. The dimers selected for **1,2F** were a C—H···π interaction (referred to below as contact A, see Figure A.11 in Section C.5 of the Appendix) and a bridged C—H···F interaction (contact C). The coordinates were taken from the DFT-optimised structure. For the purposes of comparison, the same procedure was applied to the H-bonded water dimer with a geometry taken from the Benchmark Energy and Geometry Database from dataset S22.[31,32]

Calculation of the force constants was accomplished by evaluating the energy of a dimer as its geometry was distorted. Two methods were applied. In the first method, dimers were distorted along the eigenvectors of Γ-point phonons using the program ModeFollow.[33] Modes were selected from within the low energy envelope of vibrations below 200 cm^{-1}, which contribute most to entropy.[34,35] For the bridged C—H···F interaction (contact C) three separate modes were selected, two modes provide the highest distortion of the C—H···F interaction and the third as the highest energy mode below the 200 cm^{-1} limit for external vibrations. These modes are numbers 16, 22 and 24 in Table A.9 in Section C.9 of the Appendix. The dimer of interest was isolated from the unit cell at each step along the distortion and the energy calculated for the isolated dimer (see below). The extent of distortion (R) for each mode was defined numerically as the average distance between atoms in their distorted state relative to their optimised positions. The same modes were studied for the C—H···π interaction (contact A) for comparison. Mode 16 for contact A

and mode 24 for contact C produced very flat potentials, indicating that the energies of these modes are determined by the distortions of interactions with other molecules in the crystal structure; these modes were excluded from further analysis.

The second method of distortion followed the procedure of Carlucci *et al*.[36] This method was used for **1,2F** contacts A, C, and the water dimer. The interaction axis of each dimer was aligned along the Cartesian z axis using Avogadro.[21] For both contacts A and C the centroids of the aromatic rings were calculated to define this axis. For the water dimer the axis was defined along the shortest H···O intermolecular distance. The separation distance of the two molecules were then linearly distorted along z by +1.0 and −1.0 Å from its optimised value in steps of 0.1 Å. Energies were calculated at each point.

The dimer energies in both sets of calculations were evaluated using symmetry adapted perturbation theory (SAPT) in the program Psi4.[37,38] Convergence testing was carried out using contact C. All combinations of double, triple and quadruple zeta basis sets with the augmentation levels of August, July and June from the 'calendar' set (reducing backwards in month for each diffuse shell removed from the calculation) were tested at all levels of SAPT truncation.[39] On the basis of these tests, SAPT2+3 was selected with aug-cc-pVDZ basis set for further calculations; further data are available in Section C.4 of the Appendix.

The above calculations yielded potential curves $E(R)$ for each interaction studied. The force constants, k^0, for the dimers were calculated as in Carlucci *et al*.[36] using Equation 5.1 after fitting polynomials to the potential curves; it was found found that order six polynomials yielded better fits than the fourth order set used in Carlucci *et al*.[36]

$$k^0 = \frac{d^2E}{dR^2}. \tag{5.1}$$

5.3 Results and Discussion

5.3.1 Intermolecular Interactions

Descriptions of the packing in all the fluorobenzene derivatives studied here (Figure 5.1) have been provided by Thalladi *et al*.[10] The energies and centroid-centroid distances of the interactions formed in each structure, as determined by the Pixel method are given in Tables 5.1 (**1F** and **1,2F**) and A.8 in Section C.5 of the Appendix (all derivatives). The structures contain 12 (**1,2,4,5F**), 13 (**1,2F**) or 14 (**1F**, **1,4F** and **1,3,5F**) molecules in the first coordination sphere, linked by C—H···F, C—H···H and C—F···F contacts which are in some cases bifurcated or bridged. The structures of **1F**, **1,2F** and **1,4F** also contain C—H···π interactions, whereas those of **1,3,5F** and **1,2,4,5F** feature offset π···π interactions.

The lattice energies are similar, varying between −45.8 and −52.0 kJ mol^{-1}. The modest values[27] reflect the low melting points of these materials. The π···π or C—H···π interactions are more stabilising than single, bridged, or bifurcated C—H···F interactions in all structures except **1,4F** where a bridged C—H···F interaction provides 0.5 kJ mol^{-1} greater stabilisation than a C—H···π interaction.

The largest contribution to the C—H···F interactions is dispersion rather than electrostatics. This finding is in agreement with work by Sudheendranath *et al.* on α-fluoroketones, where the dispersive term was found to outweigh the electrostatic by 3-4 times, and also with the survey of C, H and F-containing structures by Gavezzotti and Lo Presti.[40,41] Desiraju *et al.* have similarly reported that C—H···F interactions, which they have described as weak hydrogen bonds, have less pronounced electrostatic contributions.[42]

Although the aromatic nature of structures studied here contribute towards dispersion terms in these contacts, a search of the Cambridge Structural Database (CSD)[16] reveals that average contact lengths for C—H···F bonds hardly vary for aromatic (2.512 Å) and non-aromatic (2.519 Å) structures (see Section C.2 of the Appendix), indicating that formation of these contacts is not simply a feature of optimisation of aromatic interactions. Relevant interactions were defined in this search as those shorter than the sum of the contributing van der Waals radii. Database entries without coordinates, *R*-factors above 5%, or those containing disorder, errors, powder data or ions were filtered out. Hydrogen bond lengths were normalised to neutron values. Searches were conducted on the CSD version 2023.2.0.

The low energies of C—H···F interactions, as well as their lower or similar energies to π···π or C—H···π contacts, suggests that it is difficult on energy grounds alone to support their description[9-11] as structure directing interactions. Nevertheless, their persistence in the crystal structures of organo-fluorine derivatives implies that they must make a significant stabilising contribution.[4] The aim of the following sections is to explore the role that entropy has in resolving this apparent paradox.

Table 5.1: Symmetry unique interactions within the first coordination spheres of **1F** and **1,2F**. All energies are given in kJ mol^{-1}. U = lattice energy. Data for other derivatives and diagrams of each dimer are available in the Appendix (Section C.5, Table A.8, Figures A.10-14).

Structure	Dimer	Symmetry	Contact Type	Centroid separation	Coulombic	Polarisation	Dispersion	Repulsion	Total	Total per molecule
1F, FACFAQ	A	$y \pm 1/2, -x + 5/2, z + 1/4$ $-y + 5/2, x \pm 1/2, z - 1/4$	H···π	4.759	−3.6	−1.8	−14.5	11.0	−8.9	−4.5
$P4_32_12$	B	$x, y \pm 1, z; x \pm 1, y, z$	H···F, H···H	5.752	−1.7	−0.8	−9.1	4.5	−7.0	−3.5
Z = 4,	C	$y \pm 1/2, -x + 3/2, z + 1/4$ $-y + 3/2, x \pm 1/2, z - 1/4$	H···F, H···H	6.542	−2.6	−1.2	−7.6	5.0	−6.4	−3.2
Z' = ½	D	$x \pm 1, y \pm 1, z$	H···F	8.134	−1.3	−0.2	−1.5	0.2	−2.8	−1.4
U = −52.0										
1,2F,	A	$-x + 3/2, y \pm 1/2, -z + 3/2$	H···π	4.686	−5.1	−1.8	−15.4	11.2	−11.1	−5.6
FACFOE	B	$-x + 5/2, y \pm 1/2, -z + 3/2$	H···π	4.843	−4.1	−1.3	−12.6	7.4	−10.6	−5.3
$P2_1/n$	C	$x \pm 1/2, -y + 1/2, z \pm 1/2$	H···F, H···F	6.782	−4.7	−1.1	−6.9	4.3	−8.5	−4.3
Z = 4	D	$-x + 2, -y + 1, -z + 1$	F···F, F···F	5.864	−0.7	−0.8	−8.5	3.2	−6.8	−3.4
Z' = 1	E	$-x + 2, -y + 1, -z + 2$	H···H···H	5.954	−1.3	−0.7	−7.6	3.4	−6.2	−3.1
U = −49.8	F	$x, y \pm 1, z$	H···F	6.007	−0.6	−0.9	−8.6	4.2	−5.9	−3.0
	G	$x \pm 1/2, -y + 3/2, z \pm 1/2$	H···F, H···H	6.935	−1.4	−0.7	−5.2	2.8	−4.5	−2.3
	H	$-x + 2, -y, -z + 1$	F···F	8.268	1.5	−0.2	−1.7	0.6	0.2	0.1

5.3.2 Phonon Calculations

5.3.2*a* Inelastic Neutron Scattering Experiments

The entropy of crystalline solids is determined by access to low frequency vibrational (or phonon) energy levels. In disordered solids there is also a contribution from alternative configurations but since all structures studied here are ordered this was not considered in the present Chapter. In order to explore the role of entropy in the stabilisation of crystal structures containing C—H···F interactions it is thus necessary to calculate vibrational frequencies. In crystalline solids, internal vibrations occur within molecules and external vibrations occur between molecules. The relative phases of vibrational modes between unit cells defines different points in the Brillouin zone. The point where the vibrations of all unit cells are in phase defines the Γ-point; the opposite zone edge, which is given a variety of different symbols depending on crystal symmetry, occurs where vibrations in neighbouring unit cells are out of phase.

Calculation of the vibrational properties of solids is computationally expensive, and approximations such as the assumption of harmonicity are usually invoked. This assumption was made in the present Chapter. The effects of thermal expansion were also neglected. More elaborate models could be based on the quasi-harmonic approximation to account for thermal expansion, or explicit inclusion of anharmonicity. Destabilising (positive) contributions to free energy, namely the increase in internal energy due to the occupation of vibrational modes as temperature increases and zero-point energy, were also neglected. These quantities can differ significantly in crystal structures with very strong intermolecular interactions,[34] but differences

between polymorphs are usually very small;[13] further comments are available in Section C.10 of the Appendix.

Inelastic neutron scattering data, which provides very useful benchmarking data for phonon spectra calculated by periodic DFT,[25] were collected in order to determine the impact of these assumptions on the accuracy of the phonon calculations. Calculated and experimental INS spectra in the range 0-1000 cm^{-1} are presented in Figure 5.2 for **1F**. Spectra for all structures and spectra over larger ranges can be found in Figures A.15-19 in Section C.6 of the Appendix. The intensities of measured INS spectra depend on experimental parameters such as the sample volume and incident beam intensity as well as the characteristics of the vibrations, and the simulated intensities were scaled visually to match the peaks near 1000 cm^{-1}. In all cases, the simulated spectra show a level of agreement with the experimental data that is typical in comparable studies employing periodic DFT and the harmonic approximation.[43,44] The frequencies and intensities of the internal modes are well reproduced. The external modes below 200 cm^{-1} show small frequency offsets which reflect the assumptions described above and which tend to lead to overestimation of the frequencies.[45] Nevertheless, the essential features of the INS spectra below 200 cm^{-1} have been reproduced. Inclusion of thermal expansion, for example by use of a quasi-harmonic model,[46] would be expected to decrease the frequencies,[47] while inclusion of anharmonic effects usually, though not always, also leads to a decrease in calculated frequencies when compared to the harmonic approximation.[48-51]

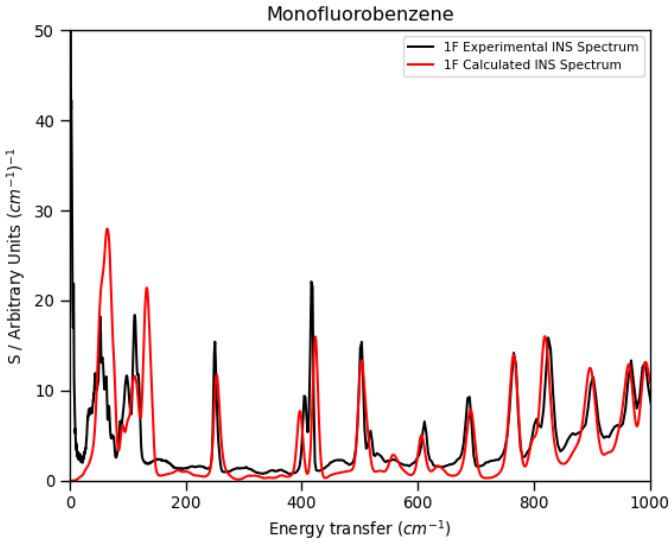


Figure 5.2: Experimental and simulated INS spectra for **1F**. Spectra for other compounds and larger spectral ranges can be found in Figures A.15-19 in SectionC.6 of the Appendix.

5.3.2*b* Entropy Calculations

The significance of the offsets between the experimental and DFT simulations in the low frequency vibrational modes can be assessed by comparison of the calculated entropies with experimental values which are available from heat capacity (C_P) data for **1F**, **1,2F** and **1,2,4,5F**.[52-54]

The values of $-TS$ (T = temperature, S = entropy) terms for all five compounds at 150 K are listed in Table 5.2 alongside experimental values where available. The value of 150 K used here and below is representative of the temperatures of the experimental structure determinations. The experimental and simulated values of $-TS$ in Table 5.2 all agree to 2 kJ mol^{-1}, the underestimation in the simulated values being consistent with the slight over-estimation in the positions of the low frequency bands shown in Figure 5.2. The calculations also recover the increasing trend with fluorine content.

Table 5.2: Values of −*TS* at 150 K for fluorobenzenes calculated using periodic DFT at the Γ-point and across the Brillouin zone compared to available data determined experimentally from heat capacity measurements. All values in kJ mol^{-1}.

	Γ Phonon	Full Phonon	Experimental
1F	−8.63	−10.31	−12.42[42]
1,2F	−9.99	−11.73	−13.72[43]
1,4F	−8.58	−12.33	
1,3,5F	−9.83	−13.85	
1,2,4,5F	−11.40	−15.05	−15.77[44]

Thus, although the agreement is not perfect, the INS and entropy benchmarking data indicate that the level of theory applied in this work is suitable for the purpose of estimating the magnitude of entropy effects in weakly-bound crystal structures such as those of fluorobenzenes.

5.3.2*c* Partial Density of States Calculations

The values of −10 to −15 kJ mol^{-1}forthe −*TS* contribution to the stabilisation of fluorobenzenes at 150 K shown in Table 5.2 is comparable to the enthalpic stabilisation afforded by the N—H···N and O—H···O hydrogen bonds in the crystal structure of ammonia (−9.3 kJ mol^{-1}) and the optimised water dimer (−20.6 kJ mol^{-1}).[55,56] But what contribution is made by C—H···F interactions towards this stabilisation?

The phonon density of states, shown in Figure 5.3 for **1F**, were decomposed into contributions from each atom type at the Γ-point; plotting these partial density of states reveals how individual modes are distributed over different atoms.[14] The advantage of using the Γ-point is that the contribution of entropy by bands of increasing frequency is easily calculated

by hand (Section C.9 and Table A.9 of the Appendix), providing useful insight into the importance of low energy vibrations; note the entropy calculated in this way is underestimated because the contributions of acoustic phonons are not calculated (see Table 5.2).

The partial density of states for **1F** are shown for F, C and H in Figure 5.3, with the remaining compounds shown in Figures A.25-29 in Section C.8 of the Appendix. For all structures, the fluorine atoms contribute significantly more per atom to the modes below 200 cm^{-1} than C or H, as indicated by the higher intensity of the peaks within this range for the fluorine atoms.

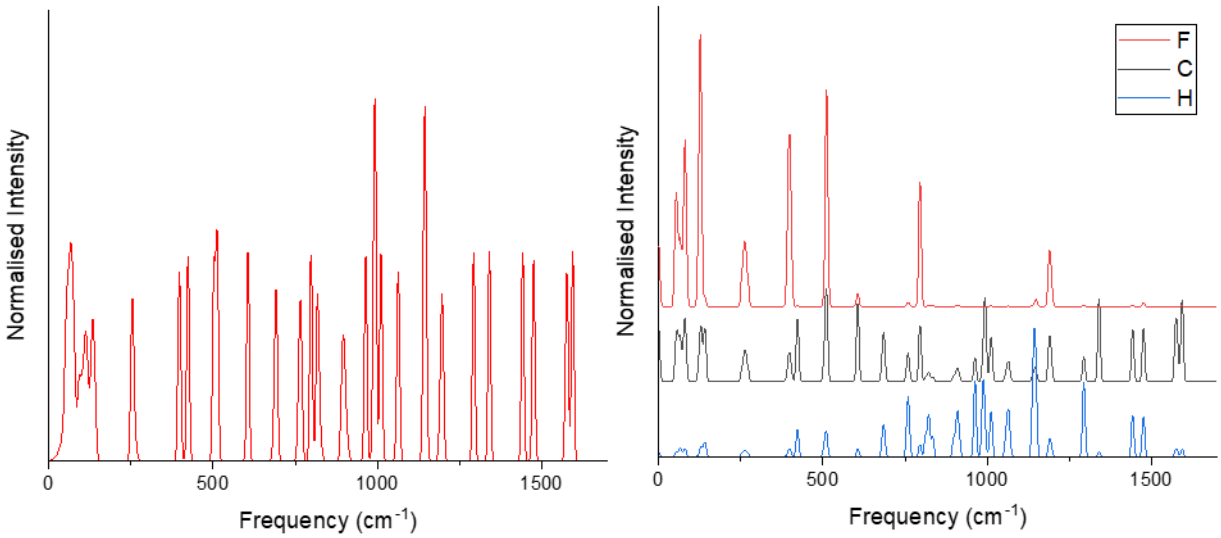


Figure 5.3: Left) Density of states for **1F** in the range 0-1700 cm^{-1}. Right) Partial density of states per atom type per atom for **1F**. Red is fluorine, black is carbon and blue is hydrogen. Data for individual atom types have been offset for clarity. Plots for all other structures and larger spectral ranges are available in Figures A.20-29 in Sections C.7 and C.8 of the Appendix.

5.3.2*d* Force Constant Calculations

The thermodynamic contribution made by an intermolecular interaction to the stability of a phase depends on the characteristics of its potential. The enthalpic significance depends on the depth of the potential. The entropic significance

depends on the shape of the potential at its minimum. A low curvature at the potential minimum is indicative of a deformable interaction with a low force constant (Equation 5.1). An interaction with a low force constant is associated with a low vibrational frequency, and hence a high entropy. However, vibrational frequencies also depend on mass, and low frequencies are also associated with vibrations which involve heavy atoms.

The partial density of states analysis indicates that the motions of the fluorine atoms contribute most to the lowest energy vibrations in fluorobenzenes and therefore contribute most to entropy. The question addressed in the following section is whether this is simply because fluorine is the heaviest atom present, or whether the low energy of vibrations involving fluorine can also be ascribed to the low force constants, *i.e.*, the deformability, of C—H···F interactions. Answering this question requires explicit calculation of the potentials of intermolecular interactions involving C—H···F contacts, along with those of other classes of interaction for the purposes of comparison.

Potentials were calculated for contacts selected from the crystal structure of **1,2F**. Contact C (Table 5.1, Figure A.11 in Appendix C.5) is formed across an inversion center and contains two bridged C—H···F interactions; for comparison the same calculations were carried out for the C—H···π interaction, contact A. The force constants for both contacts were calculated by two methods as described in Section 5.2.7. For comparison, the second method was applied to the isolated water dimer; the same method has been applied to the pentafluoroiodobenzene-pyrazine C—I···N halogen bond in ref. [36] The potentials and force constants are shown in Figure 5.4 and Table 5.3, which also reports separation distances and energies from the optimised structure alongside equilibrium distances and cohesive energies.

For the **1,2F** modes, the minimum of the fitted polynomial was slightly shorter than the separation within the optimised structure, a reflection of interaction distances in crystal structures being the result of a competition between the pushes and pulls of all contacts present, so that none adopt exactly the same geometry as an isolated dimer.[57] The C—H···π interaction (A) is stronger than the bridged C—H···F interaction (C) with equilibrium energies of −15.8 and −9.4 kJ mol^{-1}, respectively (Figure 5.4; the corresponding values in the crystal structure calculated by the Pixel method are −11.1 and −8.5 kJ mol^{-1}). The potentials for contact C in Figure 5.4 are distinctly shallower than for contact A, producing substantially lower force constants. The average of the of the force constants calculated by method 1 are 162.4 kJ mol^{-1}Å^{-2} and 40.8 kJ mol^{-1}Å^{-2} for contacts A and C respectively. Although the results of method 2 are different (they correspond to different distortions), they present a consistent picture: 55.6 and 31.7 kJ mol^{-1}Å^{-2} for A and C, respectively. Comparative figures for the hydrogen and halogen bonded dimers in Table 5.3 are both of the order of 100 kJ mol^{-1}Å^{-2}.

The energies of the interactions discussed here lie between about 5 and 20 kJ mol^{-1}; the energetic differences are accompanied by a distribution in the range of values for the force constants, with those for C—H···F interactions being significantly lower than those of other contacts. These data indicate that the entropic contribution of C—H···F interactions stems not only from the mass of the fluorine atom, but also the deformability of the interactions it forms.

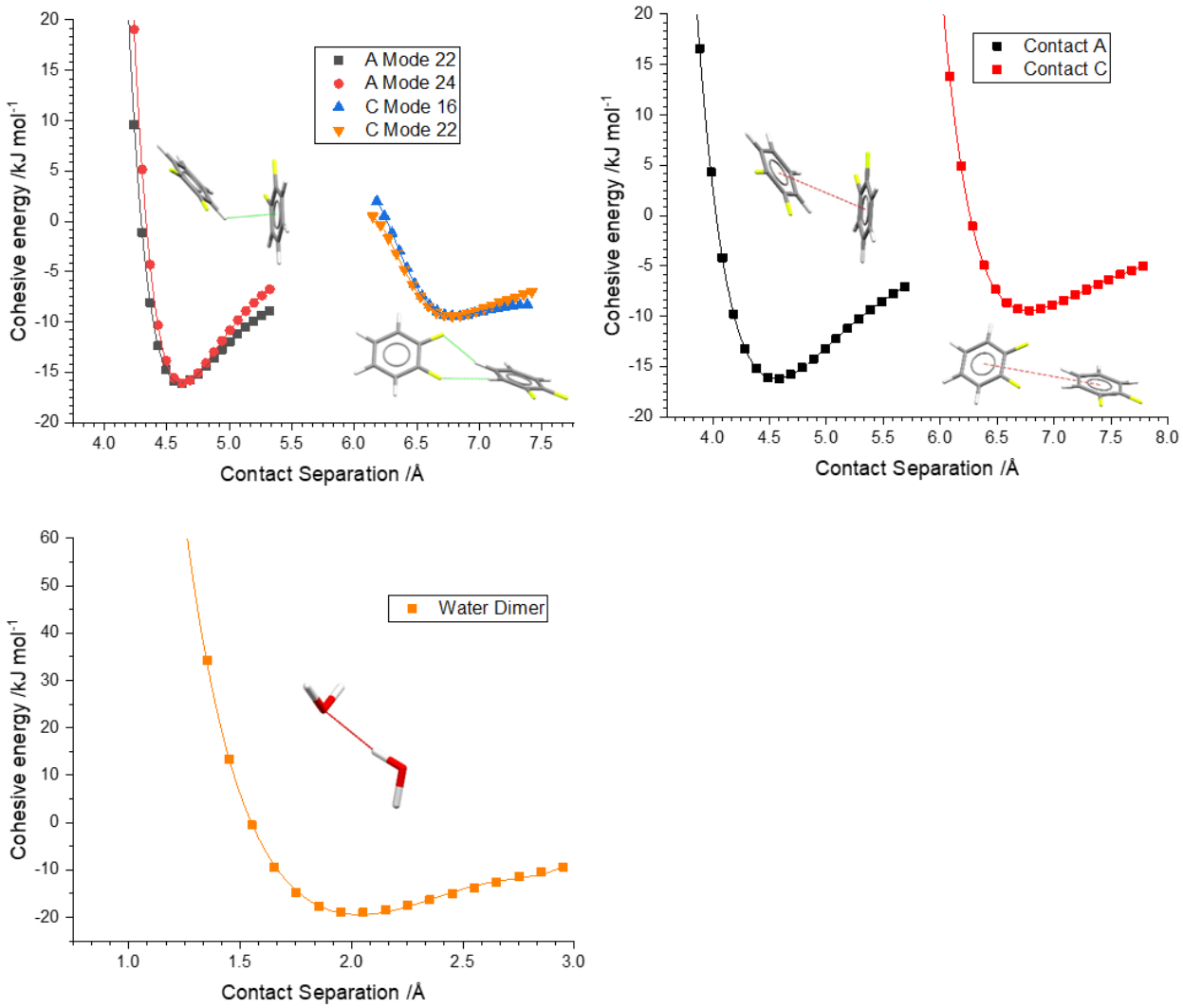


Figure 5.4: Energy potentials for: Top left) Distortions along phonon modes for contacts A and C for **1,2F** (method 1). Top right) Distortions of centroid-centroid distances for contact A and C in **1,2F** (method 2). Bottom) Distortion of the O···H distance in the water dimer (method 2). Red dashed lines represent defined interaction axes for method 2.

Table 5.3: The equilibrium distances (R_{opt}) and energies (E_{opt}) from the DFT-optimised structures and the minimum distances (R^0), energies (E^0), and force constants (k^0) for selected contacts in **1,2F** and water from the fitted potentials. Data for the C—I···N interaction in the pentafluoroiodobenzene-pyrazine dimer (**FBz-I**) taken from Carlucci *et al.* are listed for comparison.[36] Distances are in Å, energies in kJ mol^{-1} and force constants in kJ mol^{-1}Å^{-2}.

Contact/Mode	R_{opt}	E_{opt}	R^0	E^0	k^0
Method 1					
A/22	4.686	−15.813	4.615	−16.097	136.6
A/24	"	"	4.625	−16.103	188.2
Average			4.620	−16.100	162.4
C/16	6.782	−9.421	6.806	−9.445	38.5
C/22	"	"	6.787	−9.440	43.0
Average			6.797	−9.443	40.8
Method 2					
A	4.686	−15.813	4.556	−16.211	55.6
C	6.782	−9.421	6.772	−9.421	31.7
Water	1.952	−18.867	2.036	−19.455	94.2
FBz-I[36]	-	-	3.17	−9.9	97

5.4 Conclusions

The current work was inspired by a comment made by Taylor in a paper entitled *It Isn't, It Is: The C–H···X (X = O, N, F, Cl) Interaction Really Is Significant in Crystal Packing*.[4] The final paragraph of that paper reads: *It took me 30 years to be persuaded that C–H···F–C … contacts matter, but I remain of this view. Whether this belief is accepted by others or not, one thing is clear: any explanation of the crystal packing of the structures… must account for the fact that they contain many more X···H interactions than would be expected by chance.*

Although C—H···F interactions are weak, they occur frequently, 3.5 times more likely than would be expected by chance alone. Taylor reasoned that '*The simplest explanation is that X···H* (X being a halogen) *interactions are sufficiently favorable in themselves to ensure their over-representation in crystal structures*'. The hypothesis explored in the present Chapter is that the favourable feature of C—H···F interactions lies in their contribution to entropy.

The vibrational, enthalpic and entropic properties of five fluorobenzene derivatives have been investigated using periodic DFT, SAPT and Pixel calculations. The phonon calculations, which were carried out in the harmonic approximation and neglected the effects of thermal expansion, nevertheless adequately reproduced the experimental entropies, determined from heat capacity data, and INS spectra. The molecule-molecule energies associated with individual C—H···F interactions lie between −5 and −10 kJ mol^{-1}. The interactions are largely dispersive, rather than electrostatic in character. The total contribution of entropy is between −10 and −15 kJ mol^{-1} at 150 K, more akin to the enthalpic stabilisation made by a moderate hydrogen bond.

Entropy in ordered, crystalline solids is derived from access to low energy vibrational modes. Each mode is distributed over many different

interactions in solids, and one of the difficulties encountered in the work described here has been to assess the entropic contribution of interactions involving C—H···F contacts. One technique has been to examine partial phonon density of states plots, which show that the lowest energy modes in fluorobenzenes are dominated by the motions of fluorine atoms. Another has been to investigate the properties of the potentials of dimers mediated by C—H···F interactions and to compare them with those for other classes of contact. This analysis showed that the low vibrational energies associated with C—H···F interactions reflect the relatively high mass of the halogen atom but also the deformability of the interaction, with potentials which are flatter than other contacts such as C—H···π interactions and hydrogen and halogen bonds. The prevalence C—H···F interactions thus stems from a combination of the entropic advantage conferred by these two features as well as from the energies of the interactions themselves.

Taylor's study shows that many weak interactions form randomly and tend not to be structure directing. For example, enthalpically weak C—O···C and C—N···C interactions were found to both occur 0.5 times less than would be expected by chance. But C—H···F interactions do matter, the results presented here implying that entropy can have significant role to play in structures dominated by enthalpically weak interactions. The results of this work suggest that entropy can be an important consideration when interpretating the structural significance of weak interactions.

5.5 References

(1) Desiraju, G. R., Supramolecular Synthons in Crystal Engineering—a New Organic Synthesis. *Angew. Chem., Int. Ed. Engl.* **1995,** 34, 2311-2327.

(2) Dunitz, J. D.; Gavezzotti, A., Supramolecular Synthons: Validation and Ranking of Intermolecular Interaction Energies. *Cryst. Growth Des.* **2012,** 12, 5873-5877.

(3) Desiraju, G. R.; Steiner, T., *The Weak Hydrogen Bond*. Oxford University Press, **1999**.

(4) Taylor, R., It Isn't, It Is: The C–H··· X (X= O, N, F, Cl) Interaction Really Is Significant in Crystal Packing. *Cryst. Growth Des.* **2016,** 16, 4165-4168.

(5) Hathwar, V. R.; Chopra, D.; Panini, P.; Guru Row, T. N., Revealing the Polarizability of Organic Fluorine in the Trifluoromethyl Group: Implications in Supramolecular Chemistry. *Cryst. Growth Des.* **2014,** 14, 5366-5369.

(6) Das, P.; Rao, G. B. D.; Bhandary, S.; Mandal, K.; Seth, S. K.; Chopra, D., Quantitative Investigation into the Role of Intermolecular Interactions in Crystalline Fluorinated Triazoles. *Cryst. Growth Des.* **2024,** 24, 703-721.

(7) Dey, D.; Seth, S. K.; Mohan, T. P.; Chopra, D., Quantitative analysis of intermolecular interactions in crystalline substituted triazoles. *J. Mol. Struct.* **2023,** 1273, 134380.

(8) Chopra, D.; Row, T. N. G., Role of Organic Fluorine in Crystal Engineering. *CrystEngComm.* **2011,** 13, 2175-2186.

(9) Thakur, T. S.; Kirchner, M. T.; Bläser, D.; Boese, R.; Desiraju, G. R., C–H··· F–C Hydrogen Bonding in 1, 2, 3, 5-Tetrafluorobenzene and Other Fluoroaromatic Compounds and the Crystal Structure of Alloxan Revisited. *CrystEngComm.* **2010,** 12, 2079-2085.

(10) Thalladi, V. R.; Weiss, H.-C.; Bläser, D.; Boese, R.; Nangia, A.; Desiraju, G. R., C− H···F Interactions in the Crystal Structures of Some Fluorobenzenes. *J. Am. Chem. Soc.* **1998,** 120, 8702-8710.

(11) Desiraju, G. R., C–H···O and Other Weak Hydrogen Bonds. From Crystal Engineering to Virtual Screening. *Chem. Commun.* **2005**, 2995-3001.

(12) Wilson, C. J. G.; Cervenka, T.; Wood, P. A.; Parsons, S., Behavior of Occupied and Void Space in Molecular Crystal Structures at High Pressure. *Cryst. Growth Des.* **2022,** 22, 2328-2341.

(13) Nyman, J.; Day, G. M., Static and lattice vibrational energy differences between polymorphs. *CrystEngComm.* **2015,** 17, 5154-5165.

(14) Kieslich, G.; Skelton, J. M.; Armstrong, J.; Wu, Y.; Wei, F.; Svane, K. L.; Walsh, A.; Butler, K. T., Hydrogen Bonding versus Entropy: Revealing the Underlying Thermodynamics of the Hybrid Organic–Inorganic Perovskite $[CH_3NH_3]$ $PbBr_3$. *Chem. Mater.* **2018,** 30, 8782-8788.

(15) 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*. **2010,** 12, 1065-1078.

(16) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C., The Cambridge Structural Database. *Acta Crystallogr.* **2016,** B72, 171-179.

(17) Clark, S. J.; Segall, M. D.; Pickard, C. J.; Hasnip, P. J.; Probert, M. I. J.; Refson, K.; Payne, M. C., First principles methods using CASTEP. *Z. Kristallogr.* **2005,** 220, 567-570.

(18) BIOVIA, *Materials Studio*. Dassault Systèmes, **2021**.

(19) Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996,** 77, 3865.

(20) Macrae, C. F.; Sovago, I.; Cottrell, S. J.; Galek, P. T. A.; McCabe, P.; Pidcock, E.; Platings, M.; Shields, G. P.; Stevens, J. S.; Towler, M.; Wood, P. A., Mercury 4.0: From Visualization to Analysis, Design and Prediction. *J. Appl. Crystallogr.* **2020,** 53, 226-235.

(21) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R., Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminformatics.* **2012,** 4, 1-17.

(22) Refson, K.; Tulip, P. R.; Clark, S. J., Variational Density-Functional Perturbation Theory for Dielectrics and Lattice Dynamics. *Phys. Rev.* **2006,** B73, 155114.

(23) Michalchuk, A. A. L.; Rudić, S.; Pulham, C. R.; Morrison, C. A., Vibrationally Induced Metallisation of the Energetic Azide α-NaN_3. *Phys. Chem. Chem. Phys.* **2018,** 20, 29061-29069.

(24) Ibberson, R. M., A Simple Technique for Preparing Low-Melting-Point Samples for Neutron Powder Diffraction. *J. Appl. Crystallogr.* **1996,** 29, 498-500.

(25) Dymkowski, K.; Parker, S. F.; Fernandez-Alonso, F.; Mukhopadhyay, S., AbINS: The modern software for INS interpretation. *Physica.* **2018,** B551, 443-448.

(26) Gavezzotti, A., *Molecular Aggregation*. Oxford University Press, **2007**.

(27) Gavezzotti, A., Calculation of lattice energies of organic crystals: the PIXEL integration method in comparison with more traditional methods. *Z. Kristallogr. Cryst. Mater.* **2005,** 220, 499-510.

(28) Gavezzotti, A., Efficient computer modeling of organic materials. The atom–atom, Coulomb–London–Pauli (AA-CLP) model for intermolecular

electrostatic-polarization, dispersion and repulsion energies. *New J. Chem.* **2011,** 35, 1360-1368.

(29) Reeves, M. G.; Wood, P. A.; Parsons, S., MrPIXEL: automated execution of Pixel calculations via the *Mercury* interface. *J. Appl. Crystallogr.* **2020,** 53, 1154-1162.

(30) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, 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.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; 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, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., Gaussian 09, Revision E. 01, Gaussian. *Inc:, Wallingford CT* **2009,** 201.

(31) Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P., Benchmark Database of Accurate (MP2 and CCSD (T) Complete Basis Set Limit) Interaction Energies of Small Model Complexes, DNA Base Pairs, and Amino Acid Pairs. *Phys. Chem. Chem. Phys.* **2006,** 8, 1985-1993.

(32) Řezáč, J.; Jurečka, P.; Riley, K. E.; Černý, J.; Valdes, H.; Pluháčková, K.; Berka, K.; Řezáč, T.; Pitoňák, M.; Vondrášek, J.; Hobza, P., Quantum Chemical Benchmark Energy and Geometry Database for Molecular Clusters and

Complex Molecular Systems (www.begdb.com): A Users Manual and Examples. *Collect. Czech. Chem. Commun.* **2008,** 73, 1261-1270.

(33) Refson, K., *MODE_FOLLOW. A routine for generating structures distorted along phonon modes calculated using CASTEP*, Chilton, OXON, UK: ISIS Facility, Rutherford-Appleton Laboratory, **2014**.

(34) Rivera, S. A.; Allis, D. G.; Hudson, B. S., Importance of Vibrational Zero-Point Energy Contribution to the Relative Polymorph Energies of Hydrogen-Bonded Species. *Cryst. Growth Des.* **2008,** 8, 3905-3907.

(35) Armstrong, J.; Banerjee, S.; Schu̇nemann, V.; Wolny, J. A.; Sadler, P. J., Vibrational Motions Make Significant Contributions to Sequential Methyl C–H Activations in an Organometallic Complex. *J. Phys. Chem. Lett.* **2021,** 12, 658-662.

(36) Carlucci, L.; Gavezzotti, A., A Quantitative Measure of Halogen Bond Activation in Cocrystallization. *Phys. Chem. Chem. Phys.* **2017,** 19, 18383-18388.

(37) Jeziorski, B.; Moszynski, R.; Szalewicz, K., Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chem. Rev.* **1994,** 94, 1887-1930.

(38) Smith, D. G. A.; Burns, L. A.; Simmonett, A. C.; Parrish, R. M.; Schieber, M. C.; Galvelis, R.; Kraus, P.; Kruse, H.; Di Remigio, R.; Alenaizan, A.; James, A. M.; Lehtola, S.; Misiewicz, J. P.; Scheurer, M.; Shaw, R. A.; Schriber, J. B.; Xie, Y.; Glick, Z. L.; Sirianni, D. A.; O'Brien, J. S.; Waldrop, J. M.; Kumar, A.; Hohenstein, E. G.; Pritchard, B. P.; Brooks, B. R.; Schaefer, H. F., III; Sokolov, A. Y.; Patkowski, K.; DePrince, A. E., III; Bozkaya, U.; King, R. A.; Evangelista, F. A.; Turney, J. M.; Crawford, T. D.; Sherrill, C. D., PSI4 1.4: Open-Source Software for High-Throughput Quantum Chemistry. *J. Chem. Phys.* **2020,** 152, 184108.

(39) Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D., Levels of Symmetry Adapted Perturbation Theory (SAPT). I. Efficiency and Performance for Interaction Energies. *J. Chem. Phys.* **2014,** 140, 094106.

(40) Sudheendranath, A.; Sandeep; Avinash, K.; Venugopalan, P.; Kumar, A.; Thomas, S. P., Fluorine versus Hydroxyl in Supramolecular Space: Effect of Organic Fluorine in Molecular Conformation, Lattice Cohesive Energies, and Receptor Binding in a Series of α-Fluoroketones. *Cryst. Growth Des.* **2023,** 23, 4013–4024.

(41) Gavezzotti, A.; Lo Presti, L., Building Blocks of Crystal Engineering: A Large-Database Study of the Intermolecular Approach between C–H Donor Groups and O, N, Cl, or F Acceptors in Organic Crystals. *Cryst. Growth Des.* **2016,** 16, 2952-2962.

(42) Thakur, T. S.; Dubey, R.; Desiraju, G. R., Intermolecular Atom–Atom Bonds in Crystals– A Chemical Perspective. *IUCrJ.* **2015,** 2, 159-160.

(43) Hudson, M. R.; Allis, D. G.; Hudson, B. S., The Vibrational Spectrum of Parabanic Acid by Inelastic Neutron Scattering Spectroscopy and Simulation by Solid-State DFT. *J. Phys. Chem.* **2010,** A114, 3630-3641.

(44) Plazanet, M.; Fukushima, N.; Johnson, M. R.; Horsewill, A. J.; Trommsdorff, H. P., The Vibrational Spectrum of Crystalline Benzoic Acid: Inelastic Neutron Scattering and Density Functional Theory Calculations. *J. Chem. Phys.* **2001,** 115, 3241-3248.

(45) Armstrong, J.; O'Malley, A. J.; Ryder, M. R.; Butler, K. T., Understanding Dynamic Properties of Materials Using Neutron Spectroscopy and Atomistic Simulation. *J. Phys. Commun.* **2020,** 4, 072001.

(46) Baroni, S.; Giannozzi, P.; Isaev, E., Density-Functional Perturbation Theory for Quasi-Harmonic Calculations. *Rev. Mineral. Geochem.* **2010,** 71, 39-57.

(47) Schauer, A., Thermal expansion of solids and the temperature dependence of lattice vibration frequencies. *Can. J. Phys.* **1964,** 42, 1857-1864.

(48) Reilly, A. M.; Middlemiss, D. S.; Siddick, M. M.; Wann, D. A.; Ackland, G. J.; Wilson, C. C.; Rankin, D. W. H.; Morrison, C. A., The Phonon Spectrum of Phase-I Ammonia: Reassignment of Lattice Mode Symmetries from Combined Molecular and Lattice Dynamics Calculations. *J. Phys. Chem.* **2008,** A112, 1322-1329.

(49) Biczysko, M.; Panek, P.; Scalmani, G.; Bloino, J.; Barone, V., Harmonic and Anharmonic Vibrational Frequency Calculations with the Double-Hybrid B2PLYP Method: Analytic Second Derivatives and Benchmark Studies. *J. Chem. Theory Comput.* **2010,** 6, 2115-2125.

(50) Jacobsen, R. L.; Johnson, R. D., III; Irikura, K. K.; Kacker, R. N., Anharmonic Vibrational Frequency Calculations are not Worthwhile for Small Basis Sets. *J. Chem. Theory Comput.* **2013,** 9, 951-954.

(51) Maillard, R.; Sethio, D.; Hagemann, H.; Lawson Daku, L. M., Accurate Computational Thermodynamics Using Anharmonic Density Functional Theory Calculations: The Case Study of B–H Species. *ACS Omega.* **2019,** 4, 8786-8794.

(52) Scott, D. W.; McCullough, J. P.; Good, W. D.; Messerly, J. F.; Pennington, R. E.; Kincheloe, T. C.; Hossenlopp, I. A.; Douslin, D. R.; Waddington, G., Fluorobenzene: Thermodynamic Properties in the Solid, Liquid and Vapor States; A Revised Vibrational Assignment. *J. Am. Chem. Soc.* **1956,** 78, 5457-5463.

(53) Scott, D. W.; Messerly, J. F.; Todd, S. S.; Hossenlopp, I. A.; Osborn, A.; McCullough, J. P., 1, 2-Difluorobenzene: Chemical Thermodynamic Properties and Vibrational Assignment. *J. Chem. Phys.* **1963,** 38, 532-539.

(54) Andon, R. J. L.; Martin, J. F., Thermodynamic properties of fluorine compounds. Part 11.—Low-temperature heat capacities of the three tetrafluorobenzenes. *J. Chem. Soc., Faraday Trans. 1.* **1973,** 69, 761-770.

(55) Morrison, C. A.; Siddick, M. M., Determining the Strengths of Hydrogen Bonds in Solid-State Ammonia and Urea: Insight from Periodic DFT Calculations. *Chem. Eur. J.* **2003,** 9, 628-634.

(56) Takatani, T.; Hohenstein, E. G.; Malagoli, M.; Marshall, M. S.; Sherrill, C. D., Basis Set Consistent Revision of the S22 Test Set of Noncovalent Interaction Energies. *J. Chem. Phys.* **2010,** 132, 144104.

(57) Dunitz, J. D.; Gavezzotti, A., How Molecules Stick Together in Organic Crystals: Weak Intermolecular Interactions. *Chem. Soc. Rev.* **2009,** 38, 2622-2633.

This page has been left intentionally blank

Chapter 6: Dominant Driving Forces at Pressure: A Study of New Polymorphs of Glyphosate Under Varying Conditions.

Overview

This Chapter describes the formation of potentially four new high-pressure phases of glyphosate under varying conditions. In Chapter 4 the integrity of a single crystal of glyphosate was lost on slow compression above 5.6 GPa. Within this Chapter attempts to employ more rapid rates of compression in different pressure transmitting media, to 5.9 GPa in a 4:1 mixture of methanol and ethanol and to 7.0 GPa in neon, generated two new single-crystal phases. The phase at 7.0 GPa, named phase II, has a higher volume than the phase obtained at 5.9 GPa, named phase III, indicating formation of phase II from phase III is thermodynamically forbidden. In an attempt to recover the destructive transition seen previously and to potentially drive its transition to phase II or III, neutron powder diffraction data on glyphosate-d_4 was completed at ISIS neutron and muon source. This revealed a far more complex structural landscape and the formation of two further phases, named IV and V, which did not fit the structural parameters of phases I (Chapter 4), II or III. So far, the structures of phases IV and V have not been characterised further and it is the aim of future work to attempt to recover structural parameters for these phases, possibly through crystal structure prediction. Free energy calculations involving internal energy, volume and entropy are used to establish the driving forces of the single crystal transitions. Extrapolation of the parameters of phase I to the pressures both phase II and III were observed reveals that both transitions are driven by volume as the only significant stabilising contribution.

6.1 Introduction

The thermodynamic aspects of high-pressure phase transitions have been discussed extensively in Chapters 1, 3 and 4, but the role of kinetics has largely been neglected. Employing different rates of compression can modify phase behaviour. For example, *L*-serine has been shown to undergo different phase transitions depending on the rate of compression. *L*-serine I transforms to phase II under rapid compression to ~5 GPa and to phase III under rapid compression to ~8 GPa. However, under slow compression phase IV is formed at 5.6 GPa.[1-8] Ridout *et al*. have also shown that different rates of compression can result in three different polymorphs for 2-fluorophenylaccetylene when crystallised *in situ* by isothermal freezing.[9]

For in-house experiments, the slowest rate of compression is largely limited by the coarseness of pressure control available for DACs. Generally, the smallest steps possible are of the order 0.2 - 0.3 GPa, or about 3-6 times the typical precision of the ruby pressure measurement (0.05 GPa). At the other end of the scale, the time between steps is limited by the length of a high-pressure data collection, generally of the order of at least one or two hours. Central facilities can achieve measurements in significantly shorter times due to the intensity of the beam. O'Bannon and co-workers have utilised this to achieve millisecond time-resolved X-ray diffraction with dynamic DACs and compression rates in excess of 1000 GPa s^{-1}.[10]

In Chapter 4, a high-pressure crystallographic study on the compound glyphosate, N-(phosphonomethyl)glycine, was presented up to a pressure of 5.6 GPa. Pressure points were taken in 0.2-0.3 GPa steps at a maximum rate of once every 5 hours. This was necessary to allow for equilibration of pressure point and sample measurement. In this range two subtle second order phase transitions were seen. These were not denoted unique identifiers and are

referred to throughout this chapter as phase I. At 5.6 GPa, data quality deteriorated and further compression to 5.98 GPa led to a proliferation of diffraction spots which was interpreted as a destructive transition to a new polycrystalline phase. Although a similar result had been apparent in an earlier Raman study,[11] attempts to characterise this new phase were unsuccessful. The aim of the present chapter is to attempt to access this phase by utilising higher rates of compression and different pressure transmitting media which may enable the integrity of a single crystal to persist through the transition. The phase behaviour is also studied by neutron powder diffraction to provide a quantitative comparison between the effects of rapid and gradual compression and the role of deuteration.

On recovery of structural data from high pressure phases, free energy calculations involving internal energy, volume and entropy will be used to establish the thermodynamic ordering of all phases as a function of pressure and the driving forces behind all transitions. The energy difference between polymorphs is usually less than 1 kcal mol^{-1} (4.184 kJ mol^{-1}).[12] Subtle changes in internal energy, volume or entropy contributions can have significant impacts on the structures which form.

6.2 Materials and Methods

6.2.1 Single Crystal X-ray Diffraction

Crystals of glyphosate (Sigma Aldrich) were grown by the same procedure outlined in Chapter 4, slow evaporation of an aqueous solution of concentration 10 mg mL^{-1}. Single crystals were screened for diffraction quality using a Bruker AXS D8 Venture diffractometer with Ag Kα radiation (λ = 0.56083 Å) at 150 K. High pressure single crystal studies were subsequently completed both in house, on the same diffractometer and the same source, and at the Deutsches Elektronen-Synchrotron (DESY) on the P02.2 extreme conditions beamline.

For in house studies a single crystal was loaded into a Merrill−Bassett diamond-anvil cell (DAC) with half opening angle of 38°, 600 μm Boehler−Almax-cut diamonds and tungsten carbide backing plates.[13,14] A tungsten gasket was used of thickness 300 μm and indented by 200 μm with a hole diameter of 300 μm. A 4:1 mixture of methanol and ethanol was used as the pressure transmitting medium.[15] Three ruby spheres were also included in the sample loading and the ruby fluorescence method was used to measure pressure; the values quoted have an uncertainty of 0.05 GPa.[16] Pressure was increased gradually to 1 GPa and unit cell parameters determined to confirm the same trend was being followed as in Chapter 4 for the ambient pressure phase (phase I).

Pressure was then increased rapidly to 5.9 GPa in a single step, *i.e.*, to just above where the destructive transition was observed in Chapter 4. The crystal was observed to go through a single crystal to single crystal phase transition accompanied by a space group change from $P2_1/c$ to $P2_1/n$ with a complex relationship between the basis vectors of the two phases (Section

6.3.1b). This phase is referred to hereafter as phase III in the light of the phase development described below.

At DESY, a single crystal was loaded into a Short Symmetric DAC (described in greater detail in Chapter 2). A rhenium gasket was used to provide improved pressure control. The cell was loaded with neon gas as the pressure transmitting medium[15] and ruby spheres were again used as a pressure marker with the same pressure uncertainty of 0.05 GPa.[16] Phase I persisted under gradual compression to 5.6 GPa, where the destructive transition was previously seen. The sample was then compressed rapidly to 7.0 GPa. Again, a single crystal to single crystal phase transition occurred but to a third phase, phase II hereafter, with a much closer translational symmetry relationship to phase I than phase III. Although the standard setting of phase II is in $P2_1/n$ it is described here in $P2_1/c$; the difference in β is only 2.47° and the change facilitates comparison with phase I. Analysis of the relationship between phases I and II reveals that the translational symmetry of the lattice is unchanged over the course of the transition, that is, the relationship between the basis vectors of the two phases is expressed by the identity matrix.

6.2.2 Structure Solution and Refinement of X-ray Data

Diffraction patterns of in-house data were indexed using APEX4 and integrated using SAINT.[17] Diffraction patterns from DESY were indexed and integrated using CrysAlisPRO.[18] Corrections for the gasket shading, absorption and other systematic errors were applied using the multiscan procedures in SADABS and CrysAlisPRO respectively.[18,19]

The structures were refined against $|F|$ by full-matrix least-squares in Crystals using data with $|F| > 4\sigma(|F|)$.[20] Improved refinement statistics were found by refinements on $|F|$ over $|F|^2$, reflecting the quality of data obtained after rapid compression and the phase transitions. For both structures, atoms were modelled with anisotropic displacement parameters except for hydrogen which was modelled with isotropic displacement parameters to improve the data to parameter ratio of the high-pressure data sets. Hydrogens on the backbone of the molecule (hydrogens H2A/B, H3A/B and H4A/B, Figure 4.1) were placed in calculated positions for both structures. For the structure of phase II at 7.0 GPa hydroxyl hydrogens (H1 and H3, Figure 4.1) were located from residual electron density peaks. The same functional hydrogen atoms of phase III could not be located and were placed in positions that resulted in the formation of favourable hydrogen bonds and dimers. Structures were visualised throughout using Mercury.[21] Tables of crystal and refinement data for both phases are available in Appendix D.1.

6.2.3 Neutron Powder Diffraction

High-pressure, ambient temperature, neutron powder diffraction data were measured at the ISIS Neutron and Muon spallation source on the PEARL time-of-flight (TOF) instrument. A deuterated sample was prepared by the deuteration lab at ISIS starting from a fully protonated sample. Carbon bound hydrogens were not exchanged, and deuteration was only achieved at O1, O3 and N4 positions. Nevertheless, the semi deuterated sample provided a significantly lower background than the protonated sample, though higher than expected for a fully deuterated sample.

Approximately 50 mm^3 of sample was gently ground to a fine powder and loaded into a Paris-Edinburgh (PE) press (see Chapter 2).[22] Sintered

diamond anvils were used with a null-scattering Ti-Zr gasket to contain the sample with a lead pellet of roughly 5 mm^3 for the measurement of pressure.[23] A perdeuterated 4:1 mixture of methanol and ethanol was used as the hydrostatic medium.[15] Diffraction was detected in the transverse geometry in the 90° detector banks. Diffraction pattern intensities were corrected for the wavelength and scattering angle dependence of the neutron attenuation of the press and gasket with instrument models available for PEARL.[7]

Pressure was applied with a computer controlled hydraulic system and data collected at loadings 5, 15, 25, 35, 45, 50, 55, 60, 70, and 80 tonnes accounting for a pressure range of 0-7.83 GPa. Shorter scans, 2.5 hours, were collected in the range 5-45 tonnes (0.09-4.78 GPa), the range previously characterised for phase I in Chapter 4, and these patterns had poorer statistics than those collected at higher pressure. Loss of peaks between 4.78 and 5.12 GPa was interpreted as the sample undergoing a gradual phase transition and in the higher-pressure range 50-80 tonnes (5.15-7.83 GPa), longer scans, 12 hours, were completed in an attempt to provide higher quality data.

Peaks showed a substantial broadening with increasing pressure and became weaker against the background such that by 7.83 GPa, reliable refinements could not be obtained. This dataset is not included in further analysis.

Note: Refinements of neutron powder data were carried out by Professor Simon Parsons of the University of Edinburgh. The details are included here for completeness.

6.2.4 Analysis of Neutron Data

Powder-pattern indexing and Pawley refinements were carried out using TOPAS-Academic.[24] Refinement models were initially based on those obtained using X-ray models of phases I, II and III, and included lead (the pressure marker) and sintered diamond (from the anvil). The refined cell dimensions of the lead were used to calculate the pressure applied to the press using a Birch-Murnaghan equation of state with $V_0 = 121.418$ Å^3, $K_0 = 41.725$ GPa and $K_0' = 5.394$.[25] In the range 0.09-4.78 GPa Pawley refinements recovered cell parameters that corresponded to phase I. The structure of phase I at 4.78 GPa was refined by the Rietveld method. Glyphosate-I has a plate-like morphology developed in (100) (see Appendix D.3), and a small amount of preferred orientation in this direction was included in the Rietveld refinement using the March-Dollase model.[26]

In the range 5.15-6.10 GPa Pawley refinements starting from the cell parameters of phase II failed to model the patterns. The dataset at 5.61 GPa (55 tonnes) was indexed using simulated annealing to obtain cell parameters that were similar but not identical to phases I and II. Attempts to introduce structural models for Rietveld refinement based on the coordinates of phases I, II or III failed to match the diffraction data. The data sets at 5.61 and 6.10 GPa (55 and 60 tonnes) could be Pawley-fitted using the new indexing parameters, but those at 5.15 GPa (50 tonnes) required addition of phase I with a Rietveld model. Refinement of the dataset at 6.88 GPa (70 tonnes) was more complex still and this has not been successfully modelled. These observations suggest that neither phase II nor III were observed in the neutron study, which instead yielded a further two new phases, IV and V.

6.2.5 DFT Geometry Optimisation

Crystal structures were geometry-optimised using procedures described in Chapter 5. Experimental single crystal X-ray data for phase I at 0 and 5.18 GPa were taken from Chapter 4. These structures were optimised along with phase II at 7.03 GPa and phase III at 5.93 GPa. Additionally, unit cell parameters determined for phase I from Chapter 4 were used to extrapolate the unit cell dimensions of phase I to 5.93 and 7.03 GPa using 'linear' equation of state fitting in EoSFit.[27,28] (Further details of the fittings are available in Appendix D.2). The structure of phase I at 5.18 GPa from Chapter 4 was then optimised into these new unit cells for direct comparison between phases I and II at 7.03 GPa and phase I and III at 5.93 GPa.

Periodic density functional theory (DFT) was used throughout using CASTEP and implemented in the Materials Studio interface.[29,30] The Perdew-Burke-Ernzerhof (PBE)[31] exchange-correlation functional and norm-conserving pseudopotentials[29] were used with a basis set cut-off energy of 1100 eV and a **k**-point spacing of 0.04 $Å^{-1}$, which converged the total energy to <0.1 meV per atom. Dispersion was treated using the Tkatchenko-Scheffler method.[32] The unit cell dimensions were fixed in all calculations. The convergence criteria were: 1×10^{-8} eV $atom^{-1}$ (energy), 0.002 eV $Å^{-1}$ (force) and 1×10^{-4} Å (displacement).

6.2.6 Phonon Calculations

Following geometry optimisation, vibrational frequencies were calculated in the harmonic approximation at the Γ-point using the linear response method. The results of Chapter 5 demonstrated that Γ-point calculations make a reasonable estimation of calculated entropy differences for similar phases

while recovering significant computational expense when compared to calculations across the full Brillouin zone. For fluorobenzene and 1,2-difluorobenzene the difference in the *TS* term calculated from Γ-point calculations in Chapter 5 was −1.36 kJ mol^{-1}, whereas the same value from calculations across the full Brillouin zone of significantly greater computational expense is −1.42 kJ mol^{-1}, only a −0.06 kJ mol^{-1} difference. This is in line with results reported by Nyman and Day.[33] The frequencies of the acoustic modes at the Γ-point were within 0.06 cm^{-1} of 0 cm^{-1}; all other frequencies were positive.

The frequency output of the phonon calculations was used for the calculation of vibrational entropies at the Γ-point using the routines available in CASTEP. For comparative purposes, entropic contributions to free energy are estimated at 298 K for all structures.

6.2.7 Pixel Calculations

Intermolecular interaction energies were calculated via the semi-empirical Pixel method[34,35] using the MrPixel interface and the coordinates obtained after DFT geometry optimisation.[36] Gaussian-09[37] was used for electron density calculations at the MP2 level of theory using the 6-31G** basis set. Molecular electron density was calculated on a grid of 0.08 × 0.08 × 0.08 Å^3 applied with a condensation level of 4 and a cluster radius of 30 Å. Only results from within the first coordination sphere were used, a larger radius was used to capture the full extent of the electrostatic potential of the zwitterionic molecule. Component energies were calculated for electrostatic, polarisation, dispersion and repulsion interactions with total energy taken as the sum of all contributions. Contacts within the first coordination sphere are identified as those with non-zero repulsive contributions. These are labelled A, B, C… for

all structures and are directly comparable to the labelled contacts provided in Chapter 4.

6.2.8 Volume Calculations

Occupied ('network') and unoccupied ('void') volumes in crystal structures were evaluated in a Monte Carlo procedure using the CellVol code introduced in Chapter 3 for all optimised structures. This was additionally used to provide packing coefficients for all structures.

6.3 Results and Discussions

6.3.1 Single Crystal X-ray Diffraction

The aim of the present chapter is to rationalise and characterise the destructive phase transition described in Chapter 4 for glyphosate on slow compression to above 5.6 GPa. The aim was to attempt to bypass the destruction of the crystal by utilising rapid compression through the transition pressure. Separate experiments using 4:1 methanol/ethanol and neon as pressure transmitting media yielded different phases.

Rapid compression using a Merrill-Bassett DAC to 5.93 GPa using a 4:1 methanol, ethanol mixture (the same media as the slow compression study in Chapter 4) yielded a new phase (III) with a substantially lower volume, Table 6.1. A second loading of glyphosate in a short symmetric DAC with neon as the hydrostatic medium and rapid compression to 7.03 GPa yielded a transition to a different phase (II). The unit cell parameters of this phase were closer than those of phase III to those of phase I, Table 6.1.

Table 6.1: The different phases of glyphosate observed under different conditions. All pressures are measured via the ruby fluorescence method and carry the uncertainty 0.05 GPa. Data for phase I was taken from Chapter 4.

Pressure /GPa	Phase	Space group	Cell Volume /Å^3	*a* /Å	*b* /Å	*c* /Å	*β* /°
0	I	$P2_1/c$	658.51 (12)	8.6693 (13)	7.9704 (6)	9.8968 (7)	105.644 (9)
5.18	I	$P2_1/c$	547.3 (4)	8.496 (5)	7.510 (2)	9.249 (3)	111.96 (3)
7.03	II	$P2_1/c$	521.0 (3)	8.2046 (13)	7.536 (2)	9.473 (3)	117.18 (3)
5.93	III	$P2_1/n$	516.94 (8)	4.7277 (3)	7.1064 (4)	15.5653 (19)	98.689 (6)

6.3.1*a* Comparison of Phase I and II

Intermolecular interactions in phase I, which were discussed extensively in Chapter 4, are dominated by hydrogen bonding interactions, with 13 molecules in the first coordination sphere at ambient pressure. These molecules are arranged to form layers roughly parallel to the (1 0 −2) plane. This packing arrangement persists up to 5.18 GPa, the highest pressure measured in Chapter 4, except for the ingress of a 14th, weakly destabilising, interaction.

The packing schemes of both phases I and II are directly comparable, Figure 6.1. As defined by the relationship between the orientation matrices, the relationship between the two phases is the identity matrix. Phase II features all 14 of the dimers seen for phase I within its first coordination sphere. This includes the same 6 hydrogen bonded interactions (A, B, E, E′, F and F′) and further 7 more general electrostatic interactions (C, C′, D, D′, G, H and H′) described in Chapter 4 with the additional repulsive interaction (labelled J). The total energies of the interactions in phase II at 7.03 GPa are compared to their energies in phase I at 5.18 GPa in Table 6.2 below. A full breakdown of these energies into their constituent terms is given in Appendix D.4.

On transition from phase I to phase II a substantial compression in the molecular backbone occurs. This results in a reduction in the long axis of the molecule by 0.30 Å (calculated using the same method as Chapter 4, further details available in the Appendix B.1). To accommodate this compression the magnitudes of both torsion angles along the backbone decrease. The C1-C2-N3-O4 torsion angle changes from −82.21° to −81.82° and the C2-N3-O4-P5 torsion angle changes from 165.54° to 151.23° (values are given without ESDs

as these are taken from geometry optimised structures). Similar behaviour has previously been noted for glycine with increasing pressure.[38]

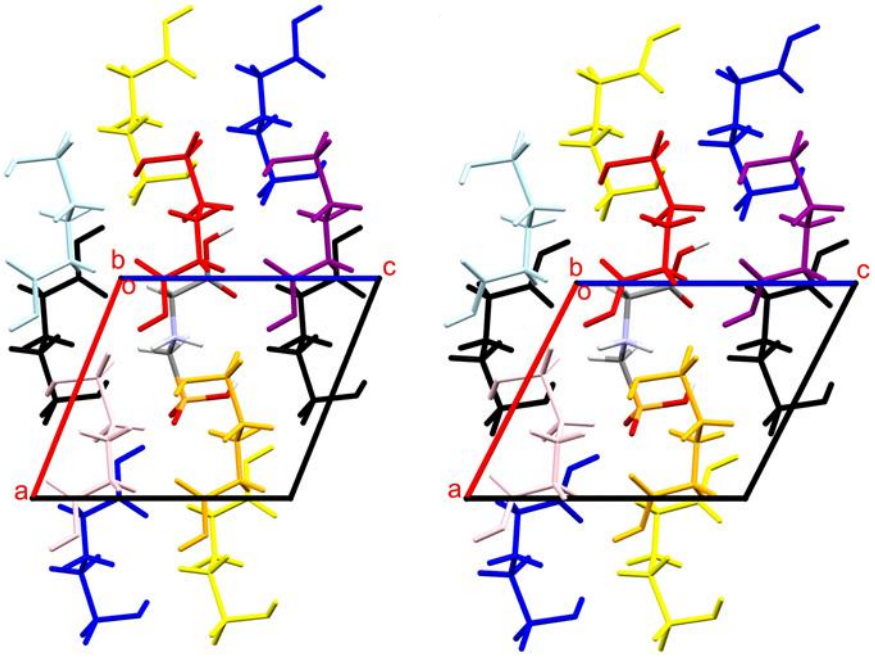


Figure 6.1: A comparison of the packing schemes in phase I at 5.18 GPa (left) and phase II at 7.03 GPa (right) viewed down the ***b*** axis. Contacts are colour coded. B-orange, C-red, D-yellow, E-pink, F-dark blue, G-purple, H-black, J-light blue. Contact A is hidden behind contact B by symmetry in both structures.

To maintain contacts F and F′, with a modest stabilisation, the carboxylate group undergoes a significant rotation. The torsion angle O1-C1-C2-N3 changes from −175.21° to −143.91° and the torsion angle H1-O1-C1-C2 changes from 170.68° to −171.05°. This rotation also modifies contacts D and D′ which both form a head-to-tail electrostatic interactions between a carboxylic acid group and the phosphate hydrogen, a lattice translation of contact F/F′. The rotation of the carboxylic acid group increases the separation of an O3 and O1 short contact between the molecules from 2.652 to 2.740 Å. Both are bonded to hydrogen and so similarly polarised, the increase in separation provides a stabilisation of −36.1 kJ mol^{-1}.

The changes in energy of the remaining contacts can be largely rationalised by the changes in centroid separation of all contacts on the transition. All remaining H-bonded interactions, A, B and E/E′, reduce in centroid separation. For contact A the 0.4 Å reduction in centroid separation results in a destabilisation by 4.7 kJ mol^{-1}. The reduction actually results in an increase in the length of the hydrogen bond, acceptor-donor and the H3···O4 distance by 0.046 and 0.054 Å, respectively. The greater overlap of the molecules pushes the hydrogen bonds in the ring motif beyond an idealised parallel geometry which can be seen in the O3-H3-O4-P5 torsion angle which reduces in magnitude from 75.23° to 61.70°.

For contact B and E/E′ the reduction results in a stabilisation of the energy of the contacts. The reduction produces a complementary reduction in the hydrogen bond donor-acceptor distance for contact B by 0.02 Å, but by only 0.004 Å for E/E′. The more modest reduction in contact E/E′ is additionally accompanied by an increase in the hydrogen bond angle N3—H3A···O4 from 155.78° to 162.19°, closer to the idealised geometry of a hydrogen bond.[39] The stabilisation of contact B is roughly double that of contact E/E′ (−36.3 vs −16.1 kJ mol^{-1}) as contact B contains two hydrogen bonds, whereas E/E′ contain one.

The remaining interactions destabilise at the transition. For contacts C and C′ this destabilisation is 11.7 kJ mol^{-1} for a reduction of 0.09 Å in centroid separation. At the transition the electrostatic term goes from destabilising to stabilising with a decrease of 20.0 kJ mol^{-1}, Table A.12 and A.13 in Appendix D.4. However, this is accompanied by significant increase in repulsion of 39.5 kJ mol^{-1} and overall, the interaction destabilises. For contact G a destabilisation of 17.3 kJ mol^{-1} is seen for an apparent increase in separation of an interaction between two like charged carboxylate groups. This apparently paradoxical

behaviour is similar to what was seen above for hydrogen bonds in contact A due to separation being defined in terms of centroids. Despite an increase in centroid separation of 0.263 Å, the separation between H1 atoms decreases by 1.055 Å and the carboxylate groups form a more parallel arrangement due to the rotation described above. Similar comments apply to contact H/H', where an increase in centroid separation results in increased repulsion between like charged moieties. For contact J the reduction in separation of the C2-H2A/H2B groups results in a large positive coulombic contribution which provides an overall destabilisation of the contact.

Table 6.2: Interactions in the first coordination sphere for phase I at 5.18 GPa and phase II at 7.03 GPa. All energies are in kJ mol^{-1}. Classifications are relative to the interaction with the central molecule.

		Phase I at 5.18 GPa		**Phase II at 7.03 GPa**		
Label	**Symmetry**	**Centroid distance (Å)**	**Total energy**	**Centroid distance (Å)**	**Total energy**	**Contact description**
					H-bonded	
Contact A	$1-x, 2-y, 1-z$	6.426	−128.6	6.022	−123.9	Pairs of O3H3···O4 H-bond ring motifs across inversion centres.
Contact B	$1-x, 1-y, 1-z$	4.909	−264.1	4.603	−300.4	Two N3H3B···O5 H-bond ring motifs across inversion centres.
Contact E/E′	$1-x, y \mp \frac{1}{2}, \frac{1}{2}-z$	6.390	−164.4	6.293	−180.5	Pair of N3H3A···O4 H-bonds.
Contact F/F′	$x \pm 1, y, z$	8.496	−92.1	8.205	−93.0	Lattice translations of the central molecule in ***a***. O1H1···O5 H-bond through a head to tail arrangement.
					Electrostatic Interactions	
Contact C/C′	$-x, y \mp \frac{1}{2}, \frac{1}{2}-z$	5.436	5.1	5.349	16.8	Lattice translation of contact E/E′ along ***a***. Composed principally of destabilising electrostatic and stabilising dispersion components at 5.18 GPa. By 7.03 GPa the electrostatic component has become stabilising.
Contact D/D′	$x \mp 1, \frac{3}{2}-y, z \mp \frac{1}{2}$	8.093	−35.1	7.433	−71.2	Lattice translation of contact F/F′ along ***a***. Longer range electrostatic interaction.
Contact G	$-x, 1-y, 1-z$	6.497	−1.6	6.760	15.7	Destabilising electrostatic interaction as a result of the juxtaposition of like charged moieties.
Contact H/H′	$x, \frac{3}{2}-y, z \mp \frac{1}{2}$	4.764	53.9	4.841	60.4	Destabilising electrostatic interaction as a result of the juxtaposition of like charged moieties.
Contact J	$-x, 1-y, -z$	6.715	18.0	6.581	27.8	Destabilising electrostatic interaction as the result of the short separation of molecules.

6.3.1*b* Comparison of Phase I and III

The quality of the diffraction data obtained for phase III did not allow H-atom positions to be located from Fourier maps. Locations of the carboxylic acid and phosphate-based H-atoms (H1 and H3) were placed on the longest C—O and P—O bonds, initial positions being estimated for structure refinement on the basis of favourable H-bond formation. The structure was then optimised using periodic DFT. This procedure led to rotation of H3 about the P5-O3 bond, Figure 6.2. This rotation enables the formation of a O3—H3···O2 hydrogen bond and a ring motif with the O1—H1···O4 hydrogen bond. The change in the position of H3 results in a closer approach of the adjacent O1—H1···O4 hydrogen bond. The donor-acceptor distance reduces by 0.023 Å resulting in formation of a bridging, symmetrical hydrogen bond. The same feature was observed in optimisations based on different pseudo-potentials (Norm conserving, OFTG Norm conserving and OTFG ultrasoft) and functionals (PBE and PW91) as well as the optimisation of the unit cell.

The presence of the shuttling O1—H1···O4 hydrogen bond makes the optimised structure of phase III unsuitable for both Pixel and SAPT calculations. The significant rotation of the H3 hydrogen means experimental structures will likely have very different energy values. As such the energies of individual dimers were not calculated and breakdowns into constituent energies are not available.

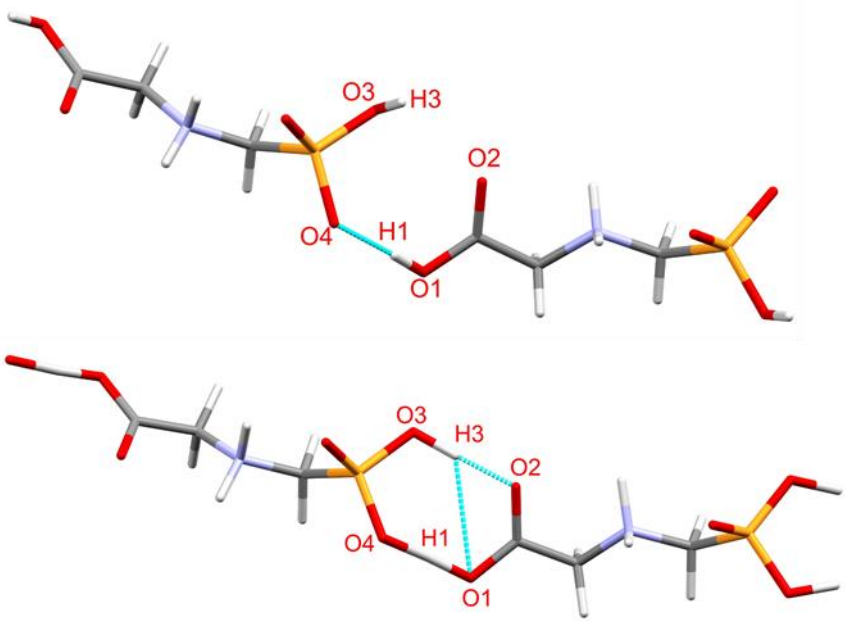


Figure 6.2: The result of optimisation of phase III, hydrogen bonds are shown in light blue. The H3 hydrogen in the experimental structure (top) undergoes a significant rotation on optimisation. In the optimised structure (bottom) this allows a closer approach of neighbouring molecules, the formation of a ring motif and shuttling hydrogen atom.

The structure of phase III is substantially different to both phases I and II. For both phase I and II, increasing pressure increases the value of β with the largest strain tensor (shown previously in Chapter 4 to make angles 27.1, 90.0 and 132.8° with the unit cell axes respectively) aligned to compression along this vector. Phase III however has a substantially smaller angle of β, 98.689°. Calculation of the transformation matrix from the orientation matrices of phases I and III reveals a more complex phase relationship than that between phase I and II.

$$R = \begin{pmatrix} -0.002 & -0.0125 & -0.493 \\ 0.025 & -0.908 & 0.022 \\ -1.872 & -0.050 & -0.273 \end{pmatrix}$$

Beyond small deviations from rational elemental values which are present due to the compression of the unit cell of phase III, the $\boldsymbol{a}$ axis is related to the $\boldsymbol{c}$ axis by approximately half, the $\boldsymbol{b}$ axes are reversed and the $\boldsymbol{c}$ axis is

roughly double the ***a*** axis with an additional less well defined component in ***c*** (−0.27), Figure 6.3. This additional component accounts for the lower angle of ***β***.

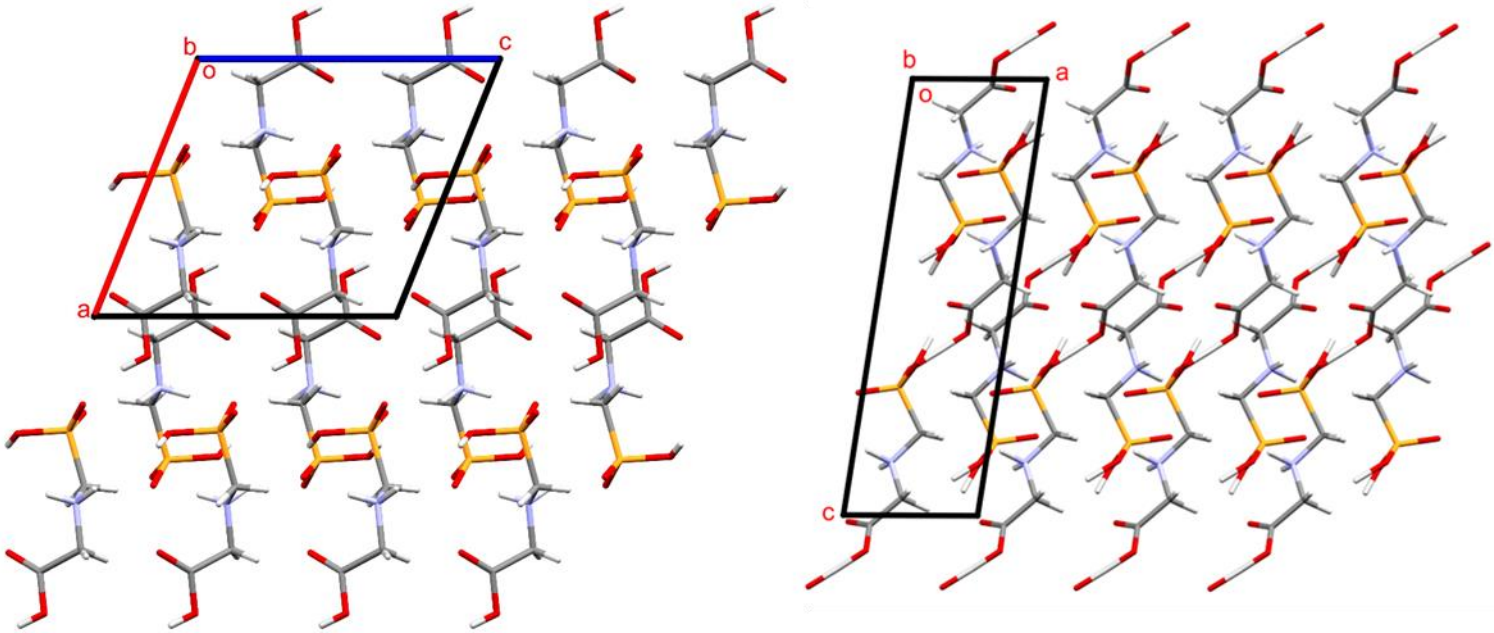


Figure 6.3: Comparison of phase I at 5.18 GPa (left) viewed down positive ***b***, and phase III at 5.93 GPa (right) viewed down −***b***. The unit cell of phase III has been translated for easier comparison between the phases. The relationship between the phases identified by the translation matrix is apparent.

Comparison of phases in Figure 6.3 using the relationship identified by the translation matrix allows better characterisation of the phase transition. At the transition a far greater overlap occurs along the long axis of the molecules. This overlap breaks the pair of O3—H3···O4 hydrogen bonds that form a ring motif (contact A in phase I with an energy of −128.6 kJ mol^{-1} at 5.18 GPa). In the absence of this motif, a significant rotation occurs to 'untwist' the molecule to form a new ring motif. The O3—H3 bond in the phosphate moity, rotates, decreasing the O3-P5-C4-N3 torsion angle from −73.32° to −131.34°. This enables the formation of a new O3—H3···O2 hydrogen bond at each end of the

molecule and breaks the pre-existing O1—H1···O5 hydrogen bond formed in phase I at each end of the molecule (contact F/F′) to form the shorter, shuttling O1—H1···O4 hydrogen bond with the now available O4 oxygen from the broken ring motif. Together these hydrogen bonds form a new head-to-tail ring motif at each end of the molecule, contact A/A′ for phase III.

The rotation of the phosphate moiety also results in the rotation of the NH_2 group. The torsion angle P5-C4-N3-C2 reduces from 165.54° to −154.49° and the torsion angle C4-N3-C2-C1 reduces from −82.21° to -177.61°. This rotation breaks the original two N3—H3A···O4 hydrogen bonds (contact E/E′ for phase I) but forms a new N3—H3A···O4 hydrogen bond ring motif with a different molecule due to the rotation of the phosphate moity and the greater overlap of the long axes of the molecules. These new hydrogen bonds form contact G for phase III. Rotation of the NH_2 group also breaks the N3—H3B···O5 hydrogen bond ring motif (contact B in phase I). However, the rotation results in a similar contact to contact G on the other side of the molecule with a new ring motif of N3—H3B···O3 hydrogen bonds, contact D.

The deviations of the axial relationships in phase I and III from rational values and reorganisation of the hydrogen bonds indicates that the phase I to III transition is likely a reconstructive transition. In the absence of calculated energies of the hydrogen bonds in phase III, the geometric parameters of the hydrogen bonds lost from phase I and the new hydrogen bonds formed in phase III are listed in Table 6.3. Making the assumption that shorter hydrogen bonds are stronger and taking the idealised hydrogen bond angle as 180°,[39] the hydrogen bonds formed in the two new head-to-tail ring motifs are stronger than the phosphate ring motif and head-to-tail hydrogen bonds lost from phase I. On the other hand, the nitrogen-based hydrogen bonds from phase I are stronger than those formed in phase III.

The overall molecular packing of phase III consists of 16 molecules within the first coordination sphere. The head-to-tail ring motifs featuring the shuttling hydrogen bond form polymeric chains along the long axis of the molecules. This aligns along the ***c***-axis, forming layers along the (0 1 0) planes, Figure 6.4. The layers stack in an AB arrangement with molecules arranged in alternating directions in adjacent layers. The layers consist of 6 dimers; contact A/A′ described above and contacts B/B′ and C/C′ which form translations in the direction of ***a*** within the layers of the central molecule and contact A/A′ respectively. Contacts B/B′ and C/C′ forming an analogous ring motif with each other and longer-range interactions with the central molecule. A further 10 interactions exist between the layers. These include contacts D and G described above with the remaining general longer-range electrostatic interactions, largely formed as symmetry equivalents of hydrogen bonded interactions.

Table 6.3: Comparison of the hydrogen bonds broken in phase I and formed in phase III at the transition. Hydrogen bond distances are defined as the donor-acceptor distance and angles are defined along the donor-hydrogen-acceptor direction.

Geometric Parameters	
Broken from phase I	**Formed in phase III**
Two O3—H3···O4: 2.525 Å, 170.98°	Two O3—H3···O2: 2.495 Å, 178.23°
Two O1—H1···O5: 2.479 Å, 168.37°	Two O1—H1···O4: 2.413 Å, 176.32°
Two N3—H3A···O4: 2.612 Å, 155.78°	Two N3—H3A···O4: 2.764 Å, 153.71°
Two N3—H3B···O5: 2.712 Å, 167.28°	Two N3—H3B···O3: 2.851 Å, 131.05°

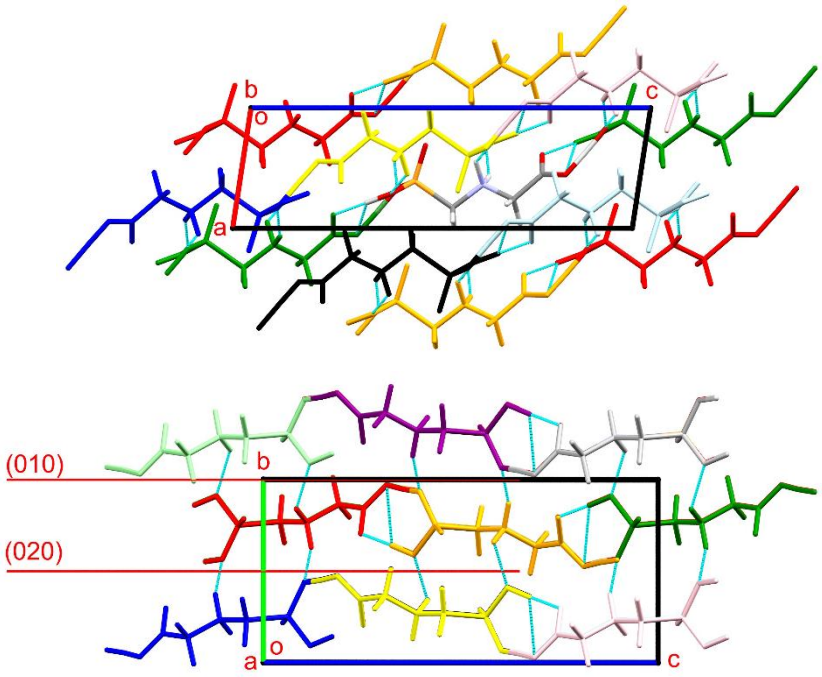


Figure 6.4: The packing diagrams of phase III. Layers are formed in the (0 1 0) planes with hydrogen bonds between the layers. Molecular contacts in Table 6.4 are colour coded. A-green, B-orange, C-red, D-yellow, E-pink, F-dark blue, G-purple, H-black, J-light blue, K-dark grey, L-light green.

Table 6.4: Interactions in the first coordination sphere of phase III at 5.93 GPa. All energies are in kJ mol^{-1}. Classifications are relative to the interaction with the central molecule.

Label	Symmetry	Contact description
		H-bonded
Contact A/A′	$x \pm \frac{1}{2}, \frac{3}{2} - y, z \mp \frac{1}{2}$	Related to the central molecule in a head-to-tail arrangement through the shuttling hydrogen bond O1H1···O4 and a O3H3···O2 hydrogen bond.
Contact D	$1 - x, 1 - y, 1 - z$	Related to the central molecule by an inversion centre. Two N3H3B···O3 hydrogen bonds are formed with the central molecule between the layers.
Contact G	$1 - x, 2 - y, 1 - z$	Related to the central molecule through inversion. Two N3H3A···O4 hydrogen bonds are formed with the central molecule.
		Electrostatic Interactions
Contact B/B′	$x \pm 1, y, z$	Translations of the central molecule in ***a*** within the layers.
Contact C/C′	$x \pm \frac{1}{2}, \frac{3}{2} - y, z \pm \frac{1}{2}$	Translations of contact A in ***a***. These molecules form polymeric bonds with contact B.
Contact E	$\frac{1}{2} - x, y - \frac{1}{2}, \frac{3}{2} - z$	Connected to contact D through the polymeric bond. Longer range interaction with the central molecule.
Contact F	$\frac{3}{2} - x, y - \frac{1}{2}, \frac{1}{2} - z$	Connected to contact D through the polymeric bond. Longer range interaction with the central molecule.
Contact H	$2 - x, 1 - y, 1 - z$	Related to contact D through a translation in ***a***. Longer range interaction with the central molecule.
Contact J	$\frac{3}{2} - x, y - \frac{1}{2}, \frac{3}{2} - z$	A translation in ***a*** of contact E. Longer range interaction with the central molecule.
Contact K	$\frac{1}{2} - x, \frac{1}{2} + y, \frac{3}{2} - z$	A translation in ***b*** of contact E. Longer range interaction with the central molecule between layers.
Contact L	$\frac{3}{2} - x, \frac{1}{2} + y, \frac{1}{2} - z$	A translation in ***b*** of contact F. Longer range interaction with the central molecule between layers.
Contact M	$2 - x, 2 - y, 1 - z$	A translation in ***a*** of contact G Longer range interaction with the central molecule between layers.
Contact N	$\frac{3}{2} - x, \frac{1}{2} + y, \frac{3}{2} - z$	Connected to contact G through the polymeric bond. Longer range interaction with the central molecule between layers.

6.3.2 Neutron Powder Diffraction

Neutron powder diffraction data were measured with the aim of determining whether either of the phases obtained on rapid compression could be identified as the phase that forms as a result of the destructive transition. It is clear from the X-ray data alone that phase III cannot convert to phase II with increasing pressure. Phase III is observed 1.1 GPa lower than phase II but also has a volume 1.023 Å^3 per molecule lower. The first derivative of Gibbs free energy, G, with respect to pressure, P, is volume, V,

$$\left(\frac{\partial G}{\partial P}\right)_T = V \tag{6.1}$$

Therefore, an increase in volume with increasing pressure from phase III to II is thermodynamically forbidden. The expected result of neutron measurements was therefore to observe the sequence of a phase transition from phase I to phase II and possibly to phase III on increasing pressure. The results found however were, somewhat more complex.

The neutron powder patterns for glyphosate as a function of pressure are presented in Figure 6.5. All significant peaks between 0.3 Å and the large peak at 2.1 Å are from either the sintered diamond anvils or the lead pressure marker. Note the peaks for the anvils do not change in intensity or d-spacing with increasing pressure as these are outside of the pressure chamber. The peaks at d = 2.8 Å and 2.4 Å in the 5.15 GPa data point (the peaks labelled 7 and the larger of the peaks in peak 10 in Figure 6.5) are also peaks from the lead pressure marker. The remaining peaks are from the sample.

With increasing pressure all sample peaks move to lower values of d-spacing, as expected. At 5.15 GPa a subtle but clear change occurs in several of the peaks, and this was identified as a phase transition. Most notably, there is a sudden drop in intensity of peaks 2 and 4 at this point and also a loss of the

group of peaks labelled 11. Peak 2 is the most intense peak at ambient pressure and corresponds to the (0 2 1) miller plane. The plane does not align with the (1 0 −2) slip planes or the strain axes identified in Chapter 4. It also does not appear to align with any significant interaction identified via the Pixel method for phase I. Significant intensity persists for peaks 2 and 4 until 5.61 GPa where these peaks then disappear entirely. At 5.61 GPa two new peaks 12 and 13 are observed. This is indicative of a gradual transition, with contamination by the lower pressure phase still present at 5.15 GPa and the transformation to the higher-pressure phase not complete until 5.61 GPa. Peaks 3, 5, 6, 8 and 9 all broaden and are lost to the background as pressure increases. These are no longer visible beyond the lower background above the transition point. Peak number 1 remains intense throughout all pressures as an indication that crystallinity in our sample endures. This peak corresponds to the (0 1 2) miller planes but again does not appear to align with slip planes, strain axes or significant interactions identified in Chapter 4.

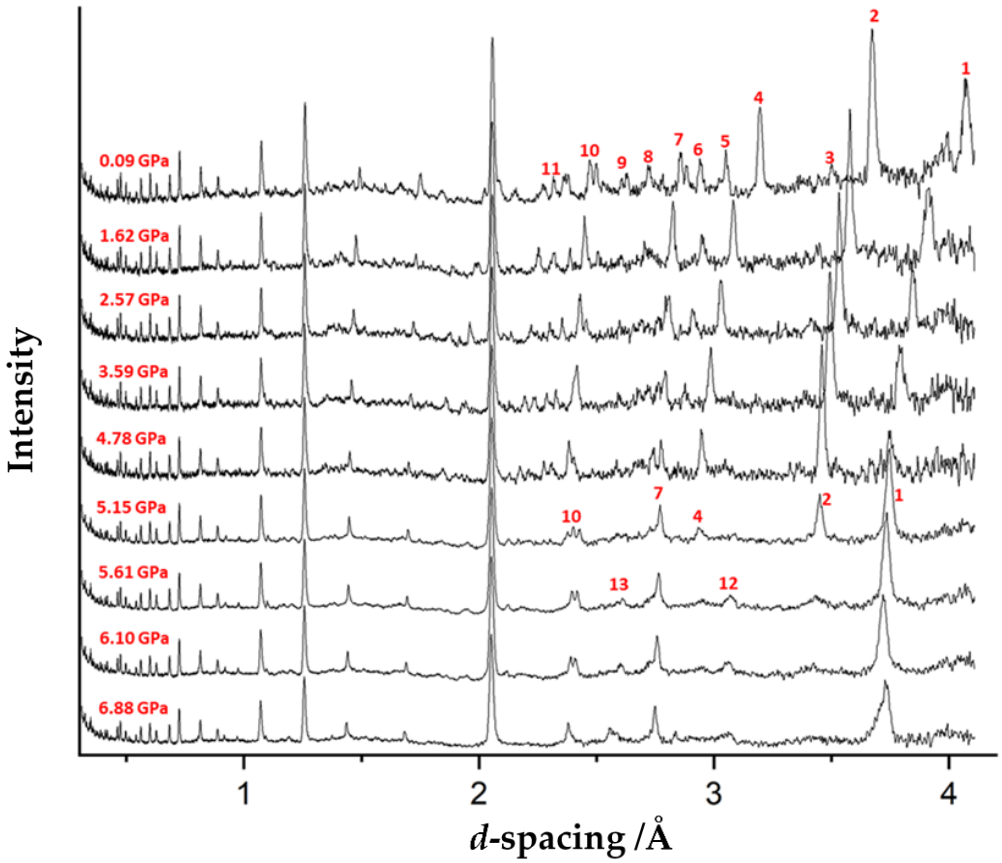


Figure 6.5: The trend in *d*-spacing from the neutron powder study. Peaks discussed in the main text have been labelled. Pressures measured from the unit cell parameters of the lead marker feature above each pattern.

Pawley refinements against these data yield good fitting statistics of phase I for all datasets between 0.09-4.78 GPa. Data below 5 μs was not fitted in all refinements. This region corresponds to the body of the cell only. Rietveld refinement of phase I was completed for the 4.78 GPa dataset converging to R_{wp} = 5.66%, Figure 6.6.

At 5.15 GPa, the onset of a gradual transition to a new phase, phase IV, was observed, Figure 6.7, as indicated by the *d*-spacing plot. The pattern was fitted using a Pawley model for phase IV (indexed using the 5.61 GPa data set) and a Rietveld model for phase I, converging to R_{wp} = 2.72%. At 5.61 GPa the proportion of phase I refined to zero on application of a similar model, confirming that the transition to phase IV was complete at this pressure, Figure 6.8, R_{wp} = 2.89%. The data obtained at 6.10 GPa could also be modelled as pure phase IV, Figure 6.9, R_{wp} = 3.03%. The lattice parameters from these fits are listed in Table 6.5.

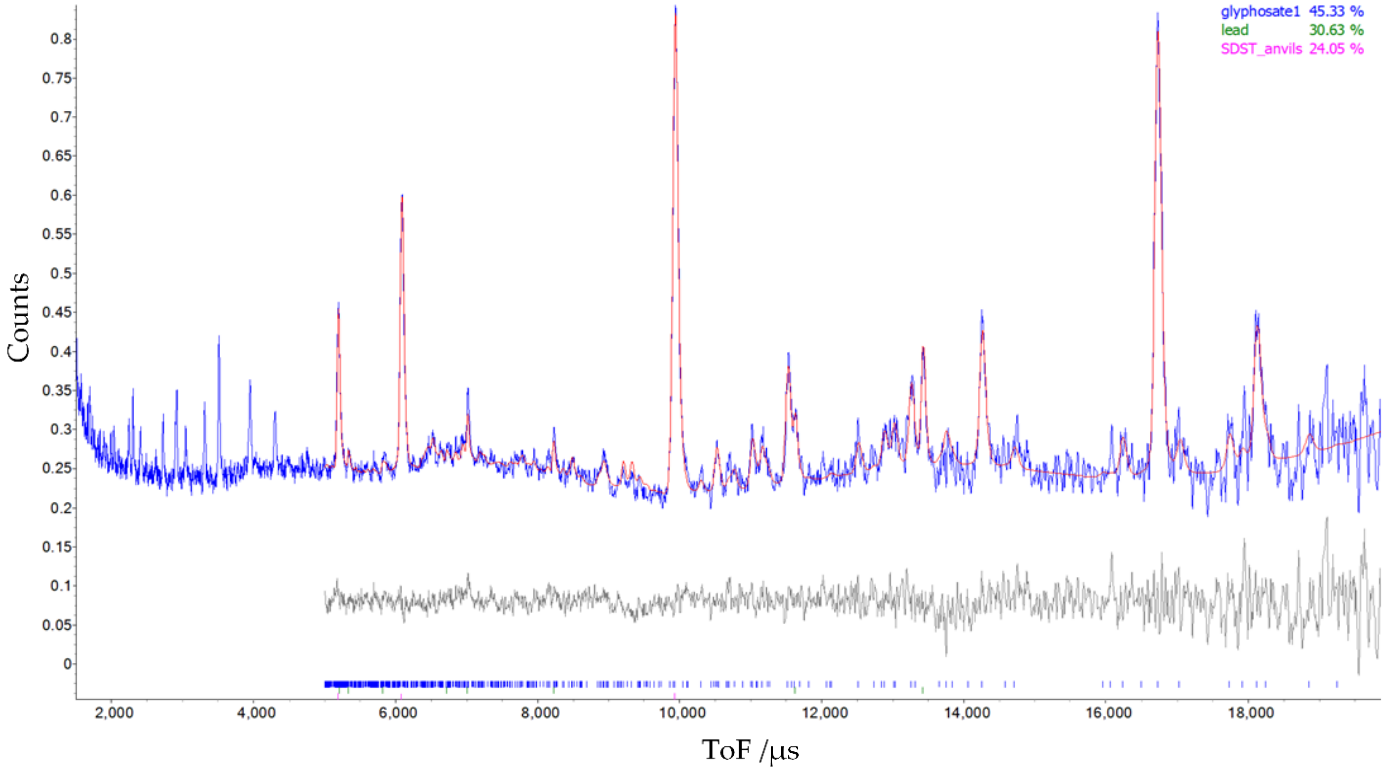


Figure 6.6: Rietveld refinement of the 4.78 GPa dataset starting from phase I, R_{wp} = 5.66%.

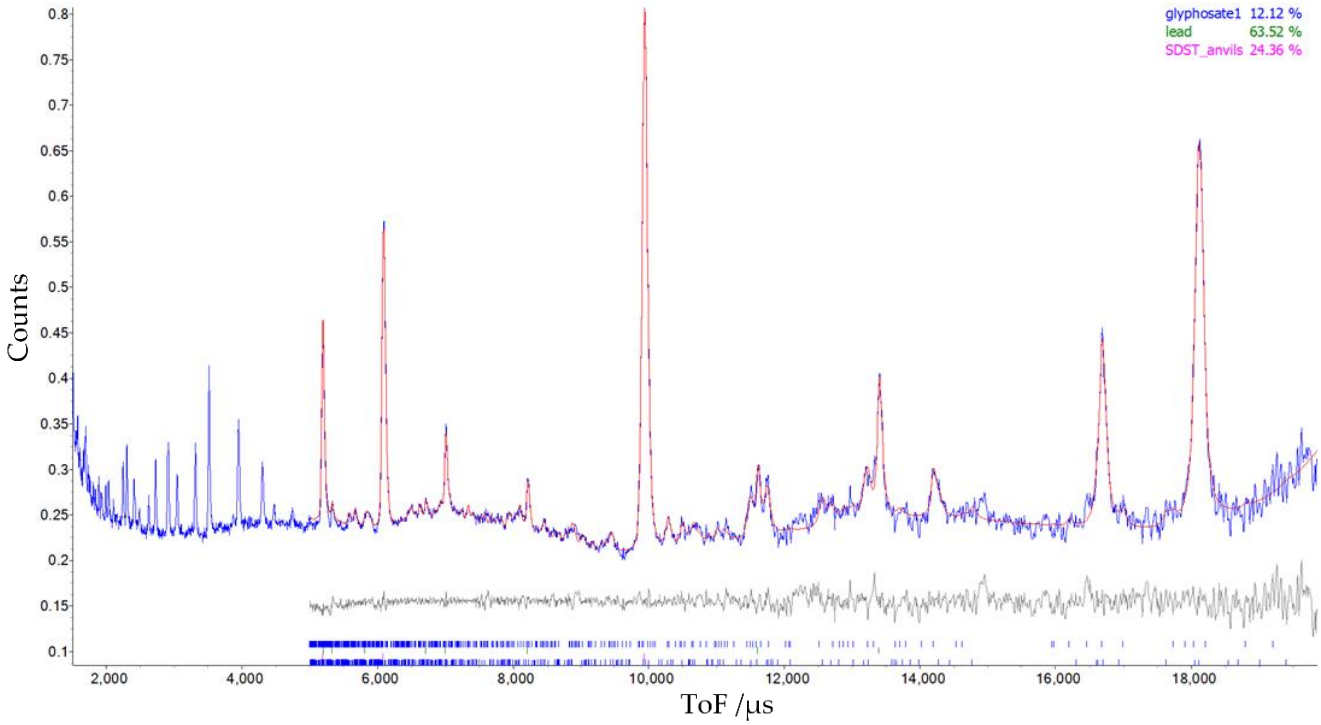


Figure 6.7: Combined Rietveld refinement starting from phase I and Pawley refinement for phase IV for the 5.15 GPa dataset, R_{wp} = 2.72%.

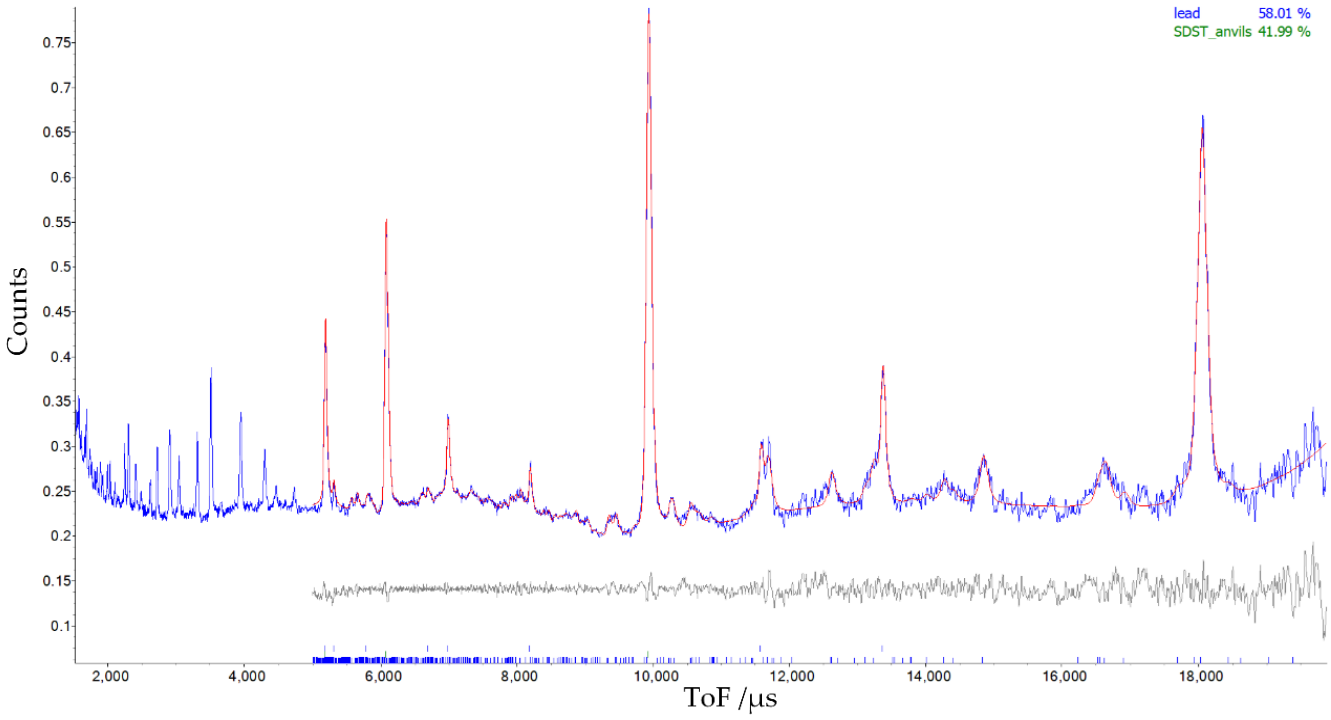


Figure 6.8: Pawley refinement for phase IV at 5.61 GPa, R_{wp} = 2.89%. The proportion of phase I refined to zero in earlier cycles of refinement.

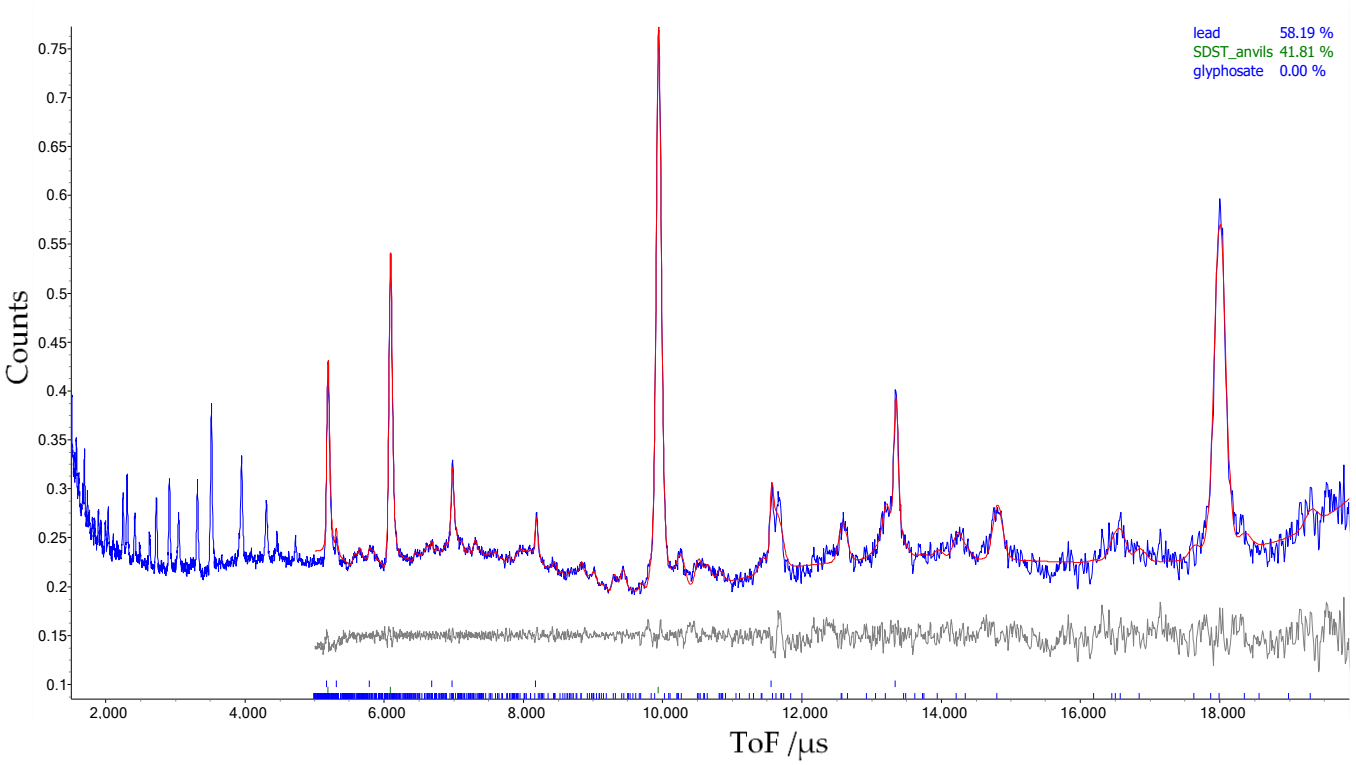


Figure 6.9: Pawley refinement of phase IV at 6.10 GPa dataset, R_{wp} = 3.03%. Starting parameters for the refinement were taken from the results of the refinement at 5.61 GPa.

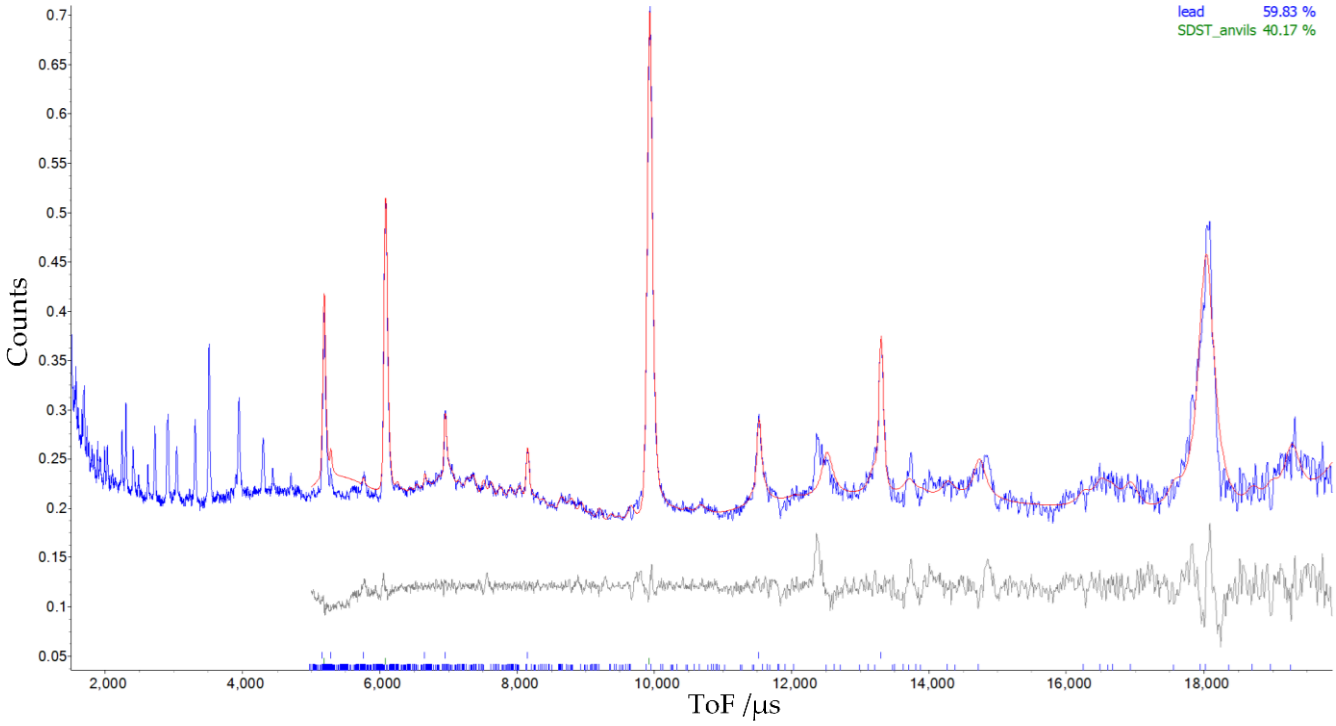


Figure 6.10: Pawley refinement starting from the 6.10 GPa dataset parameters for the 6.88 GPa dataset, R_{wp} = 4.36%. Large residual peaks can be seen as an indication of a poor fit.

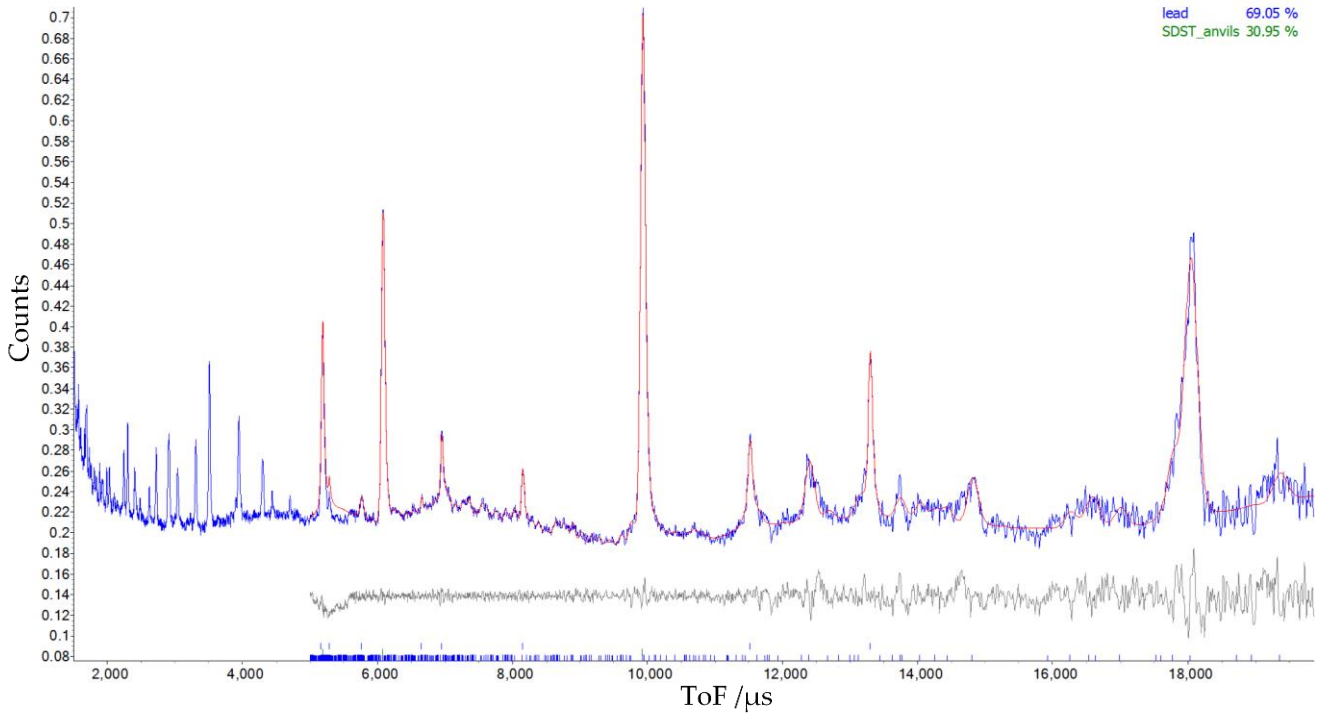


Figure 6.11: Pawley refinement starting from phase II for the 6.88 GPa dataset, R_{wp} = 3.29%. The fit is better than in Figure 6.10, but residual peaks can still be seen as an indication of a poor fit.

For the 6.88 GPa dataset, Pawley refinements starting from the cell refined at 6.10 GPa resulted in a poor fit, Figure 6.10, R_{wp} = 4.36%, an indication that another change had occurred in the sample. Pawley refinements of the dataset at 6.88 GPa starting from the cell dimensions of phase II determined at 7.03 GPa resulted in improved fitting (Figure 6.11, R_{wp} = 3.29%), but with cell dimensions which are clearly different from those of phase II, Table 6.5. The fitting is not as close as for the patterns obtained between 5.15 and 6.10 GPa, and addition of a Rietveld model of phase II did not yield any improvement (Figure 6.12, R_{wp} = 3.51%). These observations suggest that yet another phase, V, was formed at 6.88 GPa, but the values of the lattice parameters should at this stage be regarded as preliminary results.

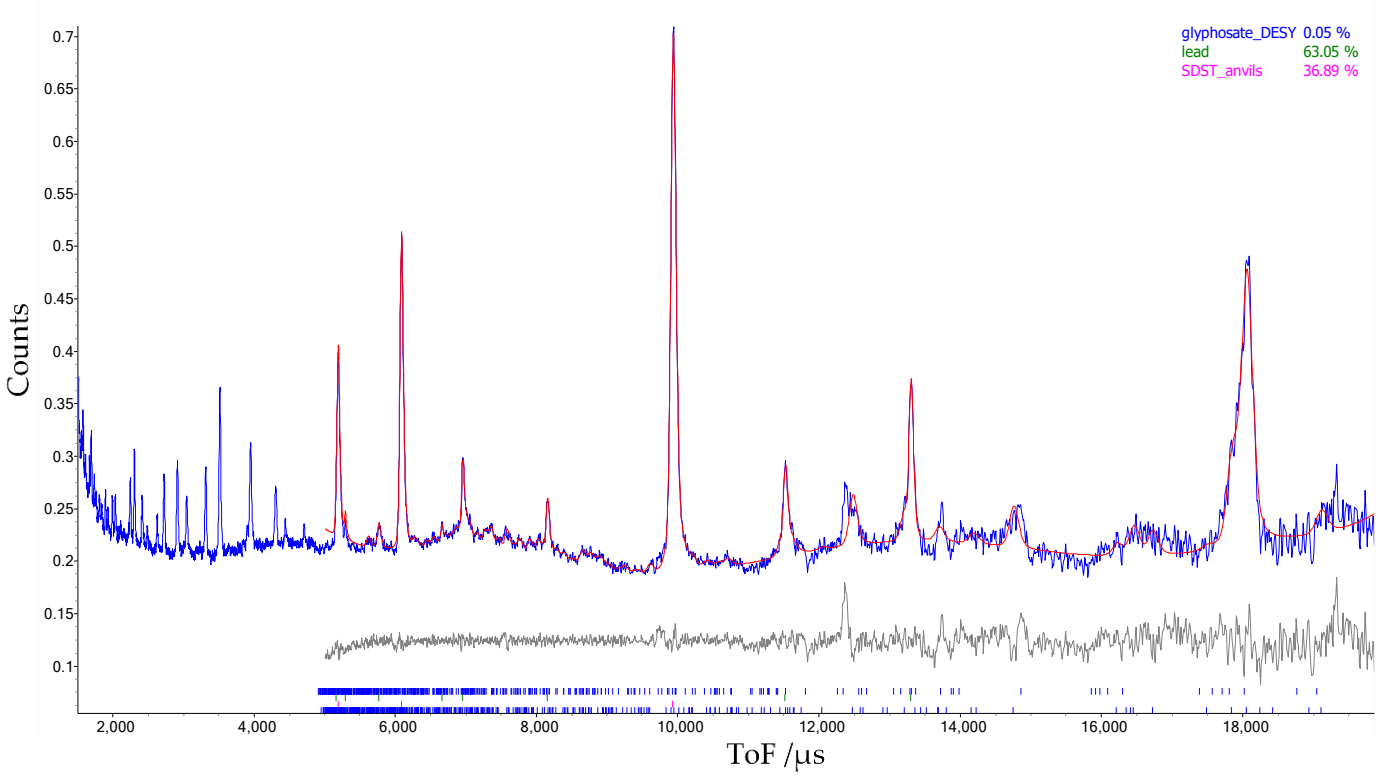


Figure 6.12: Combined Rietveld refinements for phase II and Pawley refinements starting from the 6.10 GPa parameters for the 6.88 GPa dataset, R_{wp} = 3.51%. The Rietveld scale refines to zero, residual peaks can still be seen as an indication of a poor fit.

Table 6.5: Cell dimensions recovered from the Rietveld refinement of phase I at 4.78 GPa, Pawley refinements of phase IV at 5.15, 5.61 and 6.10 GPa and Pawley refinements of phase V at 6.88 GPa. These are compared to the cell dimensions of phases II and III from the X-ray data.

Pressure /GPa	Phase	Cell Volume /Å^3	*a* /Å	*b* /Å	*c* /Å	*β* /°
4.78 (6)	I	557.6 (2)	8.554 (2)	7.5500 (5)	9.273 (2)	111.38 (2)
5.15 (7)	IV	545.8 (6)	8.535 (5)	7.732 (4)	9.421 (7)	118.62 (5)
5.61 (8)	IV	542.8 (7)	8.470 (4)	7.706 (5)	9.437 (7)	118.20 (5)
6.10 (9)	IV	536.4 (8)	8.433 (4)	7.677 (8)	9.391 (7)	118.08 (5)
6.88 (12)	V	531.4 (7)	8.212 (4)	7.747 (7)	9.425 (5)	117.61 (3)
7.03 (5)	II	521.0 (3)	8.204 (1)	7.536 (2)	9.473 (3)	117.18 (3)
5.93 (5)	III	516.94 (8)	4.7277 (3)	7.1064 (4)	15.5653 (19)	98.689 (6)

Figure 6.13 shows the development of the cell dimensions of all phases observed in this study as a function pressure. All plots show a clear discontinuity at 5.15 GPa, consistent with a first order transition between phase I and IV. The cell volume for the points below 4.78 GPa conform to a 3rd order Birch-Murnaghan equation of state (BM3 EoS) with a bulk modulus of 15 (1) (V_0=661 (1), K_0'=7.1, χ^2=2.56). The points for 5.15, 5.61 and 6.10 GPa reflect a consistent trend for phase IV in most plots. The unusual hump seen in the ***c***-axis is possibly a reflection of the sample at 5.15 GPa being mixed phase.

The ***a***, ***b*** and ***c***-axes all show a second discontinuity for the point at 6.88 GPa on the formation of phase V, this difference aligns with the reduced quality of Pawley refinements. The cell volume, ***b*** and ***c***-axes and, to a lesser extent, $\boldsymbol{\beta}$, all show further discontinuities between this point and the X-ray data point for phase II at 7.03 GPa (black triangle in all plots). Moreover, attempts to complete Rietveld refinements based on phase II structural models for the 5.15, 5.61, 6.10 and 6.88 GPa datasets all failed to reproduce the measured patterns. This indicated that although phases IV and V have similar cell dimensions to phase II, they are structurally distinct phases.

The appearance of phases IV and V was unexpected and it is clear that the rapid and slow compression of glyphosate results in significantly different phase behaviour. Although their cell parameters have been used in Pawley-modelling of the diffraction data, we have been unable to solve the structures. Structural models based on molecular models of phases I or II placed into the cells listed in Table 6.5 and optimised by periodic DFT failed to fit the data. Introduction of phase III never yielded acceptable fits for any of the patterns.

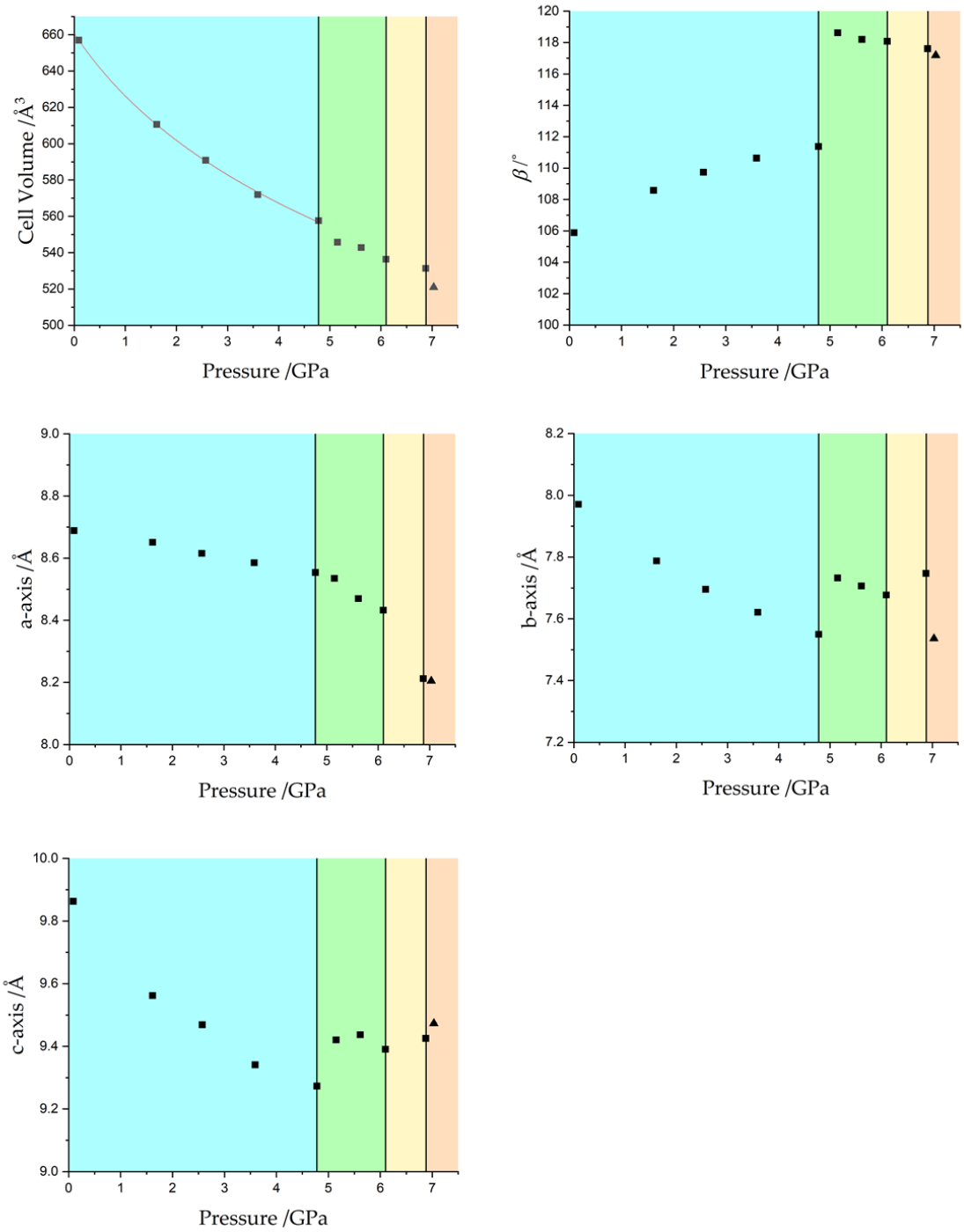


Figure 6.13: Trends in unit cell parameters from Pawley and Rietveld refinements. The black triangle in all plots corresponds to the X-ray phase II structure at 7.03 GPa. Top left: Trends in unit cell volume. Points up to 4.78 GPa have been fitted with a BM3 EoS. Top right: Trends in the β angel. Middle left: trends in the a-axis length. Middle right: trends in the b-axis length. Bottom: trends in the c-axis length. Discontinuities at both I → IV and IV → V transitions can be seen in the plots and are further indicated by black lines.

6.3.3 Thermodynamic Driving Forces of Transitions

To investigate the driving forces behind the transitions to phases II and III, a computational study has been completed to assess the relative contributions towards Gibbs free energy at the transitions. Volume, entropy and internal energy contributions have all been considered, Table 6.6. As the structures of phases IV and V are unknown, the structures considered are from X-ray studies only. All discussions of volumes and energies in this section have been scaled to the difference per molecule for direct comparison between phases and computational methods.

6.3.3*a* The Phase I to II Transition

Extrapolation of the cell dimensions of phase I to 7.03 GPa allows direct comparison between phases I and II. As would be expected from a transition driven by increasing pressure, the unit cell volume reduces at the transition by 1.957 Å^3 per molecule. The equivalent energy contribution towards the Gibbs free energy (Chapter 1) as a result of the volume change, $P\Delta V$, can be calculated by

$$P\Delta V = P \times (V_2 - V_1) \times 0.6 \quad (6.2)$$

where P is the pressure of the transition in GPa, $V_1$and V_2 are the volumes per molecule of the lower and higher-pressure phase in Å^3 respectively, and 0.6 is a conversion factor to kJ mol^{-1}. The volume change contributes −8.25 kJ mol^{-1} to the free energy change for the I → II transition at this pressure.

The final energies of the DFT geometry optimisations show that there is an increase in internal energy at the transition of 4.83 kJ mol^{-1}, a significant

contribution but roughly half the energy reduction conferred by the reduction in volume.

The difference in *TS* on the transition is negligible compared to other energy contributions, a small increase of 0.06 kJ mol^{-1}. It is clear that the rotation of the carboxylate group that occurs at the transition has little impact on the vibrational characteristics of the compound.

The total Gibbs energy change on the I → II transition is −3.36 kJ mol^{-1}. As the only negative contribution comes from the volume term, this indicates that the transition is driven by the ability of phase II to occupy space more efficiently. On the transition the occupied network volume (calculated by the CellVol code introduced in Chapter 3) actually increases by 0.418 $Å^3$ while the void space reduces by 2.374 $Å^3$. This produces an equivalent increase in the packing coefficient from 0.875 to 0.891. Thus, it is more efficient use of void space that drives the transition.

6.3.3*b* The Phase I to III Transition

The extrapolation of the cell parameters of phase I to 5.93 GPa again allows direct comparison of the free energies of phases I and III. The transition is accompanied by a large reduction in volume of 5.761 $Å^3$. This is equivalent to a significant energy difference of −20.50 kJ mol^{-1}. However, over this same range the internal energy increases by 17.76 kJ mol^{-1} and the *TS* term decreases by 0.01 kJ mol^{-1}. The total contribution to the Gibbs energy is therefore a more modest −2.75 kJ mol^{-1}.

Table 6.6: The thermodynamic contributions towards the phase transitions observed. All values have been scaled to per molecule.

Pressure /GPa	Phase	Space group	Cell Volume /Å^3	Network Volume /Å^3	Void Volume /Å^3	Packing coefficient	$-TS$ / kJ mol^{-1}	DFT energy / kJ mol^{-1}	ΔG / kJ mol^{-1}
0	I	$P2_1/c$	164.629	122.570 +/− 0.045	42.058 +/− 0.045	0.745	−52.48	-308817.7793	
5.18	I	$P2_1/c$	136.829	117.343 +/− 0.026	19.486 +/− 0.026	0.858	−45.42	-308761.7569	
5.93	I	$P2_1/c$	134.997	116.335 +/− 0.072	18.662 +/− 0.072	0.862	−45.00	-308752.4191	
5.93	III	$P2_1/n$	129.236	117.167 +/− 0.029	12.069 +/− 0.029	0.907	−45.01	-308734.6588	−2.75
7.03	I	$P2_1/c$	132.216	115.641 +/− 0.046	16.575 +/− 0.046	0.875	−44.39	-308736.7452	
7.03	II	$P2_1/c$	130.259	116.059 +/− 0.038	14.201 +/− 0.038	0.891	−44.33	-308731.9195	−3.37

Again, the volume is the only significant stabilising contribution at the transition. Similar to the transition from phase I to II, phase III forms under pressure due to its significantly improved ability to utilise void space within the structure. Between phase I and phase III at 5.93 GPa, the occupied, network volume increases slightly by 0.832 $Å^3$ while the void volume reduces by 6.593 $Å^3$. This results in an increase in the packing coefficient from 0.862 to the remarkably high 0.907. The formation of this denser polymorph is favourable as pressure increases and the need to minimise volume becomes paramount.

6.4 Conclusions

The aim of the present chapter was to characterise and rationalise the destructive transition seen previously for glyphosate. In attempting to do so we have uncovered a wealth of information about a far more complex structural landscape.

Potentially four new phases have been discovered, but only two of which have been solved with the data available. It is possible that formation of phase IV is the destructive transition seen in Chapter 4. This phase then transforms to phase V with increasing pressure in powder studies. Future attempts to access the structure of phases IV and V could make use of crystal structure prediction or molecular dynamics techniques. A similar procedure has been used to determine the phase of ζ-glycine,[40] a short-lived intermediate phase of the ε and γ phases formed on decompression. Without knowing the structures of phases IV and V it is difficult to analyse the changes in intensity seen for indexed reflections in the *d*-spacing plot. Once these are known however, this could provide valuable insight into the driving forces of the transitions.

Phases II and III were obtained via rapid compression, but the different pressure-transmitting medium used to obtain phase II may also play a role. It is clear from analysis of volumes that phases II and III exist on separate kinetic pathways but it is not yet clear why. Phase III has a lower volume than phase II and appears to form the thermodynamic minimum, it remains unclear why phase II therefore forms preferentially under certain circumstances. Repeated observation of these crystal phases under a greater range of conditions would be highly beneficial.

The driving force in both the phase I → II and phase I → III transitions is volume minimisation. This forms the only significant stabilising

contribution in both cases. For the phase I → II transition the volume contribution is twice as large as the positive difference in internal energy. For the phase I → III transition the case is more extreme, with a large stabilising volume contribution outweighing a large destabilising internal energy contribution. This is in line with what we may expect from the previous chapters. As pressure increases, the need to minimise volume becomes paramount. For both transitions the *TS* contribution negligibly small. This is in line with Nyman and Day's observations that polymorphs can be expected to have similar entropy contributions.[33]

The overall Gibbs energy change on the transition to phase II and III is −3.37 kJ mol^{-1} and −2.75 kJ mol^{-1} respectively. Cruz-Cabeza and co-workers report that the energy differences between polymorphs are usually under 4.2 kJ mol^{-1}.[12] These transitions seem to fall into this range. In both cases the ability to provide the drop in volume that provides the required reduction in energy to drive the transition is through a more efficient use of the space available. A reduction in void space and a higher packing coefficient. For phase II this is completed by a reduction in the backbone of the molecule and a rotation of the carboxylate group. For phase III this is completed by a reconstructive mechanism involving the formation of polymeric shuttling hydrogen bonds.

6.5 References

(1) Fisch, M.; Lanza, A.; Boldyreva, E.; Macchi, P.; Casati, N., Kinetic control of high-pressure solid-state phase transitions: A Case Study on L-serine. *J. Phys. Chem.* **2015**, 119 (32), 18611-18617.

(2) Giordano, N.; Beavers, C. M.; Kamenev, K. V.; Marshall, W. G.; Moggach, S. A.; Patterson, S. D.; Teat, S. J.; Warren, J. E.; Wood, P. A.; Parsons, S., High-pressure polymorphism in L-threonine between ambient pressure and 22 GPa. *CrystEngComm.* **2019**, 21 (30), 4444-4456.

(3) Moggach, S. A.; Allan, D. R.; Morrison, C. A.; Parsons, S.; Sawyer, L., Effect of pressure on the crystal structure of L-serine-I and the crystal structure of L-serine-II at 5.4 GPa. *Acta Crystallogr.* **2005**, B61 (1), 58-68.

(4) Wood, P. A.; Francis, D.; Marshall, W. G.; Moggach, S. A.; Parsons, S.; Pidcock, E.; Rohl, A. L., A study of the high-pressure polymorphs of L-serine using ab initio structures and PIXEL calculations. *CrystEngComm.* **2008**, 10 (9), 1154-1166.

(5) Boldyreva, E. V.; Kolesnik, E. N.; Drebushchak, T. N.; Sowa, H.; Ahsbahs, H.; Seryotkin, Y. V., A comparative study of the anisotropy of lattice strain induced in the crystals of DL-serine by cooling down to 100 K, or by increasing pressure up to 8.6 GPa. A comparison with L-serine. *Z. Kristallogr. Cryst. Mater.* **2006**, 221 (2), 150-161.

(6) Rychkov, D. A.; Stare, J.; Boldyreva, E. V,. Pressure-driven phase transition mechanisms revealed by quantum chemistry: L-serine polymorphs. *Phys. Chem. Chem. Phys.* **2017**, 19 (9), 6671-6676.

(7) Moggach, S. A.; Marshall, W. G.; Parsons, S., High-pressure neutron diffraction study of L-serine-I and L-serine-II, and the structure of L-serine-III at 8.1 GPa. *Acta Crystallogr.* **2006**, B62 (5), 815-825.

(8) Boldyreva, E. V.; Sowa, H.; Seryotkin, Y. V.; Drebushchak, T. N.; Ahsbahs, H.; Chernyshev, V.; Dmitriev, V., Pressure-induced phase transitions in crystalline L-serine studied by single-crystal and high-resolution powder X-ray diffraction. *Chem. Phys. Lett.* **2006**, 429 (4-6), 474-478.

(9) Ridout, J.; Price, L. S.; Howard, J. A. K.; Probert, M. R., Polymorphism arising from differing rates of compression of liquids. *Cryst. Growth Des.* **2014**, 14 (7), 3384-3391.

(10) O'Bannon Iii, E. F.; Husband, R. J.; Baer, B. J.; Lipp, M. J.; Liermann, H.-P.; Evans, W. J.; Jenei, Z., Dynamic compression of Ce and Pr with millisecond time-resolved X-ray diffraction. *Sci. Rep.* **2022**, 12 (1), 17294.

(11) Holanda, R. O.; da Silva, C. B.; Vasconcelos, D. L. M.; Freire, P. T. C., High pressure Raman spectra and DFT calculation of glyphosate. *Spectrochim. Acta.* **2020**, A242, 118745.

(12) Cruz-Cabeza, A. J.; Reutzel-Edens, S. M.; Bernstein, J., Facts and fictions about polymorphism. *Chem. Soc. Rev.* **2015**, 44 (23), 8619-8635.

(13) Moggach, S. A.; Allan, D. R.; Parsons, S.; Warren, J. E., Incorporation of a new design of backing seat and anvil in a Merrill–Bassett diamond anvil cell. *J. Appl. Crystallogr.* **2008**, 41 (2), 249-251.

(14) Merrill, L.; Bassett, W. A., Miniature diamond anvil pressure cell for single crystal X-ray diffraction studies. *Rev. Sci. Instrum.* **1974**, 45 (2), 290-294.

(15) Klotz, S.; Chervin, J. C.; Munsch, P.; Le Marchand, G., Hydrostatic limits of 11 pressure transmitting media. *J. Phys.* **2009**, D42 (7), 075413.

(16) Mao, H. K.; Bell, P. M.; Shaner, J. W. t.; Steinberg, D. J., Specific volume measurements of Cu, Mo, Pd, and Ag and calibration of the ruby R_1 fluorescence pressure gauge from 0.06 to 1 Mbar. *J. Appl. Phys.* **1978**, 49 (6), 3276-3283.

(17) *APEX4*; Bruker AXS Inc.: Madison, Wisconsin, USA, **2021**.

(18) CrysAlisPro, Oxford Diffraction/Agilent Technologies UK Ltd. Yarnton, England: **2014**.

(19) Krause, L.; Herbst-Irmer, R.; Sheldrick, G. M.; Stalke, D., Comparison of silver and molybdenum microfocus X-ray sources for single-crystal structure determination. *J. Appl. Crystallogr.* **2015**, 48 (1), 3-10.

(20) Betteridge, P. W.; Carruthers, J. R.; Cooper, R. I.; Prout, K.; Watkin, D. J., CRYSTALS version 12: software for guided crystal structure analysis. *J. Appl. Crystallogr.* **2003**, 36 (6), 1487-1487.

(21) Macrae, C. F.; Sovago, I.; Cottrell, S. J.; Galek, P. T. A.; McCabe, P.; Pidcock, E.; Platings, M.; Shields, G. P.; Stevens, J. S.; Towler, M., Mercury 4.0: From visualization to analysis, design and prediction. *J. Appl. Crystallogr.* **2020**, 53 (1), 226-235.

(22) Bull, C. L.; Funnell, N. P.; Tucker, M. G.; Hull, S.; Francis, D. J.; Marshall, W. G., PEARL: the high pressure neutron powder diffractometer at ISIS. *High Press. Res.* **2016**, 36 (4), 493-511.

(23) Kuznetsov, A.; Dmitriev, V.; Dubrovinsky, L.; Prakapenka, V.; Weber, H. P., FCC–HCP phase boundary in lead. *Solid State Commun.* **2002**, 122 (3-4), 125-127.

(24) Coelho, A. A., TOPAS and TOPAS-Academic: an optimization program integrating computer algebra and crystallographic objects written in C++. *J. Appl. Crystallogr.* **2018**, 51 (1), 210-218.

(25) Fortes, A. D., A revised equation of state for in situ pressure determination using fcc-Pb (0< P< 13GPa, T> 100 K); STFC: **2019**.

(26) Zolotoyabko, E., Determination of the degree of preferred orientation within the March–Dollase approach. *J. Appl. Crystallogr.* **2009**, 42 (3), 513-518.

(27) Gonzalez-Platas, J.; Alvaro, M.; Nestola, F.; Angel, R., EosFit7-GUI: a new graphical user interface for equation of state calculations, analyses and teaching. *J. Appl. Crystallogr.* **2016**, 49 (4), 1377-1382.

(28) Angel, R. J.; Alvaro, M.; Gonzalez-Platas, J., EosFit7c and a Fortran module (library) for equation of state calculations. *Z. Kristallogr. Cryst. Mater.* **2014**, 229 (5), 405-419.

(29) Clark, S. J.; Segall, M. D.; Pickard, C. J.; Hasnip, P. J.; Probert, M. I. J.; Refson, K.; Payne, M. C., First principles methods using CASTEP. *Z. Kristallogr. Cryst. Mater.* **2005**, 220 (5-6), 567-570.

(30) BIOVIA, *Materials Studio*. Dassault Systèmes, **2021**.

(31) Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, 77 (18), 3865.

(32) Tkatchenko, A.; Scheffler, M., Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data. *Phys. Rev. Lett.* **2009**, 102 (7), 073005.

(33) Nyman, J.; Day, G. M., Static and lattice vibrational energy differences between polymorphs. *CrystEngComm.* **2015**, 17 (28), 5154-5165.

(34) Gavezzotti, A., *Molecular aggregation: structure analysis and molecular simulation of crystals and liquids*. Oxford University Press, **2006**.

(35) Gavezzotti, A., Calculation of lattice energies of organic crystals: the PIXEL integration method in comparison with more traditional methods. *Z. Kristallogr. Cryst. Mater.* **2005**, 220 (5-6), 499-510.

(36) Reeves, M. G.; Wood, P. A.; Parsons, S., MrPIXEL: automated execution of Pixel calculations via the *Mercury* interface. *J. Appl. Crystallogr.* **2020**, 53, 1154-1162.

(37) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, 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.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; 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, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., Gaussian 09, Revision E. 01, Gaussian. *Inc:, Wallingford CT* **2009,** 201.

(38) Dawson, A.; Allan, D. R.; Belmonte, S. A.; Clark, S. J.; David, W. I. F.; McGregor, P. A.; Parsons, S.; Pulham, C. R.; Sawyer, L., Effect of high pressure on the crystal structures of polymorphs of glycine. *Cryst. Growth Des.* **2005**, 5 (4), 1415-1427.

(39) Wood, P. A.; Allen, F. H.; Pidcock, E., Hydrogen-bond directionality at the donor H atom—analysis of interaction energies and database statistics. *CrystEngComm.* **2009**, 11 (8), 1563-1571.

(40) Bull, C. L.; Flowitt-Hill, G.; De Gironcoli, S.; Küçükbenli, E.; Parsons, S.; Pham, C. H.; Playford, H. Y.; Tucker, M. G., ζ-Glycine: insight into the mechanism of a polymorphic phase transition. *IUCrJ.* **2017**, 4 (5), 569-574.

This page has been left intentionally blank

Chapter 7: Concluding Remarks

6.1 Concluding Remarks

The aim of this thesis has been to provide a better understanding of the forces that drive phase transitions. Commonly described solely in terms of the energies of internal interactions, the aim has been to assess to what extent the neglect of volume and entropy is an oversimplification. We have shown how entropy can stabilise interactions with weak internal energy contributions, how volume can drive phase transitions which would otherwise seem very unfavourable and the importance of studying component volumes. This has been completed with a particular focus on elevated pressures. Both computational and experimental methods have been utilised.

Following two introductory chapters, Chapter 3 introduced the CellVol code. The CellVol code enables calculation of occupied and void space in crystal structures which individually provide a spectroscopic level of sensitivity for identification of phase transitions. The code was used to signal a first order phase transition in histidine, identify second order phase transitions in naphthalene, and to identify premonitory behaviour to a destructive phase transition in Cu(L-Asp)(H_2O)$_2$. Chapter 4 extended the use of the CellVol code to identify two second order phase transitions in glyphosate.

Component volumes were used in Chapter 3 to reveal the underlying similarities in the response of molecular materials to pressure. Initially pressure is accommodated by loss of void space, as demonstrated by the narrow range of void bulk moduli calculated for these compounds. The loss of this void space was shown to drive second order phase transitions in both naphthalene and glyphosate. The network volume was shown to have a far greater variation with different ranges for H-bonded and non-H-bonded compounds.

The CellVol code requires atom positions to be well defined and so currently cannot handle disorder. Design of a process for handling disorder would be highly valuable. A recent addition to the Mercury interface (from builds 2023.2.0, the current build at the time of writing) is a new disorder handing tool which allows the various disorder components to be toggled in the visualisation window. The design of CellVol compatibility to run on these individual ordered components should be significantly more manageable than it was previously.

The CellVol code is limited by the quality and accuracy of van der Waals radii available. Under certain conditions the radii used become less appropriate. A key example is for highly ionic compounds where there is an argument for using ionic radii instead. We have found success in the past at identifying phase transitions in ionic compounds using ionic radii rather than van der Waals radii. But does this still provide the same level of information about the molecular volume as atomic van der Waals radii? More generally, knowing what to do in the presence of highly charged centres is complex. Ideally, we would instead use van der Waals radii of ions and calculate these in a similar way to the Alvarez set. The accuracy of this data would depend on the number of observations in the database of different ions. The commonality of Na^+ for example probably already dominates the van der Waals radius of Na. A limited set for the most commonly observed ions could therefore be made. As the database continues to grow these will become better defined.

Applying the CellVol code to porous materials could provide useful information on pore volumes and guest uptake by studying component volumes. Combinations of the MOF (metal-organic framework) and high-pressure subsets in the CSD using filters for *R*-factors under 5, no errors, no disorder, no ions, 3D coordinates and single crystal data, yields 161 structures

(build 2023.2.0). These feature low numbers of observations for many different compounds with only 11 compounds featuring sufficient data for trends to be extracted. Although the data field is currently too limited to make meaningful general observations, as the field of high-pressure crystallography grows, this could be an interesting area.

The role of entropy was investigated in Chapter 5. The entropy contributions of five fluorobenzenes were calculated in an attempt to rationalise the high frequency of C—H···F interactions in the Cambridge Databse. The presence of these interactions was shown to confer high entropy contributions on the scale of the enthalpic contribution of a hydrogen bond. This was shown to have origins in the C—H···F interactions these compounds contain, in part due to the relative weight of fluorine but also because of the deformability of the interactions it forms. This can to some extent explain the high occurrence of C—H···F interactions in the CSD through an entropic stabilisation. The results of this chapter indicate analysis of entropy contributions should be more common, especially in weakly-bound crystal structures.

The main limitation of these calculations is their computational expense. The modest phonon calculation completed for 1,2,4,5-tetrafluorobenzene in Chapter 5 took in excess of three weeks on 36 processors (a very large PC). The increase in computational expense for higher level calculations is exponential. There have been a number of recent advances in the use of semiempirical codes in combination with DFT for calculations of vibrational spectra of reduced computational expense. One example of this is Density Functional Tight Binding (DFTB) as a tight-binding approximation of DFT, a modification of the finite displacement method. These approaches are used commonly in crystal structure prediction but DFTB based methods still

generally underperform in phonon calculations. Advancements in these methods to provide reliable values at much shorter computation times would be very valuable.

The aim of Chapter 5 was to find a way to quantify the entropy contribution of a fragment. This was approximated with two force constant based methods. Method 1 followed calculated vibrational modes with ModeFollow but was not sensitive exclusively to the interaction of interest and is likely to include an intramolecular component. Method 2 probed directly the interaction of interest but through a vibration with no physical basis. A way to directly probe the entropy of a fragment would be very valuable. Our ideas in Chapter 5 of entropy as a description of the deformability of an interaction implies that a measure of entropy could be conferred by the spread of geometries of certain interactions in the CSD. This could enable a structural correlation to be established as an indication of high entropy for interactions. This would be a significant but very interesting project.

The setup of phonon calculations is also an involved process when compared to, for example, an energy calculation on an isolated molecule. There is significant scope to improving the accessibility of these calculations by the design of an automated process such as a comparable program to the MrPixel routine. The design of calculations of improved computational expense will improve the ease of design of such a program.

Finally, in Chapter 6 the importance of the rate of compression in driving high-pressure phase transitions was investigated in an attempt to rationalise the destructive transition seen previously under slow compression of glyphosate in Chapter 4. Single crystal studies revealed phase transitions to two new polymorphs under rapid compression. Compression to 7.03 GPa in neon resulted in phase II and compression to 5.93 GPa in a 4:1 methanol-

ethanol mixture resulted in phase III. Analysis of the thermodynamic driving forces of these transitions from phase I revealed that they are both driven by volume as the only significant stabilising contribution at the transition. The transition from phase III to II is thermodynamically forbidden by an increase in Gibbs free energy and the transition from phase II to III has not yet been observed.

High-pressure neutron powder diffraction was completed under slow compression in an attempt to determine whether either of these new phases was the phase formed in the destructive transition in Chapter 4 and whether phase II could be driven to phase III. This revealed a far more complex structural landscape and appeared to indicate the presence of two further phases (IV, V) that did not correspond to the structural parameters of phase I, II or III. It is believed that the first of these phases is that formed on the destructive transition but so far, further characterisation of either of these phases has not been possible.

It is clear that the rate of compression forms different structural pathways but the full story of glyphosate remains incomplete. The conditions under which phase II or III are formed preferentially under rapid compression are currently unclear. A greater number of observations under changing conditions would be highly beneficial for mapping out the condition dependence of these transitions.

Attempts should be made to recover structural parameters for the two phases seen under slow compression of the powder by using crystal structure prediction or molecular dynamics methods. As noted in Chapter 6, there has been success recovering structures from similar structural relationships for ζ-glycine. By comparison with the structure of phase I this would allow assessment of the structural rearrangement that results in the destructive

transition seen for single crystals. Recovery of these structures would also allow similar analysis of the thermodynamics driving forces of the transitions to be assessed and rationalisation of whether a transition at higher pressure to phases II or III is likely from either of these phases. This would reveal whether phases II and III are only accessible under rapid compression for single crystals by 'jumping' over the destructive transition.

Molecular compounds are highly prone to the formation of multiple forms. Understanding why certain forms exist and the conditions under which they interconvert is vital to understanding the world we live in. Within this thesis attempts have been made to highlight structural behaviour that cannot be rationalised by consideration of the internal energy of intermolecular interactions alone. In fact, for the high-pressure phases of glyphosate, internal energy changes would indicate that their formation is a highly unfavourable process.

For the vast majority of compounds consideration of internal energies IS sufficient. The aim of this thesis was therefore to provide a caveat that when internal energies do not seem to appropriately describe phase behaviour, it is also worth considering volumes and entropies.

This page has been left intentionally blank

Appendices

Appendix A- Chapter 3

A.1 Optimising the Density of Points for CellVol Calculations

In order to optimise the number of random points required for calculations, the variation of the volume was calculated for different numbers of points using structures with small and large unit cell volumes. Figure A.1 shows extreme examples of the variance for the lowest and highest unit cell volumes likely to be encountered by the program. The graphs featured are for formamide (FORMAM02), one of the smallest-volume molecular solids (unit cell 224 $Å^3$) and for a porous chromium terephthalate MIL-101 (OCUNAC, unit cell 701860 $Å^3$) the largest non-disordered structure on the Cambridge Database with coordinates. For both structures network and void volumes have converged by 1 million points.

The time taken for calculations using 1 million points per run that achieve a precision of <0.1% for the network volume is shown for a range of materials with increasing unit cell volume in Figure A.2. The plot is overlaid on a histogram of unit cell volumes for structures in the database. This figure shows that below 1000 $Å^3$ the calculation time is relatively independent of unit cell size but becomes linear beyond this. Calculations on ~97% of structures in the Cambridge Database are completed within 5 minutes. The largest non-disordered structure in the database which contains coordinates, OCUNAC, completed within 24 minutes. It is important to note that calculation wall-clock times presented here are a demonstrative guide only and will be dependent on the computer used for the calculations. The values quoted were obtained on a modest desktop PC with an Intel® Core™ i7-9700 CPU with a base speed of 3.00 GHz.

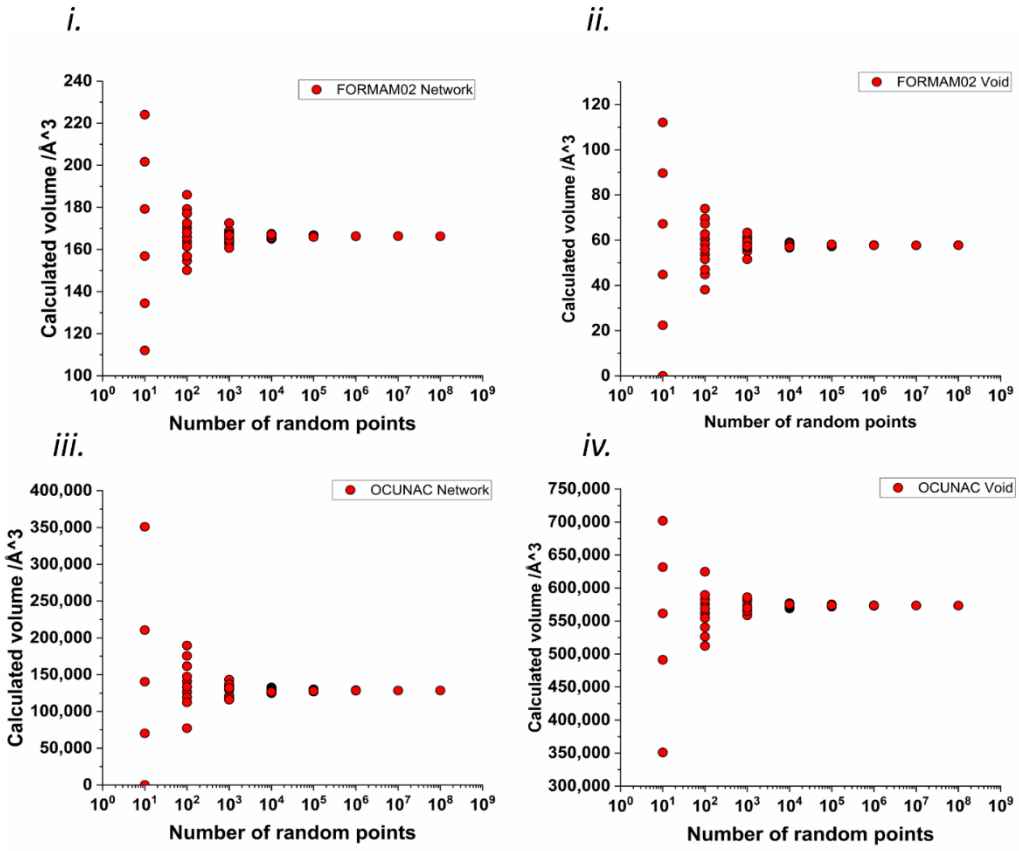


Figure A.1: The convergence of the network and void volumes of FORMAM02 and OCUNAC with increasing numbers of random points.

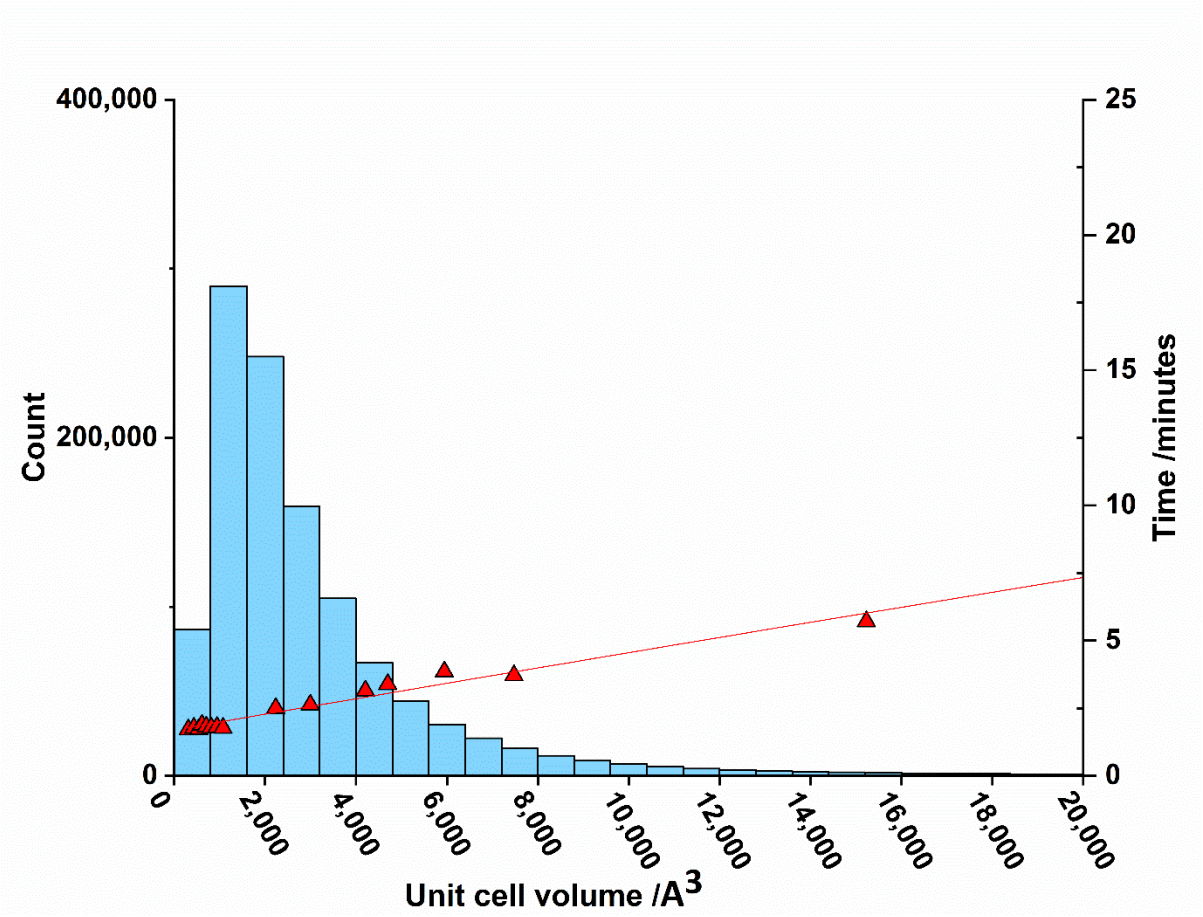


Figure A.2: Calculation times for a range of unit cell volumes overlaid on a histogram of unit cell volumes in the CSD. The volume range covers 99.1% of all structures in the CSD.

A.2 Equation of State Fitting

Equations of state have been fitted to the variation of network and void volume with pressure for all structures on the Cambridge Database which were determined above 0.1 GPa before 2019. The results are presented in Table 1 of the electronic supplementary information. Pressures are given in a variety of units in refcode entries on the CSD, but they were all converted to GPa for this study. At least five points in a pressure series were necessary for inclusion, structures which contained no coordinates, had unhandled errors or contained disorder were also removed. The final dataset contained 1472 separate refcodes on 129 different refcode families. Systematic errors were seen when

data from several separate studies or even separate diamond anvil cell loadings were combined, and where possible data were taken from contiguous sets of measurements.

The EoSFit program described in Chapter 2 was used to fit volume-pressure data using Birch-Murnaghan or Vinet equations of states (EoSs) described in Chapter 1. Network data were generally fitted well with a 2nd or 3rd order Birch-Murnaghan EoS, whereas void data were modelled better with a (3rd order) Vinet EoS. For some structures network volume was fitted poorly by EoS and for these structures Equation A.1 was used,

$$K = -\tilde{V} \times \frac{\Delta P}{\Delta V} \qquad (A.1)$$

where K is the bulk modulus, $\tilde{v}$ is the average volume, ΔP is the change in pressure and ΔV is the change in volume.

Where multiple polymorphs were present for one compound, fitting has been attempted for each of the forms individually. However, for several compounds either a complex structural landscape or incompatible measurement techniques has meant no meaningful fitting was possible. Ambient pressure data points have been used to fix V_0 where possible, and for these entries V_0 is given without a standard deviation. However, for many compounds a significantly improved fit was found by allowing V_0 to vary during fitting and standard deviations are quoted for these entries.

In some cases, a poor EoS fit results from discontinuous changes of gradient which may imply some form of structural rearrangement such as a change in compression mechanism or phase transition. For the purposes of this list, discontinuities (especially in the network curves) that are observed in the absence of an identified first order phase transition are described as possible second order transitions. Comments have been added in these cases. This behaviour seems to be quite common but difficult to recognise in plots of

total unit cell volume versus pressure. The changes in some network volumes for example that of ABUMOZ span only a small range of volume increasing the relative importance of precision. It is important to note the standard deviation used for fitting were taken from the Monte Carlo estimates of network and void volume and do not include any experimental uncertainties and this could account for the generally high values of χ^2. Some fits also yield negative values of K' implying a slightly convex form to the volume-pressure curve.

A.3 Comparison of Network and Void Bulk Moduli

Figure A.3 shows all network bulk moduli plotted against all void bulk moduli from this study. There is no clear correlation in these plots, indicating that a high or low network bulk modulus has very little impact on the compressibility of the voids. No matter how incompressible a network, the way in which it 'uses' void space at elevated pressure to minimise volume is consistent across a very broad range of molecular solids.

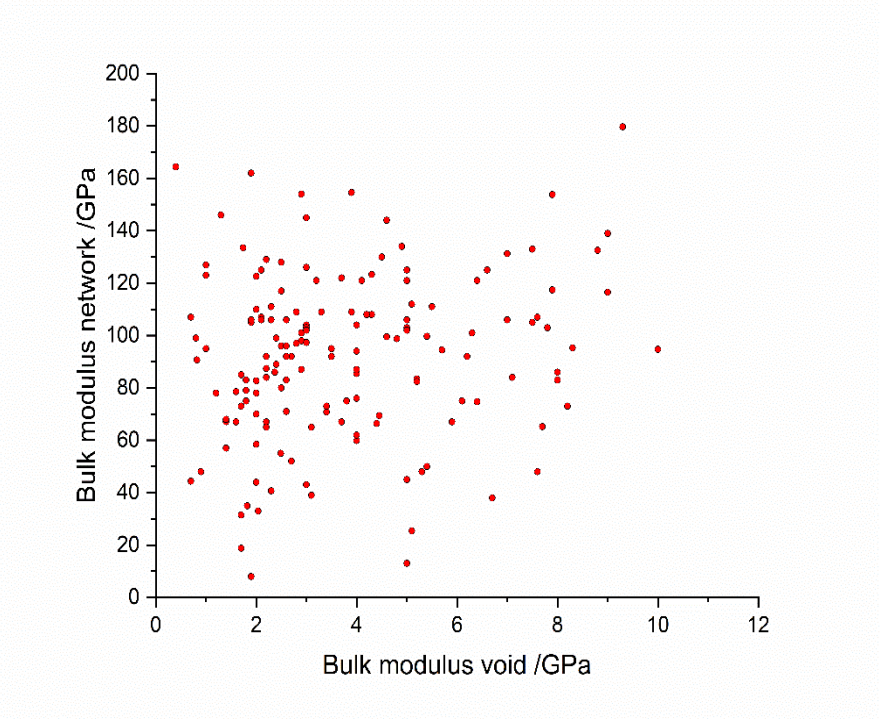


Figure A.3: Network bulk moduli vs void bulk moduli for all applicable structures. No correlation is seen.

Appendix B- Chapter 4

B.1 Occupied Volume Calculations

The uncertainties in the experimental structural parameters were propagated as described in Chapter 3 with Gaussian random numbers taken from the distribution defined by each coordinate's standard uncertainty to produce 50 structures. 50 CellVol calculations yielded a distribution of V_{net} and V_{void} from which the standard uncertainties of each was evaluated. Since precision is not defined for DFT-optimised structures, the standard uncertainty of V_{net} and V_{void} given for these structures reflect only the reproducibility of each Monte Carlo run, typically ≤0.1% of the network volume.

In an effort to identify which intermolecular interactions undergo the largest changes at the transitions, the volumes of the glyphosate molecule, each of the intermolecular dimers, A-H, and layer components were calculated by placing a box around them. The box is defined by producing a reference frame based on the three principal inertial axes of the molecules using the algorithm described by Gavezzotti.[2] The atomic coordinates of all the points within the molecules were used to construct an inertial matrix, ignoring atomic masses, and the eigenvalues of this matrix used to define the principal moments of inertia. The origin of our inertial reference frame is then taken as the centre of all atomic coordinates with the eigenvectors of the inertial matrix defining the rotation matrix between the original and inertial reference frame. By expressing the atomic coordinates within the inertial reference frame, the extreme points occupied by atoms of the molecule in each of the three inertial axes can then be located. These are used to define the box extended in all three dimensions beyond these values by a set length (2.5 Å in this work). When calculated for a single isolated molecule, the longest inertial axis is defined as

the long axis of the molecule, used for discussions of the second transition within the main body of the text. The occupied volume was then evaluated using a similar Monte Carlo procedure to that described in Chapter 3 for unit cells. The layer elements used in layer occupied volume calculations are shown in Figure A.4. Occupied volume plots for each of the dimers is presented in Figure A.5.

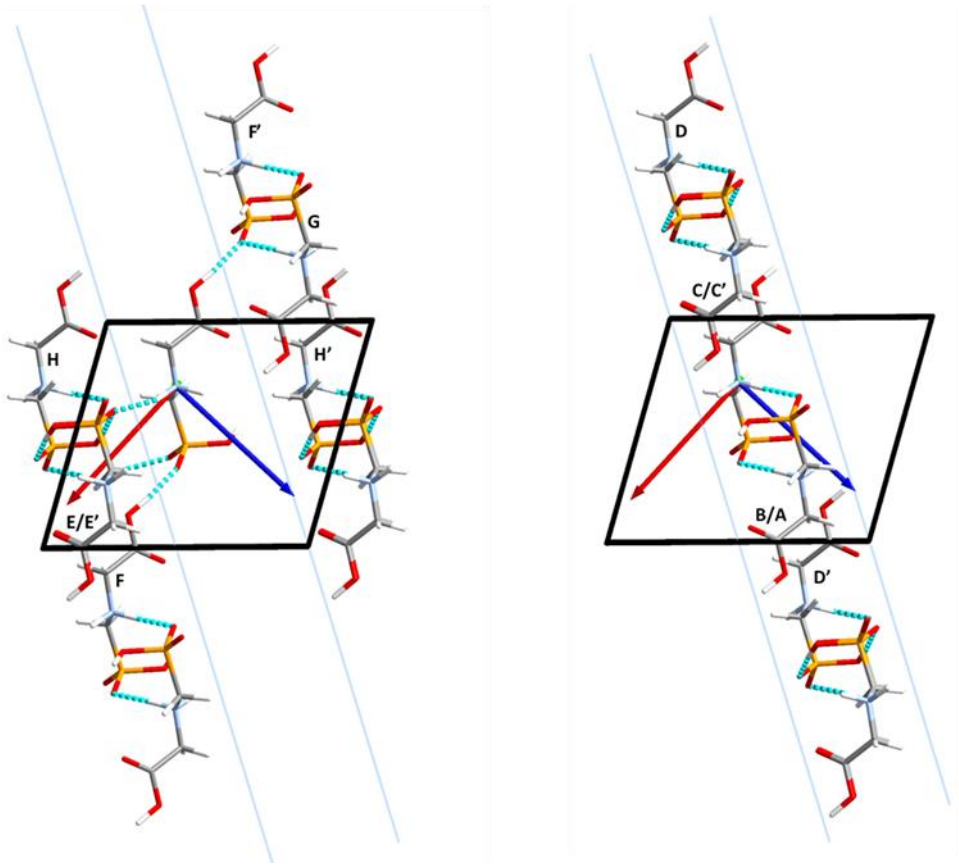
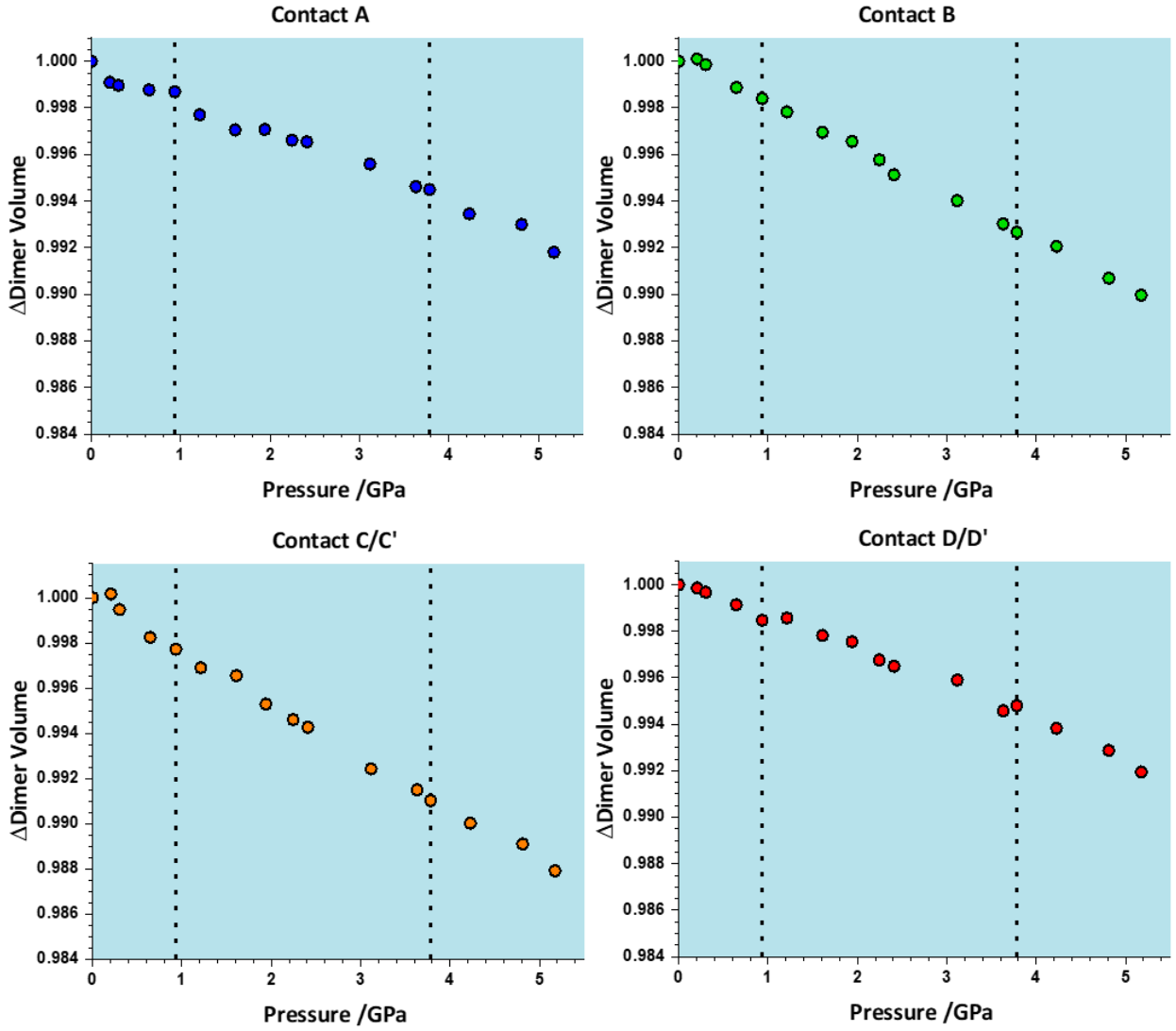


Figure A.4: Contacts which make up the inter (left) and intra (right) layer volumes presented in Chapter 4. Note that labels refer to molecule - molecule interactions, as listed in Table 4.1, rather than individual contacts such as H-bonds.


Contact A
Contact B
Contact C/C'
Contact D/D'
ΔDimer Volume
Pressure /GPa
1.000
0.998
0.996
0.994
0.992
0.990
0.988
0.986
0.984
0
1
2
3
4
5

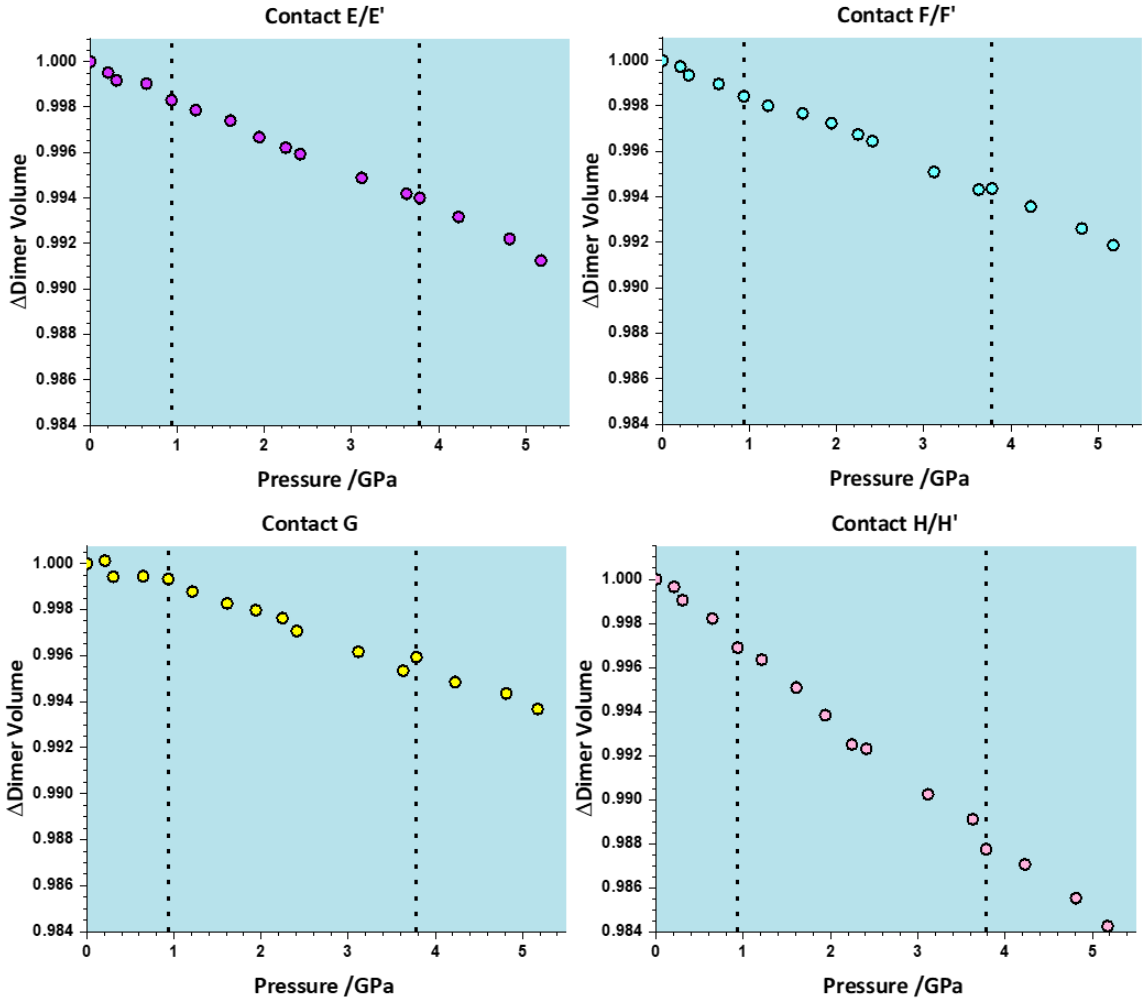


Figure A.5: Molecular volume in each contact listed in Table 4.1. A consistent set of vertical axes is used across all contacts to aid comparison.

B.2 Crystallography Tables

Data tables for high pressure crystal structures with refinement statistics are given below.

Table A.1: Experimental parameters of high-pressure single-crystal measurements.

Pressure /GPa	0, DAC	0.206	0.302	0.645	0.934	1.210	1.609	1.940
				Crystal data				
Temperature (K)	293	293	293	293	293	293	293	293
Spacegroup	$P2_1/c$	$P2_1/c$	$P2_1/c$	$P2_1/c$	$P2_1/c$	$P2_1/c$	$P2_1/c$	$P2_1/c$
a, b, c (Å)	8.6693(13) 7.9704(6) 9.8968(7)	8.6720(6) 7.9508(3) 9.8213(3)	8.6626(10) 7.9208(4) 9.7820(5)	8.6567(16) 7.8828(7) 9.7114(9)	8.6512(13) 7.8496(6) 9.6513(8)	8.6419(10) 7.8197(4) 9.6036(6)	8.6334(14) 7.7827(7) 9.5468(9)	8.6285(9) 7.7531(4) 9.4944(5)
β (°)	105.644(9)	106.235(4)	106.599(7)	107.244(11)	107.785(9)	108.174(7)	108.729(10)	109.104(5)
V (Å^3)	658.51(12)	650.17(6)	643.22(9)	632.91(15)	624.08(12)	616.61(9)	607.49(13)	600.17(8)
Z	4	4	4	4	4	4	4	4
μ (mm^{-1})	0.203	0.206	0.208	0.211	0.214	0.217	0.220	0.223
Crystal size (mm)	0.15 × 0.135 × 0.05	0.19 × 0.17 × 0.05	0.15 × 0.135 × 0.05	0.15 × 0.15 × 0.05	0.15 × 0.135 × 0.05	0.15 × 0.135 × 0.05	0.15 × 0.135 × 0.05	0.19 × 0.17 × 0.05
				Data Collection				
Wavelength (Å)	0.56086	0.56086	0.56086	0.56086	0.56086	0.56086	0.56086	0.56086
Absorption correction (wR^2_{bef}, wR^2_{aft}, max : min transmission)	0.0895, 0.0623, 0.6587	0.0581, 0.0477, 0.8968	0.0856, 0.0589, 0.6493	0.0912, 0.0666, 0.7006	0.0833, 0.0627, 0.7371	0.0884, 0.0645, 0.6571	0.0855, 0.0635, 0.7825	0.0589, 0.0464, 0.8885
T_{min}, T_{max}	0.4904, 0.7445	0.6677, 0.7445	0.4834, 0.7445	0.5216, 0.7445	0.5488, 0.7445	0.4892, 0.7445	0.5826, 0.7445	0.6615, 0.7445
Measured, independent and observed reflections	3682, 517, 365	5852, 511, 435	5682, 517, 376	5611, 506, 375	5430, 490, 383	5409, 494, 382	5243, 471, 373	5320, 468, 416
R_{int}	0.0762	0.0422	0.0703	0.0783	0.0738	0.0724	0.0762	0.0363
$(\sin\theta/\lambda)_{max}$ (Å^{-1})	0.625	0.627	0.625	0.625	0.625	0.627	0.625	0.626
				Refinement				
$R[F^2>2\sigma(F^2)]$, $wR(F^2)$, S	0.0413, 0.1108, 1.069	0.0253, 0.0642, 1.071	0.0369, 0.1008, 1.052	0.0352, 0.0937, 1.097	0.0375, 0.0820, 1.052	0.0365, 0.0946, 1.042	0.0298, 0.0720, 1.067	0.0236, 0.0622, 1.103
Data completeness	0.394	0.391	0.404	0.400	0.395	0.399	0.391	0.393
No. of reflections	517	511	517	506	490	494	471	468
No. of parameters	93	93	93	93	94	93	93	93
No. of restraints	82	82	82	82	82	82	82	82
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.216, -0.242	0.103, -0.144	0.208, -0.283	0.214, -0.238	0.165, -0.168	0.189, -0.212	0.134, -0.132	0.120, -0.138

Pressure /GPa	2.244	2.410	3.117	3.630	3.783	4.228	4.813	5.176
				Crystal data				
Temperature (K)	293	293	293	293	293	293	293	293
Spacegroup	$P2_1/c$	$P2_1/c$	$P2_1/c$	$P2_1/c$	$P2_1/c$	$P2_1/c$	$P2_1/c$	$P2_1/c$
	8.6185(7)	8.616(5)	8.5889(9)	8.5740(13)	8.562(2)	8.5435(9)	8.526(2)	8.496(5)
a, *b*, *c* (Å)	7.7238(3)	7.706(2)	7.6517(4)	7.6158(5)	7.6058(9)	7.5706(4)	7.5312(9)	7.510(2)
	9.4573(4)	9.438(3)	9.3709(5)	9.3347(7)	9.3306(12)	9.2993(5)	9.2654(12)	9.249(3)
β (°)	109.454(4)	109.78(3)	110.256(5)	110.657(7)	110.868(13)	111.200(6)	111.625(13)	111.96(3)
V (Å^3)	593.61(6)	589.6(4)	577.77(8)	570.35(11)	567.79(18)	560.77(8)	553.06(17)	547.3(4)
Z	4	4	4	4	4	4	4	4
μ (mm^{-1})	0.225	0.227	0.231	0.234	0.235	0.238	0.242	0.244
Crystal size (mm)	0.19 × 0.17 × 0.05	0.15 × 0.135 × 0.05	0.19 × 0.17 × 0.05	0.19 × 0.17 × 0.05	0.15 × 0.135 × 0.05	0.15 × 0.135 × 0.05	0.15 × 0.135 × 0.05	0.15 × 0.135 × 0.05
				Data Collection				
Wavelength (Å)	0.56086	0.56086	0.56086	0.56086	0.56086	0.56086	0.56086	0.56086
Absorption correction (wR^2_{bef}, wR^2_{aft}, max : min transmission)	0.0580, 0.0473, 0.8547	0.0823, 0.0662, 0.7907	0.0562, 0.0466, 0.8880	0.0759, 0.0600, 0.7511	0.0830, 0.0617, 0.8156	0.0732, 0.0614, 0.8109	0.0679, 0.0568, 0.8423	0.0697, 0.0560, 0.8176
T_{min}, T_{max}	0.6363, 0.7445	0.5887, 0.7445	0.6611, 0.7445	0.5592, 0.7445	0.6072, 0.7445	0.6037, 0.7445	0.6271, 0.7445	0.6087, 0.7445
Measured, independent and observed reflections	5389, 468, 411	4649, 479, 376	5234, 466, 418	5032, 446, 399	4804, 532, 408	4847, 527, 412	4936, 526, 420	4692, 551, 426
R_{int}	0.0378	0.0740	0.0371	0.0502	0.0726	0.0701	0.0619	0.0587
$(\sin\theta/\lambda)_{max}$ (Å^{-1})	0.627	0.625	0.627	0.626	0.630	0.625	0.625	0.625
				Refinement				
$R[F^2>2\sigma(F^2)]$, $wR(F^2)$, S	0.0242, 0.0658, 1.097	0.0348, 0.0760, 1.038	0.0236, 0.0624, 1.093	0.0262, 0.0659, 1.141	0.0397, 0.0920, 1.045	0.0366, 0.0883, 1.092	0.0333, 0.0797, 1.106	0.0377, 0.0985, 1.139
Data completeness	0.390	0.402	0.397	0.388	0.464	0.467	0.471	0.497
No. of reflections	468	479	466	446	532	527	526	551
No. of parameters	93	93	93	93	93	93	93	93
No. of restraints	82	82	82	82	82	82	82	82
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.111, -0.144	0.188, -0.188	0.127, -0.137	0.117, -0.139	0.208, -0.250	0.202, -0.234	0.206, -0.274	0.222, -0.279

Pressure /GPa	0, fibre
Crystal data	
Temperature (K)	270
Spacegroup	$P2_1/c$
a, *b*, *c* (Å)	8.6764(5) 7.9722(5) 9.8883(6)
β (°)	105.691(2)
V (Å^3)	658.48(7)
Z	4
μ (mm^{-1})	0.384
Crystal size (mm)	0.15 × 0.135 × 0.05
Data Collection	
Wavelength (Å)	0.71073
Absorption correction (wR^2_{bef}, wR^2_{aft}, max : min transmission)	0.0923, 0.0443, 0.9144
T_{min}, T_{max}	0.6822, 0.7461
Measured, independent and observed reflections	12666, 1981, 1747
R_{int}	0.0318
$(\sin\theta/\lambda)_{max}$ (Å^{-1})	0.715
Refinement	
$R[F^2>2\sigma(F^2)]$, $wR(F^2)$, S	0.0285, 0.0805, 1.112
Data completeness	0.982
No. of reflections	1981
No. of parameters	93
No. of restraints	0
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.419, -0.324

B.3 Strain Tensor Results

The Strain tensor was calculated as described in Hazen, R. M., Finger, L. W., *Comparative Crystal Chemistry*, John Wiley & Sons, **1982**. Eigenvalues and vectors were calculated using JACOBI from Press, W. H., Flannery, B. P., Teukolsky S. A. and Vetterling, W. T., *Numerical recipes in fortran: The art of scientific computing*, Cambridge University Press, **1992**. Results at each pressure are represented below in Table A.2. Table A.3 features these results scaled by GPa which are also plotted in Figure A.6. Glyphosate shows remarkably 2D compression with strain axes 1 barely varying from zero. After a minor

increase, this strain axes decreases indicating minor negative linear compressibility.

Table A.2: Calculated strain axes in glyphosate.

Pressure /GPa	1	2	3	a	b	c	α	β	γ
0	0	0	0	8.669	7.97	9.897	90	105.644	90
0.206	0.001613	-0.00246	-0.01187	8.672	7.951	9.821	90	106.235	90
0.302	0.001687	-0.00622	-0.01886	8.663	7.921	9.782	90	106.599	90
0.645	0.002734	-0.01099	-0.03112	8.657	7.883	9.711	90	107.244	90
0.934	0.003467	-0.01516	-0.0415	8.651	7.85	9.651	90	107.785	90
1.21	0.003363	-0.01891	-0.04944	8.642	7.82	9.604	90	108.174	90
1.609	0.003856	-0.02355	-0.05981	8.633	7.783	9.547	90	108.729	90
1.94	0.003874	-0.02726	-0.06788	8.628	7.753	9.494	90	109.104	90
2.244	0.003668	-0.03094	-0.07463	8.618	7.724	9.457	90	109.454	90
2.41	0.004416	-0.03319	-0.07967	8.616	7.706	9.438	90	109.782	90
3.117	0.002475	-0.03999	-0.0904	8.589	7.652	9.371	90	110.256	90
3.63	0.00201	-0.04449	-0.09778	8.574	7.616	9.335	90	110.657	90
3.783	0.001865	-0.04574	-0.10068	8.562	7.606	9.331	90	110.868	90
4.228	0.000852	-0.05016	-0.10708	8.543	7.571	9.299	90	111.2	90
4.813	0.000292	-0.0551	-0.11468	8.526	7.531	9.265	90	111.625	90
5.176	-0.00119	-0.05781	-0.12043	8.496	7.51	9.249	90	111.961	90

Table A.3: Calculated strain axes in glyphosate per GPa.

Pressure /GPa	1	2	3	*a*	*b*	*c*	*α*	*β*	*γ*
0	0	0	0	8.669	7.97	9.897	90	105.644	90
0.206	0.007829	-0.01194	-0.05763	8.672	7.951	9.821	90	106.235	90
0.302	0.005587	-0.02061	-0.06246	8.663	7.921	9.782	90	106.599	90
0.645	0.004239	-0.01704	-0.04825	8.657	7.883	9.711	90	107.244	90
0.934	0.003712	-0.01623	-0.04443	8.651	7.85	9.651	90	107.785	90
1.21	0.002779	-0.01563	-0.04086	8.642	7.82	9.604	90	108.174	90
1.609	0.002396	-0.01464	-0.03717	8.633	7.783	9.547	90	108.729	90
1.94	0.001997	-0.01405	-0.03499	8.628	7.753	9.494	90	109.104	90
2.244	0.001635	-0.01379	-0.03326	8.618	7.724	9.457	90	109.454	90
2.41	0.001832	-0.01377	-0.03306	8.616	7.706	9.438	90	109.782	90
3.117	0.000794	-0.01283	-0.029	8.589	7.652	9.371	90	110.256	90
3.63	0.000554	-0.01226	-0.02694	8.574	7.616	9.335	90	110.657	90
3.783	0.000493	-0.01209	-0.02661	8.562	7.606	9.331	90	110.868	90
4.228	0.000201	-0.01186	-0.02533	8.543	7.571	9.299	90	111.2	90
4.813	0.000061	-0.01145	-0.02383	8.526	7.531	9.265	90	111.625	90
5.176	-0.00023	-0.01117	-0.02327	8.496	7.51	9.249	90	111.961	90

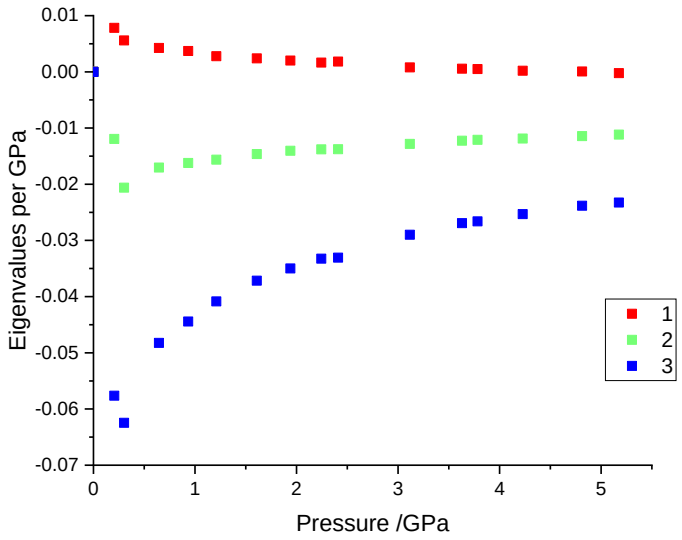


Figure A.6: Eigenvalues of the strain tensor per GPa as a function of pressure.

Appendix C- Chapter 5

C.1 Cell Optimisation

To provide a suitable level of geometrical convergence for the phonon calculations, experimental crystal structures were geometry-optimised as described in Chapter 5. The optimised unit cell dimensions are compared to experimental values in Table A.4. The structures are all monoclinic except for **1F**, which is tetragonal.

Table A.4: Comparison between the experimental and optimised unit cells for all structures studied.

	FACFAQ (**1F**)		FACFOE(**1,2F**)		FACGEV(**1,4F**)		PVVAWA01(**1,3,5F**)		FACJAU(**1,2,4,5F**)	
	Exp	Opt	Exp	Opt	Exp	Opt	Exp	Opt	Exp	Opt
a/Å	5.799(2)	5.7515	7.4806(11)	7.3941	5.809(2)	5.7979	6.160(2)	6.0959	4.4719(11)	4.4285
b/Å	5.799(2)	5.7515	5.9608(9)	6.0073	6.530(2)	6.3318	11.909(3)	11.9534	10.285(2)	10.2540
c/Å	14.530(7)	14.4875	11.725(2)	11.8097	7.190(2)	7.0582	7.504(2)	7.4034	6.342(2)	6.4289
β/°	90.00	90.00	103.815(11)	103.9730	101.89(2)	101.9271	95.47(2)	95.1573	107.97(2)	107.8677

C.2 Comparison of Aromatic H···F Interactions

Searches of the Cambridge Structural Database were conducted to investigate whether the formation of C—H···F interactions is simply a consequence of the optimisation of aromatic interactions. Relevant interactions were defined as those shorter than the sum of the contributing van der Waals radii. Data without 3D coordinates, R-factors above 5%, containing disorder, errors, powder data or ions were filtered out of searches. Hydrogen bond lengths were normalised to neutron values. Searches were conducted on the CSD 2023.2.0 dataset. The average distances are shown in Table A.5, no significant

systematic difference between the lengths of C—H···F contacts in aromatic and non-aromatic systems are seen.

Table A.5: A comparison of the average bond length of intermolecular H···F interactions in the CSD and the aromatic and non-aromatic subsets of this.

Interaction		Length /Å
1, All Intermolecular	X—H···F—X	2.51750
2, Aromatic	Ar—H···F—Ar	2.51252
3, Non-Aromatic	1 and not 2	2.51933

C.3 Convergence of Phonon Calculations

Phonon frequencies for all structures were calculated in the harmonic approximation using the linear response method implemented in CASTEP (as opposed to the finite displacement method based on supercells) using a grid of **q**-points with a spacing of 0.04 $Å^{-1}$. The results were then interpolated onto a finer grid of 16 x 16 x 16 **q**-points. [Note: **q**-points define the positions in reciprocal space where phonon frequencies are calculated.] Our approach was based on the following convergence tests.

Calculations of increased computational expense were applied to monofluorobenzene, **1F**. An increase in the density of **q**-points from 0.04 $Å^{-1}$ to 0.03 $Å^{-1}$, yielding an increase from 3 to 12 **q**-points, was applied. Levels of interpolation from 1 x 1 x 1 to 16 x 16 x 16 were then applied to the results obtained at 0.04 and 0.03 $Å^{-1}$ For even-numbered interpolations (*e.g.* 2 x 2 x 2), a grid shift was applied to include the Γ-point in the interpolation grid. The value of *TS* at 150 K was calculated for each interpolated **q**-point set. The results are tabulated in Table A.6.

The value of *TS* converged with a grid of 16 × 16 × 16 to within 0.01 kJ mol^{-1} for calculations based on the 0.04 and 0.03 $Å^{-1}$ grids. At this level of

interpolation, the difference between the values of *TS* for the two sets of calculations was of the order of 0.1 kJ mol^{-1}.

A comparison of the DoS and calculated INS spectra for the 0.04 and 0.03 Å^{-1} **q**-point sets at the highest level of interpolation are provided in Figure A.7 and Figure A.8 respectively; note that the integrals of both DoS plots are normalised to three times the number of atoms in one unit cell. The differences are very small.

The minor impacts on calculated entropies, DoS plots and INS spectra led to a **q**-point spacing of 0.04 Å^{-1} interpolated to 16 x 16 x 16 to be applied for all further calculations in Chapter 5.

Table A.6: Effect of **q**-point density on the value of *TS* at 150 K for **1F**.

Interpolation level	3 **q**-points (0.04 Å^{-1})	12 **q**-points (0.03 Å^{-1})
1 × 1 × 1	8.62568	8.62568
2 × 2 × 2	10.23244	10.48119
3 × 3 × 3	10.21952	10.34219
4 × 4 × 4	10.30388	10.42043
5 × 5 × 5	10.28744	10.39312
6 × 6 × 6	10.30705	10.41631
7 × 7 × 7	10.30138	10.40712
8 × 8 × 8	10.31065	10.41693
9 × 9 × 9	10.30564	10.41235
10 × 10 × 10	10.30979	10.41631
11 × 11 × 11	10.30758	10.41490
12 × 12 × 12	10.31058	10.41637
13 × 13 × 13	10.30893	10.41499
14 × 14 × 14	10.31087	10.41852
15 × 15 × 15	10.30932	10.41554
16 × 16 × 16	10.31041	10.41719

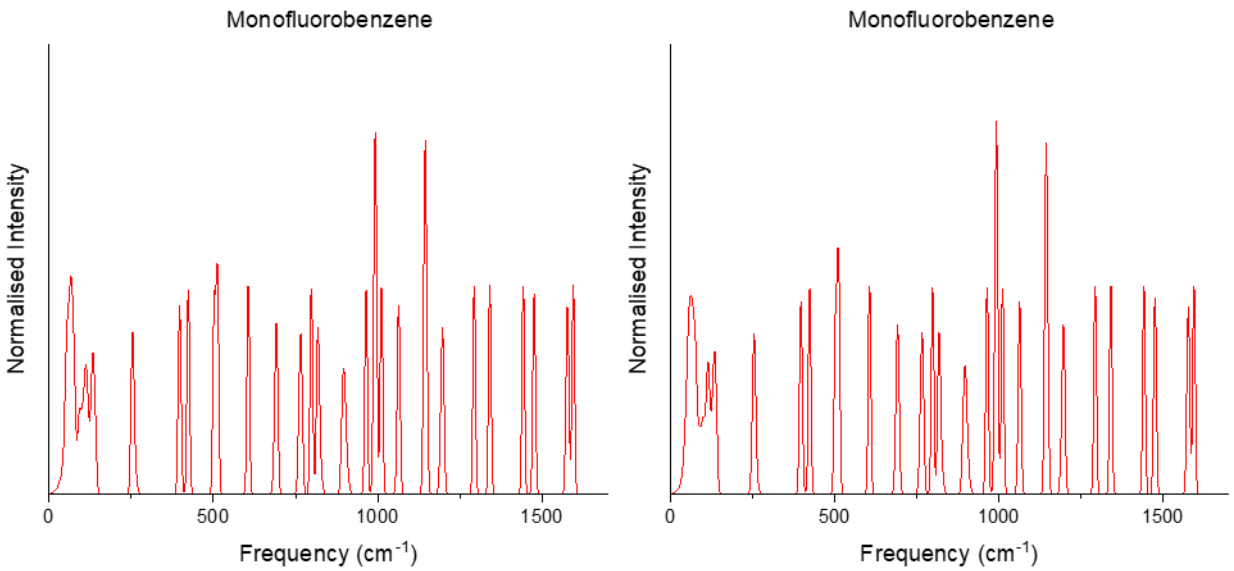


Figure A.7: The impacts of **q**-point density on DoS plots of **1F**. Left: 3 **q**-points (0.04 Å^{-1} spacing), right: 12 **q**-points (0.03 Å^{-1} spacing) both are interpolated to a 16 × 16 × 16 grid.

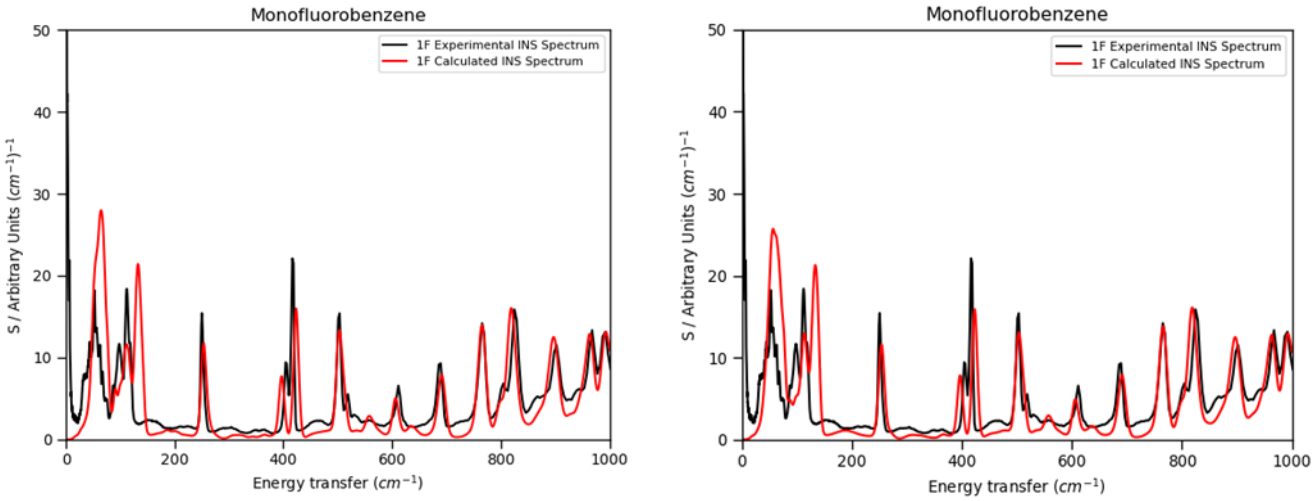


Figure A.8: The impacts of **q**-point density on INS spectra of **1F**. Left: 3 **q**-points (0.04 Å^{-1} spacing), right: 12 **q**-points (0.03 Å^{-1} spacing) both are interpolated to a 16 × 16 × 16 grid.

C.4 SAPT Calculation Convergence

Dimer energies were calculated for mapping potentials using symmetry adapted perturbation theory (SAPT) through the program Psi4. Convergence testing of the most appropriate level of theory for these calculations was completed using a dimer of 1,2-difluorobenzene, **1,2F**, contact C, a bridged C—H···F interaction (below). Combinations of the double, triple and quadruple zeta basis sets with the augmentation levels of August, July and June from the 'calendar' set (reducing backwards in month for each diffuse shell removed from the calculation) were tested at all levels of SAPT truncation. The results are shown in Table A.7. The wall times refer to the SAPT2+3 calculations, which were given 150 GB of memory and 30 processors. Convergence of SAPT energies against time taken for the calculations is shown in Figure A.9.

SAPT2+3 was considered the most reliable for all computations. From Table A.7 it is clear that, especially the for the lower basis sets, the augmentation level of the calculation has a comparatively modest impact on wall times whilst giving significant improvements to calculated energies. Increases in basis set quality however are very computationally expensive. For the August augmentation level, the shift from a double zeta to quadruple zeta basis set would change the calculation time of a 21-point potential well from ~1 day to ~85 days. From Figure A.9 an initial significant improvement to the energy quickly flattens to an area where modest energy improvements require multiple hour increases in wall time. It is from the beginning of this flattening that the aug-cc-pVDZ basis was chosen.

Table A.7: Wall times and calculated energies for different levels of SAPT theory.

Augmentation level	Basis set	Wall time /s	Wall time /h	SAPT0	SAPT2	SAPT2+	SAPT2+(3)	SAPT2+3
	aDZ	4159	1.155277778	-9.03150157	-8.31043818	-7.99142424	-7.43396411	-7.65941191
aug-cc-pVXZ	aTZ	44983	12.49527778	-9.55937917	-9.17881487	-8.80731815	-8.11903590	-8.40289603
	aQZ	347775	96.60416667	-9.73965335	-9.48913342	-9.03262549	-8.27017400	-8.58993156
heavy-aug-	haDZ	2401	0.666944444	-8.79201228	-8.06491298	-7.66813311	-7.14694053	-7.36329872
OR jul-cc-	haTZ	36462	10.12833333	-9.44328157	-9.06042716	-8.64521165	-7.98983586	-8.26182458
pVXZ	haQZ	324877	90.24361111	-9.67965261	-9.42998841	-8.95197416	-8.21810102	-8.52649965
	jaDZ	1441	0.400277778	-6.14861685	-5.55568244	-5.03569801	-4.62352795	-4.80044752
jun-cc-pVXZ	jaTZ	23925	6.645833333	-9.14506180	-8.76504080	-8.24003207	-7.64289891	-7.89102228
	jaQZ	214196	59.49888889	-9.63292755	-9.38230719	-8.88186333	-8.16804894	-8.46803956

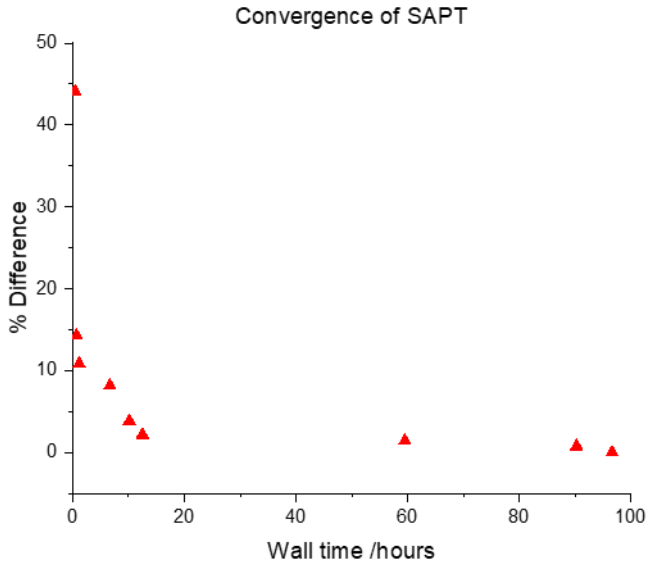
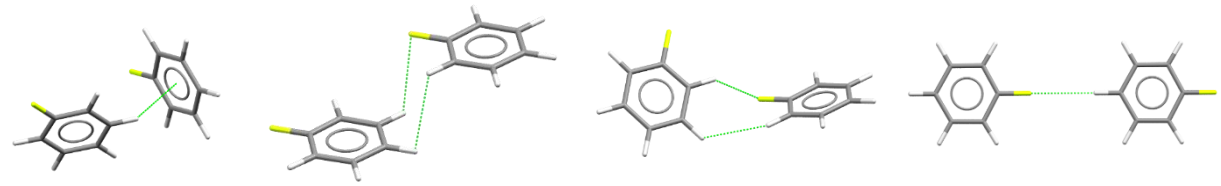


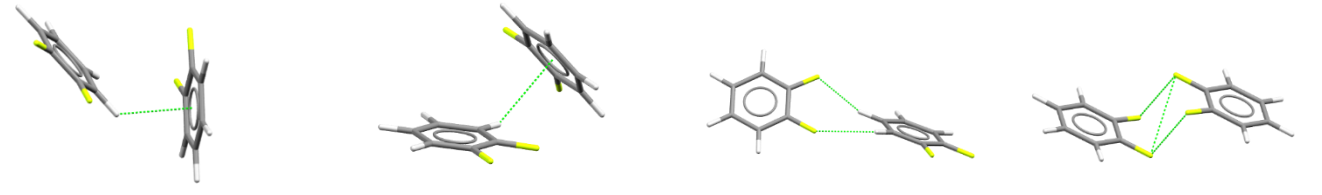
Figure A.9: Convergence of relative SAPT energies. The y-axis has been extended towards negative % differences for clarity of values which would otherwise lie very close to the x-axis.

C.5 Diagrams and Energies of Dimers for All Structures

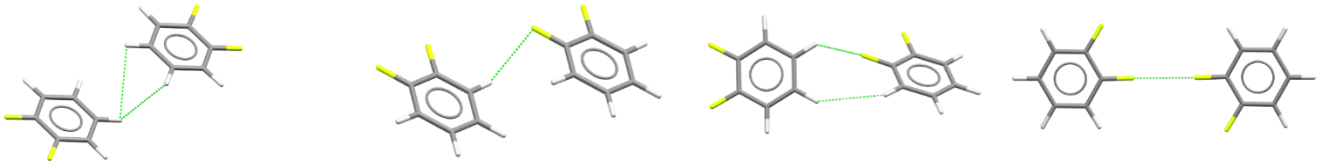


A B C D

Figure A.10: The symmetry inequivalent dimers within **1F** (monofluorobenzene), FACFAQ.



A B C D



E F G H

Figure A.11: The symmetry inequivalent dimers within **1,2F** (1,2-difluorobenzene), FACFOE.

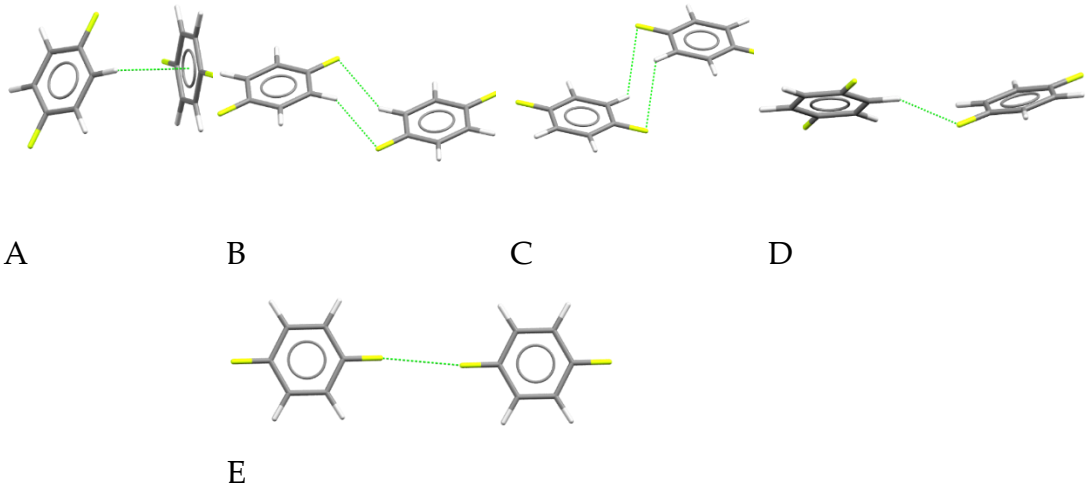


Figure A.12: The symmetry inequivalent dimers within **1,4F** (1,4-difluorobenzene), FACGEV.

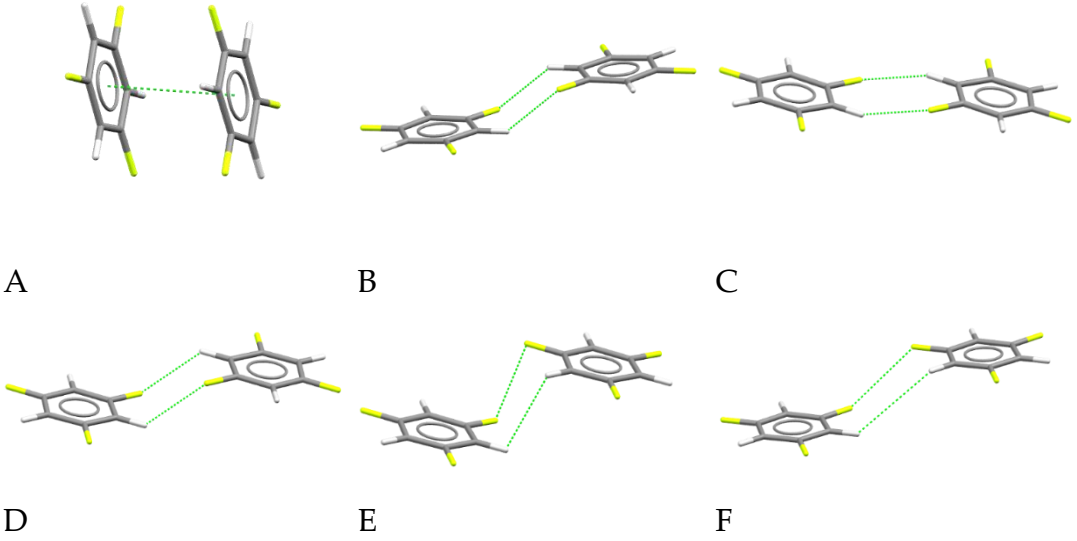
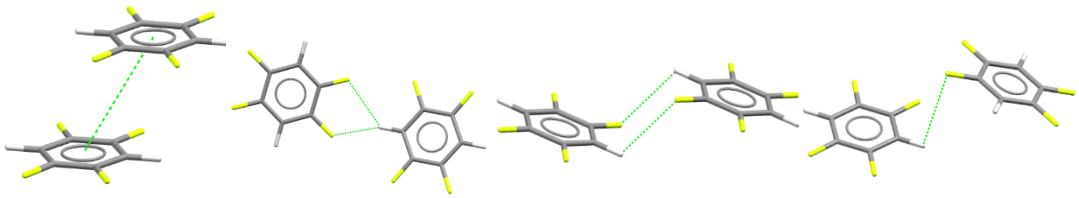


Figure A.13: The symmetry inequivalent dimers within **1,3,5F** (1,3,5-trifluorobenzene), PVVAWA01.



A B C D

Figure A.14: The symmetry inequivalent dimers within **1,2,4,5F** (1,2,4,5-tetrafluorobenzene), FACJAU.

Table A.8: Symmetry unique interactions within the first coordination spheres of fluorobenzenes. All energies are given in kJ mol^{-1}. U = lattice energy.

Structure	Dimer	Symmetry	Contact Type	Centroid separation	Coulombic	Polarisation	Dispersion	Repulsion	Total	Total per molecule
1F, FACFAQ	A	$y \pm 1/2, -x + 5/2, z + 1/4$ $-y + 5/2, x \pm 1/2, z - 1/4$	H···π	4.759	−3.6	−1.8	−14.5	11.0	−8.9	−4.5
$P4_32_12$	B	$x, y \pm 1, z; x \pm 1, y, z$	H···F, H···H	5.752	−1.7	−0.8	−9.1	4.5	−7.0	−3.5
Z = 4,	C	$y \pm 1/2, -x + 3/2, z + 1/4$ $-y + 3/2, x \pm 1/2, z - 1/4$	H···F, H···H	6.542	−2.6	−1.2	−7.6	5.0	−6.4	−3.2
Z′ = ½	D	$x \pm 1, y \pm 1, z$	H···F	8.134	−1.3	−0.2	−1.5	0.2	−2.8	−1.4
U = −52.0										
1,2F,	A	$-x + 3/2, y \pm 1/2, -z + 3/2$	H···π	4.686	−5.1	−1.8	−15.4	11.2	−11.1	−5.6
FACFOE	B	$-x + 5/2, y \pm 1/2, -z + 3/2$	H···π	4.843	−4.1	−1.3	−12.6	7.4	−10.6	−5.3
$P2_1/n$	C	$x \pm 1/2, -y + 1/2, z \pm 1/2$	H···F, H···F	6.782	−4.7	−1.1	−6.9	4.3	−8.5	−4.3
Z = 4	D	$-x + 2, -y + 1, -z + 1$	F···F, F···F	5.864	−0.7	−0.8	−8.5	3.2	−6.8	−3.4
Z′ = 1	E	$-x + 2, -y + 1, -z + 2$	H···H···H	5.954	−1.3	−0.7	−7.6	3.4	−6.2	−3.1
U = −49.8	F	$x, y \pm 1, z$	H···F	6.007	−0.6	−0.9	−8.6	4.2	−5.9	−3.0
	G	$x \pm 1/2, -y + 3/2, z \pm 1/2$	H···F, H···H	6.935	−1.4	−0.7	−5.2	2.8	−4.5	−2.3
	H	$-x + 2, -y, -z + 1$	F···F	8.268	1.5	−0.2	−1.7	0.6	0.2	0.1

Structure	Dimer	Symmetry	Contact Type	Centroid separation	Coulombic	Polarisation	Dispersion	Repulsion	Total	Total per molecule
1,4F, FACGEV *P*2₁/*c* *Z* = 2, *Z*′ = ½ *U* = −50.0	A	$-x+1, y\pm1/2, -z+3/2$ $-x+1, y\pm1/2, -z+5/2$	H···π	4.741	−3.3	−1.3	−13.9	9.2	−9.4	−4.7
	B	$x, y\pm1, z$	H···F, H···F	6.332	−5.3	−1.3	−8.6	5.0	−10.3	−5.2
	C	$x\pm1, y, z$	H···F, H···F	5.798	−2.8	−0.9	−9.3	4.4	−8.5	−4.3
	D	$-x, y\pm1/2, -z+3/2$ $-x+2, y\pm1/2, -z+5/2$	H···F	6.902	−3.1	−1.1	−5.6	4.2	−5.6	−2.8
	E	$x\pm1, y\pm1, z$	F···F	8.585	1.8	−0.1	−1.1	0.1	0.8	0.4
1,3,5F, PVVAWA01 *I*2/*a* *Z* = 4, Z′ = ½ *U* = −45.8	A	$x, -y+3/2, z\pm1/2$	π···π	3.706	−4.9	−1.6	−21.5	17.4	−10.6	−5.3
	B	$x\pm1/2, -y+2, z$	H···F, H···F	6.547	−6.2	−1.6	−8.4	5.9	−10.2	−5.1
	C	$x\pm1, -y+3/2, z\pm1/2$	H···F, H···F	6.844	−7.9	−1.6	−7.7	7.8	−9.5	−4.8
	D	$x\pm1/2, -y+1, z$	H···F, H···F	6.872	−3.2	−0.6	−6.2	3.1	−6.9	−3.5
	E	$x\pm1, y, z$	F···F, H···H	6.096	0.2	−0.3	−6.0	1.5	−4.6	−2.3
	F	$x-1/2, y\pm1/2, z-1/2$ $x+1/2, y\pm1/2, z+1/2$	F···F, H···H	7.529	1.2	−0.1	−2.0	0.1	−0.8	−0.4

Structure	Dimer	Symmetry	Contact Type	Centroid separation	Coulombic	Polarisation	Dispersion	Repulsion	Total	Total per molecule
1,2,4,5F,	A	$x \pm 1, y, z$	π···π	4.428	−3.0	−1.5	−16.7	11.3	−9.9	−5.0
FACJAU	B	$-x+2, y \pm 1/2, -z+1/2$	F···H···F	6.892	−4.9	−1.5	−7.4	5.7	−8.1	−4.1
$P2_1/c$	C	$-x+4, y \pm 1/2, -z+3/2$ $x \pm 1, y, z \pm 1$	H···F, H···F	6.593	−3.2	−0.5	−6.3	2.6	−7.4	−3.7
$Z = 2$	D	$-x+3, y \pm 1/2, -z+1/2$	H···F	6.051	−0.1	−0.5	−7.5	2.5	−5.6	−2.8
$Z' = ½$		$-x+3, y \pm 1/2, -z+3/2$								
$U = -49.1$										

C.6 Inelastic Neutron Scattering Plots

Inelastic neutron scattering plots were made using AbINS through the Mantid workbench GUI. Temperatures used for simulations were taken from the experimental data log and were: **1F**- 24 K, **1,2F**- 26 K, **1,4F**- 20 K, **1,3,5F**- 30 K and **1,2,4,5F**- 20 K. AbINS applies a semi-empirical powder averaging model to the relative atomic displacements and frequencies from the phonon calculations. In this model every phonon is treated as an independent quantum harmonic oscillator. Atoms are weighted by their neutron scattering cross-sections. Fundamental quantum modes up to 10 were considered, each increasing order having diminishing contribution. The specific instrument resolution function for the TOSCA instrument is built into AbINS and this was used to convolute the theoretical spectra. AbINS also removes experimentally observed overtones and accounts for the lower scattering atom types in the experimental data. Scaling of intensities was completed visually for peaks closest 1000 cm^{-1}. Data for averages of the front and back detector were used. Comparisons of experimental and simulated INS for all structures over smaller and larger spectral ranges are given below.

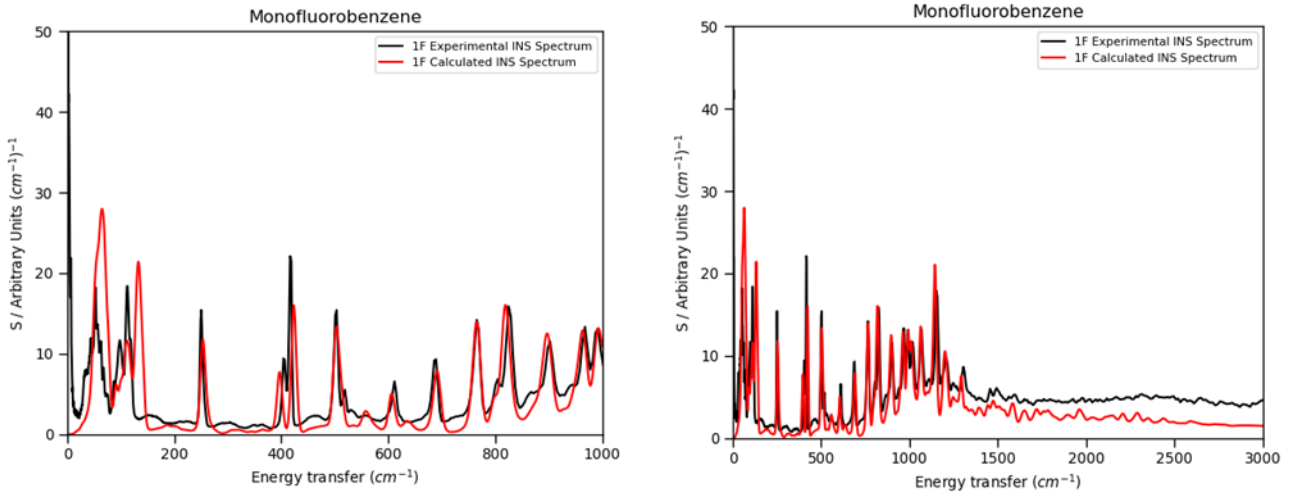


Figure A.15: Comparison of experimental and calculated INS spectra for **1F** (monofluorobenzene), FACFAQ.

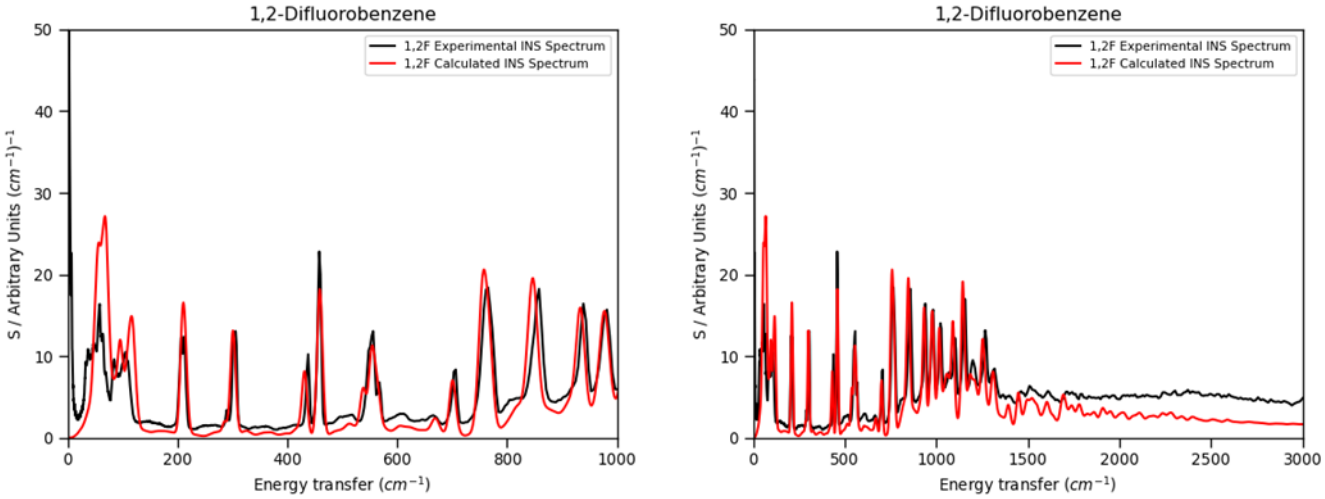


Figure A.16: Comparison of experimental and calculated INS spectra for **1,2F** (1,2-difluorobenzene), FACFOE.

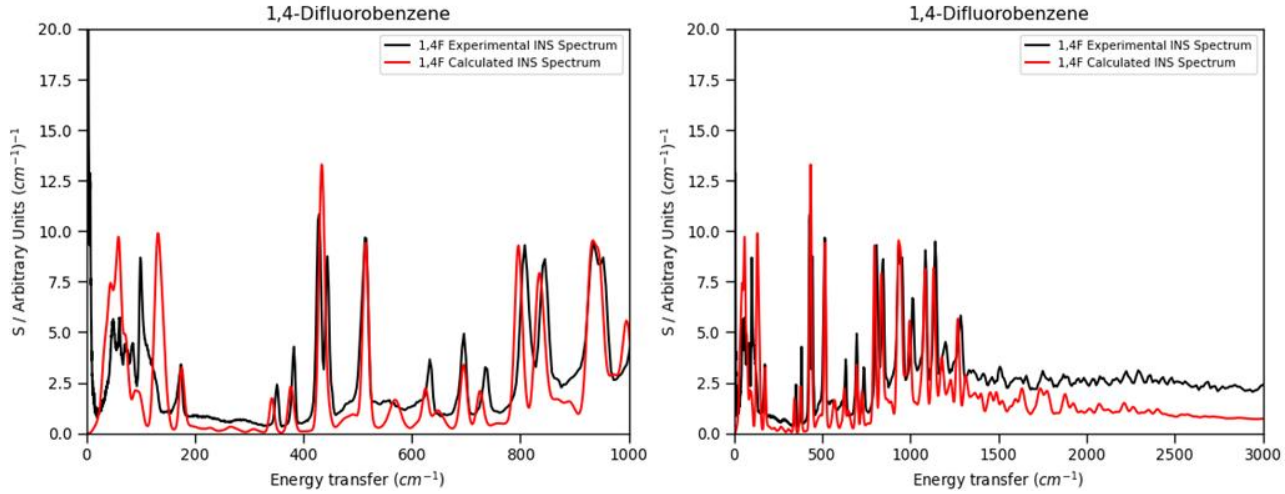


Figure A.17: Comparison of experimental and calculated INS spectra for **1,4F** (1,4-difluorobenzene), FACGEV.

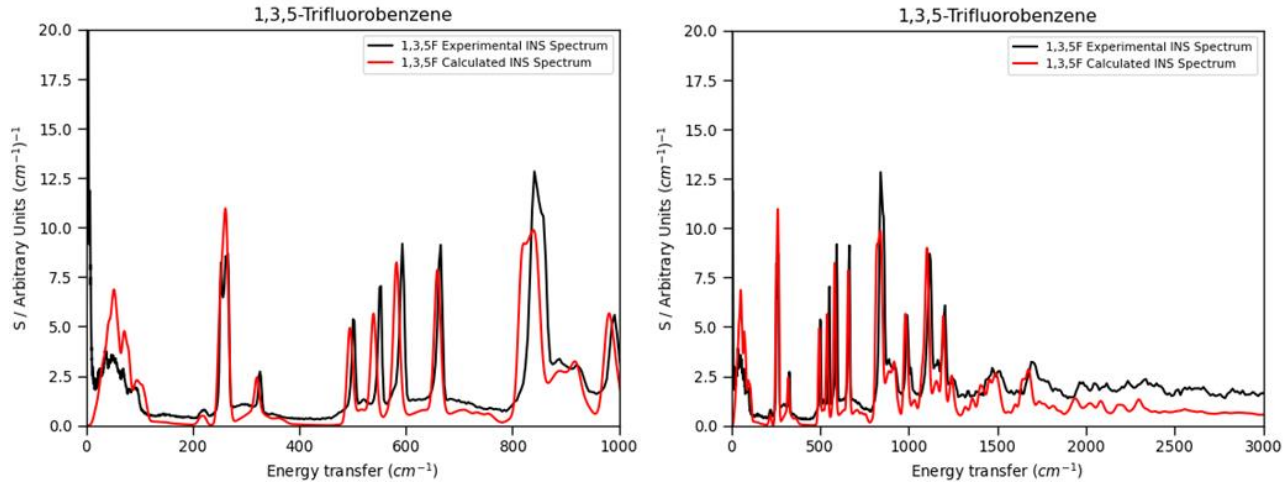


Figure A.18: Comparison of experimental and calculated INS spectra for **1,3,5F** (1,3,5-trifluorobenzene), PVVAWA01.

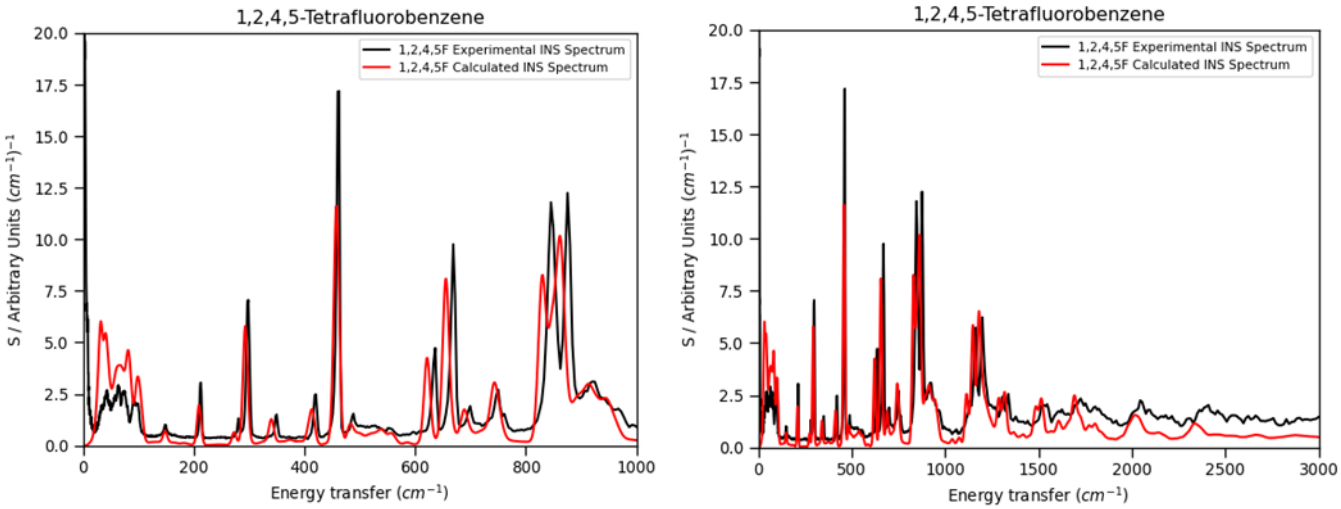


Figure A.19: Comparison of experimental and calculated INS spectra for **1,2,4,5F** (1,2,4,5-tetrafluorobenzene), FACJAU.

C.7 Density of States Plots

Phonon density of states plots for all structures over ranges of 0-1500 and 0-3250 cm^{-1} are presented below. These plots were produced through Gaussian smearing in Materials Studio with a defined smear width, *c*, of 0.1 THz (~3.3 cm^{-1}). Intensities of all plots are normalised such that the total integral is equal to *3N*, where *N* is the number of atoms.

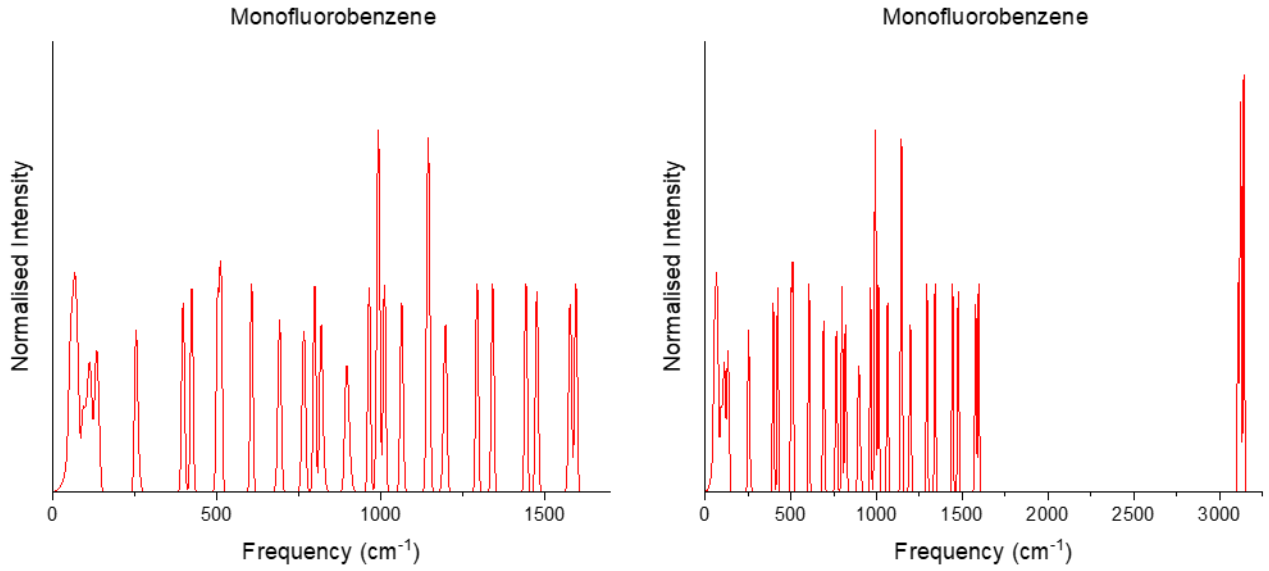


Figure A.20: Density of states plots for **1F** (monofluorobenzene), FACFAQ.

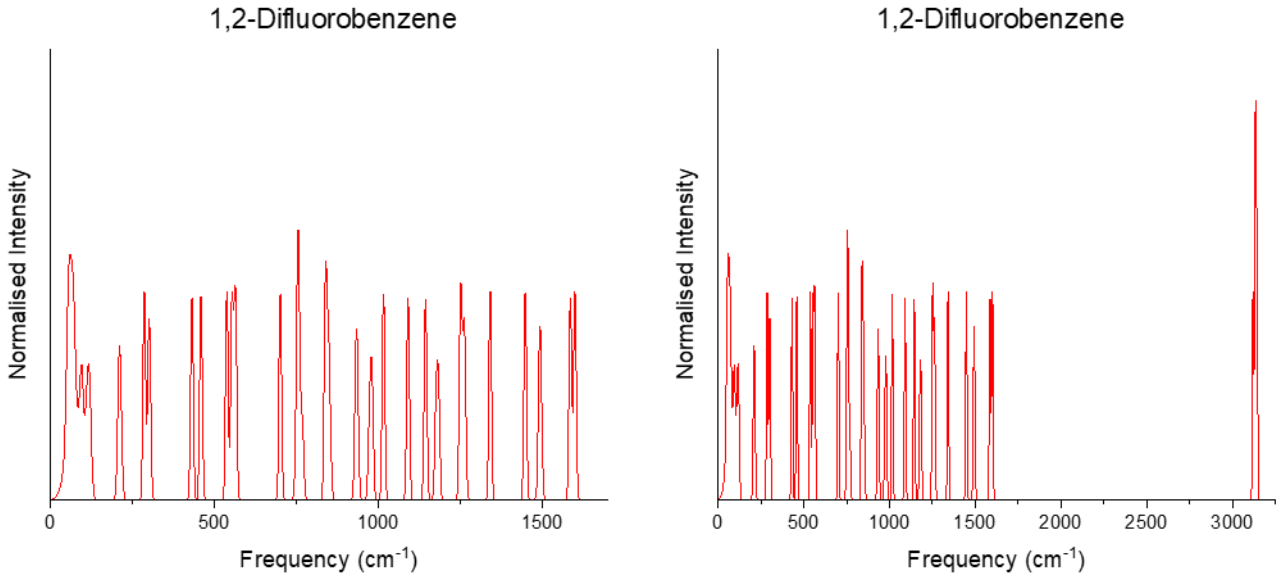


Figure A.21: Density of states plots for **1,2F** (1,2-difluorobenzene), FACFOE.

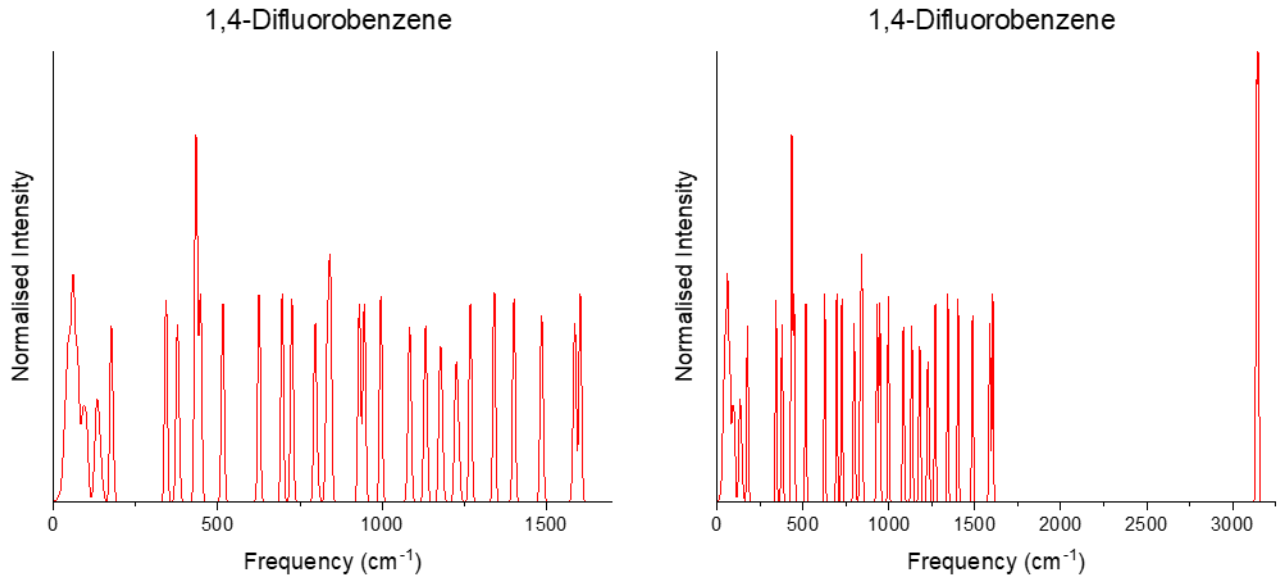


Figure A.22: Density of states plots for **1,4F** (1,4-difluorobenzene), FACGEV.

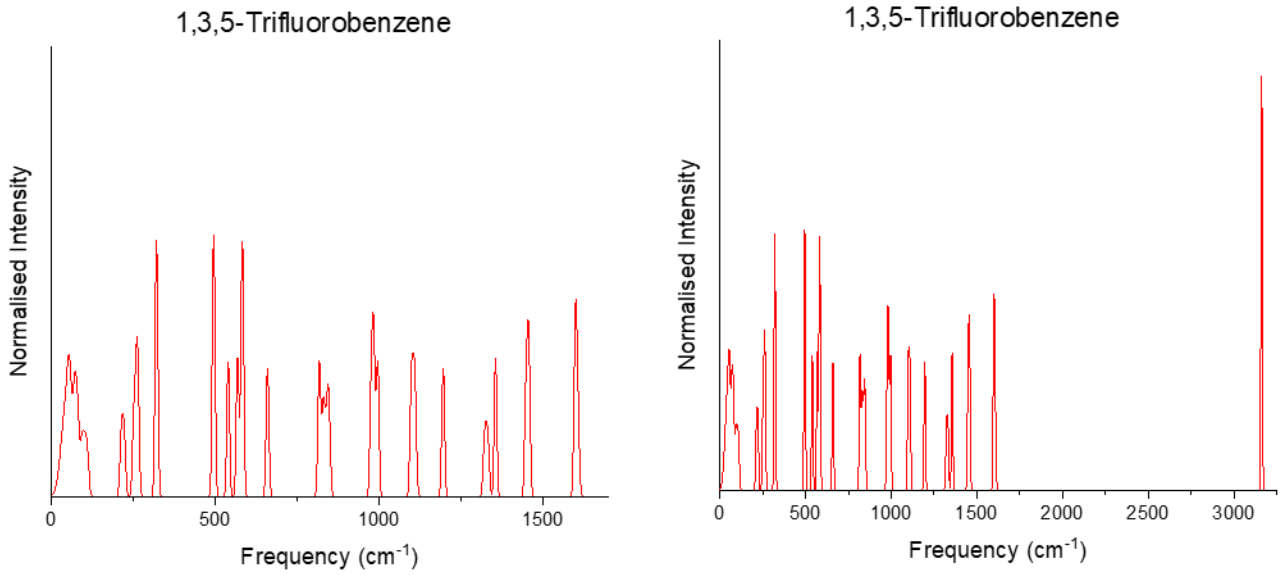


Figure A.23: Density of states plots for **1,3,5F** (1,3,5-trifluorobenzene), PVVAWA01.

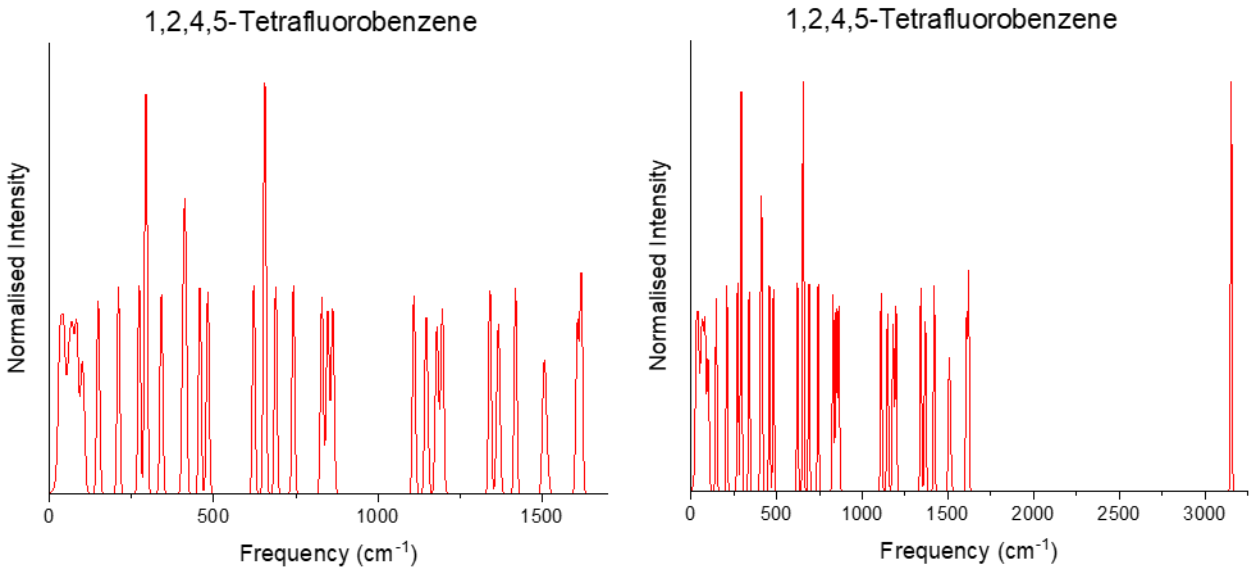


Figure A.24: Density of states plots for **1,2,4,5F** (1,2,4,5-tetrafluorobenzene), FACJAU.

C.8 Partial Density of States Plots

Partial density of states plots for selected atom(s) were derived from phonons at the Γ-point.

The output of the phonon calculation contains N, where N is the number of atoms in the model, eigenvector descriptions for x, y and z perturbations in atom positions for each of the $3N$ vibrational modes. For each eigenvector the fraction of the total motion attributed to the selected atom(s) is calculated and used to provide a total fractional contribution towards each vibrational mode. The partial density of states plots were produced via the same Gaussian smearing regime described in Section C.7. The height of the resulting peaks represents the fraction of motion attributed to the selected atom type. Partial density of states plots for all structures at the Γ-point are presented below in the same ranges as the total density of states plots above.

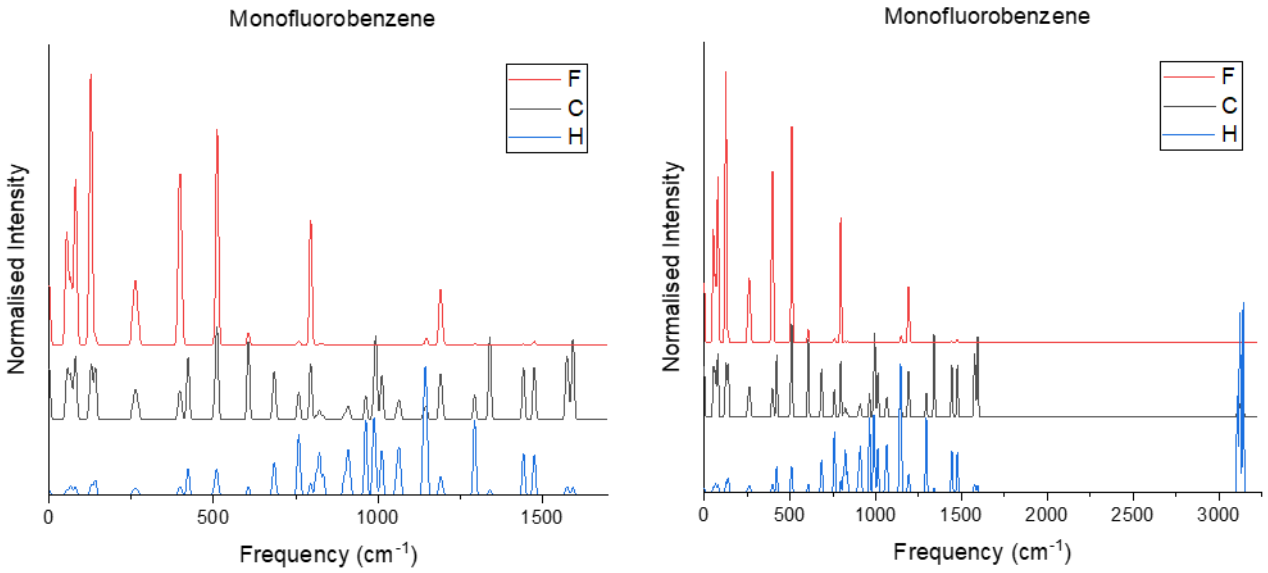


Figure A.25: Partial density of states plots for **1F** (monofluorobenzene), FACFAQ.

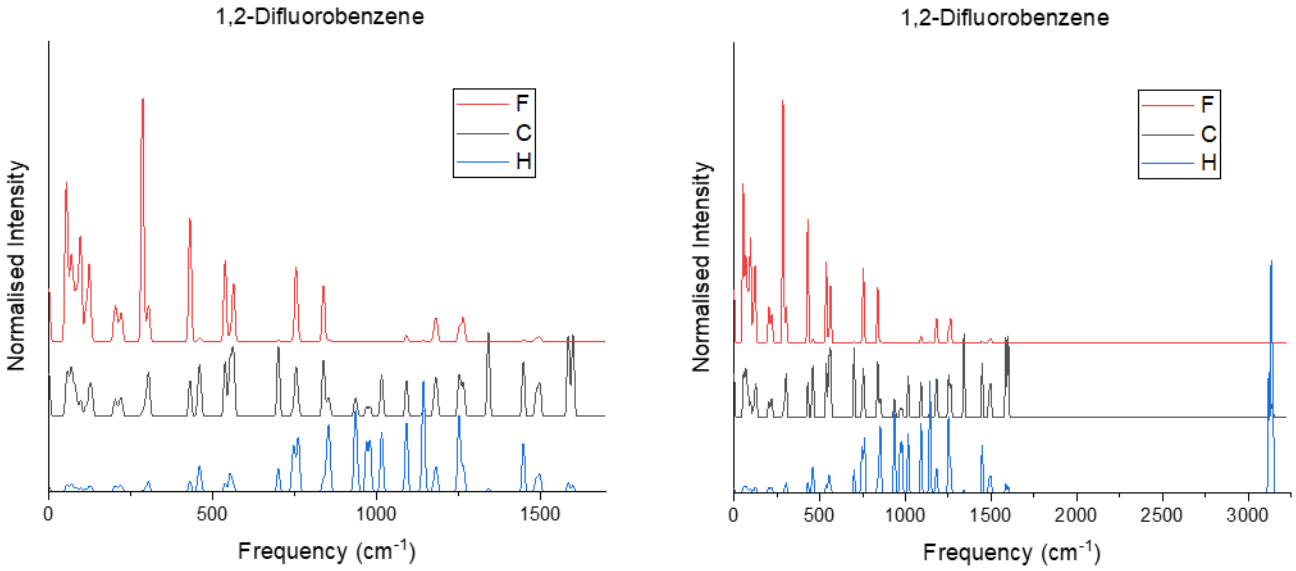


Figure A.26: Partial density of states plots for **1,2F** (1,2-difluorobenzene), FACFOE.

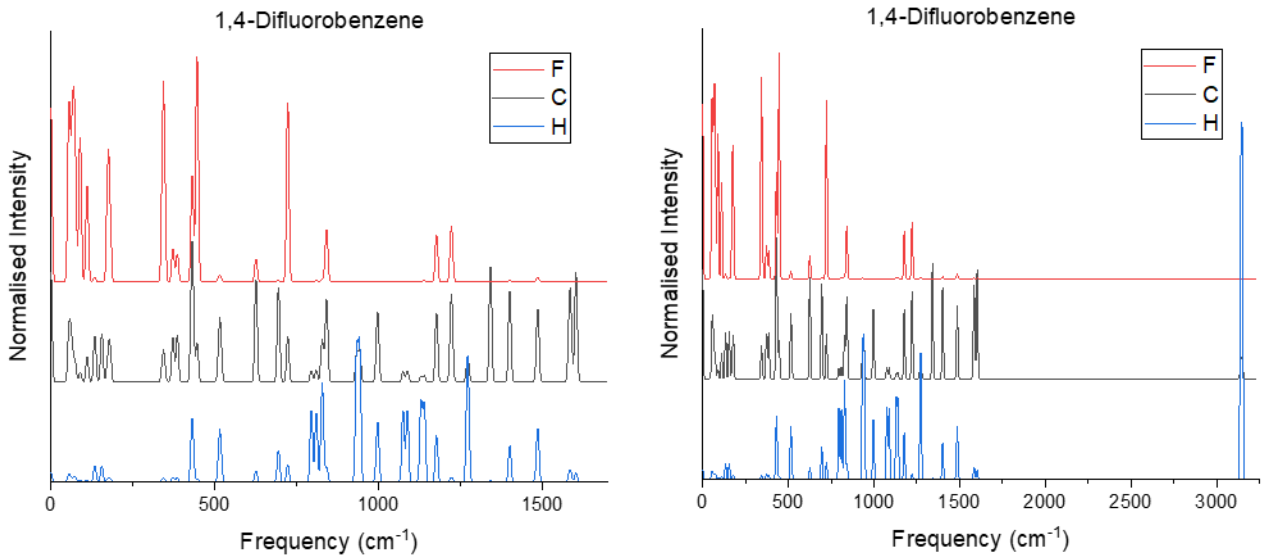


Figure A.27: Partial density of states plots for **1,4F** (1,4-difluorobenzene), FACGEV.

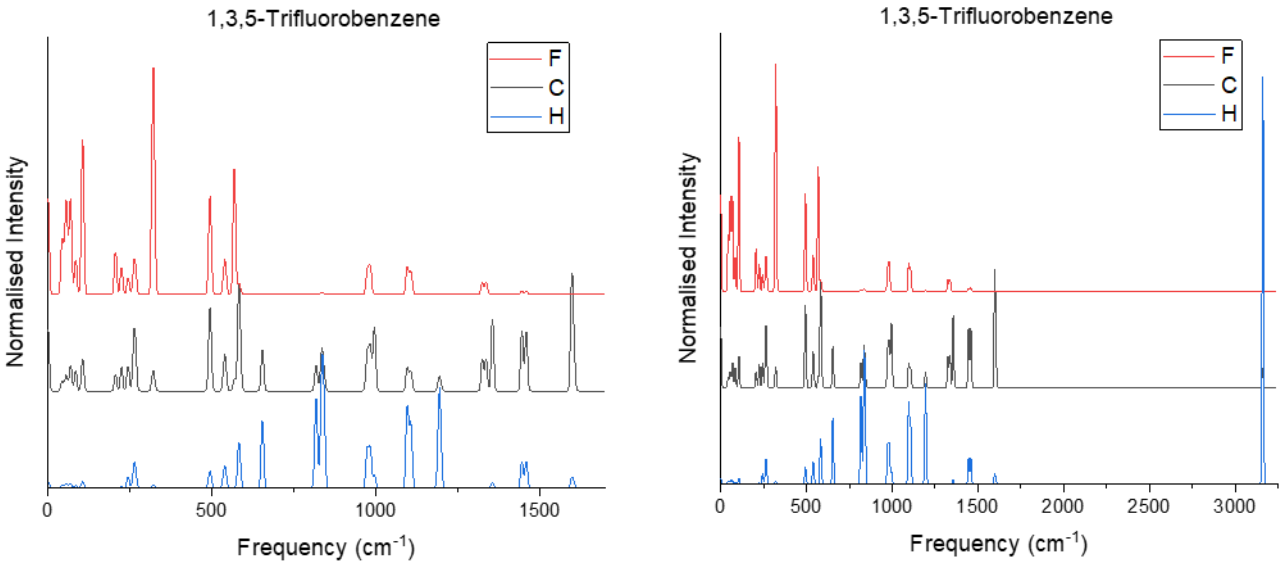


Figure A.28: Partial density of states plots for **1,3,5F** (1,3,5-trifluorobenzene), PVVAWA01.

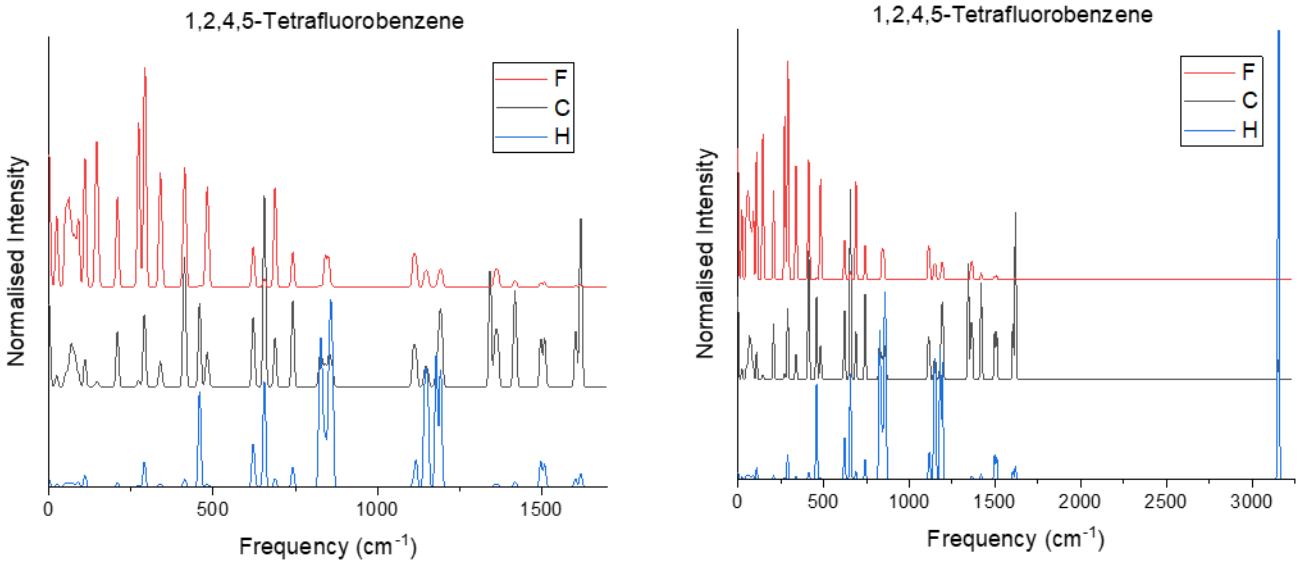


Figure A.29: Partial density of states plots for **1,2,4,5F** (1,2,4,5-tetrafluorobenzene), FACJAU.

C.9 Example Calculation of the Contributions to Internal Energy and Entropy by Occupation of Energy Levels at 150 K in 1,2F

The vibrational frequencies ν at the Γ-point were obtained by periodic DFT as described in Section 5.2.3 of Chapter 5. Calculation of the contribution to the entropy (S) and thermal energy (U) from each vibration at a temperature T = 150 K is accomplished using the following equations, where k, h and R are Boltzmann's constant, Plank's constant and the gas constant and the value of ν in Hz is obtained from the mode wavenumber (ω in cm^{-1}) as equal to $100c\omega$, where c is the speed of light:

$$\theta = \frac{h\nu}{k}$$

$$U = \frac{R\theta}{\exp(\theta / T) - 1}$$

$$\frac{S}{R} = \frac{\theta / T}{\exp(\theta / T) - 1} - \ln(1 - \exp(-\theta / T))$$

The values of *U* and *TS* were calculated for each of the 141 non-acoustic modes as shown in Table A.9. The magnitude of the total entropy calculated at the Γ-point (Table 5.2) is obtained by summing the values in the *TS* column.

Table A.9: Calculation of vibrational contributions to internal energy (U) and entropy (S) at temperature T = 150 K. The zero-frequency acoustic modes (1-3) have been omitted from the calculation.

	ω/ cm^{-1}	U/ $kJmol^{-1}$	TS/ $kJmol^{-1}$		ω/ cm^{-1}	U/ $kJmol^{-1}$	TS/ $kJmol^{-1}$
4	50.387	0.970	2.166	27	216.475	0.371	0.538
5	52.680	0.958	2.112	28	223.340	0.355	0.511
6	54.537	0.949	2.069	29	284.909	0.237	0.321
7	55.257	0.946	2.053	30	285.158	0.237	0.320
8	57.727	0.934	2.000	31	285.516	0.236	0.319
9	62.390	0.911	1.906	32	285.785	0.236	0.319
10	65.167	0.898	1.853	33	295.341	0.221	0.297
11	68.598	0.882	1.791	34	301.758	0.211	0.282
12	69.426	0.878	1.777	35	305.260	0.206	0.275
13	72.310	0.864	1.728	36	305.625	0.206	0.274
14	77.293	0.841	1.649	37	429.554	0.085	0.105
15	80.153	0.829	1.605	38	430.007	0.085	0.105
16	86.037	0.803	1.522	39	431.235	0.084	0.104
17	92.702	0.774	1.434	40	432.977	0.083	0.102
18	97.617	0.753	1.374	41	456.531	0.069	0.085
19	99.402	0.746	1.353	42	456.764	0.069	0.085
20	112.082	0.695	1.215	43	461.892	0.067	0.082
21	120.809	0.661	1.131	44	462.272	0.066	0.081
22	124.301	0.648	1.099	45	536.993	0.037	0.045
23	125.096	0.645	1.092	46	537.001	0.037	0.045
24	132.490	0.618	1.029	47	538.370	0.037	0.044
25	200.273	0.411	0.609	48	538.978	0.037	0.044
26	206.649	0.395	0.580	49	550.182	0.034	0.040

	ω/ cm^{-1}	*U*/ $kJmol^{-1}$	*TS*/ $kJmol^{-1}$		ω/ cm^{-1}	*U*/ $kJmol^{-1}$	*TS*/ $kJmol^{-1}$
50	551.971	0.033	0.040	77	933.550	0.001	0.002
51	555.564	0.032	0.038	78	933.686	0.001	0.002
52	556.017	0.032	0.038	79	936.992	0.001	0.002
53	561.972	0.031	0.036	80	939.419	0.001	0.002
54	562.823	0.031	0.036	81	968.447	0.001	0.001
55	563.466	0.030	0.036	82	969.878	0.001	0.001
56	564.468	0.030	0.036	83	980.077	0.001	0.001
57	699.186	0.010	0.012	84	980.241	0.001	0.001
58	699.642	0.010	0.012	85	1013.351	0.001	0.001
59	699.814	0.010	0.012	86	1014.100	0.001	0.001
60	699.985	0.010	0.012	87	1016.232	0.001	0.001
61	743.165	0.007	0.008	88	1017.029	0.001	0.001
62	748.325	0.007	0.008	89	1089.389	0.000	0.000
63	752.538	0.007	0.008	90	1089.934	0.000	0.000
64	753.831	0.007	0.007	91	1091.601	0.000	0.000
65	754.133	0.007	0.007	92	1092.953	0.000	0.000
66	756.503	0.006	0.007	93	1140.712	0.000	0.000
67	761.356	0.006	0.007	94	1141.424	0.000	0.000
68	761.549	0.006	0.007	95	1143.964	0.000	0.000
69	836.760	0.003	0.004	96	1145.049	0.000	0.000
70	837.175	0.003	0.004	97	1174.767	0.000	0.000
71	838.342	0.003	0.004	98	1179.551	0.000	0.000
72	839.018	0.003	0.004	99	1181.671	0.000	0.000
73	847.352	0.003	0.003	100	1184.284	0.000	0.000
74	850.503	0.003	0.003	101	1249.477	0.000	0.000
75	855.783	0.003	0.003	102	1250.856	0.000	0.000
76	856.147	0.003	0.003	103	1251.636	0.000	0.000

	ω/ cm^{-1}	*U*/ $kJmol^{-1}$	*TS*/ $kJmol^{-1}$		ω/ cm^{-1}	*U*/ $kJmol^{-1}$	*TS*/ $kJmol^{-1}$
104	1251.882	0.000	0.000	124	1585.517	0.000	0.000
105	1255.445	0.000	0.000	125	1598.034	0.000	0.000
106	1260.825	0.000	0.000	126	1598.538	0.000	0.000
107	1265.812	0.000	0.000	127	1599.391	0.000	0.000
108	1266.429	0.000	0.000	128	1600.502	0.000	0.000
109	1340.613	0.000	0.000	129	3118.791	0.000	0.000
110	1340.917	0.000	0.000	130	3118.986	0.000	0.000
111	1341.586	0.000	0.000	131	3119.141	0.000	0.000
112	1341.688	0.000	0.000	132	3119.381	0.000	0.000
113	1447.824	0.000	0.000	133	3131.218	0.000	0.000
114	1447.854	0.000	0.000	134	3131.230	0.000	0.000
115	1447.893	0.000	0.000	135	3131.704	0.000	0.000
116	1448.364	0.000	0.000	136	3131.721	0.000	0.000
117	1484.981	0.000	0.000	137	3134.187	0.000	0.000
118	1490.357	0.000	0.000	138	3134.643	0.000	0.000
119	1496.781	0.000	0.000	139	3134.727	0.000	0.000
120	1498.169	0.000	0.000	140	3135.236	0.000	0.000
121	1584.175	0.000	0.000	141	3141.774	0.000	0.000
122	1585.079	0.000	0.000	142	3142.085	0.000	0.000
123	1585.394	0.000	0.000	143	3142.512	0.000	0.000
				144	3143.433	0.000	0.000
				Total/cell		21.669	39.985
				Total/mol		5.417	9.996

C.10 Justification for the Neglect of U and Zero-Point Energy

The equations for U and TS given in Section C.9 can be plotted as a function of ω for T = 150 K to allow the influence of these quantities on the free energy to be compared (Figure A.30). The plot shows that TS is much more sensitive to the values of the frequencies of low-energy vibrational modes than is U. Small differences in external mode vibrational frequencies make little difference to the contribution of mode to U, and alternative packing arrangements will have similar values of U. This reflects the observation made by Nyman and Day that the difference in heat capacity between polymorphs are very small.

In the harmonic approximation the zero-point energy is given by $\frac{1}{2}\Sigma h\nu$ and is dominated by the highest energy internal vibrational modes. These modes can vary between polymorphs, but usually not by much and differences are a minor contributor to the free energy differences between polymorphs, though they can assume significance in the cases where strong intermolecular interactions such as H-bonds are present as these can perturb internal bonds more substantially than the weak interactions that are considered in Chapter 5.

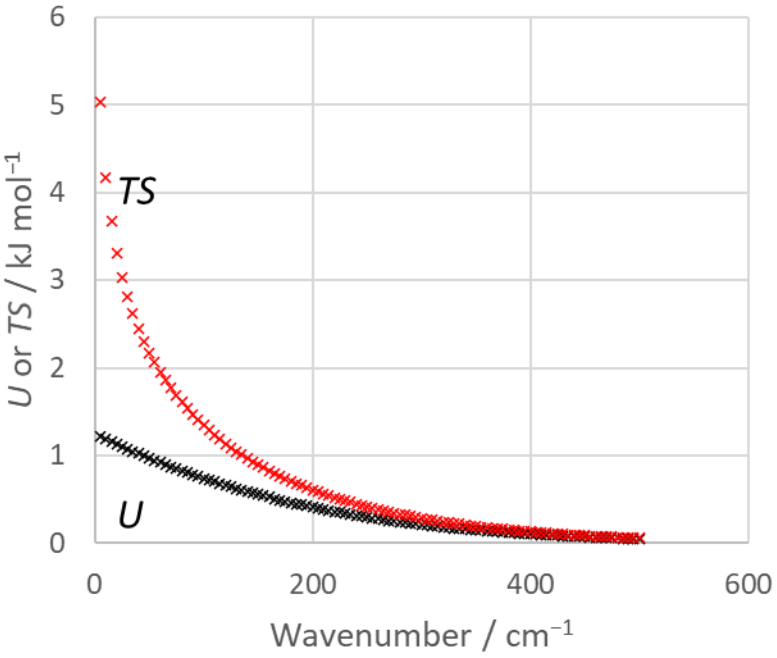


Figure A.30. Comparison of the contributions of low-energy vibrational modes to entropy and internal energy at T = 150 K.

Appendix D- Chapter 6

D.1 Crystallographic Tables

Table A.10: Full crystallographic tables for high pressure single crystal phases of glyphosate formed under rapid compression.

	glyphosate_II_7p0GPa	glyphosate_III_5p9GPa
	Crystal data	
Temperature	293 K	293 K
Crystal system, space group	Monoclinic, $P2_1/c$	Monoclinic, $P2_1/n$
	8.2046(13)	4.7277(3)
a, b, c (Å)	7.536(2)	7.1064(4)
	9.473(3)	15.5653(19)
β (°)	117.18(3)	98.689 (6)
V (Å^3)	521.0(3)	516.94 (8)
Z	4	4
Mr	169.07	169.07
Radiation type	Synchrotron, λ = 0.29040 Å	Ag $K\alpha$, λ = 0.56086 Å
μ (mm^{-1})	0.07	0.26
Crystal size (mm)	0.10 × 0.10 × 0.01	0.20 × 0.20 × 0.01
	Data collection	
Diffractometer	Synchrotron DESY beamline P02.2	Bruker Advance
Absorption correction	Multi-scan *CrysAlis PRO* 1.171.43.67a (Rigaku Oxford Diffraction, 2023) Empirical absorption correction using spherical harmonics, implemented in SCALE3 ABSPACK scaling algorithm.	Multi-scan *SADABS* (Siemens, 1996)
T_{min}, T_{max}	0.33, 1.00	0.90, 1.00
No. of measured, independent and observed [$I > 2.0\sigma(I)$] reflections	1106 601 333	3681 394 327
R_{int}	0.028	0.083
θ_{max} (°)	9.3	20.5
$(\sin\theta/\lambda)_{max}$ (Å^{-1})	0.556	0.623
	Refinement	
$R[F^2 > 2\sigma(F^2)]$, $wR(F^2)$, S	0.099 0.109 1.13	0.079 0.095 1.23
No. of reflections	333	327
No. of parameters	91	99
No. of restraints	98	141
H-atom treatment	H-atom parameters not refined	H atoms treated by a mixture of independent and constrained refinement
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.27, −0.23	0.67, −0.39

D.2 Fitting of Phase I to 5.93 and 7.03 GPa

For direct comparison between phases I and III, and phases I and II, the structure of phase I was extrapolated to 5.93 and 7.03 GPa respectively. This was completed by fitting individual Equations of State (EoS) to the volume, a, b and c parameters with a Birch-Murnaghan and Vinet EoS. The fitting parameters are given below. These EoS were then extrapolated to 5.93 and 7.03 GPa using the EoSCalc functionalisation. The value of β was then calculated from the other parameters.

Table A.11: The EoS fitting parameters used for the extrapolation of phase I to 5.93 GPa and 7.03 GPa.

Parameter	EoS	K0	K0′	K0′′	χ^2	Value at 5.93 GPa	Value at 7.03 GPa
V	BM3	14.9 (4)	7.3 (4)	-1.2	0.73	539.9876	528.8637
a	V3	436 (16)	-53 (5)	-1.4	3.87	8.4169	8.3328
b	BM2	60.1 (3)	12	-0.58	1.11	7.4597	7.3972
c	V3	21.8 (8)	37 (1)	-18.4	1.45	9.2025	9.1517
β						110.8437798	110.3597

D.3 Morphology of Glyphosate Crystals

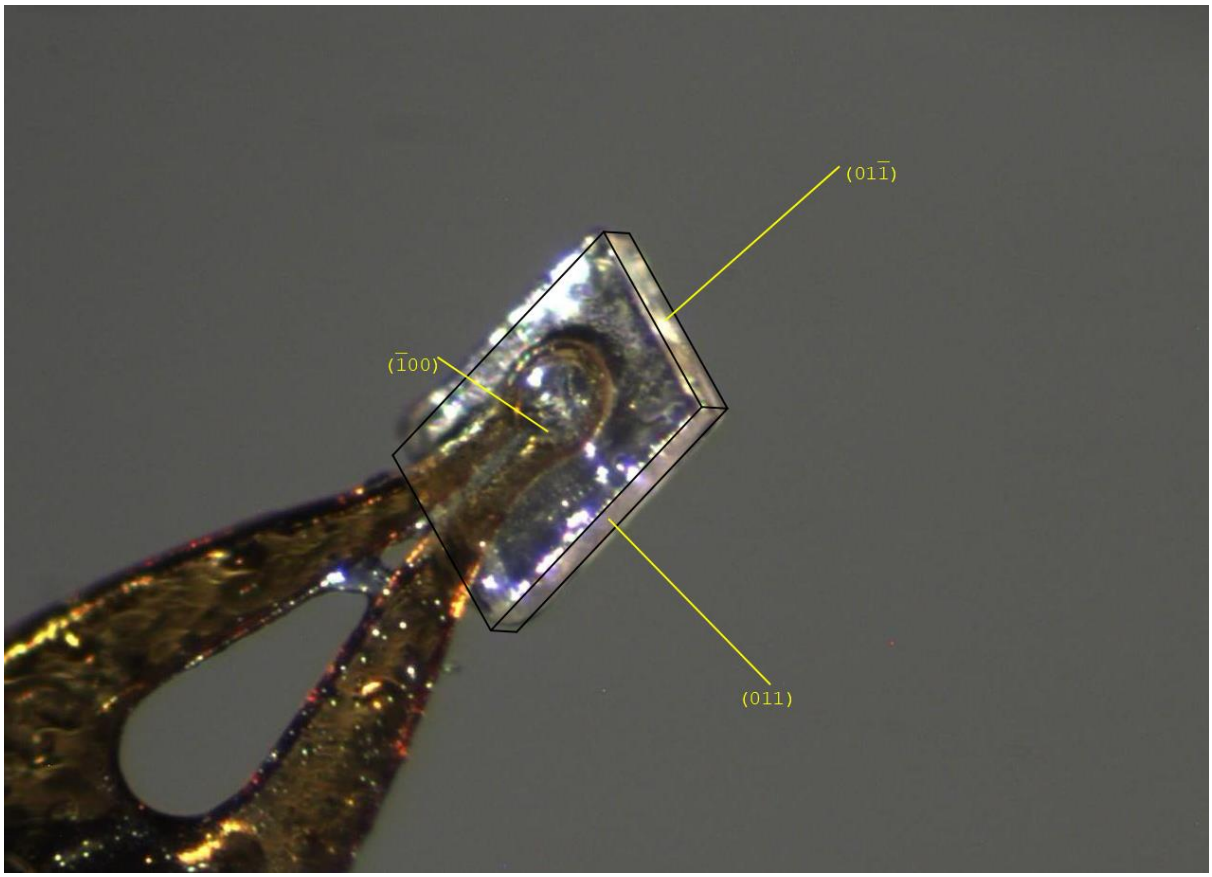


Figure A.31: The morphology of crystals of glyphosate are plates with the largest face in the ($\bar{1}$ 0 0) plane.

D.4 Pixel Tables for Phase I and II

Table A.12: Interactions in the first coordination sphere of phase I at 5.18 GPa. All energies are in kJ mol^{-1}. Classifications are relative to the interaction with the central molecule.

Label	Symmetry	Centroid distance (Å)	Electrostatic	Polarisation	Dispersion	Repulsion	Total	Contact description
				H-bonded				
Contact A	$1-x, 2-y, 1-z$	6.426	−292.7	−142.7	−40.3	347.1	−128.6	Pairs of O3H3···O4 H-bond ring motifs across inversion centres.
Contact B	$1-x, 1-y, 1-z$	4.909	−303.8	−124.0	−48.6	212.4	−264.1	Two N3H3B···O5 H-bond ring motifs across inversion centres.
Contact E/E′	$1-x, y \mp \frac{1}{2}, \frac{1}{2}-z$	6.390	−185.1	−68.9	−22.2	111.9	−164.4	Pairs of N3H3A···O4 H-bonds.
Contact F/F′	$x \pm 1, y, z$	8.496	−157.8	−86.3	−14.7	166.7	−92.1	Lattice translations of the central molecule in **a**. O1H1···O5 H-bond through a head to tail arrangement.
				Electrostatic Interactions				
Contact C/C′	$-x, y \mp \frac{1}{2}, \frac{1}{2}-z$	5.436	3.1	−19.5	−25.9	47.5	5.1	Lattice translation of contact E/E′ along **a**. Composed principally of destabilising electrostatic and stabilising dispersion components at 5.18 GPa. By 7.03 GPa the electrostatic component has become stabilising.
Contact D/D′	$x \mp 1, \frac{3}{2}-y, z \mp \frac{1}{2}$	8.093	−46.7	−8.4	−10.8	30.9	−35.1	Lattice translation of contact F/F′ along **a**. Longer range electrostatic interaction.
Contact G	$-x, 1-y, 1-z$	6.497	9.1	−8.7	−12.1	10.1	−1.6	Destabilising electrostatic interaction as a result of the juxtaposition of like charged moieties.
Contact H/H′	$x, \frac{3}{2}-y, z \mp \frac{1}{2}$	4.764	37.0	−29.2	−31.9	77.9	53.9	Destabilising electrostatic interaction as a result of the juxtaposition of like charged moieties.
Contact J	$-x, 1-y, -z$	6.715	27.2	−5.1	−5.9	1.8	18.0	Destabilising electrostatic interaction as the result of the short separation of like uncharged moieties.

Table A.13: Interactions in the first coordination sphere of phase II at 7.03 GPa. All energies are in kJ mol^{-1}. Classifications are relative to the interaction with the central molecule.

Label	Symmetry	Centroid distance (Å)	Electrostatic	Polarisation	Dispersion	Repulsion	Total	Contact description
				H-bonded				
Contact A	$1-x, 2-y, 1-z$	6.022	−260.2	−125.6	−45.8	307.7	−123.9	Pairs of O3H3···O4 H-bond ring motifs across inversion centres.
Contact B	$1-x, 1-y, 1-z$	4.603	−336.5	−130.9	−52.0	219.0	−300.4	Two N3H3B···O5 H-bond ring motifs across inversion centres.
Contact E/E′	$1-x, y \mp \frac{1}{2}, \frac{1}{2}-z$	6.293	−202.0	−74.7	−21.9	118.1	−180.5	Pairs of N3H3A···O4 H-bonds.
Contact F/F′	$x \pm 1, y, z$	8.205	−155.6	−88.8	−14.7	166.2	−93.0	Lattice translations of the central molecule in **a**. O1H1···O5 H-bond through a head to tail arrangement.
				Electrostatic Interactions				
Contact C/C′	$-x, y \mp \frac{1}{2}, \frac{1}{2}-z$	5.349	−16.9	−24.8	−28.5	87.0	16.8	Lattice translation of contact E/E′ along **a**. Composed principally of destabilising electrostatic and stabilising dispersion components at 5.18 GPa. By 7.03 GPa the electrostatic component has become stabilising.
Contact D/D′	$x \mp 1, \frac{3}{2}-y, z \mp \frac{1}{2}$	7.433	−73.8	−14.4	−16.1	33.2	−71.2	Lattice translation of contact F/F′ along **a**. Longer range electrostatic interaction.
Contact G	$-x, 1-y, 1-z$	6.760	27.1	−9.4	−12.0	9.9	15.7	Destabilising electrostatic interaction as a result of the juxtaposition of like charged moieties.
Contact H/H′	$x, \frac{3}{2}-y, z \mp \frac{1}{2}$	4.841	43.5	−31.8	−34.0	82.6	60.4	Destabilising electrostatic interaction as a result of the juxtaposition of like charged moieties.
Contact J	$-x, 1-y, -z$	6.581	39.6	−6.9	−6.8	1.9	27.8	Destabilising electrostatic interaction as the result of the short separation of like uncharged moieties.

This page has been left intentionally blank