Determining Structure and Atomic Properties of Materials using Resonant X-ray Diffraction

Karim J. Sutton

Worcester College

A thesis submitted for the degree of

Doctor of Philosophy

Department of Chemistry

University of Oxford

Hilary Term 2015

THE UNIVERSITY OF OXFORD

DEPARTMENT OF CHEMISTRY

The work reported in this thesis was carried out in the Chemistry Research Laboratory in Oxford between October 2011 and February 2015 under the supervision of Professor Richard Cooper. All the work is my own, unless stated to the contrary, and has not been submitted, either in full or part, for any degree in this or any other university.

Karim Jeremy Sutton

For my parents

Summary

X-ray crystallography is a widely used analytical technique for the structure solution of small molecules. Since the determination of the rock salt structure in 1913 by Henry and Lawrence Bragg the technique has developed allowing the solution of larger and more complex structures. The information that can be determined about these structures has increased as X-ray sources, detectors, and computational methods have improved. However, certain properties of molecules cannot always be directly determined from single wavelength X-ray diffraction. These include, *inter alia*: the site specific oxidation or spin state of an element in compounds where more than one state of the same element exist; discrimination between consecutive heavy elements in the periodic table. As the size of molecules being studied increases, reduced data resolution also becomes a problem. The aim of this research was to determine whether these problems can be addressed by measuring the changing anomalous scattering contribution of heavy atoms within structure through careful selection of the X-ray energy.

Firstly, I report an investigation into the problem of discriminating oxidation state, spin state and elements of near identical scattering by exploiting their anomalous signal. I first present DetOx, [1] a program written during the course of the project to deconvolute the fluorescence signal from materials containing more than one state of the same element into their respective spectra. This allows the calculation of anomalous scattering factors for both atomic states of an atom, which can subsequently be used to refine the occupancy of the different states at ambiguous sites within the crystal structure. The approach taken here, to determine differences due to relatively small anomalous signals, is analogous to the refinement of the Flack parameter whereby small changes in many hundreds or thousands of observations can be used

to fit a parameter with a high degree of precision and accuracy. I show the application of this technique to the mixed oxidation state compound, $GaCl_2$, and the two-step spin crossover material, $Fe(btr)_3(ClO_4)_2$. Refinement of the occupancy of charged ions on multiple sites using data at a single, carefully selected wavelength proved successful for these compounds, although upon extension to materials containing a larger number of anomalous scatterers, the absorption became a major issue in the data along with problems associated with simultaneously refining occupancies at more sites in the structure. We have demonstrated that calculations can be made to select specific experimental data to collect in order to improve the measured signal. However, due to limitation of the current collision model on the diffractometer used we have not yet been able to construct data collection strategies to take advantage of this. [2]

I next present a new ratio refinement technique to overcome this absorption problem due to the increased number of scatterers. By using ratios between datasets close in energy, but below the absorption edge, we were able to exploit small changes in f' without encountering absorption problems associated with the increase in f''. These ratio values were then refined against a lab structure using a modified version of CRYSTALS [3] to reveal the site specific occupancies of different atomic species within a given structure. For mixed-valence compounds, *e.g.* Mn_3 and Mn_6 clusters, the difference in anomalous signal between the different states proved too small for a stable least-squares refinement solution. However, we have shown that using a simulated annealing algorithm (to refine only occupancies), we can consistently obtain the expected structure. For mixed-metal structures *e.g.* the Mn_5Co_4 cluster, there was enough contrast in the data to refine occupancies with a least-squares approach, and these results were supported using simulated annealing.

Lastly, I describe the application of structure solution techniques based on methods used in macromolecular crystallography to 'large' small molecules. Traditionally these have been reserved for non-centrosymmetric protein structures, however with

the trend of synthesising larger and larger small molecules, problems encountered in macromolecular crystallography leading to low resolution datasets are becoming increasingly common. I have shown that it is possible to solve the structure of centrosymmetric structures by exploiting the anomalous signal in multiple wavelength diffraction experiments. The technique is applied successfully to two relatively small molecules, however the results are promising for moving to larger structures in the future.

References

[1] K. J. Sutton, S. A. Barnett, K. E. Christensen, H. Nowell, A. L. Thompson, D. R. Allan, R. I. Cooper, *Journal of Synchrotron Radiation* **2013**, *20*, 200–4.

[2] P. Parois, K. J. Sutton, R. I. Cooper, *Journal of Chemical Crystallography* **2015**, in press.

[3] P. W. Betteridge, J. R. Carruthers, R. I. Cooper, K. Prout, D. J. Watkin, *Journal of Applied Crystallography* **2003**, *36*, 1487.

Table of Contents

Summary **iv**

Acknowledgements **xix**

Abbreviations **xx**

1 Introduction **1**

1.1 X-Ray Diffraction 1

1.1.1 X-ray Scattering 1

1.1.2 Bragg's Law 9

1.1.3 Anomalous Scattering 10

1.1.4 Determination of f' and f'' 15

1.1.5 Phase Problem 16

1.2 Structure Solution and Refinement 18

1.2.1 Small-Molecule 18

1.2.2 Macromolecular 24

1.2.3 Structure Refinement 26

1.3 Structure Solution of 'Large' Small Molecules 32

1.3.1 Multi-Wavelength Anomalous Dispersion 33

1.3.2 Phase Equations for MAD 34

1.3.3 Optimal Choice of Wavelength 36

1.3.4 Solution of the Phase Equations 38

1.3.5 Determination of Heavy Atom Substructure 39

1.3.6 Maximum-Likelihood and Bayesian Estimates 40

1.3.7 Complete Structure Solution 46

1.3.8 Current Software . . . 47
1.4 Discrimination of Elements and their Electronic States . . . 48
1.4.1 Mixed Oxidation Compounds . . . 48
1.4.2 Spin Crossover Compounds . . . 50
1.4.3 Near-Identical Element Discrimination . . . 52
1.4.4 Current Techniques . . . 56
1.4.5 Underlying Theory . . . 61
1.4.6 EXAFS *vs.* XANES . . . 65
1.5 Thesis Overview . . . 69

2 Experimental Apparatus and Methods 71

2.1 Beamline I19 . . . 71
2.1.1 Undulator . . . 73
2.1.2 Optics . . . 75
2.1.3 Experimental Hutch 1 . . . 76
2.1.4 Experimental Hutch 2 . . . 78
2.2 In-House Diffractometer . . . 81
2.3 Syntheses . . . 82
2.3.1 4,4′-bis-1,2,4-triazole - (btr) Ligand . . . 82
2.3.2 $Fe(btr)_3(ClO_4)_2$ . . . 83
2.3.3 $Mn_{12}O_{12}(O_2CCH_3)_{16}(H_2O)_4 \cdot 4\,H_2O \cdot 2\,CH_3CO_2H$ . . . 84
2.3.4 $Mn_6(O)_2(O_2CCMe_3)_{10}(C_4H_8O)_4$ . . . 84
2.3.5 $Co_4{}^{III}Mn_5{}^{II}Cl_6(C_8NO_5H_{15})_2(C_8NO_5H_{16})_2(H_2O)(OH)_2$ . . . 84
2.3.6 $Mn_3O(O_2CPh)_6(pyr)_2(H_2O)[2\,MeCN]$ . . . 84
2.3.7 $TBAB[CO(NH(4-NO_2-C_6H_4))_2]_2$ - Bromide Urea Salt . . . 85
2.3.8 Bromide Rotaxane . . . 86
2.4 Crystallisation Techniques . . . 87
2.4.1 Vapour Diffusion . . . 87
2.4.2 Layered Diffusion . . . 88

3 Discrimination of Elements and their Electronic States I **90**

3.1 DetOx: Extraction of overlapped absorption edges 91

3.1.1 Description of Program 91

3.1.2 Experimental 96

3.2 $GaCl_2$ - A Mixed Oxidation State Compound 99

3.2.1 Experimental and AS Factor Calculations 100

3.2.2 Results and Discussion 100

3.3 $[Fe(btr)_3](ClO_4)_2$ - A Spin Crossover Compound 103

3.3.1 Two-Step Spin Crossover 104

3.3.2 Proof-of-Concept 105

3.3.3 Experimental 108

3.3.4 Results and Discussion 109

3.4 Discussion 112

4 Discrimination of Elements and their Electronic States II **114**

4.1 Ratio Refinement 114

4.1.1 Implementation 114

4.2 Simulated Annealing 119

4.3 Single Molecule Magnets 121

4.3.1 Origin of Magnetism 121

4.3.2 Justification 124

4.4 Proof of Concept - simulated data 125

4.4.1 Mixed Valence Discrimination 125

4.4.2 Near Identical Atom Discrimination 127

4.4.3 Comparison of Occupancy Refinement Methods 130

4.5 Initial Testing 131

4.5.1 $Mn_6(O)_2(O_2CCMe_3)_{10}(C_4H_8O)_4$ 131

4.5.2 Limitations 132

4.6 Method Developments 133

4.6.1 Improvements . . . 134
4.6.2 $Mn_3O(O_2CPh)_6(pyr)_2(H_2O)$[2 MeCN] . . . 135
4.6.3 $Co_4{}^{III}Mn_5{}^{II}Cl_6(C_8NO_5H_{15})_2(C_8NO_5H_{16})_2(H_2O)(OH)_2$ . . . 139
4.6.4 Discussion . . . 143

5 Structure Solution of 'Large' Small Molecules 147

5.1 Bromide Urea Salt . . . 147
5.1.1 Experimental . . . 148
5.1.2 Substructure Determination . . . 149
5.1.3 Complete Structure . . . 151
5.1.4 EH1 Limitations - Bromide Rotaxane . . . 153
5.2 *o*-Xylylenebis(triphenylphosphonium Bromide) . . . 156
5.2.1 Experimental . . . 156
5.2.2 Data Collection . . . 158
5.2.3 Substructure Determination . . . 159
5.2.4 Complete Structure . . . 161
5.3 Discussion . . . 163

6 Conclusions and future work 166

6.1 Conclusions . . . 166
6.2 Future work . . . 167

References 169

A Crystallographic Information 184

A.1 $GaCl_2$ . . . 185
A.2 Btr Ligand . . . 187
A.3 $Fe(btr)_3(ClO_4)_2$ . . . 188
A.4 $Mn_{12}O_{12}(O_2CCH_3)_{16}\cdot 4\,H_2O\cdot 2\,CH_3CO_2H$ . . . 192
A.5 $Mn_6(O)_2(O_2CCMe_3)_{10}(C_4H_8O)_4$ . . . 193
A.6 $Mn_3O(O_2CPh)_6(pyr)_2(H_2O)$[2 MeCN] . . . 194

A.7 $Co_4{}^{III}Mn_5{}^{II}Cl_6(C_8NO_5H_{15})_2(C_8NO_5H_{16})_2(H_2O)(OH)_2$ 195

A.8 $TBAB[CO(NH(4-NO_2-C_6H_4))_2]_2$ 196

A.9 Bromide Rotaxane 197

A.10 o-Xylylenebis(triphenylphosphonium bromide) 198

List of Tables

1.1 Anomalous scattering factors for a selection of elements 12

1.2 Constraints placed on the electron density in direct methods 21

2.1 Specific energy ranges provided by different stripes on the focusing mirrors 76

2.2 Beamline details for EH1 78

2.3 Beamline details for EH2 79

2.4 Technical specifications of the Pilatus 300 K detector 81

3.1 Anomalous scattering factors output from DetOx for $GaCl_2$ 100

3.2 Refined $GaCl_2$ occupancies from CRYSTALS 102

3.3 FeA-FeB and Fe-N bond lengths (Å) at different temperatures 105

3.4 Anomalous scattering factors output from DetOx for $Fe(btr)_3(ClO_4)_2$ . . . 109

3.5 Calculated and experimental anomalous scattering factors for $Fe(btr)_3(ClO_4)_2$ at 7119 eV 109

3.6 Iron site occupancies at 200 K using both calculated and experimental anomalous scattering factors for $Fe(btr)_3(ClO_4)_2$ at 7119 eV 111

4.1 Anomalous scattering factors output from DetOx for Mn_6 cluster 126

4.2 Manganese site occupancies as calculated from a ratio refinement in CRYSTALS using simulated data 127

4.3 Anomalous scattering factors output from DetOx for $[Co_4{}^{III}Mn_5{}^{II}]$ cluster 129

4.4 Manganese and cobalt site occupancies for the $[Co_4{}^{III}Mn_5{}^{II}]$ cluster . . . 130

4.5 Comparison of occupancy refinement techniques on Mn_6 using simulated data 130

4.6 Mn_6 site occupancies as calculated from a least-squares ratio refinement using experimental data 132
4.7 Anomalous scattering factors output from DetOx for Mn_6 cluster 135
4.8 Anomalous scattering factors output from DetOx for Mn_3 cluster 137
4.9 Mn_3 site occupancies as calculated from both a least-squares and simulated annealing ratio refinement using experimental data 138
4.10 Anomalous scattering factors output from DetOx for $[Co_4{}^{III}Mn_5{}^{II}]$ cluster 140
4.11 Simulated *vs.* experimental manganese and cobalt site occupancies for the $[Co_4{}^{III}Mn_5{}^{II}]$ cluster 141

5.1 Anomalous scattering factors for bromide urea salt at selected wavelengths 149
5.2 Bromine atom coordinates refined using data at different resolutions . . 151
5.3 Anomalous scattering factors for bromide rotaxane at selected wavelengths 154
5.4 Refined bromide coordinates for the bromide rotaxane 155
5.5 Anomalous scattering factors for bromide urea salt at selected wavelengths 157
5.6 Maximum resolution (in Å) and completeness of datasets collected at different detector positions and wavelengths 157
5.7 Xylylene bromine substructure coordinates as refined using data at different resolutions 159

List of Figures

1.1 Diagram to show a particle at origin, O, being irradiated by a plane monochromatic wave propagating along the x-axis 2

1.2 Diagram to show the scattering of the wave vector $\boldsymbol{s}_0$ along $\boldsymbol{s}$ by scatterers O and O′. XO = - $\boldsymbol{r} \cdot \boldsymbol{s}_0$, YO = $\boldsymbol{r} \cdot \boldsymbol{s}$ 5

1.3 Reflection of X-rays from two planes belonging to the same Miller planes 10

1.4 Atomic scattering factor split into its real and imaginary parts when anomalous scattering is present 10

1.5 Argand diagram showing the relationship between $F_{\boldsymbol{h}}$ and $F_{-\boldsymbol{h}}$ 14

1.6 Argand diagram showing the relationship between $F_{\boldsymbol{h}}$ and $F^*_{-\boldsymbol{h}}$ 14

1.7 AS factors for a selenium atom 17

1.8 The effect of EXAFS and XANES on the imaginary part of AS factors 17

1.9 Observed Patterson map for Copper (II) Acetate in the *ab* plane 20

1.10 Flow diagram of the charge flipping algorithm 23

1.11 Argand diagram of $\boldsymbol{F}_{\mathrm{P}}$, $\boldsymbol{F}_{\mathrm{H}}$ and $\boldsymbol{F}_{\mathrm{PH}}$ 25

1.12 Absorption-fluorescence spectrum showing wavelength selections for the MAD phasing experiment 37

1.13 Argand diagram to represent the vector relationship between $\mathbf{F}_T$, $\mathbf{F}_A$, and $\Delta\varphi$ 38

1.14 Bijovet Difference Patterson map for a Bromo-Urea Salt 41

1.15 Anomalous Difference Patterson map for a Bromo-Urea Salt 41

1.16 Bayesian-Anomalous Patterson map for a Bromo-Urea Salt 46

1.17 Representation of the chemical structure of Prussian Blue 49

1.18 Representation of the principal types of spin transition curves 51

1.19 Atomic scattering factors for Co, Fe, O and H *vs.* $\sin\theta/\lambda$ - including temperature factor 53
1.20 Atomic scattering factors for corrected Co and Fe *vs.* $\sin\theta/\lambda$ - excluding temperature factor 55
1.21 Recoil of a nucleus upon absorption or emission of a γ-ray 58
1.22 Periodic table showing Mössbauer active elements (blue) 59
1.23 X-ray photoelectron spectra recorded for a Ag-Au doped zeolite sample . 62
1.24 Left: Absorption of X-ray by atom, Center: Ejection of photo-electron from core, Right: Fluorescence emission as electron drops to fill core hole 64
1.25 K and L absorption edge energies as a function of atomic number Z . . . 66
1.26 Fluorescence-absorption spectrum of a zinc porphyrin compound showing the EXAFS and XANES regions of interest 66
1.27 Cr K-edge XANES spectrum 68
1.28 Fe K-edge XANES spectrum 68

2.1 Schematic drawing of Beamline I19 72
2.2 Schematic drawing of an undulator 73
2.3 Graph to show the maximum flux as a function of energy and ID gap . . 74
2.4 Photograph of the DCrM on Beamline I19 75
2.5 Photograph of the focusing mirrors on Beamline I19 77
2.6 Sample position in EH1 77
2.7 Robot in EH1 78
2.8 Sample position in EH2 79
2.9 Vortex output showing photon counts as a function of energy 80
2.10 GDA output showing the integrated fluorescence peak area as a function of wavelength 81
2.11 Photograph of the in-house SuperNova machine 82
2.12 Reaction scheme for the synthesis of N,N-Dimethylformamide Azine Dihydrochloride 83

2.13 Reaction scheme the for synthesis of 4,4′-bis-1,2,4-triazole 83
2.14 Reaction scheme for the synthesis of a bromide urea salt 85
2.15 Reaction scheme for the synthesis of a bromide rotaxane 86
2.16 Crystallisation techniques 88

3.1 Graph to show the raw fluorescence data collected on $GaCl_2$ before normalisation 93
3.2 The effect on EXAFS on the peak position in the real scattering factors . . 94
3.3 Imaginary structure containing two atomic sites each occupied by two atoms 97
3.4 Comparisons between DetOx and CHOOCH for $GaCl_3$ 98
3.5 Crystal structure of $GaCl_2$ 99
3.6 Output from DetOx for a sample of $GaCl_2$ 101
3.7 Overlay of the two deconvoluted curves superimposed onto the original spectrum for $GaCl_2$ 101
3.8 Crystal Structure of the btr ligand 103
3.9 Crystal structure of $Fe(btr)_3(ClO_4)_2$ 104
3.10 Spin Transition Curve for $Fe(btr)_3(ClO_4)_2$ 105
3.11 A plot of $\chi_M T$, per iron(II), *vs.* temperature for $[Fe_2(ddpp)_2(NCS)_4] \cdot x CH_2Cl_2$ complexes 106
3.12 Absorption-fluorescence spectrum of $Fe(btr)_3(ClO_4)_2$ collected at 130 K, 200 K and 250 K 107
3.13 Plot of the absorption-fluorescence spectrum of $Fe(btr)_3(ClO_4)_2$ collected at 200 K *vs.* the combined spectrum from the 130 K and 250 K datasets . 107
3.14 Overlay of the two deconvoluted curves superimposed onto the original spectrum for $Fe(btr)_3(ClO_4)_2$ 108
3.15 Overlay of the two deconvoluted curves superimposed onto the spectra collected at 130 K and 250 K for $Fe(btr)_3(ClO_4)_2$ 110

3.16 Absorption-fluorescence spectrum of $Fe(btr)_3(ClO_4)_2$ collected at 130 K and 250 K 111

4.1 Graphs to show the scaling between datasets using **(a)** calculated data and **(b)** observed data 117

4.2 Normal equations used during the least-squares refinement 119

4.3 Simulated annealing cooling regime 121

4.4 Crystal structure of $Mn_{12}O_{12}(O_2CCH_3)_{16}\cdot 4\,H_2O\cdot 2\,CH_3CO_2H$ 122

4.5 Spin alignments in the ground state of Mn_{12} complexes that gives their S = 10 ground states 122

4.6 Plot of the potential energy *vs.* the magnetisation direction for a single molecule magnet 123

4.7 Crystal structure of $Mn_6(O)_2(O_2CCMe_3)_{10}(C_4H_8O)_4$ 126

4.8 Manganese substructure of Mn_6 cluster 127

4.9 Crystal structure of $[Co_4{}^{III}Mn_5{}^{II}]$ cluster 128

4.10 Manganese and cobalt substructure of $[Co_4{}^{III}Mn_5{}^{II}]$ cluster 129

4.11 Deconvoluted fluorescence graph for Mn_6 cluster showing the energies used for initial data collections 132

4.12 Deconvoluted fluorescence graph for Mn_6 cluster showing the energies used for improved-method data collections 134

4.13 Crystal structure of Mn_3 cluster. Hydrogen atoms and solvent molecules have been removed for clarity 136

4.14 Deconvoluted fluorescence graph for Mn cluster showing the energies used for data collections 137

4.15 Manganese substructure of Mn_3 cluster 138

4.16 Deconvoluted fluorescence graph for $[Co_4{}^{III}Mn_5{}^{II}]$ cluster showing the energies used for data collections 140

4.17 Comparison of calculated and simulated occupancies for $[Co_4{}^{III}Mn_5{}^{II}]$ cluster at different occupancies 142

4.18 Comparison of signals available for **(a)** refining the Flack parameter and **(b)** resolving oxidation state in the Mn_3 cluster 143

5.1 Crystal structure of TBAB[$CO(NH(4-NO_2-C_6H_4))_2]_2$ 148
5.2 Normalised absorption-fluorescence spectrum for bromide urea salt . . . 149
5.3 Bromine urea salt Bayesian Patterson maps at different resolutions 150
5.4 Electron density map of bromine urea salt at **(a)** full resolution and **(b)** 2.0 Å 152
5.5 Overlay of bromide urea salt solved by MAD phasing and by Superflip . . 153
5.6 Bromide rotaxane crystal structure. Bromide atom is orange and the iodine is purple 154
5.7 Bromide rotaxane Bayesian Patterson map at full resolution. Map viewed at an oblique angle; further details can be found in Figure 1.9 155
5.8 Crystal structure of *o*-Xylylenebis(triphenylphosphonium Bromide) . . . 157
5.9 *o*-Xtpb Bayesian Patterson maps using data collected at different detector positions 160
5.10 Overlay of *o*-Xtpb solved by MAD phasing and by Superflip 161
5.11 Overlay of *o*-Xtpb solved from different resolution data 161
5.12 Overlay of *o*-Xtpb solved from different resolution data including restraints on the MAD solution 163
5.13 Graph of radiation intensity *vs.* energy for different insertion devices . . . 165

Acknowledgements

This thesis is a culmination of my work over the past three and a half years, but it would not have been possible without the help of numerous people along the way who I would now like to thank.

Firstly I would like to thank my supervisor, Professor Richard Cooper (a.k.a. Coops/ Scooby/ Scuba Steve), not only for the opportunity but also for his continued support and guidance throughout the project. Ever the optimist, he kept things going through frustrating periods of experiments not quite going to plan. I will certainly miss the all-nighters at Diamond watching Netflix and listening to dubstep at 4am (whilst collecting data of course).

A huge thanks also to all members of the chem-cryst group both past and present. Thanks for making the office a great place to work, not forgetting the cheesetallography, quiztallography and all other '-tallography' meetings we've enjoyed throughout the years. I'm sure I'll be back from time to time...

I would also like to thank everyone who has directly helped with the project: the beamline staff at Diamond; various collaborators for supplying samples, saving me unwanted time in the lab; and Amber, Kirsten, Pascal and Jerome for giving up their time to help out at Diamond experiments. Special thanks to Pascal 'Poirot' for his help on the computing side of things. His input in the project has been invaluable right up until the very end. I would also like to give a special mention to Jerome and Andrew for the unforgettable 'post-conference' holidays, the numerous coffee breaks, the West Wing binges and so much more.

Last but by no means least I would like to thank my parents, friends and family for their encouragement and support over the past few years.

Abbreviations

°C	degree(s) Celsius
Å	Ångström(s), 10^{-10} m
Ac	acetyl
aq.	aqueous
AS	anomalous scattering
BVS	bond valence sums
Bz	benzoyl
cm^{-1}	wavenumber(s), inverse centimetre(s)
CCD	charge coupled device
DCrM	double crystal monochromator
DCM	dichloromethane
DMF	dimethyl formamide
EH1	experimental hutch 1
EH2	experimental hutch 2
EtOAc	ethyl acetate
EXAFS	extended X-ray aborption fine structure
$\boldsymbol{F}$	vector of quantity F
F	magnitude of quantity F
FWHM	full width half maximum
g	gram(s)
$\boldsymbol{h}$	vector of hkl
HFM	horizontal focussing mirror
HS	high spin
ID	insertion device

IR	isomorphous replacement
IUPAC	International Union of Pure and Applied Chemistry
K	Kelvin
LS	low spin
m	metre(s)
M	molarity, $\mathrm{mol\,dm^{-3}}$
MAD	multiple-wavelength anomalous diffraction
Mcps	mega counts per second
m.p.	melting point
Me	methyl
MeOH	methanol
mL	millilitre(s)
mol	mole(s)
NMR	nuclear magnetic resonance spectroscopy
s	second(s)
SAD	single-wavelength anomalous diffraction
SCO	spin-crossover
SQUID	superconducting quantum interference device
TBAB	tetrabutylammonium Bromide
TBAF	tetrabutylammonium Fluoride
THF	tetrahydrofuran
μrad	microradian(s)
μ	micro
VFM	vertical focussing mirror
XAFS	X-ray absorption fine structure
XANES	X-ray absorption near-edge spectroscopy
XPS	X-ray photoelectron spectroscopy
λ	wavelength

Chapter 1

Introduction

This chapter introduces the theory behind the project. Firstly an overview of X-ray scattering is described including the anomalous scattering phenomenon, determination of f' and f'', solution of the phase problem and subsequently structure solution. The project is divided into two main objectives linked by their dependence on exploiting the anomalous signal of certain elements: the structure solution of 'large' small molecules and the discrimination of elements and their electronic states. Sections 1.3 and 1.4 discuss the theory behind how we can maximise and exploit the anomalous signal in order achieve the objectives stated.

1.1 X-Ray Diffraction

In 1895, Wilhelm Röntgen became the first person to produce and detect electromagnetic radiation that he named X-rays – a discovery that would later earn him a Nobel prize in 1901. However, equipment at this time did not allow Röntgen to observe the interference and diffraction of X-rays. In 1912, Max von Laue and coworkers (later awarded the Nobel prize in 1914) demonstrated that crystals could be used as a diffraction grating for X-rays and since then numerous scientists have followed pioneers in the field such as William Lawrence Bragg and William Henry Bragg (jointly awarded the Nobel prize in 1915), to use this phenomenon to solve the three-dimensional structure of molecules.

1.1.1 X-ray Scattering

X-ray scattering occurs as a result of the interaction between the incident beam of X-rays and matter. This interaction occurs as a result of one of two processes. Firstly,

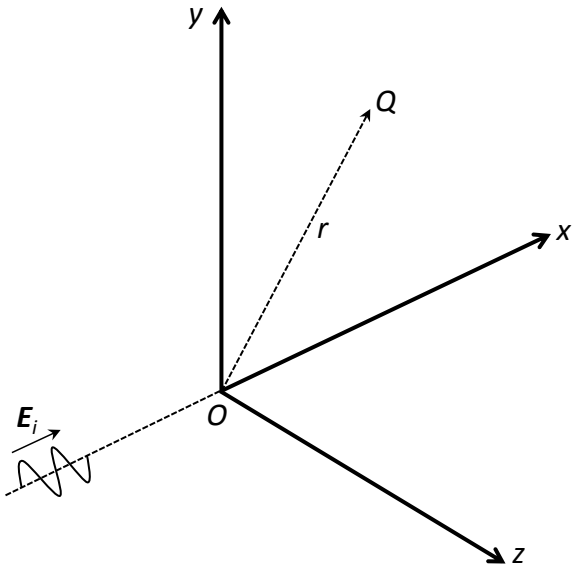


Figure 1.1 Diagram to show a particle at origin, O, being irradiated by a plane monochromatic wave propagating along the x-axis. Q represents the observation point for the system.

the X-ray beam is scattered by the electron density of the material. This process can either occur without a loss of energy *i.e.* the incident and scattered photons have the same wavelength, or with a slight loss of energy (known as Compton radiation whereby the scattered beam has a slightly longer wavelength than the incident radiation). [4] The second process by which X-rays are scattered is via an absorption of the X-ray by atoms in the crystal lattice. When an X-ray is absorbed, its energy removes an electron from the inner-shell of the atom. Upon returning to its original energy state, the atom emits a second photon with a wavelength characteristic of the absorbing atom (this process is known as fluorescence). [5]

Thomson Scattering

Thomson scattering is the classical scattering of X-rays from a charged free particle. [6]

Figure 1.1 shows a particle of charge e and mass m at position O being intersected with a plane monochromatic electromagnetic wave (frequency v propagating along the x-axis). If this is the case, the electric field at position x and time t is defined by,

$$\mathbf{E}_i = \mathbf{E}_{0i} \exp[2\pi i \nu (t - x/c)], \tag{1.1}$$

where $\mathbf{E}_{0i}$ is the amplitude of the wave and c is the speed of light. The force exerted by the electric field on the particle causes the particle to accelerate in the direction of the oscillating electric field. As described by the classical theory, this accelerated particle is now a source of electromagnetic radiation with an electronic field $\mathbf{E}_d$. It follows that if we measure the electronic field of the scattered radiation at a point Q (see Figure 1.1), it can be described by the equation

$$\mathbf{E}_d = \mathbf{E}_{0d} \exp[\pi i \nu (t - r/c) - i\alpha]. \tag{1.2}$$

where α is a phase difference between the incident radiation and the re-emitted wave. In 1898, Thomson showed that if we take the angle between the observer and the direction of acceleration of the free particle to be φ, Eqn. 1.2 can be written as [7]

$$\mathbf{E}_{0d} = \frac{1}{r}\mathbf{E}_{0i}\left(\frac{e^2}{mc^2}\right)\sin\varphi. \tag{1.3}$$

$\sin\varphi$ is a polarization term since we only observe the vibration component parallel to the observer. The value of $\mathbf{E}_d$ decreases as a function of r (the distance between the origin and the observation point) since scattering occurs from the charged particle in all directions. Finally, the $\frac{e^2}{mc^2}$ term in Eqn. 1.3 is the scattering amplitude from a single electron (equal to 2.818×10^{-13} cm). We can write Eqn. 1.3 in terms of intensity,

$$I_{eTh} = I_i\left(\frac{e^4}{m^2 r^2 c^4}\right)\sin^2\varphi, \tag{1.4}$$

where I_{eTh} is the intensity of the scattered beam and I_i, the intensity of the incident radiation. It is very hard to verify the Thomson effect experimentally since it would require a sample composed solely of free electrons. The closest equivalent would be a sample of light elements where electrons are weakly bound to their

nuclei, however the use of light elements instead reveals a different effect known as the Compton effect.

Compton Scattering

Compton scattering is the process by which a decrease in energy (or increase in wavelength) of an X-ray beam is observed upon interaction with matter. An inverse Compton effect is also possible whereby a photon gains energy (decreases in wavelength). In Compton scattering, an incident photon is scattered by a free electron along with the transfer of part of its energy to the electron. The change in wavelength as a result (the Compton shift, $\Delta\lambda$) between the incident and scattered radiation can be written as:

$$\Delta\lambda = {}^{h}/_{mc}(1-\cos 2\theta) = 0.024(1-\cos 2\theta). \tag{1.5}$$

$\Delta\lambda$ is independent of the energy of the incident radiation and it has a maximum value of 0.048 when 2θ is at a maximum of π. Finally, the Compton scattering is incoherent since there is no phase relationship between the incident and scattered light and as such, it is impossible to calculate the interference effects for Compton radiation. Therefore Compton scattering is generally ignored when considering X-ray scattering.

Interference of Diffracted Waves

It is the interference of the diffracted waves that leads to the diffraction pattern we observe in X-ray crystallography. The constructive interference of waves leads to an area of observed intensity on the detector, whereas destructive interference leads to an area of zero or near zero intensity. When considering the theory behind the interaction of diffracted waves, we regard the atoms in a crystal to be stationary. As such we can omit a time variable from future calculations.

Given a system containing two scatterers at points O and O′ (Figure 1.2), when

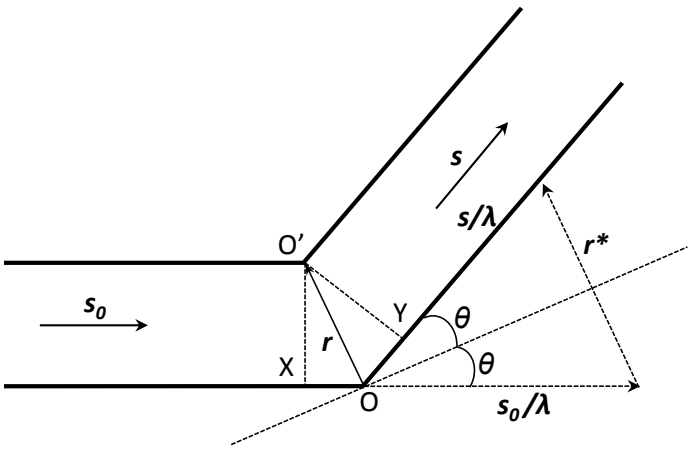


Figure 1.2 Diagram to show the scattering of the wave vector $\boldsymbol{s}_0$ along $\boldsymbol{s}$ by scatterers O and O′. XO = - $\boldsymbol{r} \cdot \boldsymbol{s}_0$, YO = $\boldsymbol{r} \cdot \boldsymbol{s}$

excited by incident radiation of wavelength λ, they become sources of secondary waves that interfere with one another. If $\boldsymbol{s}_0$ is the propagation vector of the incident X-ray beam, the phase difference, δ, of radiation scattered from points O and O′, along a propagation vector, $\boldsymbol{s}$, is calculated by

$$\delta = \frac{2\pi}{\lambda}(\boldsymbol{s}-\boldsymbol{s}_0)\cdot\boldsymbol{r} = 2\pi\boldsymbol{r}^*\cdot\boldsymbol{r}, \tag{1.6}$$

where

$$\boldsymbol{r} = O'-O, \tag{1.7}$$

$$\boldsymbol{r}^* = \lambda^{-1}(\boldsymbol{s}-\boldsymbol{s}_0). \tag{1.8}$$

When λ is much greater than $\boldsymbol{r}$, the phase difference tends to zero and so no interference can occur. For this reason we are unable to observe a diffraction pattern using visible light. However, when considering X-rays, with wavelengths comparable to inter-atomic distances (*c.* 1 - 4 Å), interference is significant.

Defining 2θ to be the angle between the incident and observed X-rays, the modulus of $\boldsymbol{r}^*$ is derived as

$$\boldsymbol{r}^* = \frac{2\sin\theta}{\lambda}. \tag{1.9}$$

If the amplitude of the wave scattered from O (phase = 0) is given by A_O, it follows that the corresponding amplitude from O′ is given by $A_{O'}\exp(2\pi i\boldsymbol{r}^*\cdot\boldsymbol{r})$. Therefore, if we consider a system of N scatterers

$$F(\boldsymbol{r}^*) = \sum_{j=1}^{N} A_j \exp(2\pi i\boldsymbol{r}^*\cdot\boldsymbol{r}_j). \tag{1.10}$$

For Thomson scattering, we can express the intensity of the scattered wave, I, in terms of intensity scattered by a free electron (I_{eTh}) – Eqn. 1.4. I/I_{eTh} is now f^2 where f is the scattering factor of an atom in this case. Therefore, when considering an atom at point O′, the term A_j in Eqn. 1.10 is replaced with f_j. However, when these scatterers are not single electrons but instead a continuum, each small volume d$\boldsymbol{r}$ contains a number of electrons equal to $\rho(\boldsymbol{r})$d$\boldsymbol{r}$, where $\rho(\boldsymbol{r})$ is the electron density. Eqn. 1.10 can now be rewritten as

$$F(\boldsymbol{r}^*) = \int_V \rho(\boldsymbol{r})\exp(2\pi i\boldsymbol{r}^*\cdot\boldsymbol{r})\,d\boldsymbol{r}. \tag{1.11}$$

This is an important result since the amplitude of the scattered wave is shown to be the Fourier transform of the density of the scatterers *i.e.* $\mathcal{F}[\rho(\boldsymbol{r})] = \mathcal{F}^{-1}[\mathrm{F}(\boldsymbol{r}^*)]$. This means that given the amplitudes and phases of the scattered waves, the electron density of the scatterer may be determined.

Temperature Factor

Atoms are bound to one another by inter-atomic forces in a way that minimises the energy of the molecule. Given a disturbance, the atoms will return to their lowest energy conformation. In doing so, the atoms gain thermal energy by oscillating around this minimum energy position. The oscillation in turn alters the electron density of a given atom (by increasing the scattering volume) and therefore its scattering

factor is affected. Since the time scale of thermal motion is much less than that of the experiment, we consider only a time-averaged description of the electron density $\rho(\boldsymbol{r}-\boldsymbol{r}')$. The function $\rho(\boldsymbol{r}')$ depends on the mass of the atom, the chemical bond forces surrounding it and temperature. If we describe the thermal motion isotropically by a mean square displacement, $U = \langle \boldsymbol{r}'^2 \rangle$, about its mean position, then the temperature factor for an atom is given by

$$B = 8\pi^2 U, \text{ [8]} \tag{1.12}$$

where U is measured in Å^2.

Scattering by a Collection of Atoms

If we consider the electron density of a thermally excited, isolated atom, $\rho_j(\boldsymbol{r})$, then the electron density of that atom at position $\boldsymbol{r}_j$ is described by $\rho_j(\boldsymbol{r}-\boldsymbol{r}_j)$. If we extend this to the electron density of a collection of N atoms which we will denote 'M', then

$$\rho_M(\boldsymbol{r}) = \sum_{j=1}^{N} \rho_j(\boldsymbol{r}-\boldsymbol{r}_j), \tag{1.13}$$

Substitution into Eqn. 1.11 gives

$$F_M(\boldsymbol{r}^*) = \int_S \sum_{j=1}^{N} \rho_j(\boldsymbol{r}-\boldsymbol{r}_j) \exp(2\pi i \boldsymbol{r}^* \cdot \boldsymbol{r})\, \mathrm{d}\boldsymbol{r}. \tag{1.14}$$

If $f_j(\boldsymbol{r}^*)$ is defined as the atomic scattering factor of the j^{th} atom with thermal motion included, the Eqn. 1.14 can be rewritten as

$$F_M(\boldsymbol{r}^*) = \sum_{j=1}^{N} f_j(\boldsymbol{r}^*) \exp(2\pi i \boldsymbol{r}^* \cdot \boldsymbol{r}_j). \tag{1.15}$$

Scattering by Crystals

An infinite three-dimensional lattice can be represented by the lattice function,

$$L(\boldsymbol{r}) = \sum_{u,v,w=-\infty}^{+\infty} \delta(\boldsymbol{r} - \boldsymbol{r}_{u,v,w}), \tag{1.16}$$

where δ is a Dirac delta function [9] and $\boldsymbol{r}_{u,v,w} = u\boldsymbol{a} + v\boldsymbol{b} + w\boldsymbol{c}$ describes the lattice vectors in the basis (**a**,**b**,**c**). The electron density of the crystal can then be calculated as a convolution of the electron density of a collection of atoms and the lattice function

$$\rho_\infty(\boldsymbol{r}) = \rho_M(\boldsymbol{r}) * L(\boldsymbol{r}). \tag{1.17}$$

By taking the product of the Fourier transforms of $\rho_M(\boldsymbol{r})$ and $L(\boldsymbol{r})$, the amplitude of a scattered wave by a crystal is shown to be

$$F_\infty(\boldsymbol{r}^*) = \frac{1}{V} F_M(\boldsymbol{h}) \sum_{u,v,w=-\infty}^{+\infty} \delta(\boldsymbol{r}^* - \boldsymbol{r}^*_{\boldsymbol{h}}), \tag{1.18}$$

where V is the unit cell volume, $\boldsymbol{h}$ represents the vector index of indices hkl and $\boldsymbol{r}^*_{\boldsymbol{h}} = h\boldsymbol{a}^* + k\boldsymbol{b}^* + l\boldsymbol{c}^*$ (lattice vectors in reciprocal space). In a crystal, the scatterers are arranged periodically on a lattice, and therefore we only observe a non-zero amplitude when $\boldsymbol{r}^*$ is coincident with a reciprocal lattice point,

$$\boldsymbol{r}^* = \boldsymbol{r}^*_{\boldsymbol{h}}. \tag{1.19}$$

From the definition of $\boldsymbol{r}^*$ given in Eqn. 1.7, we can multiply Eqn. 1.19 by vectors $\boldsymbol{a}$,$\boldsymbol{b}$ and $\boldsymbol{c}$ to give the equations

$$\boldsymbol{a} \cdot (\boldsymbol{s} - \boldsymbol{s}_0) = h\lambda, \tag{1.20}$$

$$\boldsymbol{b} \cdot (\boldsymbol{s} - \boldsymbol{s}_0) = k\lambda, \tag{1.21}$$

$$\boldsymbol{c} \cdot (\boldsymbol{s} - \boldsymbol{s}_0) = l\lambda, \tag{1.22}$$

These three equations are known as the Laue conditions and when satisfied, a

non-zero scattering amplitude will be observed. When considering the diffraction from a crystal containing N atoms in the unit cell, the value of $F_M(\boldsymbol{h})$ is referred to as the *structure factor* for a given vector index, $\boldsymbol{h}_{hkl}$,

$$F_{\boldsymbol{h}} = \sum_{j=1}^{N} f_j \exp(2\pi i \boldsymbol{r}_{\boldsymbol{h}}^* \cdot \boldsymbol{r}_j), \tag{1.23}$$

or more commonly expressed as

$$F_{\boldsymbol{h}} = \sum_{j=1}^{N} f_j \exp\left[2\pi i (hx_j + ky_j + lz_j)\right]. \tag{1.24}$$

In a crystal f_j is a continuous function of $\boldsymbol{r}_{\boldsymbol{h}}^*$. Therefore for each value of $\mathrm{F}_{\boldsymbol{h}}$, the value of f_j for each atom type at point $\boldsymbol{r}_{\boldsymbol{h}}^*$ must be calculated.

1.1.2 Bragg's Law

In 1912, William Lawrence Bragg described an elegant method for determining the conditions for diffraction to occur. [10] Bragg considered diffraction as a reflection of the X-ray beam from a set of crystal lattice planes with Miller indices hkl and spacing $\mathrm{d}_{\boldsymbol{h}}$. If we consider a wave (propagation vector $\boldsymbol{s}_0$) incident (with angle θ) on points A and B lying on given set of hkl Miller planes with spacing $\mathrm{d}_{\boldsymbol{h}}$, the path difference between the scattered beams will be BD + BC = $2\mathrm{d}_{\boldsymbol{h}} \sin\theta$ (see Figure 1.3). If the result is an integer multiple of λ then the waves will interfere constructively to give a positive intensity. This result gives rise to Bragg's Law;

$$2\mathrm{d}_{\boldsymbol{h}} \sin\theta = n\lambda. \tag{1.25}$$

In a crystal, X-rays are reflected from multiple crystal lattice planes simultaneously, and if the Bragg equation is not satisfied the waves will interfere destructively and no diffraction will be observed.

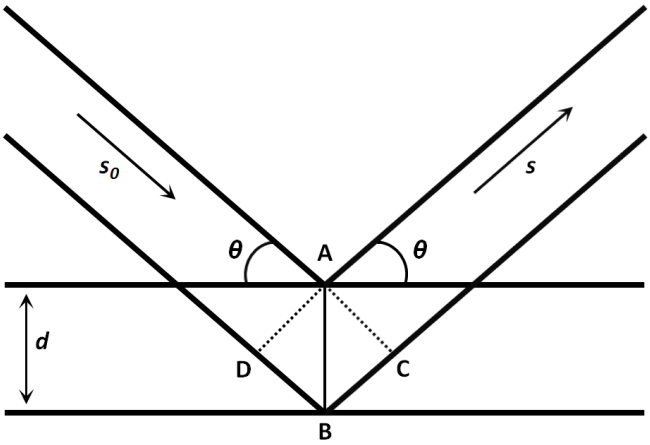


Figure 1.3 Reflection of X-rays from two planes belonging to the same Miller planes.

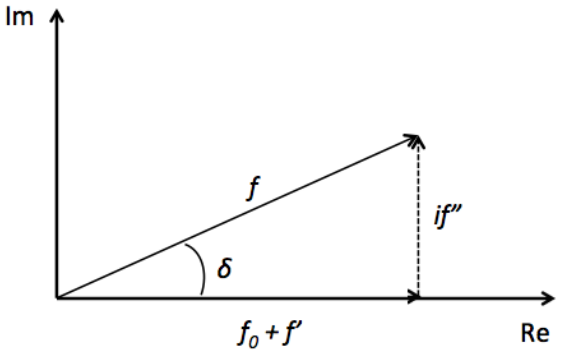


Figure 1.4 Atomic scattering factor split into its real and imaginary parts when anomalous scattering is present.

1.1.3 Anomalous Scattering

Anomalous scattering is an interaction that occurs when the energy of an incident X-ray photon is close to an electronic resonance of a bound inner-electron in a low lying atomic orbital. The resulting resonance causes a change in phase of the diffracted X-ray beam. To account for the phase change, the normal atomic (Thomson) scattering factor, f_0, is modified such that,

$$f = f_0 + f' + if'' = \mid f \mid e^{i\delta}, \tag{1.26}$$

where f' and f'' are the real and imaginary parts of the anomalous scattering respectively (see Figure 1.4).

There are three main approaches to determining how f_0 is modified. Hönl took a quantum mechanical approach to obtain oscillator strengths and from these to calculate the photoelectric absorption cross-sections. [11, 12] Parratt and Hempstead applied a semi-emiprical relation for the photoelectric absorption cross-section in order to obtain values for f' and f'', [13] and finally Cromer and Liberman used Slater-Dirac wave functions. [14] Below we only consider the scattering by a classical dipole oscillator. A more in-depth analysis is described by Wagenfeld. [15]

Firstly we consider an electron of mass m and charge e in an atom. When an electromagnetic wave hits the particle, the instantaneous electric field ($E_0 e^{iwt}$) at the dipole position causes the electron to oscillate, which can be described by the classical differential equation

$$\ddot{x} + k\dot{x} + \omega_3^2 x = \frac{eE_0}{m} \cdot e^{i\omega t}, \tag{1.27}$$

where k is a damping force proportional to the velocity and ω_3 is the natural angular frequency of the electron. The solution to Eqn. 1.27 is

$$x(t) = \frac{eE_0}{m} \cdot \frac{e^{i\omega t}}{\omega_3^2 - \omega^2 + ik\omega}, \tag{1.28}$$

and the dipole moment, M, can be calculated by multiplying the displacement by the polarisability moment such that

$$M = ex(t). \tag{1.29}$$

The dipole oscillator scatters with the same frequency of the oscillation, the amplitude of which, at a distance r, is ω^2/rc^2 times the dipole moment

$$A(r) = \frac{e^2}{mrc^2} \cdot \frac{\omega^2 E_0}{\omega_3^2 - \omega^2 + ik\omega}. \tag{1.30}$$

The scattering factor is defined as the ratio of the scattered amplitude, $A(r)$, to that of a Thomson free electron, A_{Th}. In this case A_{Th} is defined when $\omega_3 = 0 = k$

Element	f' / electrons	f'' /electrons
S	-7.76	4.10
Mn	-8.12	3.96
Fe	-8.15	3.95
Gd	-16.2	28.4

Table 1.1 Anomalous scattering factors for a selection of elements.

$$A_{\mathrm{Th}} = -\frac{e^2}{mc^2} \cdot E_0. \tag{1.31}$$

It follows that the anomalous scattering factor is given by

$$f_{\mathrm{anom}} = \frac{A}{A_{\mathrm{Th}}} = \frac{\omega^2}{\omega^2 - \omega_3^2 - i\,k\omega}. \tag{1.32}$$

Eqn. 1.32 shows that as the frequency of the incident radiation, ω, tends to ω_3, the scattering becomes imaginary. We can now define the anomalous scattering factor in terms of the real, $f'(\omega)$ and imaginary, $f''(\omega)$ parts

$$f_{\mathrm{anom}} = f' + i\,f'' = \frac{\omega^2(\omega^2 - \omega_3^2)}{(\omega^2 - \omega_3^2)^2 + k^2\omega^2} + i\frac{k\omega^2}{(\omega^2 - \omega_3^2)^2 + k^2\omega^2}. \tag{1.33}$$

Apart from the obvious dependence that scattering factor has on frequency of the incident radiation, it is generally considered that the effect of angle on f is negligible and therefore f' and f'' are constant for a given wavelength.

When using lab sources *i.e.* Cu = 1.542 Å and Mo = 0.7107 Å, the anomalous contribution to the overall scattering for most compounds composed of light elements is negligible. However, when using Synchrotron radiation, the wavelength can be tuned specifically so as to maximise the anomalous signal from a given element. The size of the signal can vary hugely depending on the element being observed. Table 1.1 shows the maximum possible anomalous signals that can be measured for a selection of elements through careful choice of wavelength.

It is this large, tunable additional signal to the normal scattering that can be exploited to provide phasing information in techniques such as multiple-wavelength anomalous diffraction. [16–18] We can also use the anomalous signal for determining oxidation states in mixed-oxidation state compounds and for identifying elements with near identical scattering factors. These methods form the basis of this project and will be discussed in more detail in Section 1.2.

Anomalous Scattering vs. Structure Factor

If we consider a molecule containing N atoms in a unit cell of which A are anomalous scatterers, it follows that the number of other normal scatterers, O, is given by $O = N - A$. Using the notation '+' and '-' as the magnitudes of vectors $\boldsymbol{h}$ and $\boldsymbol{-h}$ respectively, then

$$F^+ = F_O^+ + F_A'^+ + iF_A''^+ = F'^+ + iF_A''^+, \tag{1.34}$$

$$F^- = F_O^- + F_A'^- + iF_A''^- = F'^- + iF_A''^-. \tag{1.35}$$

Substitution of Eqn. 1.24 into Eqn. 1.34 give the overall structure factor equation for F^+, as

$$F^+ = \sum_{j=A+1}^{A+O} f_j \exp[2\pi i \boldsymbol{h}\cdot\boldsymbol{x}_j] + \sum_{j=1}^{A} f_j' \exp[2\pi i \boldsymbol{h}\cdot\boldsymbol{x}_j] + i\sum_{j=1}^{A} f_j'' \exp[2\pi i \boldsymbol{h}\cdot\boldsymbol{x}_j], \tag{1.36}$$

where $\boldsymbol{h}\cdot\boldsymbol{x}_j = (hx_j + kx_j + lx_j)$. The relationship between $F_{\boldsymbol{h}}$ and $F_{-\boldsymbol{h}}$ is represented on the Argand diagram in Figure 1.5. For clarity, Figure 1.6 shows the relationship between $F_{\boldsymbol{h}}$ and the complex conjugate of $F_{-\boldsymbol{h}}$

$$(F^-)^* = F_O^+ + F_A'^+ - iF_A''^+ = F'^+ - iF_A''^+. \tag{1.37}$$

The difference in intensity, $\Delta I = |F^+|^2 - |F^-|^2$ is known as the Bijovet difference

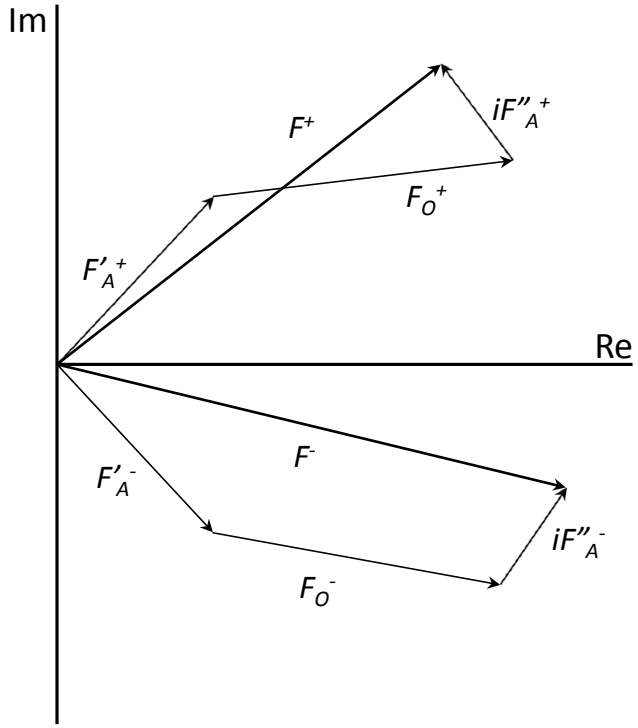


Figure 1.5 Argand diagram showing the relationship between $F_{\boldsymbol{h}}$ and $F_{-\boldsymbol{h}}$.

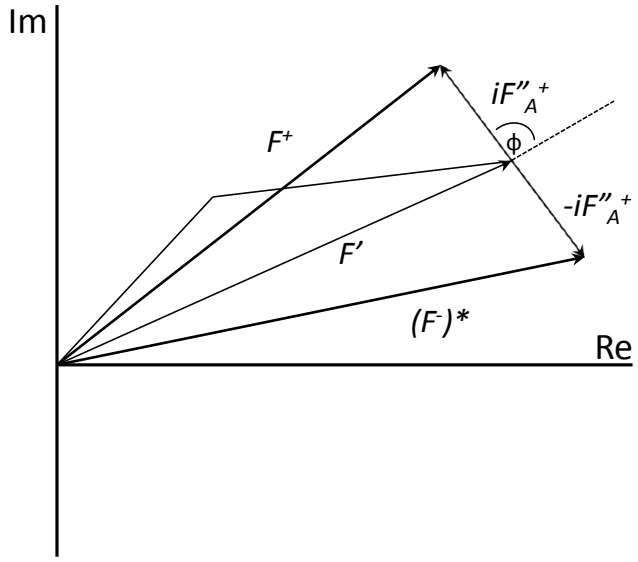


Figure 1.6 Argand diagram showing the relationship between $F_{\boldsymbol{h}}$ and $F^*_{-\boldsymbol{h}}$ – using the same contributions as Figure 1.5.

and defined by,

$$|F^+|^2 = |F'|^2 + |F''_A|^2 + 2|F'||F''_A|\cos\varphi, \tag{1.38}$$

$$|F^-|^2 = |F'|^2 + |F''_A|^2 - 2|F'||F''_A|\cos\varphi, \tag{1.39}$$

to give,

$$\Delta I = 4|F'||F''_A|\cos\varphi. \tag{1.40}$$

The important result from this shows that in the non-centrosymmetric case, $F_{\boldsymbol{h}} \neq F_{-\boldsymbol{h}}$ *i.e.* Friedel's Law ($|F^+| = |F^-|$) no longer holds. Friedel's law will only be satisfied if the molecule is centrosymmetric or the crystal is composed solely of one atom type (the anomalous scatterer).

1.1.4 Determination of f' and f''

The absorption of X-rays at resonant energies (the absorption edge) can be characterised by measuring the absorption-fluorescence with respect to the incident photon energy. From this, we are able to compute the real, f', and imaginary, f'', parts of the anomalous scattering factors at energies on and around an absorption edge.

Fluorescence-absorption experiments are routinely applied in macromolecular crystallography to find the position of the absorption edge for heavy atoms *e.g.* Selenium (Figure 1.7), that have been substituted into the crystals. [19] Experimental values for the real and imaginary parts of the anomalous scattering factor can be computed using two equations. The first equation, the Optical Theorem, [20] directly relates the absorption-fluorescence data to the imaginary, f'', part of the anomalous scattering factor by

$$f'' = \frac{m_e c \varepsilon_0 E \mu_a}{e\hbar}, \tag{1.41}$$

where m_e is the mass of an electron, ε_0 is the permittivity of free space constant, E is the incident X-ray energy, and μ_a is the atomic absorption coefficient. The real part of the anomalous scattering (AS) factor is then determined *via* a numerical integration of the Kramers-Kronig transformation, [21]

$$f'(E_0) = (2/\pi) \int_0^\infty \left[\frac{E f''(E)}{(E_0^2 - E^2)} \right] dE. \tag{1.42}$$

However, these calculations only apply to isolated atoms and therefore deviations from calculated values are observed in the experimental data. Such deviations are most noticeable in the vicinity of the absorption edge (giving rise to X-ray absorption near-edge structure (XANES) and extended absorption fine structure (EXAFS)), caused by the chemical environment around the absorbing species interfering with ejected photo-electrons (Figure 1.8). [22] Calculation of the real and imaginary components of the AS from the fluorescence experiment is implemented in the program CHOOCH [23] for cases where there are no overlapping absorption edges.

1.1.5 Phase Problem

Section 1.1.1 stated that the Fourier transform of the structure factors provides us with the electron density, ρ, of the scatterers ($\mathcal{F}[\rho(\boldsymbol{r})] = \mathcal{F}^{-1}[\mathrm{F}(\boldsymbol{r}^*)]$). We can write the Fourier transform out in full such that

$$\rho_{\boldsymbol{x}} = \frac{1}{V} \sum_{\boldsymbol{h}=-\infty}^{+\infty} |F_{\boldsymbol{h}}| \exp\left[i\varphi_{\boldsymbol{h}}\right] \exp\left[-2\pi i(hx + ky + lz)\right]. \tag{1.43}$$

This means that given the modulus of the structure factors and their respective phases, the atomic positions within the unit cell can be calculated unequivocally.

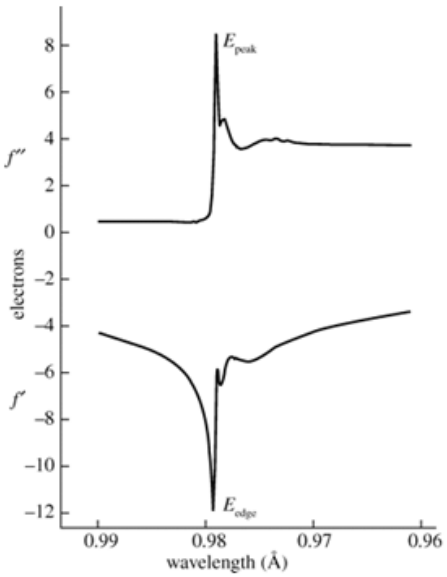


Figure 1.7 AS factors for a selenium atom.

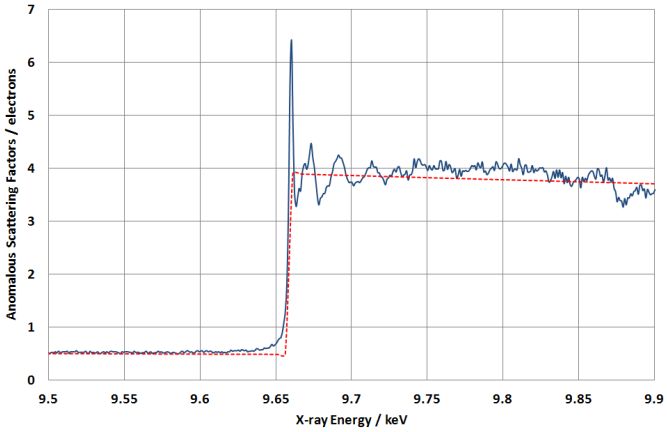


Figure 1.8 The effect of EXAFS and XANES on the imaginary part of AS factors. The red dotted line shows the theoretical values for a Zn(0) [24] species and the solid blue line shows the calculated spectrum from a sample of octaethyl-porphine zinc.

However, in a typical X-ray diffraction experiment the measured quantity is the intensity of the diffracted beams. From these we are able to calculate the structure factor moduli using

$$|F_{hkl}| \propto \sqrt{I_{hkl}}, \tag{1.44}$$

although any information pertaining to the phases is lost. In crystallography, this is known as the *phase problem*, the solution of which is essential in structure solution. Various techniques have been developed for solving the phase problem, many of which will be discussed in section 1.2.

1.2 Structure Solution and Refinement

1.2.1 Small-Molecule

The basis of all structure solution techniques is the calculation of missing information about the phases associated with the measured structure factors. Although there is no general approach to this problem, a number of reliable techniques are available to small molecule crystallographers. Originally, crystal structures were solved by a trial and error method using predicted structures. However, these techniques are difficult, time consuming, and only work for very simple structures. [25] This section describes three of the more common techniques for solving the phase problem.

Patterson Methods

The Patterson synthesis provides information about the atomic positions of just some of the atoms in the structure. The method is particularly effective when considering molecules made up of a few heavy atoms amongst a larger number of lighter atoms. This is often the case when studying inorganic molecules and especially in organometallic compounds. To do this, the Patterson method initially ignores phases from the problem completely. This is done by multiplying the observed structure

factors, F_{hkl}, by their complex conjugates, F^*_{hkl}. [26] This technique is commonly used in mathematics since it has the result of removing the imaginary part of a complex number to leave only a real number. When this is applied to Eqn. 1.43, we are still left with a Fourier transform, however there is no longer a phase dependence. This is known as the Patterson function,

$$P_{uvw} = \frac{1}{V} \sum_{\boldsymbol{h}} | F_{\boldsymbol{h}} |^2 \cos [2\pi(hu + kv + lw)] . \tag{1.45}$$

The result of the Patterson synthesis is not a direct representation of the electron density but instead a Patterson map which is closely related. To ensure a distinction between the two, we use the indices (u,v,w) instead of the conventional (x,y,z) to describe fractions of the unit cell. An example of a typical Patterson map (w = 1/2 Harker Section*) is shown in Figure 1.9.

The peaks in Figure 1.9 do not represent atoms as in an electron density map, rather they are vectors between pairs of atoms in the structure. For every atom in a given structure, the Patterson map displays all of the vectors between that atom and every other atom in the structure (including itself). For example, if a structure contained only two atoms in the unit cell, A and B, at positions (x_A,y_A,z_A) and (x_B,y_B,z_B) respectively, the Patterson Map would show peaks at $(u_{x_A-x_B},\ v_{y_A-y_B},\ w_{z_A-z_B})$ and $(u_{x_B-x_A},\ v_{y_B-y_A},\ w_{z_B-z_A})$, along with the largest peak at the origin (seen as the dark green peak in Figure 1.9) to represent the vector between each atom and itself. It therefore follows that for a structure containing n atoms, there will be n^2 peaks in the Patterson map. As a result, Patterson maps are very noisy and as the size of the structure increases, there rapidly becomes more and more overlap of peaks. One consequence of the Patterson function however is that peak sizes are proportional to the product of the atomic numbers of the atoms involved in the vector being represented. Therefore, peaks due to heavy atoms dominate the Patterson map. Given

*Due to symmetry elements within space groups, peaks appear in the Patterson map along certain lines or planes determined by those crystallographic symmetries. A Harker Section is one such plane containing a large number of Harker peaks (vectors between symmetry equivalent atoms), which are readily interpretable.

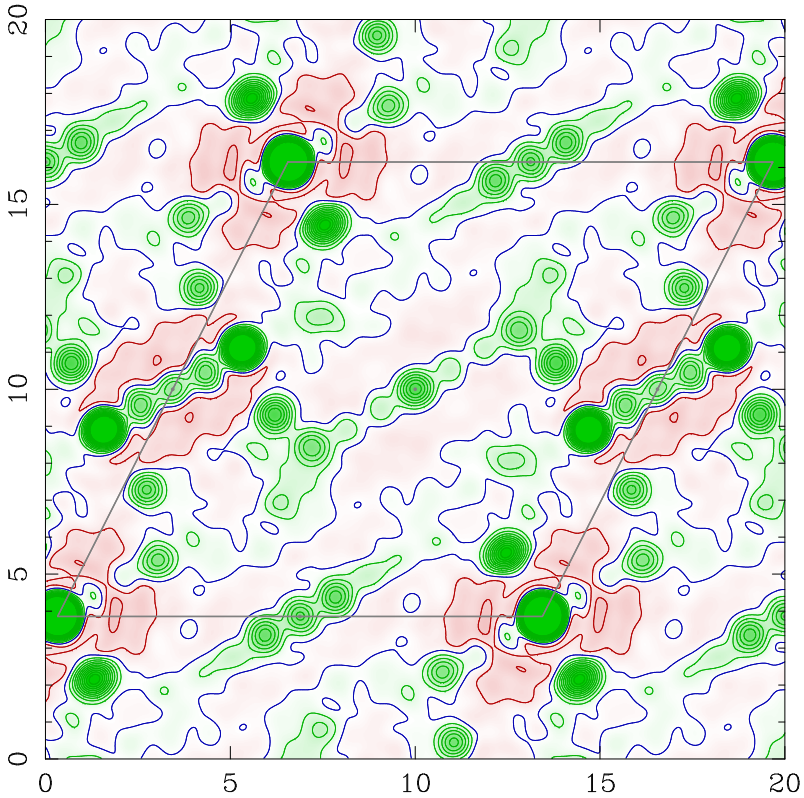


Figure 1.9 Observed Patterson map for Copper (II) Acetate, $w = 1/2$ Harker Section, parallel to the *ab* plane. Axes are measured in Ångstroms. Green = positive and red = negative electron density, where colour density is proportional to magnitude. The solid grey line shows the projection of the unit cell.

these vectors, and knowledge of the space group, atomic position for the heavy atoms in the structure can be determined. Calculated phases from these can then be used for solution of the rest of the structure and hence, Patterson maps are particularly effective for structures containing few heavy atoms. The final point to note about the Patterson map is that each pair of atoms will give two vectors that are equal and opposite *i.e.* A→B = B→A. Therefore the Patterson map will always contain a centre of inversion whether or not the crystal itself is centrosymmetric.

Direct Methods

Direct methods is the process of determining phases by searching for relationships between observed structure factors. It is deemed 'direct' since the solution requires

Constraint	Use of Constraint
$\rho_{\boldsymbol{x}} \geq 0$	Inequality relationships
Discrete atoms	Normalised structure factors
Random distribution of atoms	Phase relationships and tangent formula
$\int \rho^3 \mathrm{d}V = \max$	Tangent Formula
Equal atoms	Sayre's equation
$-\int \rho_{\boldsymbol{x}} \ln \frac{\rho_{\boldsymbol{x}}}{q_{\boldsymbol{x}}} \mathrm{d}V = \max$	Maximum entropy methods
Equal molecules	Molecular replacement methods
$\rho_{\boldsymbol{x}}$ = constant	Density modification techniques

Table 1.2 Constraints placed on the electron density in direct methods.

no prior knowledge of the structure and is solved purely mathematically.

The amplitude and phase of a given structure factor are related to one another via the electron density (Eqn. 1.43). Using chemical knowledge of what properties the electron density should have we can impose certain constraints on the electron density, $\rho_{\boldsymbol{x}}$, which in turn places constraints on the structure factors. Since we know the structure factor amplitudes, these constraints are almost purely directed at the phases. A number of such constraints are listed in Table 1.2 and are described in more detail in Crystal Structure Analysis, Principles and Practice. [27] The two most important of these constraints are;

1. the electron density, $\rho_{\boldsymbol{x}}$, can never be a negative value since this is physically impossible;

2. the electron density map contains discrete, sharp peaks near to an atomic position, and near zero values elsewhere.

In order to solve a structure using direct methods, the measured structure factors, $F_{\boldsymbol{h}}$, are first normalized to $E_{\boldsymbol{h}}$ values such that

$$E_{\boldsymbol{h}} \propto \frac{F_{\boldsymbol{h}}}{\sqrt{\sum_i f_i}}. \text{ [28]} \tag{1.46}$$

This accounts for decrease in scattering factor as a function of Bragg angle. Only the largest values of $E_{\boldsymbol{h}}$ are used for the phasing to provide the most accurate answer. Reflections are examined in triplets such that for a given triplet $(h_1 k_1 l_1)$, $(h_2 k_2 l_2)$ and $(h_3 k_3 l_3)$, $h_1 + h_2 + h_3 = k_1 + k_2 + k_3 = l_1 + l_2 + l_3 = 0$ *e.g.* $(33\bar{1})$, $(1\bar{6}0)$, $(\bar{4}31)$. To improve phase relationships between triplets and therefore the accuracy of the solution, each chosen $E_{\boldsymbol{h}}$ value should appear in as many triplets as possible. For most structures, between 20 and 300 initial phases are generally applied to reflections in order to reach a stable solution. Firstly, these are used to define an origin of the unit cell, and then phase relationships between triplets are examined. The probability of each calculated phase being correct can be determined using combined figures of merit, and only the most probable phases used to create an initial electron density map. The map can then be examined for features such as phenyl rings *etc.* which are subsequently used as an initial trial structure for further refinement. Obviously the initial phases are unlikely to be correct, so the process can be repeated using a Monte Carlo technique [29] until a stable and sensible solution is found.

Charge Flipping

Unlike direct methods, charge flipping is an iterative algorithm working in both direct and reciprocal space to give a solution. It is therefore referred to it as a dual-space technique. Firstly, random phases, φ_{random}, are applied to all of the observed structure factors, $|F_{\text{obs}}|$, setting all unobserved structure factors to zero. Once this has been done, the algorithm enters into a cycle until a convergence is reached. Since the algorithm begins with the assignment of random phases, convergence is not always possible and the process is repeated. The iterative cycle can be represented in four steps; [30, 31]

1. Using the n structure factors with their randomly assigned phases, $F^{(n)}$, the reverse Fourier transform, Eqn. 1.43, is calculated to give a starting electron density, $\rho^{(n)}$.

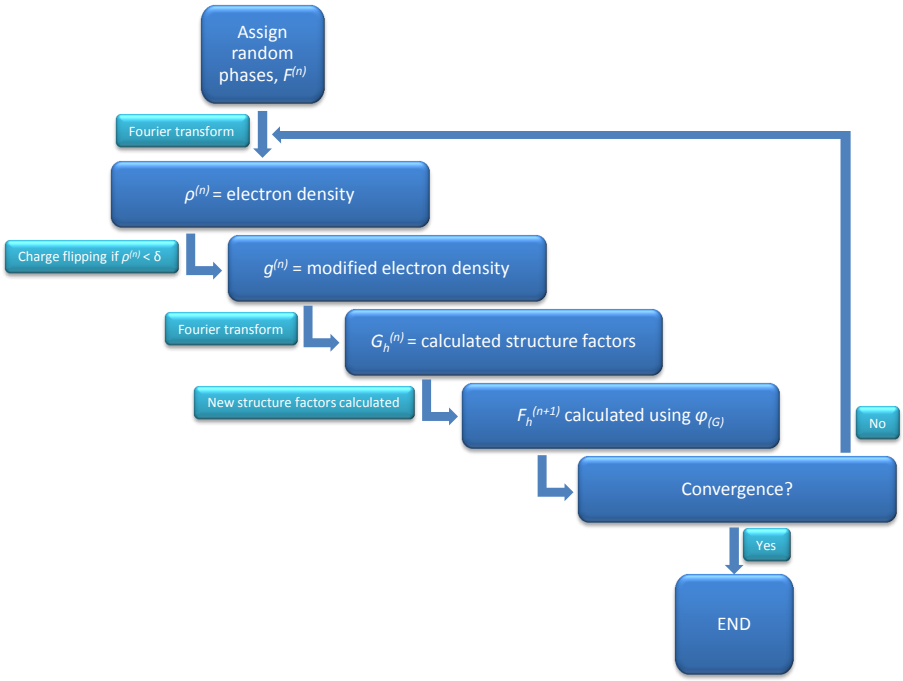


Figure 1.10 Flow diagram of the charge flipping algorithm.

2. If we imagine the 3 dimensional electron density map to be made up of N pixels, the value of the electron density is then sampled at each pixel. As described in the direct methods section, we know that the electron density must be ≥ 0 at all points in the map. The electron density is therefore modified in order to meet this constraint such that any pixel containing an electron density below a certain threshold, δ (which is close to but not exactly zero), has its value 'flipped' to become positive. We are now left with a modified electron density map, $g^{(n)}$.

3. This time the forward Fourier transform (reverse of Eqn.1.43) is applied to the modified electron density to give temporary structure factors $G_{\boldsymbol{h}}^{(n)}$.

4. From these modified structure factors, new structure factors are calculated, $F_{\boldsymbol{h}}^{(n+1)}$, by taking the original observed structure factor amplitudes and combining them with the newly calculated phases associated with the modified structure factors, $\varphi_{(G)}$. All non-measured structure factors are once again set to zero. These structure factors are then cycled back into the algorithm.

The above process is continued until a convergence on the correct answer has been reached. The algorithm has been summarised in Figure 1.10. [32] The most obvious descriptor of convergence is to look at the difference between the experimental and 'flipped' structure factor amplitudes *i.e.* $|F_{obs}| \; vs. \; |G|$. A large drop in this difference often suggests convergence on a solution. [31] It is also important to note that no symmetry operation is applied during the iteration (all structures are initially solved in the space group **P**1). This information can be calculated after convergence has been obtained by using the diffraction intensities. [33]

Charge flipping has proven successful on a wide variety of compounds and is mainly limited by data completeness. Since the algorithm heavily relies on the Fourier transform, missing a few strong reflections greatly reduce its likelihood of success. Data should also have relatively good resolution with recommended values of 1.2 Å for most organic molecules or up to 1.6 Å if the molecule contains heavy atoms.

1.2.2 Macromolecular

As with small molecule structure solution, the phase problem lies at the heart of macromolecular structure solution. However, the techniques described above are less applicable when studying proteins since they rely on good quality and high resolution data sets. Protein crystals tend to between 20 and 80% solvent, [34] are subject to radiation damage [35] (they do not last very long in the X-ray beam before breaking down), and resolution is much more limited. For this reason, macromolecular phasing takes a different approach to that of small molecule. The three most common techniques are isomorphous replacement, molecular replacement and multi-wavelength anomalous diffraction (MAD) phasing. [6]

Isomorphous Replacement

This is the simplest and most common method used in macromolecular crystallography and was used to solve some of the first ever protein structures. [36, 37] It works

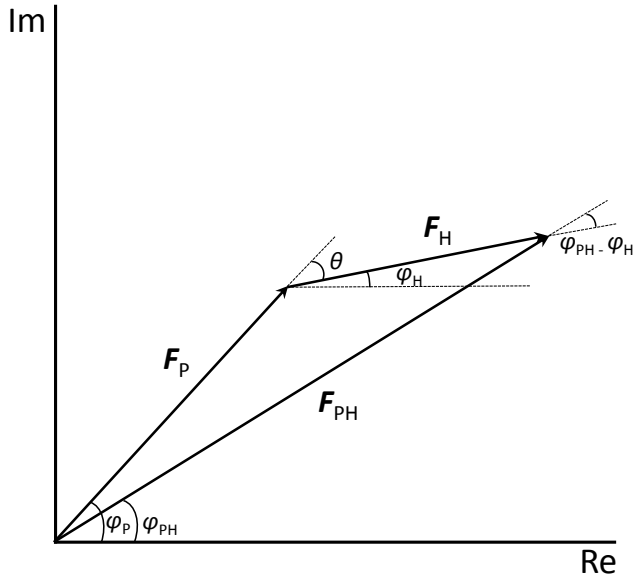


Figure 1.11 Argand diagram showing the relationship between structure factors of the native protein, $\boldsymbol{F}_P$, the heavy atoms, $\boldsymbol{F}_H$, and the heavy-atom derivative, $\boldsymbol{F}_{PH}$.

by taking isomorphous heavy-atom derivatives of the protein being studied to solve the phase problem. Changes in measured intensities between structures are used to locate the heavy atoms, and from there the whole molecule can be phased to give a solution. The heavy atom derivatives are formed by soaking the native protein in solutions such that heavy atoms interact with the surface of the protein without changing its overall structure. [38] The technique was first applied to protein structure solution by Perutz [39] and Kendrew. [40]

If F_P and φ_P represent the magnitude and phase of the native protein structure factor, and F_{PH} and φ_{PH} those of the heavy-atom derivative structure factor, assuming a perfect isomorphism between the two, their structure factors are related by

$$\boldsymbol{F}_{PH} = \boldsymbol{F}_P + \boldsymbol{F}_H. \tag{1.47}$$

In this case $\boldsymbol{F}_H$ represents the vector from the heavy atoms alone. In order to derive φ_{PH} the last quantity that must be determined is $F_{\mathbf{H}}$ (see Figure 1.11). Since it is made up from the heavy atom contributions alone, by locating the heavy-atom substructure, $F_{\mathbf{H}}$ can be calculated and therefore an estimate of φ_{PH} formulated. Deter-

mination of the heavy atom positions can either be calculated using a Patterson map (see Section 1.2 – however, due to the large number of atoms in a protein molecule these maps often cannot be interpreted) or a difference-Patterson synthesis. [41]

Molecular Replacement

Molecular replacement works by using the phases from a known structure that is similar to that being studied as a starting model for the refinement. It is suggested for proteins that the 'similar' structure should have an RMSD of <2.0 Å for the C^{α} atoms and an amino acid sequence that is > 25% identical. [42]

In order to use the phases from the similar model, it must first be rotated and translated within the unit cell to best recreate the measured structure factors from the molecule being solved. These methods were first described by Rossmann and Blow in 1962. [43] There are now a number of programs available in the CCP4 software suite to carry out these rotation and translation searches along with the MR phasing *e.g.* Molrep [44] and Phaser. [45]

MAD Phasing

Multi-wavelength anomalous dispersion exploits the anomalous signal from heavy atoms in the structure in order to solve the phase problem. Since this technique forms the basis of much of Chapter 5, the background is described in detail later (Section 1.3).

1.2.3 Structure Refinement

Once the phase problem has been solved and an initial trial structure found, the next step is to improve the structural parameters. This process is known as refinement, with the ultimate goal of finding a best fit between the observed data and the calculated data (generated by a Fourier transform from the trial structure). The most common method of refinement in small molecule crystallography is least-squares re-

finement. [46] This technique is described in more detail below since it is the method we have used for all refinement processes throughout the project using the CRYSTALS software. [3] Although least-squares refinement can be used for macromolecular structure solution, [47] more recently techniques such as maximum likelihood, [48] maximum entropy [49] and simulated annealing [50] are much more common.

Least-squares Refinement

Least-squares is a common approach taken to find a solution in an overdetermined system – a system which is described by more measurements than there are variables. The technique aims to solve the problem by minimising the sum of the squares of the errors in the results of each individual equation. In the case of diffraction data this is to minimise the difference between the observed and calculated structure factor data. We can represent this mathematically as

$$S = \sum_{\boldsymbol{h}} w(\boldsymbol{h}) \mid Y_{\mathrm{obs}}(\boldsymbol{h}) - Y_{\mathrm{calc}}(\boldsymbol{h}) \mid^2, \tag{1.48}$$

where $Y_{\mathrm{obs}}(\boldsymbol{h})$ is the observed structure factor coefficient for a given reflection and $Y_{\mathrm{calc}}(\boldsymbol{h})$ is the calculated value for the same reflection using the trial structure. [27] A weight, $w(\boldsymbol{h})$, is also applied, representing the importance of a given reflection to the overall sum. The structure factor coefficient, Y, can either refer to the intensity of a reflection, I, the magnitude of the structure factor $\mid F \mid$ or the square of the structure factor, F^2.

The trial structure consists of a number of atoms placed in the unit cell defined by a variety of parameters. By selectively changing certain parameters in the model, we can aim to reduce the difference between the observed and calculated data. These are known as 'refinable parameters' and include; atomic coordinates; atomic displacement parameters; the scale factor between the observed and calculated intensities; extinction parameters; site occupancy; and twin components. [46] A measure of the agreement between the model and the experimental data is provided by the

R-factor [51] – defined as

$$R = \frac{\sum || F_{obs} | - | F_{calc} ||}{\sum | F_{obs} |}. \tag{1.49}$$

As well as simply varying the refinable parameters, we are able to impose relationships between the parameters by modifying the general minimisation function above (Eqn. 1.48) using constraints and restraints. A constraint is defined as "an exact mathematical relationship that reduces the number of free parameters in a model".† Constraints are rigid and must be obeyed during each refinement cycle, an example of which is fixing the atomic coordinates of an atom which sits on an inversion centre. A restraint on the other hand is defined as "an additional condition that the model parameters must meet to satisfy some additional piece of knowledge appropriate to the structure".† Restraints are not strictly exact, and are subject to a slight probability distribution *e.g.* two chemically (but not crystallographically) equivalent bonds can be restrained to be approximately equal with a certain standard deviation (0.01 Å for example). Restraints are therefore equivalent to 'additional information' in the refinement on a par with the experimental measurements and do not change the number of refinable parameters.

Programs such as CRYSTALS and SHELX use the least-squares technique in order to refine small molecule structures using single crystal X-ray diffraction data. Below is a description of its implementation, which first requires a short introduction to linear least-square fitting followed by an extension to the non-linear least-squares method required in crystallography.

Linear Least-squares

The simplest case with which to describe this process is where we only have two observed data points (x_1, y_1) and (x_2, y_2). [27] The aim of the linear least-squares is to fit a straight line with the function $y = mx + c$ through these points. We can therefore

†Definition taken from the IUCr Online Dictionary of Crystallography: http://reference.iucr.org/dictionary/Main_Page

form two equations containing two unknowns, m and c

$$y_1 = mx_1 + c,$$
$$y_2 = mx_2 + c. \tag{1.50}$$

These can be written in matrix form as

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \end{pmatrix} \begin{pmatrix} m \\ c \end{pmatrix} \tag{1.51}$$

In order to re-write this matrix form in terms of the unknowns, the 2×2 matrix (also known as the *design matrix*) above is inverted and multiplied by the column matrix for y such that

$$\begin{pmatrix} m \\ c \end{pmatrix} = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \tag{1.52}$$

to give an exact solution for m and c.

Moving on from this simple case of two equations and two unknowns, we consider the overdetermined case of 5 observed data points. As before, we can write out 5 equations in matrix form, however since there are only 2 unknowns, they cannot all be satisfied simultaneously by the same values for m and c *i.e.* the points no longer lie on a straight line

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix} = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ x_4 & 1 \\ x_5 & 1 \end{pmatrix} \begin{pmatrix} m \\ c \end{pmatrix} \tag{1.53}$$

Although a solution can be found for any pair of the above equations, the answer is unlikely to be correct for a different chosen pair. Therefore, the aim now becomes to find the best (rather than exact) fitting straight line fit through the points. This is

achieved by minimising the function $(y_{\text{obs}} - y_{\text{calc}})^2$ with appropriate choice of m and c. For a straight line function through the points this minimisation function is of the form

$$S = \sum_i \left(y_i - (mx_i + c)\right)^2. \tag{1.54}$$

It can then be shown that the function S will be minimised when the partial derivatives with respect to each of the parameters in the system (*i.e.* m and c) is set to zero

$$\begin{aligned} \left(\frac{\partial S}{\partial m}\right)_c &= 2m\sum_i x_i^2 + 2c\sum_i x_i - 2\sum_i x_i y_i = 0 \\ \left(\frac{\partial S}{\partial c}\right)_m &= 2m\sum_i x_i + 2c\sum_i 1 - 2\sum_i y_i = 0 \end{aligned} \tag{1.55}$$

By solving these two linear equations containing two unknowns, the 'best' values of m and c may be determined. This approach will always work given a system described by more equations than there are linear parameters.

Non-linear Least-squares

In the case of a crystallographic least-squares refinement, the linear approach is no longer useful. The observations from the experiment are described by the structure factor equation (Eqn. 1.24) which is not linear with respect to the parameters.

The function to be minimised is described in Eqn. 1.48, in which the parameters, p_{j} are non-linear. As before, the function will be minimised when each partial derivative of each parameter is zero,

$$\left(\frac{\partial S}{\partial p_{\text{j}}}\right)_{\text{k}} = \sum_{i=1}^{N}\left[2w_i\left(y_i(\text{obs}) - y_i(\text{calc})\right)\left(\frac{\partial y_i}{\partial p_{\text{j}}}\right)_{\text{k}}\right] = 0, \tag{1.56}$$

where N is the number of observations, i. Given a system of M parameters, there will be M equations of this form. For a crystallographic model that is almost correct, the individual parameters will differ from their true values by an certain amount,

δp_{j}. Such a function at a position $(x + \Delta x)$ from its true value x can be described by a Taylor series. [52] The assumption made at this point is that as the difference Δx tends to zero, higher terms in the Taylor series may be ignored. However, if the initial values for the crystallographic model are poor, this assumption no longer holds and the result is invalid. Therefore, a crystallographic refinement requires an iterative process in order to ensure convergence on to the *best* solution. [53] Application of the Taylor series to the function y(calc) gives

$$y_i(\mathrm{calc}) = y_i(p_1 \ldots p_M) + \sum_{\mathrm{k}=1}^{M} \delta p_{\mathrm{k}} \left(\frac{\partial y_i}{\partial p_{\mathrm{k}}} \right)_{\mathrm{j}} \tag{1.57}$$

Which can then be substituted into Eqn. 1.56 and rearranged to give

$$\sum_{i=1}^{N} \sum_{\mathrm{k}=1}^{M} w_i \delta p_{\mathrm{k}} \left(\frac{\partial y_i}{\partial p_{\mathrm{k}}} \right)_{\mathrm{j}} \left(\frac{\partial y_i}{\partial p_{\mathrm{j}}} \right)_{\mathrm{k}} = \sum_{i=1}^{N} \left[w_i \left(y_i(\mathrm{obs}) - y_i(p_1 \ldots p_M) \right) \left(\frac{\partial y_i}{\partial p_{\mathrm{j}}} \right)_{\mathrm{k}} \right] \tag{1.58}$$

These are referred to as the *normal equations* since they are now of a linear form. We can therefore express Eqn.1.58 schematically in matrix form for simplicity,

$$\begin{pmatrix} \quad \mathbf{J^T} \quad \end{pmatrix} \mathbf{W} \begin{pmatrix} \mathbf{J} \end{pmatrix} \begin{pmatrix} \delta p_{\mathbf{j}} \end{pmatrix} = \begin{pmatrix} \quad \mathbf{J^T} \quad \end{pmatrix} \mathbf{W} \begin{pmatrix} \Delta \mathbf{y} \end{pmatrix}, \tag{1.59}$$

where **J** is the matrix of partial derivatives of size $N \times M$, **W** is a diagonal matrix containing the weights, δp_{j} is the vector of shifts (an $M \times 1$ column matrix) and $\Delta \boldsymbol{y}$ is the vector of observations (also an $M \times 1$ column matrix). This can be further simplified to

$$\begin{pmatrix} \quad \mathbf{A} \quad \end{pmatrix} \begin{pmatrix} \delta p_{\mathbf{j}} \end{pmatrix} = \begin{pmatrix} \mathbf{V} \end{pmatrix}, \tag{1.60}$$

where **A** is an $M \times M$ square matrix and δp_{j} and **V** are $M \times 1$ column matrices. The nature of each of these matrices is described below;

1. The $\delta\ p_{\mathrm{j}}$ matrix contains the calculated shift values to be added to each parameter during the iterative least-squares refinement cycle and is determined by inverting the **A** matrix such that $\delta p_{\mathrm{j}} = \mathbf{A}^{-1}\mathbf{V}$.

2. The **A** matrix elements are the sum of the products of the partial derivatives;

$$\mathbf{A}(\mathrm{j,k}) = \sum_{i=1}^{N} w_i \left(\frac{\partial y_i}{\partial p_{\mathrm{k}}}\right)_{\mathrm{j}} \left(\frac{\partial y_i}{\partial p_{\mathrm{j}}}\right)_{\mathrm{k}} \tag{1.61}$$

3. The **V** matrix contains all of the information pertaining to the observed data;

$$\mathbf{V}(\mathrm{j}) = \sum_{i=1}^{N} \left[w_i \left(y_i(\mathrm{obs}) - y_i(p_1 \ldots p_M)\right) \left(\frac{\partial y_i}{\partial p_{\mathrm{j}}}\right)_{\mathrm{k}} \right] \tag{1.62}$$

The result of this is that if the crystallographic model is poor but the data are good, then there will be problems when inverting the **A** matrix. Conversely if the model is good but the data are poor then, although the **A** matrix will invert, the parameters in the δp_{j} matrix will be poorly defined and the refinement is unlikely to work.

1.3 Structure Solution of 'Large' Small Molecules

As strategies are developed for assembling ever larger molecules, [54] the X-ray structure analysis of these materials becomes more challenging with problems that bear a resemblance to those encountered in macromolecular crystallography. Small molecule crystallography relies on data collected to a much higher resolution than is typical for proteins, as all structures are solved *ab initio* and more precise results are required. Indeed, for direct methods to succeed, 1.1-1.3 Å resolution data are typically required for structure solution [55] and the IUCr recommend a minimum of 98.5% completeness to 0.83 Å for refinement. [56] Examples of 'large' small molecules that have encountered great difficulties reaching these standards include conjugated porphyrin systems. [57, 58] The crystal structures of these highly conjugated π-systems are of particular interest due to their unique optical and electronic properties. [59] Porphyrins

can be found in many electrical devices including displays, biological imaging, [60] optical signalling [61] and photovoltaic materials; [62] however, the structure and property relationship between these systems is still not fully understood.

Crystals of such macrocyclic complexes often consist of greater than 50 % solvent by volume which is readily lost, giving rise to high mosaicity and leading to very poor, low resolution data. These problems are exacerbated by X-ray radiation damage during data collection, which often results in an incomplete dataset. Structure solution has been a persistent sticking point, with apparently reasonable datasets failing to lead to a solution, and demanding us to consider alternative approaches. One such approach is MAD phasing, [18, 63, 64] which exploits large changes in the magnitude of the anomalous scattering from an atom near its X-ray absorption edges (see Figure 1.7). *Ab initio* protein structure solution often employs a faster single-wavelength anomalous diffraction (SAD) experiment for obtaining initial model phases; however that experiment exploits Friedel differences and therefore requires non-centrosymmetric crystal structures. The following section gives an overview of MAD phasing for macromolecular crystallography and its application to centrosymmetric structures in small molecule crystallography .

1.3.1 Multi-Wavelength Anomalous Dispersion

Despite being the most common method of macromolecular structure solution, isomorphous replacement (IR) is not without problems. These include: non-isomorphism between crystals (unit-cell changes, reorientation of the protein, conformational changes, changes in salt and solvent ions); problems in locating all the heavy atoms; problems in refining heavy-atom positions; occupancies and thermal parameters and errors in intensity measurements. [65] Although multi-wavelength anomalous diffraction (MAD) is still prone to some of these problems, it overcomes the major problem of non-isomorphism by collecting all data on the same crystal. Instead of collecting data on a number of different, similar structures at a common wavelength (as in IR),

MAD phasing involves collecting data at a number of different wavelengths, exploiting the changes in anomalous signal from heavy atoms in the structure. Section 1.1.3 showed the wavelength dependence of anomalous scattering and Section 1.1.4 described how this quantity could be measured. This is the basis of MAD phasing and it can therefore be thought of as an *in situ* IR technique [18] *i.e.* instead of a difference between native and heavy atom derivatives it uses the differences between the below- and above-edge scattering which are entirely due to the anomalous scatterers.

1.3.2 Phase Equations for MAD

To evaluate the effect of anomalous scattering on the measured diffraction intensities, we substitute the anomalous scattering-factor expression (Eqn. 1.26) into the structure factor equation (Eqn.1.24),

$$ {}^{\lambda}F_{\text{obs}}(\boldsymbol{h}) = \sum_{i=1}^{N_{\text{atoms}}} (f^0 + f'_\lambda + i f''_\lambda)_i \exp\left[-B_i s^2(\boldsymbol{h})\right] \exp(2\pi i \boldsymbol{h}\cdot\boldsymbol{x}_i). \tag{1.63} $$

where λ denotes the wavelength of the data collection, $s(\boldsymbol{h}) = {}^{\sin\theta}/_{\lambda}$ and B_i is now included to account for the thermal parameter of the i^{th} atom, as described in Eqn. 1.12. Eqn.1.63 can then be split into its wavelength-dependent and independent parts. [66, 67] ${}^0F_T$ (with phase φ_T) represents the normal, wavelength-independent scattering from all the atoms (T), [17] and ${}^{\lambda}F'_A$ and ${}^{\lambda}F''_A$ represent the real and imaginary contributions respectively to the wavelength-dependent anomalous scattering from all of the atoms (A),

$$ {}^0F_T(\boldsymbol{h}) = \sum_{i=1}^{N_{\text{atoms}}} f_i^0 \exp\left[-B_i s^2(\boldsymbol{h})\right] \exp(2\pi i \boldsymbol{h}\cdot\boldsymbol{x}_i), \tag{1.64} $$

$$ {}^{\lambda}F'_A(\boldsymbol{h}) = \sum_{j=1}^{N_{\text{anom}}} f'_{\lambda j} \exp\left[-B_i s^2(\boldsymbol{h})\right] \exp(2\pi i \boldsymbol{h}\cdot\boldsymbol{x}_i), \tag{1.65} $$

$$ {}^{\lambda}F''_A(\boldsymbol{h}) = \sum_{j=1}^{N_{\text{anom}}} f''_{\lambda j} \exp\left[-B_i s^2(\boldsymbol{h})\right] \exp(2\pi i \boldsymbol{h}\cdot\boldsymbol{x}_i). \tag{1.66} $$

where N_{atoms} is the total number of atoms in the structure and N_{anom}, the num-

ber of anomalous scatterers. The anomalous contributions to the scattering factor (Eqns. 1.65 and 1.66) can be expressed in terms of the normal structure factors, ${}^0F_{A_k}$, for each different kind of anomalous scatterer in the structure, k, [68]

$$ {}^{\lambda}F'_A = \sum_{k=1}^{N_{\text{kinds}}} (f'_\lambda/f^0)_k \, {}^0F_{A_k}, \tag{1.67} $$

$$ {}^{\lambda}F''_A = \sum_{k=1}^{N_{\text{kinds}}} (if''_\lambda/f^0)_k \, {}^0F_{A_k}. \tag{1.68} $$

This allows the wavelength-dependent structure factor, ${}^{\lambda}F_{\text{obs}}$ to be expressed using only the normal structure factors

$$ {}^{\lambda}F_{\text{obs}} = {}^0F_T + \sum_{k=1}^{N_{\text{kinds}}} \left(\frac{f'_\lambda}{f^0} + i\frac{f''_\lambda}{f^0} \right)_k {}^0F_{A_k}. \tag{1.69} $$

This is useful since now all of the wavelength dependence is confined to only the anomalous scattering factors *i.e.* it is independent of the atomic positions, occupancies and thermal parameters. When analysing the structure factors, they are treated separately for each kind of anomalous scatterer, k. The most common case however only contains a single type of anomalous scatterer so we can ignore k from future notation. To remove the complex part of Eqn. 1.69, cosine laws are applied to ${}^0F_T$, ${}^{\lambda}F'_A$ and ${}^{\lambda}F''_A$. [17, 66] The magnitude of the observed structure factor at a given wavelength can be expressed as

$$
\begin{aligned}
|^{\lambda} F_{\text{obs}}(\pm\boldsymbol{h})|^2 &= |^0 F_T(\pm\boldsymbol{h})|^2 \\
&+ a_\lambda |^0 F_A(\pm\boldsymbol{h})|^2 \\
&+ b_\lambda |^0 F_T(\pm\boldsymbol{h})||^0 F_A(\pm\boldsymbol{h})| \cos\left[{}^0\varphi_T(\boldsymbol{h}) - {}^0\varphi_A(\boldsymbol{h})\right] \\
&+ c_\lambda |^0 F_T(\pm\boldsymbol{h})||^0 F_A(\pm\boldsymbol{h})| \sin\left[{}^0\varphi_T(\boldsymbol{h}) - {}^0\varphi_A(\boldsymbol{h})\right], \qquad (1.70)
\end{aligned}
$$

$a_\lambda = ({f'}^2 + {f''}^2)/{}^0f^2$,

$b_\lambda = 2(f'/{}^0f)$,

$c_\lambda = 2(f''/{}^0f)$,

where $|\,{}^0F_T(\pm\boldsymbol{h})\,|$ is used to represent the Friedel pairs +$\boldsymbol{h}$ and -$\boldsymbol{h}$. However, at wavelengths where the f'' has a significant contribution, Friedel's Law no longer holds and so the Friedel pairs are referred to as Bijovet mates (this is only the case for non-centrosymmetric space groups). An observation from Eqn. 1.70 is that the contributions from the real and imaginary part of the scattering are orthogonal to one another. We derive the 'dispersive' phase information from the differences in $|F_{\text{obs}}|$ at wavelengths which have different values of f' (contributing to the $\cos\left[{}^0\varphi_T(\boldsymbol{h}) - {}^0\varphi_A(\boldsymbol{h})\right]$ term), and 'Bijovet' phase information from the structure factor difference in Friedel pairs at a wavelength with a large f'' values (contributing to the $\sin\left[{}^0\varphi_T(\boldsymbol{h}) - {}^0\varphi_A(\boldsymbol{h})\right]$ term). [68] In turn, this means that by careful selection of wavelength, so as to maximise differences in the f' and f'' signals, we can improve the phase information that can be obtained for this method (see Section 1.3.3).

1.3.3 Optimal Choice of Wavelength

In order to obtain the maximum amount of information for MAD phasing from the experiment, the aim is to select wavelengths at which differences in f' and f'' are greatest. This could be obtained using values calculated from first principles, [14, 69] however these only hold true for isolated elemental atoms at wavelengths far removed from resonant energies. At energies near to the absorption edge these values de-

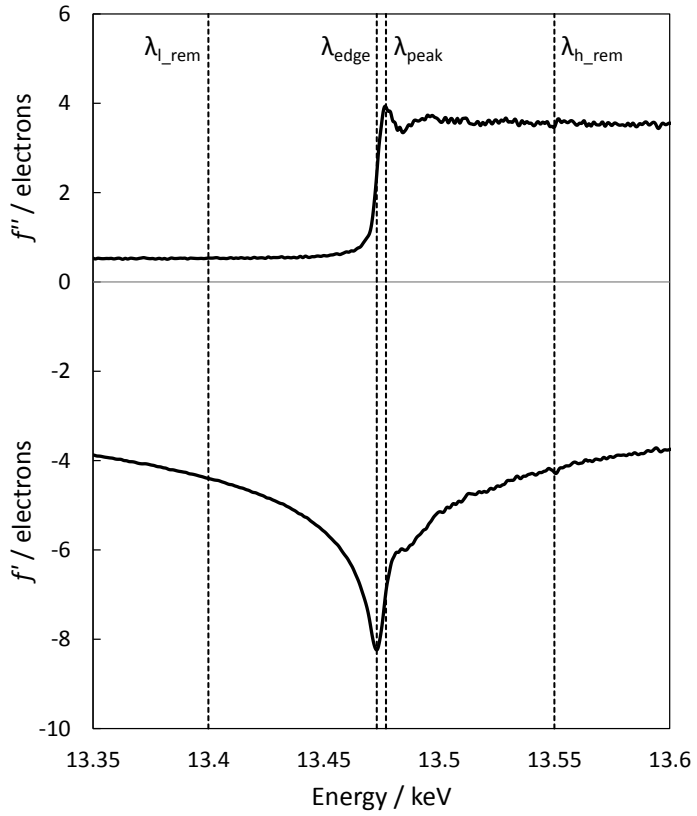


Figure 1.12 Absorption-fluorescence spectrum showing wavelength selections for the MAD phasing experiment.

viate from the theoretical and since this is the most important region for the MAD experiment (in order to maximise f' and f''), experimental values must be used. Measurement of these using absorption-fluorescence experiments are described in Section 1.1.4.

As shown in Figure 1.12, four wavelengths ($\lambda_{\text{l\_rem}}$, λ_{edge}, λ_{peak}, and $\lambda_{\text{h\_rem}}$) are specifically chosen using the absorption-fluorescence spectrum in order to maximise the anomalous signal. λ_{peak} is collected in order to optimise the Bijovet signal since f'' is at its greatest value here. The dispersive signal is maximised by collecting at the minimum f' value, λ_{edge}, which coincides with the inflection point of the f'' curve. The remaining data collections at the low and high remote ($\lambda_{\text{l\_rem}}$ and $\lambda_{\text{h\_rem}}$ respectively) are chosen at wavelengths far enough removed from the absorption edge to avoid complications with the pre- and post-edge features.

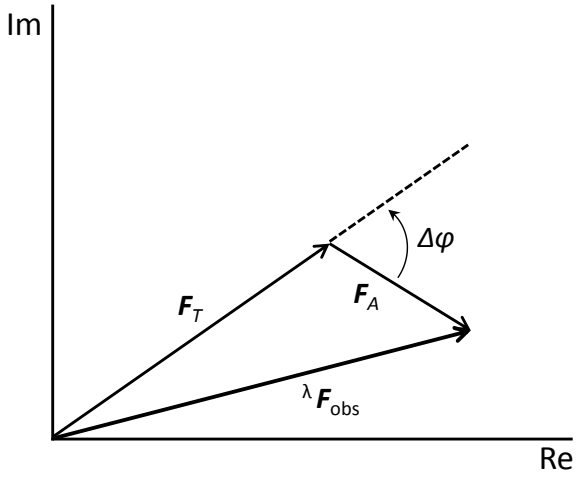


Figure 1.13 Argand diagram to represent the vector relationship between $\mathbf{F}_T$, $\mathbf{F}_A$, and $\Delta\varphi$.

1.3.4 Solution of the Phase Equations

In order to 'solve' the MAD phasing equations in Section 1.3.2, the values of F_T, F_A, and $\Delta\varphi$ must be determined for each observed structure factor, $|^{\lambda} F_{\text{obs}}(\pm\boldsymbol{h})|^2$. These values are represented on an Argand diagram in Figure 1.13. From this representation, the phases of F_T can be estimated to be $\Delta\varphi + \varphi_A$.

For clarity, we can re-write Eqn. 1.70 so that the unknowns match those shown in Figure 1.13,

$$^{\lambda}F_{\text{obs}}(\boldsymbol{h}) = F_T^2 + a_\lambda F_A^2 + b_\lambda F_T F_A \cos\Delta\varphi + c_\lambda F_T F_A \sin\Delta\varphi \tag{1.71}$$

Since Eqn. 1.71 is wavelength dependent, every measured value of $^{\lambda}F_{\text{obs}}(\boldsymbol{h})$ provides an observation to help solve the equation. Therefore if a structure factor for a given value of $\boldsymbol{h}$ is measure at four separate wavelengths, a set of simultaneous equations can be set up to solve for the three unknowns (F_T, F_A, and $\Delta\varphi$). In the case of a non-centrosymmetric molecule this case if further simplified given each observation at a particular wavelength provides two instances of Eqn 1.71 since $|F_{\boldsymbol{h}}| \neq |F_{-\boldsymbol{h}}|$. Therefore, by collecting data at four wavelengths we are able to over-determine the problem (when three wavelengths would be sufficient).

Since these equations are non-linear, a least-squares refinement can be imple-

mented. The problem however can be simplified (as in the MADSQL program [17]) by expressing it as a set linear equations with four unknowns [70]

$$F_{\text{obs}}(\boldsymbol{h}) = P_1 + a_\lambda P_2 + b_\lambda P_3 + c_\lambda P_4. \tag{1.72}$$

A further Lagrangian constraint [71] can be applied to Eqn. 1.72, which enforces the identity $\sin^2\theta + \cos^2\theta = 1$ such that

$$(P_1 \times P_2) - (P_3 \times P_3) - (P_4 \times P_4) = 0 \tag{1.73}$$

Substitution of Eqn. 1.72 into Eqn. 1.71 gives the result

$$\begin{aligned} F_T &= \sqrt{P_1} \\ F_A &= \sqrt{P_2} \\ \Delta\varphi &= \arctan {}^{P_4}/_{P_3} \end{aligned} \tag{1.74}$$

A more in depth description of this derivation is reported by Hendrickson *et al.* [17] In the case of structures with more than one anomalous scatterer, the problem is solved as above but with the addition of two extra unknowns for each new scatterer, k: F_{A_k} and $\Delta\varphi_k$.

1.3.5 Determination of Heavy Atom Substructure

As shown previously in Figure 1.13, $\varphi_T = \varphi_A + \Delta\varphi$. Therefore, after values of F_T, F_A and $\Delta\varphi$ have been obtained, all that remains is to calculate φ_A and the structure can be solved. In order to calculate these values, the positions of the heavy atoms (anomalous scatterers) in the structure must be determined. The heavy atom substructure can then be used to generate the phases required.

Section 1.2 described a number of techniques used in small-molecule crystallography for structure solution. The most applicable of these to the problem at hand is the Patterson synthesis since we are trying to locate a relatively small number of

heavy atoms in a structure comprised primarily of light atoms. Since data have been collected at four separate wavelengths, the simplest Patterson map to calculate would be a Bijovet Difference Patterson, using $|\Delta F|^2$ coefficients. The ideal wavelengths to use for such a map would be those that maximised the difference in anomalous signal between Bijovet mates *i.e.* λ_{peak} and $\lambda_{\text{l\_rem}}$ (see Figure 1.12). The Bijovet Difference Patterson ($w = 1/2$ Harker Section) for a simple urea salt (spacegroup $P2_1/c$) containing one bromine atom per asymmetric unit is shown in Figure 1.14

However, even in this simple case where the heavy atom sub-structure could be solved, the Patterson map is very noisy. Figure 1.14 shows large, poorly defined areas of both high positive and negative Patterson density. A major improvement can be made by using the calculated values of F_A instead, with Patterson map coefficients of $|F_A|^2$. Now, data from all four wavelengths are incorporated into the map and the result is a much clearer Patterson map, as shown in Figure 1.15. The background noise in the Patterson map has been greatly reduced and the two sharp peaks in the center of the map (due to inter-atomic vectors between bromide ions) are much better defined than in Figure 1.14.

As stated previously, this method is most effective when the number of anomalous scatterers is relatively small. In cases where there are over 20 anomalous atoms, direct methods and dual-space techniques [72] have proven successful for determining the anomalous substructure.

1.3.6 Maximum-Likelihood and Bayesian Estimates

The least-squares approach to solving the phase equations (described in subsection-1.3.4) does not always provide the most accurate solution to the problem. It has been shown that often values of F_A can be over-estimated, especially in cases where the data contain large experimental errors. [73] Normally these values must be removed before continuing with the solution of the heavy-atom substructure. [74] Alternatively, a Bayesian approach can be implemented to estimate values of F_A. [75, 76] The tech-

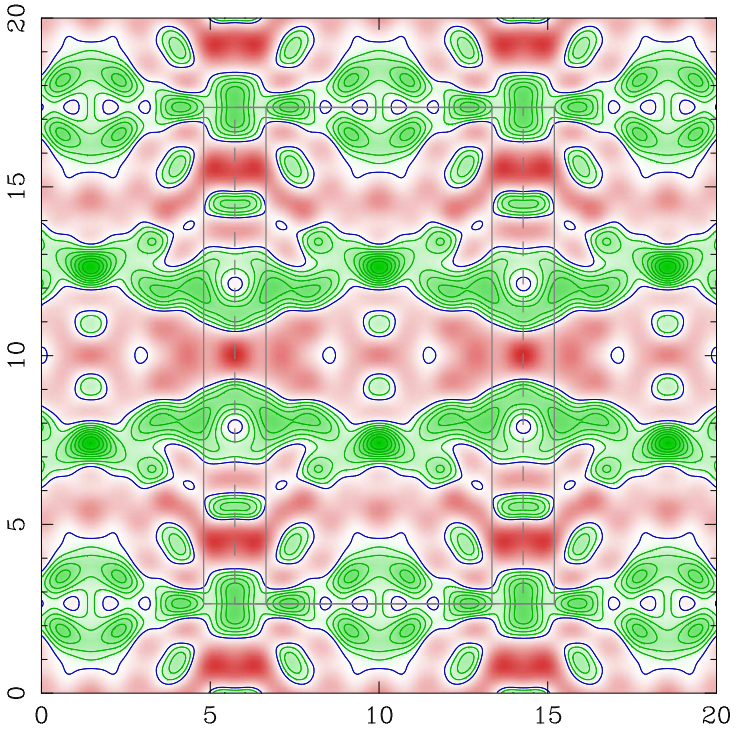


Figure 1.14 Bijovet Difference Patterson for a Bromo-Urea Salt, $w = 1/2$ Harker Section, parallel to the *ab* plane. Green indicates positive electron density and the dotted grey line is the intersection of the unit cell with the *ab* plane. Further map details can be found in Figure 1.9.

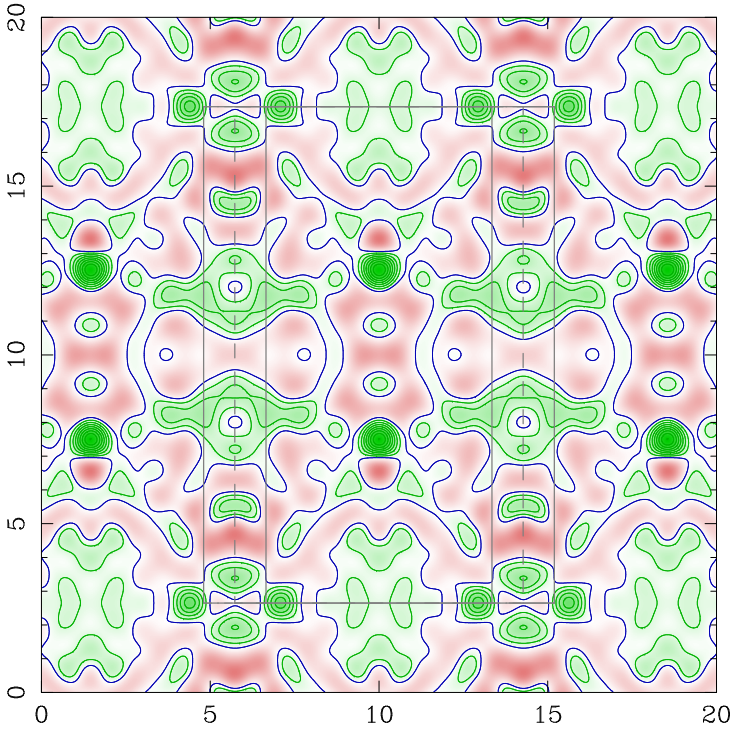


Figure 1.15 Anomalous Difference Patterson map for a Bromo-Urea Salt, $w = 1/2$ Harker Section, parallel to the *ab* planes. Map details can be found in Figure 1.9.

nique combines the data collected in the MAD experiment with information about the number and element type of the anomalous scatterers in the asymmetric unit to calculate a probability distribution of F_A values. The most likely value is then selected using an *a priori* weighting scheme. The values to be determined are defined slightly differently than those used by Hendrickson [18] (see Section 1.3.2) such that they are all independent from one another;

- F_A – structure factor magnitude corresponding to *only* the anomalous scatterers
- F_O – structure factor magnitude corresponding to *all other* atoms in the unit cell
- $\Delta\varphi$ – phase difference between F_A and F_O

This technique uses Bayesian statistics to estimate the relative probability, $P(F_A, F_O, \Delta\varphi)$, that a set of unknowns is correct, given some prior knowledge of the distribution of their values. Mathematically, Bayes' theorem can be stated as

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}. \tag{1.75}$$

In our case, B represents the measured experimental data, and A, the unknown quantities, including error estimates. A more in-depth analysis of Bayes' Theorem can be found in specialised texts. [77, 78] The best estimate of a given quantity is then calculated as the weighted average over all values of the three unknowns. Therefore, in order to carry out a Bayesian analysis, we must first generate two probability distributions – an *a priori* distribution for values of F_A and a probability distribution of the observed data given a set of values of F_A, F_O and $\Delta\varphi$. By combining these two distributions using Eqn. 1.75, we can estimate the probability of any combination of the three unknowns being correct.

***A Priori* Probability Distribution of F_A**

To overcome the problems of overestimating F_A associated with the least-squares approach, the probability distribution limits the range of possible values. This is achieved by using prior knowledge about the type and number of anomalous scatterers in the structure to produce the *a priori* probability distribution of possible scattering factors. The distribution also requires approximate knowledge of thermal factors for the scatterers and for the measured structure factors to be put on an absolute scale. With this information, the *a priori* probability distribution (for an acentric reflection) is given as

$$P_O(F_A) \propto F_A \exp - \left(\frac{F_A^2}{\Sigma^2} \right), \tag{1.76}$$

where Σ^2 is the expected mean square value of F_A within a certain resolution shell. It can be calculated directly from knowledge of the number and type of anomalous scatterers and their associated thermal parameters. [79] In most structures the number of anomalous scatterers is relatively small and thermal parameters are not well defined. Since Σ^2 will therefore have little effect on the overall probability distribution, the thermal parameters can be assumed to be the same as the average of all other atoms in the unit cell.

For the purposes of this analysis, we assume the *a priori* probabilities for F_O and $\Delta\varphi$ to be constant *i.e.* before a measurement has been made, all values of these unknowns are predicted to be equally likely. F_O is well defined by the experimental measurement (since the anomalous signal rarely accounts for more than 5-10 % of the overall structure factor). In macromolecular crystallography all values of $\Delta\varphi$ are equally likely since no prior knowledge of the arrangement of atoms in the unit cell is known. However, this case is simplified for many small molecules since the majority are centrosymmetric and therefore the possible phase angles are restricted to either 0 or 180°.

Probability Distribution of F_{obs} at a given wavelength and values of F_A, F_O and $\Delta\varphi$

Before calculating the probability distribution of F_{obs}, we can once more place conditions on it by using prior knowledge. We know that errors in Bijovet pairs are highly correlated whereas those in the anomalous difference are likely to be much less correlated. Therefore it is sensible to use average structure factor amplitudes, $\bar{F}$ and anomalous differences, Δ_{Anom}, where,

$$\bar{F} = 1/2(F^+ + F^-), \tag{1.77}$$

$$\Delta_{\text{Anom}} = (F^+ - F^-). \tag{1.78}$$

This expression of terms also allows for experimental errors in the two differences to be treated separately *i.e.* $\sigma_{\bar{F}}$ and $\sigma_{\Delta_{\text{Anom}}}$. [80] Additional errors in the two quantities are termed $E_{\bar{F}}$ and $E_{\Delta_{\text{Anom}}}$, and are discussed later. Given: values for the unknowns (F_A, F_O,$\Delta\varphi$); estimates of the errors on F_{obs}; and working with the assumption that those errors have a Gaussian distribution, the probability distribution for F_{obs} can be expressed as

$$P(\bar{F}, \Delta_{\text{Anom}} \mid F_A, F_O, \Delta\varphi) \propto \exp -\frac{1}{2}\left(\frac{\epsilon^2_{\bar{F}}}{E^2_{\bar{F}} + \sigma^2_{\bar{F}}} + \frac{\epsilon^2_{\text{Anom}}}{E^2_{\text{Anom}} + \sigma^2_{\text{Anom}}}\right), \tag{1.79}$$

where σ = experimental uncertainty, E = additional uncertainties and ϵ = difference between the observed and calculated values of $\bar{F}$ and Δ_{Anom}.

Probability Distribution of F_A, F_O and $\Delta\varphi$ given observed F and Δ_{Anom} at various wavelengths

Now that $\bar{F}$ and Δ_{Anom} have been measured at different wavelengths defined by the experiment, probability distributions for F_A, F_O and $\Delta\varphi$ can be calculated by applying Bayes' rule. This is done by taking the product of the probability distributions for F_A (Eqn. 1.76) and $\bar{F}$ & Δ_{Anom} (Eqn. 1.79) at each wavelength used ($\lambda = \lambda_1 ... \lambda_N$), to give

$$P\left(F_A, F_O, \Delta\varphi \mid \{\bar{F}, \Delta_{\text{Anom}}\}_{\lambda_1 \ldots \lambda_N}\right) \propto P_O(F_A) \prod_{\lambda_1 \ldots \lambda_N} P(\bar{F}, \Delta_{\text{Anom}} \mid F_A, F_O, \Delta\varphi). \tag{1.80}$$

The result of Eqn. 1.80 gives the relative likelihood that a given set of values for F_A, F_O and $\Delta\varphi$ are correct.

Estimation of F_A, F_A^2, F_O and $\Delta\varphi$

Eqn 1.80 can now be used to estimate the value for any of the unknowns, x, by averaging over all possible values of F_A, F_O and $\Delta\varphi$, and weighting the average by the likelihood that the set of values is correct,

$$\langle x \rangle = \frac{\int x P\left(F_A, F_O, \Delta\varphi \mid \{\bar{F}, \Delta_{\text{Anom}}\}_{\lambda_1 \ldots \lambda_N}\right) \mathrm{d}F_A \mathrm{d}F_O \mathrm{d}\Delta\varphi}{\int P\left(F_A, F_O, \Delta\varphi \mid \{\bar{F}, \Delta_{\text{Anom}}\}_{\lambda_1 \ldots \lambda_N}\right) \mathrm{d}F_A \mathrm{d}F_O \mathrm{d}\Delta\varphi}. \tag{1.81}$$

Since integration over three variables can be quite time consuming, Terwilleger [75] suggests that F_O is well defined by the MAD experimental data and so integration over this variable has a very small effect on the result. Therefore the most probable value of F_O is used as a constant in Eqn. 1.81.

Estimation of the non-experimental errors

Non-experimental errors include systematic errors such as those caused by poor data scaling and absorption problems, which are difficult to account for. However, these errors are assumed to be roughly the same for all reflections within a certain resolution range. The errors are therefore estimated by taking the weighted-average of the squared differences between the calculated and observed values of $\bar{F}$ and Δ_{Anom} ($\epsilon_{\bar{F}}$ and $\epsilon_{\Delta_{\text{Anom}}}$ respectively) as an estimate of the total mean-square error for each variable. When calculated for a particular measurement from Eqn. 1.81, $\langle \epsilon^2 \rangle$, the mean non-experimental errors can be estimated as the difference between $\langle \epsilon^2 \rangle$ and σ^2. For the anomalous difference, this can be written as

$$E^2_{\text{Anom}} = {}^1/_N \sum \left\{ \langle \epsilon^2_{\text{Anom}} \rangle - \sigma^2_{\text{Anom}} \right\} \tag{1.82}$$

Bayesian Patterson Map

Using Eqn. 1.81 we can obtain values of $|F_A|$ *via* a Bayesian treatment of the MAD phase equations, allowing a new Patterson map to be calculated using $|F_A|^2$ coefficients. The result shown in Figure 1.16 clearly shows a much improved Patterson map with very little background noise and well defined peaks (see Figures 1.14 and 1.15 for comparison).

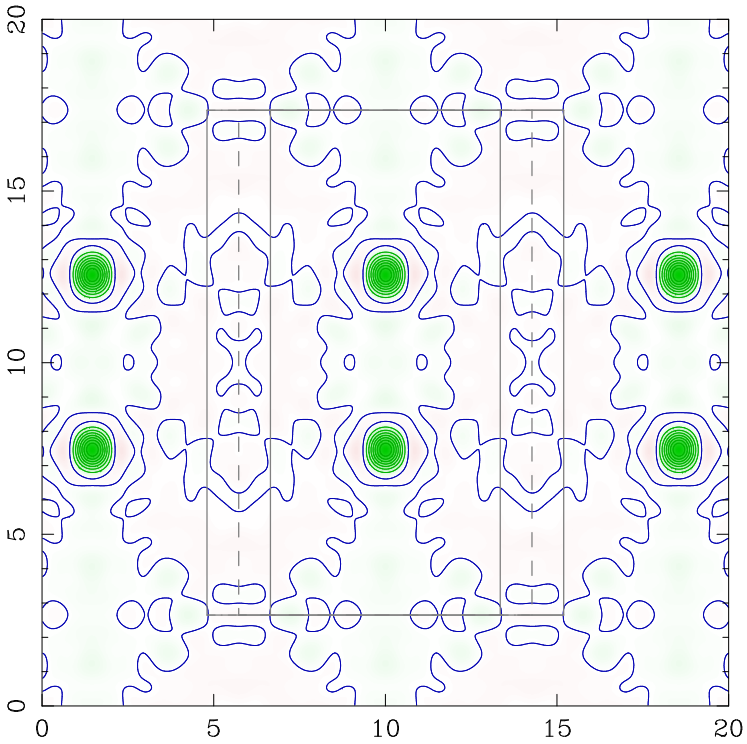


Figure 1.16 Bayesian-Anomalous Patterson map for a Bromo-Urea Salt, $w = {}^1/_2$ Harker Section, parallel to the *ab* plane. Map details can be found in Figure 1.9

1.3.7 Complete Structure Solution

From the Bayesian $|F_A|^2$ Patterson map the positions of the anomalous scatterers can be identified. Using the anomalous substructure, calculated phases for the structure, φ^{calc}_A, are readily generated in the CRYSTALS software. [3] We can then estimate the

desired set of phases for the experimental structure factors as

$$\varphi_T = \Delta\varphi^{\text{obs}} + \varphi_A^{\text{calc}}. \tag{1.83}$$

These phases are applied to the low remote data set in order to reduce errors due to absorption effects above and around the absorption edge. A Fourier transform of these experimental phases is then applied to produce an electron density map for all of the atoms in the structure. Experimental results from this step are discussed in Chapter 5.

1.3.8 Current Software

A wide variety of programs are currently available for processing data and solving macromolecular structures. Many of these programs are available through the CCP4 (Collaborative Computational Project, Number 4) software suite, which brings together programs, associated data and software libraries for the purpose of macromolecular structure determination. Descriptions of such programs can be found in the accompanying literature of the review article by Winn *et al.* [81] The software suite also provides a number of tools for converting files into compatible formats for use across programs. For example a typical phasing calculation might use the following sequence of programs availiable through the CCP4 suite:

> SCALEPACK2MTZ - takes scaled data from SCALEPACK [82] and converts the files into the .mtz format (the MTZ file format is used by CCP4 programs to store reflection data and can contain data from any number of different datasets). [83]

> TRUNCATE - converts the measured intensities into structure factors. [84]

> CAD - collects all reflection data from up to 9 separate .mtz files and merges them into a single file. [83]

> SCALEIT - scales all of the datasets to a chosen 'native' dataset. The program contains options to scale the data using (a) an overall scale factor, (b) a scale

and isotropic temperature factor or (c) a scale and anisotropic temperature factor. [83]

> MLPHARE - used to refine heavy atom parameters using the maximum likelihood method described in Section 1.3.6. These are then used to generate phase information for the structure solution. [85]

Alternatively, SHELX contains a number of programs for both small molecule and macromolecular crystal structure determination. The three programs, SHELXC, SHELXD, and SHELXE are used for macromolecular phasing. [86] Finally, there is the PHENIX (python-based hierarchical environment for integrated crystallography) software suite for the automated determination of molecular structures. This software aims to develop algorithms to automate procedures that are traditionally performed by hand in order to minimise subjective input.[87, 88]

1.4 Discrimination of Elements and their Electronic States

1.4.1 Mixed Oxidation Compounds

Oxidation state is defined by the IUPAC Gold Book [89] as:

'A measure of the degree of oxidation of an atom in a substance. It is defined as the charge an atom might be imagined to have when electrons are counted according to an agreed-upon set of rules:

1. the oxidation state of a free element (uncombined element) is zero;

2. for a simple (monatomic) ion, the oxidation state is equal to the net charge on the ion;

3. hydrogen has an oxidation state of +1 and oxygen has an oxidation state of −2 when they are present in most compounds. (Exceptions to this are that hydrogen has an oxidation state of −1 in hydrides of active metals, e.g. LiH, and oxygen has an oxidation state of −1 in peroxides, e.g. H_2O_2;

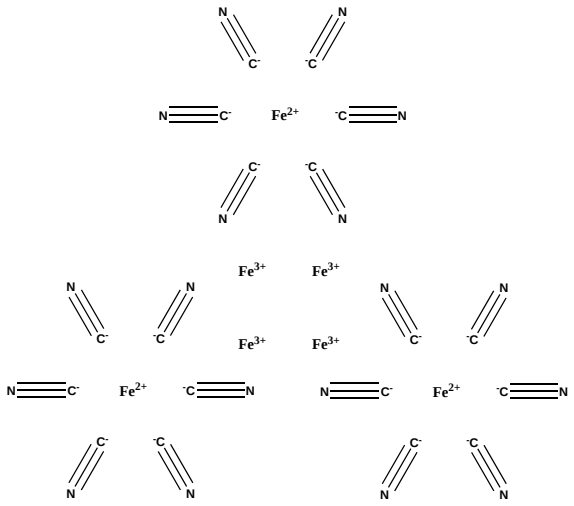


Figure 1.17 Representation of the chemical structure of Prussian Blue.

4. the algebraic sum of oxidation states of all atoms in a neutral molecule must be zero, while in ions the algebraic sum of the oxidation states of the constituent atoms must be equal to the charge on the ion. For example, the oxidation states of sulphur in H_2S, S_8 (elementary sulphur), SO_2, SO_3, and H_2SO_4 are, respectively: −2, 0, +4, +6 and +6. The higher the oxidation state of a given atom, the greater is its degree of oxidation; the lower the oxidation state, the greater is its degree of reduction.'

Compounds containing an element with more than one formal oxidation state in the structure (mixed oxidation/valence compounds) have been observed for well over a century. One of the first and most famous structures of this type was the discovery of Prussian Blue $Fe_4[Fe(CN)_6]_3 \cdot xH_2O$, a cyanide-bridged Fe(II)-Fe(III) material (Figure 1.17), in 1704. [90] They were originally of interest due to their distinct colours and unusual stoichiometry.

Classification

In 1976, two extensive reviews of mixed-valence compounds were undertaken by Allen & Hush [91] and Robin & Day, [92] and have since provided a system for classifying such compounds based on the degree of mixing between the different oxidation states

present. If a compound contains an element, E, existing in two oxidation states (A and B), and occupying two different sites within the structure (X and Y), then the extent of mixing between $E_X^A E_Y^B$ and $E_X^B E_Y^A$ will become apparent from how distinguishable the two sites are crystallographically. [93, 94]

- ***Class I*** - vastly different *e.g.* octahedral *vs.* tetrahedral sites. These compounds are said to have 'trapped' oxidation states.
- ***Class II*** - similar environments *e.g.* both octahedal but with slightly different bond lengths.
- ***Class III A*** - indistinguishable, finite clusters *e.g.* edge-sharing octahedral dimers. The charge in these compounds can be thought of as delocalised.
- ***Class III B*** - indistinguishable, continuous chains *e.g.* chain of edge-sharing octahedra.

1.4.2 Spin Crossover Compounds

Since initial reports by Cambi *et al.* [95] in 1931, spin crossover (SCO) materials have been of constant interest. They are so named due to their ability to switch between high-spin (HS) and low-spin (LS) molecular states, which can be described by ligand field theory. [96] The change in spin state is often brought about by external influences such as light, temperature, pressure and magnetic field, all of which have been well studied and reviewed. [97–99] In turn this means that SCO compounds can have applications in data storage, [100] switches [101, 102] and optical displays. [103]

The extent of spin-crossover can be detected using techniques such as X-ray crystallography, magnetic susceptibility, Mössbauer, Raman spectroscopy and optical spectroscopy (see section 1.4.4). Spin transition curves represent the fraction of high spin (γ_{HS}) *vs.* temperature, for which the most important in the solid state are shown in Figure 1.18. [104]

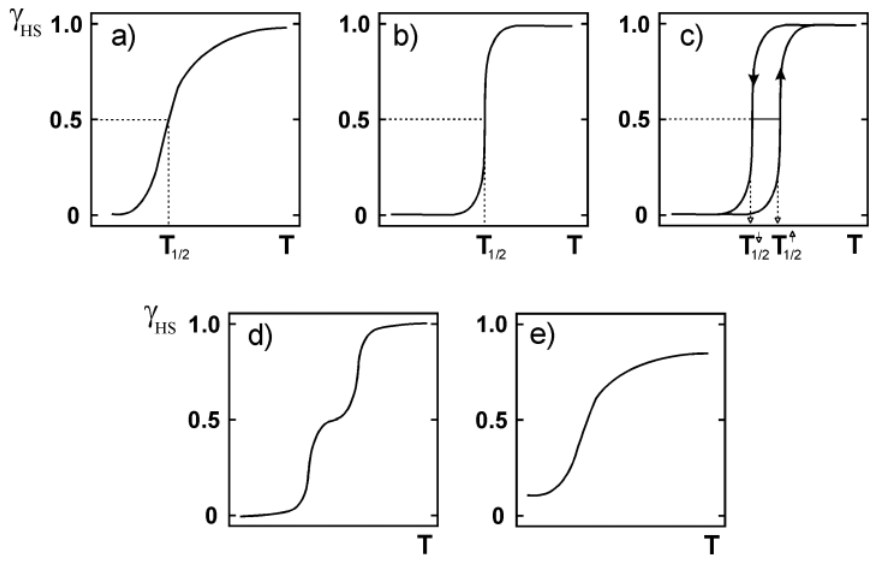


Figure 1.18 Representation of the principal types of spin transition curves: (**a**) gradual; (**b**) abrupt; (**c**) with hysteresis; (**d**) two-step; (**e**) incomplete.

The first of these is the gradual change (Figure 1.18**a**), which describes a Boltzman distribution of molecular states converting to HS slowly over a relatively large temperature range. This is the most common type of SCO material and arises from relatively weak interactions between metal centres. It follows that the second type of SCO, the abrupt case (Figure 1.18**b**), arises from strong cooperativity between centres. The third, and possibly most important type of SCO material in terms of real-life application and present research in the area, is spin hysteresis (Figure 1.18**c**). The demonstration of bistability in such compounds provides a mechanism for their use in thermal sensors, active elements in displays [105], information storage [106] and retrieval. [107] The origin of this hysteresis is often linked to either a crystallographic phase change in the lattice or a change in the intramolecular structure cooperatively causing perturbations to neighbouring atoms. Since the discovery of this hysteresis, a large number of systems have been reported showing similar behaviour – often most pronounced with Fe(II) systems. This has been achieved by designing molecules with covalent linkers between SCO centres, incorporation of H-bonding to allow direct interaction of anions with the SCO centres, and incorporation of aromatic ligands

to improve π-π interactions though stacking.

The two-step transition (Figure 1.18**d**), was first reported in 1981 and is much rarer than the rest. The origin of this is generally the presence of two crystallographically distinct SCO sites within the molecule, or binuclear complexes (even if the different centers have the same coordination). Finally, the incomplete SCO profile (Figure 1.18**e**) where some γ_{HS} is retained at low temperature could be an artefact of a number of factors. Firstly, a proportion of the SCO centres could be in a particular arrangement so as to cause a field strength significant enough to maintain some centres in a LS state. Alternatively, it could be a kinetic effect such that at low temperatures, however slow, HS↔LS conversion is still present. Therefore, upon rapid cooling a portion of HS centres can be trapped out – an effect which has been observed at liquid nitrogen temperatures.[108]

1.4.3 Near-Identical Element Discrimination

As described in Section 1.1.1, X-ray crystallography exploits the interaction of an electromagnetic beam with the electron 'cloud' around an atom. Therefore the scattering amplitude is strongly correlated to the atomic number (number of electrons in an atom). As a result X-rays are relatively insensitive to light atoms such as H and He, and for the same reason, very little contrast is provided between atoms with similar atomic numbers *e.g.* Fe (26 electrons) and Co (27 electrons).

The scattering intensity from an atom, j, is described by the atomic form factor

$$f_j(\boldsymbol{r}^*) = \int \rho_j(\boldsymbol{r}) \exp(2\pi i \boldsymbol{r}^* \cdot \boldsymbol{r}) \mathrm{d}\boldsymbol{r}, \tag{1.84}$$

where $\rho_j(\boldsymbol{r})$ is the spatial electron density of the scatterer (often assumed to be spherical) and $\boldsymbol{r}$ and $\boldsymbol{r}^*$ are vectors in real and reciprocal space, respectively. Most scattering by an atom is in the forward direction, and the scattering power decreases with increasing scattering angle. The scattering factor as a function of scattering angle can be fitted accurately by a 9 parameter equation: [109]

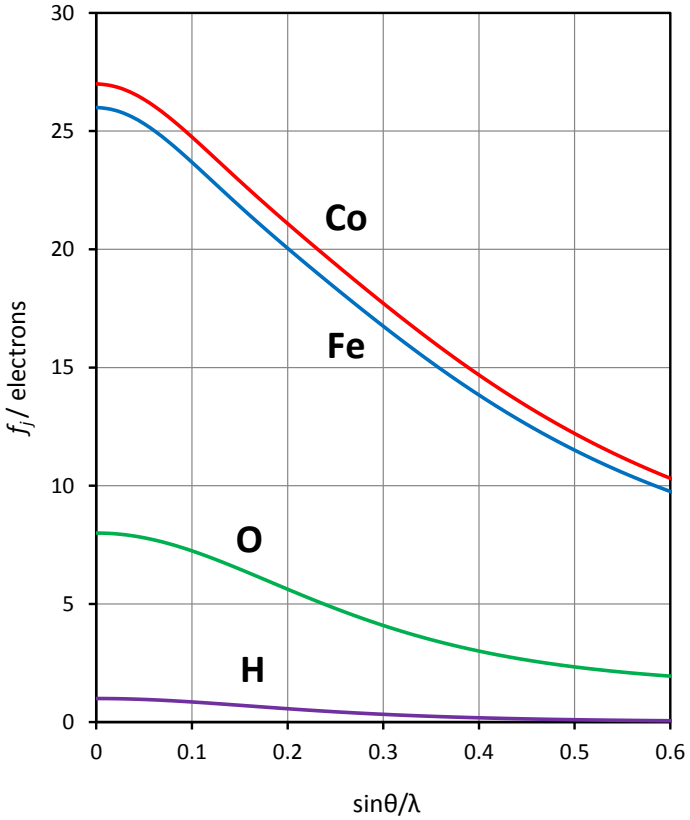


Figure 1.19 Graph to show the atomic scattering factors for Co, Fe, O and H *vs.* $\sin\theta/\lambda$. Values are calculated using Eqn. 1.85 and do not include a temperature factor.

$$f_j(\sin\theta/\lambda) = \sum_{i=1}^{i=4} a_i \cdot \exp\left[-b_i(\sin\theta/\lambda)^2\right] + c. \tag{1.85}$$

Knowledge of the 9 coefficients (a_i,b_i and c) for a particular atom allows us to calculate graphs for the scattering factor for a given atom as function of $\sin\theta/\lambda$, as shown in Figure 1.19.

One consequence of the Eqn. 1.85 is that at $\sin\theta/\lambda = 0$, the scattering factor value is equal to the number of electrons in the atom. However, this scattering only applies to the case of an isolated atom at rest. In reality, there is an associated temperature factor, (B), which has the effect of reducing the scattering power of an atom. The temperature factor is exponential and also an angle dependent variable, [110]

$$f_B = f_j \cdot \exp\left[-B(\sin\theta/\lambda)^2\right], \tag{1.86}$$

where B is related to the mean square displacement $\langle U \rangle$, as defined in Eqn. 1.12. If we convolute the temperature factor with the scattering factor in reciprocal space, the result are scattering factor curves which drop off faster with angle. This is because the angle dependence of the exponential has a greater effect on the higher angle reflections. Figure 1.20 shows the effect of the temperature factor on the atomic scattering factor. When a factor of $B = 5$ is applied to both elements, as expected, their scattering factors fall off faster with angle than in Figure 1.19. However, if the factors differ slightly and a temperature factor of $B = 5.5$ is applied to Co, the curves begin to overlap. Therefore, when trying to determine atom types when both are present in a compound, a competitive refinement would be unable to distinguish between the two.

Even in the case where both atoms have the same B factor, the problem of distinguishing atom types is not simple. The biggest difference and therefore best contrast between the curves comes at low values of $^{\sin\theta}/_{\lambda}$. However, in Eqn 1.87 we show that considering the (1,0,0) reflection in an orthorhombic system ($a \approx 5$ Å), the lowest observable value of $^{\sin\theta}/_{\lambda}$ is only 0.1 Å^{-1} *i.e.* the useful part of the curve for distinguishing Co and Fe is unobtainable.

For an orthorhombic spacegroup

$$\frac{1}{d^2} = \frac{h^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2}$$

for the (1,0,0) reflection

$$\frac{1}{d^2} = \frac{1}{a^2}$$

and given Bragg's Law

$$n\lambda = 2\mathrm{d}\sin\theta$$

then

$$\frac{\sin\theta}{\lambda} = \frac{\sqrt{\cancel{b^2}\cancel{c^2}}}{2a\cancel{b}\cancel{c}} = \frac{1}{2a} = 0.1 \text{ Å}. \qquad (1.87)$$

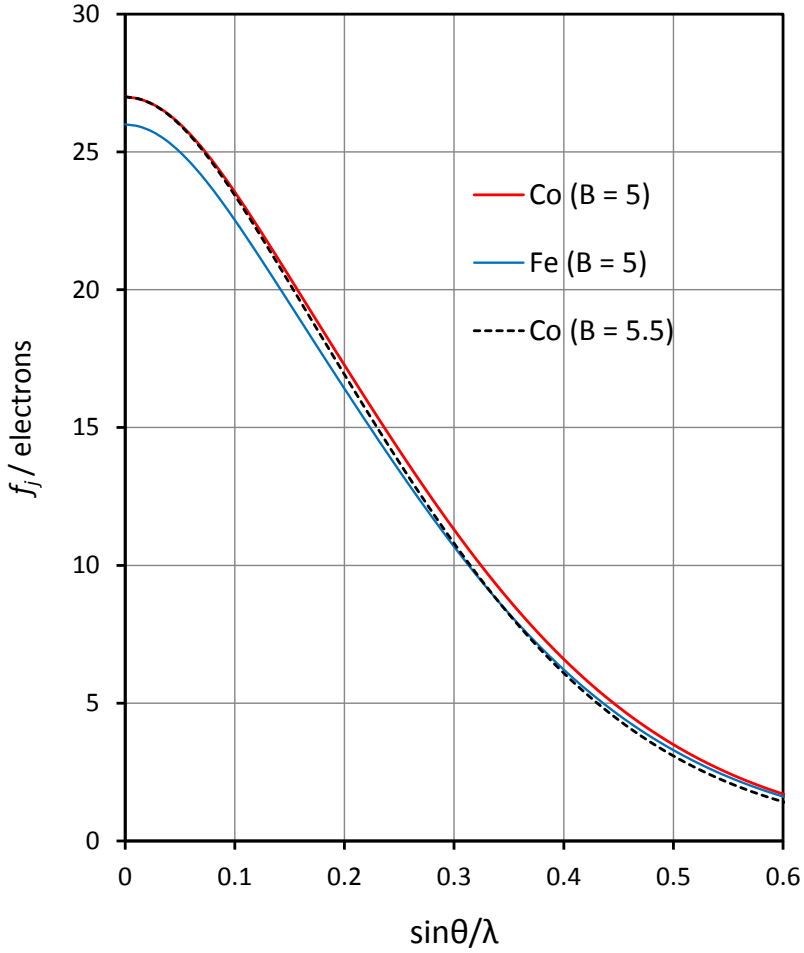


Figure 1.20 Graph to show the atomic scattering factors for Co and Fe *vs.* $\sin\theta/\lambda$. Values are calculated using Eqn. 1.85 and include temperature factors (shown in parentheses).

However, the form factor is a complex value, and the anomalous part of the scattering only becomes significant near to the absorption edge, whereas the real part of the form factor has influence over a larger energy range. Therefore, by carefully selecting the wavelength such that we are close in energy to the absorption edge, the signal from the anomalous part of the form factor can be used to gain more detailed information as to the nature of the atom by providing contrast between the scattering factor curves.

1.4.4 Current Techniques

A number of techniques for distinguishing these electronic properties of atoms within compounds currently exist, and are regularly used in chemical analysis. However, many of these techniques have limitations which, as demonstrated in this section, can be overcome by using the extra information available to us from the anomalous signal.

Bond Valence Sums

Bond Valence Sums (BVS) are often employed in coordination chemistry as a method of determining the oxidation state of an atom. An example of this is presented in work by Shoko *et al.* to determine the oxidation states of Ce ions in a number of cerium oxides. [111] The theory is a development of Pauling's Rules [112] and postulates that the oxidation state of the jth atom, z_j, can be derived from the BVS of the surrounding atoms: [113]

$$z_j = \sum_i s_{ij}, \tag{1.88}$$

$$s_{ij} = \exp[(R_0 - r_{ij})/b], \tag{1.89}$$

$$s_{ij} = (r_{ij}/R_0)^{-N}. \tag{1.90}$$

The individual bond valences, s_{ij} in Eqn. 1.88, are calculated from observed bond lengths, r_{ij}, using either Eqn. 1.89 or Eqn. 1.90, where R_0 and N are constants dependent on $i\&j$, and b is an empirical constant, (typically quoted to be 0.37 Å). R_0 is a tabulated parameter for any given ij pair and is the ideal bond length given s_{ij}=1. Therefore, when determining a BVS, an assumption must be made as to the oxidation state of the metal ion in order to use the correct R_0 value. Another limitation of the technique is that it relies on very accurate measurements of the observed bond

length. These can be measured from crystallographic data, however, there are many factors that can affect the accuracy of this measurement including: thermal motion of atoms (large anisotropic displacement parameters); steric constraints on the ligand coordination; poor quality crystal data as a result of disorder; a combination of these factors. Therefore, although BVS may be used for a good initial estimate, it does not always provide unambiguous confirmation of oxidation state.

Magnetic Susceptibility

Magnetic susceptibility, χ_m, is a measure of the degree of magnetisation in a material in response to a magnetic field. It is defined as the ratio between magnetisation, M, of the material in the magnetic field and the field intensity, H, [114]

$$M = \chi_m H \tag{1.91}$$

The observed value of χ_m can be used to classify a given material as

- $\chi_m < 0$ - diamagnetic
- $\chi_m > 0$ - paramagnetic

In the solid state, the most common technique for measuring the magnetic susceptibility is with a SQUID (superconducting quantum interference device) magnetometer, however many other techniques are available. [115, 116] When studying SCO materials, the transition between a paramagnetic HS state to a near-diamagnetic LS state can be observed as a sharp change in the magnetic susceptibility as a function of temperature, $\chi_m(T)$. With knowledge of the magnetic susceptibilities of the purely HS and LS states, the HS state (or LS) mole fraction, γ_{HS}, at any given temperature can be derived from

$$\chi_m(T) = \gamma_{\mathrm{HS}}\chi_{\mathrm{HS}} + (1-\gamma_{\mathrm{HS}})\chi_{\mathrm{LS}}, \tag{1.92}$$

where χ_{HS} and χ_{LS} are the temperature dependent contributions to the total sus-

ceptibility. [104] This information is used to produce the spin transition curves shown in Figure 1.18.

Mössbauer Spectroscopy

Mössbauer spectroscopy makes use of the Mössbauer effect to provide precise information about the chemical and structural properties of materials. The resulting spectrum is a measure of the γ-ray absorption intensity *vs.* energy for a given resonant Mössbauer isotope. For resonance to occur, a γ-ray emitted by one nucleus must be efficiently absorbed by a second nucleus. In 1957, it was discovered that for this to occur, the nuclei must be within the solid state to allow for recoilless gamma ray emission and absorption, which has since become known as the Mössbauer effect. [117]

Often when nuclei undergo energy transitions, they are accompanied by the absorption or emission of a γ-ray. In the free nuclei case, momentum is conserved upon absorption or emission by a recoil motion, analogous to a firing gun (see Figure 1.21). Consequently, due the energy of the recoil, E_R, the energy of the emitted γ-ray will be less than E_R. [118] However, in order for this γ-ray to be absorbed by a second nucleus and thereby achieve resonance, the emitted γ-ray must have an energy greater than the nuclear electronic transition due to the recoil. Therefore, resonance can only occur if the loss of energy due to recoil is overcome.

Figure 1.21 Recoil of a nucleus upon absorption or emission of a γ-ray.

The emitted γ-rays have a range of energies due to the thermal motion of the atoms. This spread is denoted E_D, since it is caused by the Doppler Effect. Mössbauer discovered that when nucleus are present in the solid state, the effective mass of a given nuclei is increased and as a result E_R and E_D are greatly reduced. Since the recoil is effectively absorbed by the whole solid lattice, it becomes an almost zero-energy event, meaning the emitted γ-ray is of the same energy as the electronic

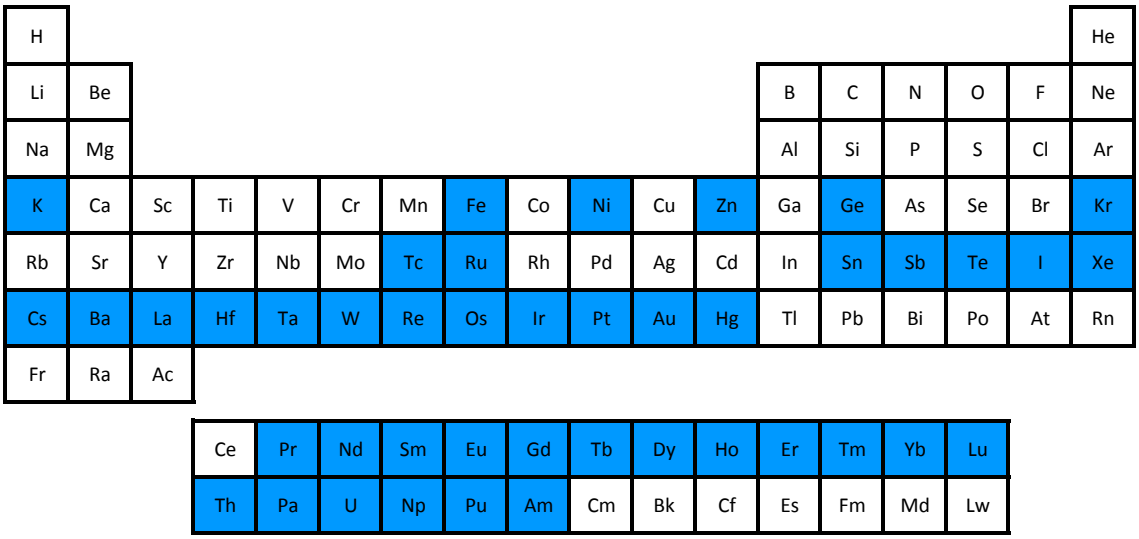


Figure 1.22 Periodic table showing Mössbauer active elements (blue).

transition and so resonance may be achieved.

Since electronic transitions within a nucleus are dependent upon chemical environment and electronic state of the atom, we can now probe these features using Mössbauer spectroscopy. The technique, however, is resolution limited due to the near zero values of E_R and E_D. In order to observe nuclei using this technique, they must have a large natural linewidth, which is related to the half life of the excited state. As a result, only relatively few nuclei (see Figure 1.22) may be probed, the most common of which is ^{57}Fe, which has both a low lying and long lived excited state. Another limitation of this technique is that although it provides information about the presence of different oxidation states within a compound, *there is no information as to the location of this charge in the molecule.*

XAFS

X-ray absorption fine structure (XAFS) is the measurement of the X-ray absorption coefficient of a given material as a function of energy. Specifically, it is *the modulation of an atom's X-ray absorption probability due to the chemical and physical state of an atom.* [119] XAFS is a very sensitive to the electronic state of a selected element and therefore we can gain much information about its oxidation state, spin state

and the coordination chemistry surrounding it, including coordination number, bond distances to neighbouring atoms and the nature of the neighbouring atoms.

Although this technique provides information about the presence of an oxidation state in a compound, it cannot provide site specific information as to the exact location of that charge within the structure. Often XAFS must be used in conjunction with one or more other techniques described, however there will still remain ambiguity in the solution. Another limitation of this technique is the presence of more than one oxidation or spin state of the same element within a compound. The spectrum observed now becomes an average of the two spectra pertaining to the two states of the element. Section 1.4.5 describes a new technique for the deconvolution of such spectra and its application to compounds containing more than one oxidation state. A more detailed explanation of XAFS is presented in Section 1.4.5, since this provides the basis of the deconvolution program, DetOx.

XPS

X-ray photoelectron spectroscopy (XPS) is a surface sensitive technique that can be used to measure the elemental composition of a material. The technique is also sensitive enough to provide an empirical formula and information into electronic states of elements within the target. The technique works on much the same principle as described in the XAFS section above *i.e.* The Photoelectric Effect. [120]

XPS works on a single photon-in/electron-out basis in which electromagnetic radiation (of energy defined by $E = h\nu$) incident on a sample is absorbed by an atom leading to the ionisation of a core electron. The kinetic energy of the ejected electron (the photoelectron) is then recorded using an electron energy analyser. The process can be described by

$$A + h\nu = A^{+} + e^{-}, \tag{1.93}$$

where A is an atom in the sample. Since the energy of the photoelectron is solely

kinetic, we can rearrange Eqn. 1.93 for its kinetic energy (KE) in terms of the energies of all the species,

$$\mathrm{KE} = h\nu - \left(E(A^+) - E(A)\right). \tag{1.94}$$

The difference in energy between the neutral and ionised atom is known as the binding energy (BE) and so Eqn. 1.94 is often expressed as

$$\mathrm{KE} = h\nu - \mathrm{BE} - \phi, \tag{1.95}$$

where ϕ is the *work function* to correct for the fact that binding energies are conventionally measured with respect to the Fermi-level of the solid whereas Eqn. 1.94 is for ionisation to the vacuum level. [121]

Since the measured binding energies are a direct representation of the energies of core atomic orbitals in an atom, each element will produce a unique set of lines in the spectrum, which can be used to identify different species in a sample. Moreover, the intensity of the lines is proportional to the concentration of the species in the sample and so ratios of different spectral lines can be used to produce an empirical formula for the composition of a sample. Finally, local chemical environment and electronic state (spin state or formal oxidation state) of an atom will affect the energies of the core atomic orbitals. Therefore, chemical shifts of peaks can be observed to discriminate between oxidation states and chemical environments in the sample *e.g.* the spectral peak for Mn^{III} will exhibit a higher binding energy than Mn^{II}. Figure 1.23 shows the XPS spectra for a Silver-Gold doped zeolite. Sharp characteristic peaks are seen for each element in the appropriate energy region. [122]

1.4.5 Underlying Theory

The unambiguous discrimination of oxidation state, spin state and atoms with near identical scattering factors, all exploit the wavelength-dependent anomalous scattering phenomenon present in all atoms. This section describes how we can measure

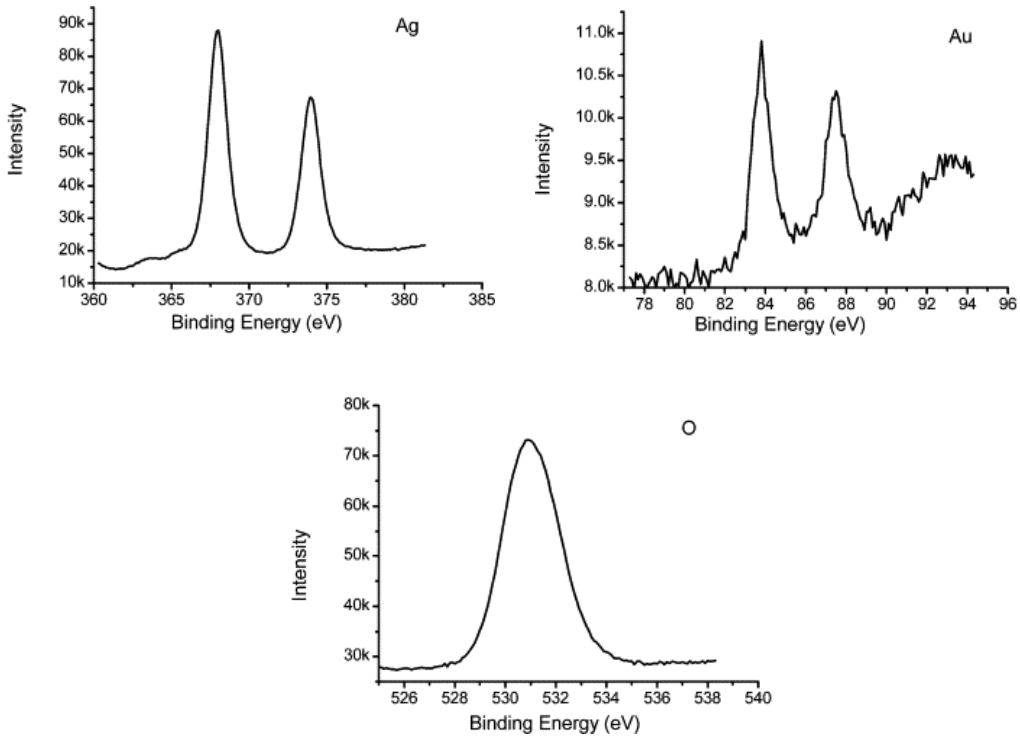


Figure 1.23 X-ray photoelectron spectra recorded for a Ag-Au doped zeolite sample.

this anomalous signal using XAFS and use this to discriminate between the electronic states of atoms.

The application of resonant scattering to determining the electronic state of an atom or the atomic species at a given site within a structure has been previously explored. However, to date these studies have almost exclusively been confined to powder X-ray crystallography. Examples of such valence contrast experiments in powder samples include: $GaCl_2$, Eu_3O_4, Fe_3O_4 and $\alpha-Fe_2PO_5$. [123–126] These experiments all involve the refinement of f' at various wavelengths around the absorption edge at ambiguous sites in the structure. The calculated values are then fitted and compared against standards for different oxidation states to confirm their identity. The technique is more commonly applied to element selective contrast studies to determine site occupancy disorder in compounds *e.g.* in doped high T_c superconductors. [127–129] Although these techniques have proven successful, contrast experiments from powder samples encounter two major difficulties compared to single crystal diffraction: Bragg reflection overlap and higher background. Corrections to the data

can be made to account for these, however, given that the anomalous signal is generally very small compared to the overall scattering of the molecule, any such corrections applied can greatly affect the analysis.

In light of previously reported work on powder samples, we aim to undertake single crystal experiments to contrast the electronic states of atoms, and distinguish neighbouring elements within compounds. Using the tunable-wavelength capabilities and fast, low background data collection available on beamline I19, the limitations of powder experiments can be overcome. This allows for the study of larger molecules as opposed to the relatively simple metal oxide structures reported. Very little work has been done in this area due to the availability of tunable-wavelength set-ups for small molecule crystallography. Ohara *et al.* [130] have demonstrated that Eu^{II} and Eu^{III} can be differentiated from single crystal data on a sample of Eu_3S_4, however, this work exploits very large anomalous signals from the Europium ions around the L_{II} absorption edge. Furthermore, the results from the refinement are compared to standard values from EuS and Eu_2O_3. This project studies valence and atomic contrast of first-row transition elements, where the anomalous signal is much smaller around the *K*-edge (since *L*-edges are not within accessible range of the synchrotron) and without any pre-assumed knowledge of the structure.

The use of such small anomalous signals compared to the total scattering of the molecule is analogous to refinement of the Flack parameter in small molecules. [131] The Flack parameter exploits the differences ($D(hkl) = | F(hkl) |^2 + | F(\overline{hkl}) |^2$) in Friedel pairs to determine the overall absolute configuration of a structural model. [132] Although the signal is very small, by using hundreds or thousands of data points, a high degree of precision and accuracy can be reached. For example, the absolute structure of Ibuprofen ($C_{13}H_{18}O_2$) can be determined from data collected using Cu K_α radiation by exploiting the anomalous signal from the oxygen atoms. The F_{000} for Ibuprofen is $(13 \times 6) + (18 \times 1) + (2 \times 8) = 112$. The anomalous contribution to the scattering from the two oxygen atoms at the Cu K_α wavelength can be calculated by,

$$f'_O + f''_O = \sqrt{(f')^2 + (f'')^2} = \sqrt{(0.05)^2 + (0.03)^2} = 0.058 \text{ electrons per O atom.} \quad (1.96)$$

As a fraction of the F_{000}, the anomalous signal from the oxygen atoms only accounts for $^{(2\times0.058)}/_{112} = 0.1\%$. It follows that an anomalous signal of $\geq 0.1\%$ in the data should be enough to contrast different electronic states of elements within a structure.

X-Ray Absorption and Fluorescence

X-ray absorption occurs via the photo-electric effect, discovered by Einstein in 1905. When an incident X-ray photon has sufficient energy, it is absorbed by a core electron which in turn is emitted from the atom. The atom is now said to be in an excited state containing a *core hole*, which must be filled in order to return the atom to its ground state. When a core electron from a higher state drops down to fill the core hole, an X-ray is emitted and it is this which we measure as fluorescence (Figure 1.24). Depending on which level the electron falls from to fill the core hole, we label the fluorescence differently *e.g.* decay from the L ($n = 2$) to K ($n = 1$) level is labelled K_α, whereas from the M ($n = 3$) level to K is labelled K_β.

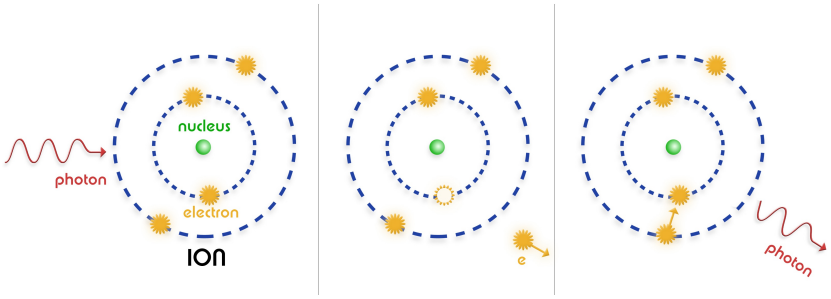


Figure 1.24 Left: Absorption of X-ray by atom, Center: Ejection of photo-electron from core, Right: Fluorescence emission as electron drops to fill core hole. Image courtesy of NASA/CXC/SAO.

We measure the absorption by an atom quantitatively using the Beer-Lambert

Law, [133]

$$I = I_0 e^{-\mu t}, \tag{1.97}$$

where I_0 is the intensity of the incident X-ray, I, is the intensity of the transmitted X-ray beam through the sample, t, is the sample thickness and μ, is the absorption coefficient for the observed atom. The absorption coefficient is strongly dependent on the physical factors of the atom itself and the energy of the incident X-ray;

$$\mu \approx \frac{\rho Z^4}{AE^3}, \tag{1.98}$$

where ρ is the sample density, Z, is the atomic number, A, is the atomic mass and E, the X-ray energy. The Z^4 dependence in the equation allows for strong contrast between atomic species to be drawn.

For the purpose of this project, we are particularly interested in the absorption behaviour of an atom when the energy of the incident X-ray beam is near to that of the binding energy of core electrons: we observe a rapid increase in X-ray absorption known as the absorption edge. We are able to follow this absorption as a function of energy using a fluorescence detector (described in Section 2.1.4). The binding energy of core electrons is unique to each element, thus we can probe individual atom types by tuning the wavelength to the appropriate absorption edge. Since most elements have either a K or an L edge within the range 5-35 keV, most element types can be probed by XAFS using a synchrotron source (Figure 1.25).

1.4.6 EXAFS *vs.* XANES

Figure 1.26 is a typical fluorescence-absorption spectrum, showing the absorption coefficient as a function of X-ray energy. Extended X-Ray Absorption Fine Structure, EXAFS, is the study of oscillations in the absorption coefficient well above the absorption edge, defined as

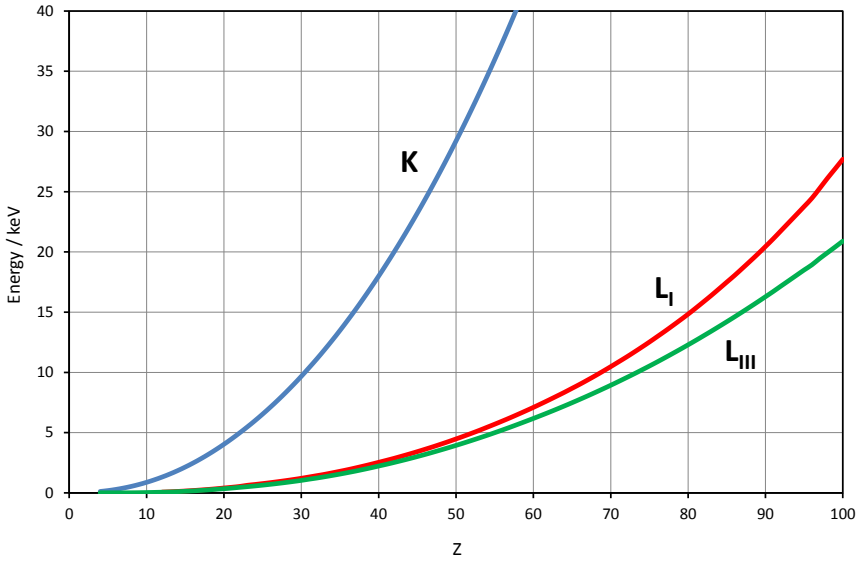


Figure 1.25 K and L absorption edge energies as a function of atomic number Z.

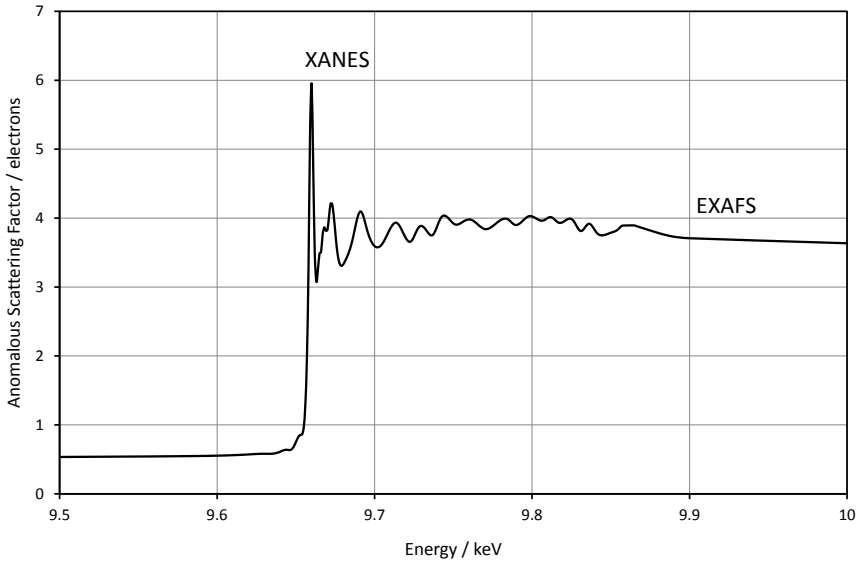


Figure 1.26 Fluorescence-absorption spectrum of a zinc porphyrin compound showing the EXAFS and XANES regions of interest.

$$\chi(E) = \frac{\mu(E) - \mu_o(E)}{\Delta\mu_0(E)}, \tag{1.99}$$

where $\mu(E)$ is the observed absorption coefficient, $\mu_0(E)$, a smoothed function representing the absorption from an isolated atom and $\Delta\mu_0$, the change in absorption coefficient at the absorption edge. Using this equation, or variations on the equation such that the signal is weighted to increase resolution at high energy, the coordination number and distance to neighbouring atoms can be determined. Information about the atomic species of the neighbouring atom can also be determined using EXAFS.

X-ray Absorption Near Edge Structure, XANES, is a much more intense region of the spectrum, as seen in Figure 1.26, and is very sensitive to properties of the structure such as coordination number and oxidation state of the element. The EXAFS equations break down at energies near to the absorption edge and so currently the XANES region of the spectrum tends to be treated qualitatively rather than quantitatively. The shape of the 'pre-edge' features and the position of the absorption edge can give us information about coordination number, symmetry, geometry, the molecular orbitals (and their hybridisation) and also the band structure. The position of the absorption edge can provide insight into the oxidation state (Figure 1.27) and if more than one oxidation state exists in the structure, the ratio of heights of the pre-edge to edge features equates to the ratio of oxidation states in the sample.

Finally, we have shown that the position of the absorption edge will also move as a result of changing spin state from low↔high (Figure 1.28). The shape of the edge also changes as the structure undergoes a change in geometry to account for the change in spin state. The example shown in Figure 1.28 is discussed in more detail in Chapter 3.

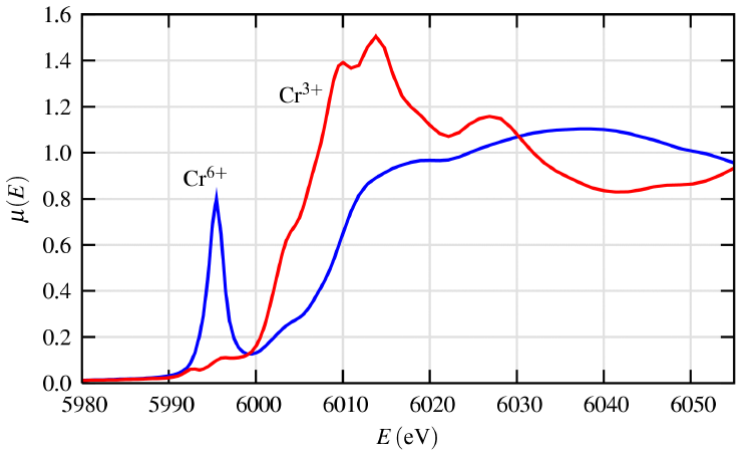


Figure 1.27 Cr K-edge XANES spectrum showing the difference between a tetrahedrally coordinated Cr^{+6} oxide giving rise to a sharp pre-edge feature (blue) and an octahedrally coordinated Cr^{+3} oxide (red). [119]

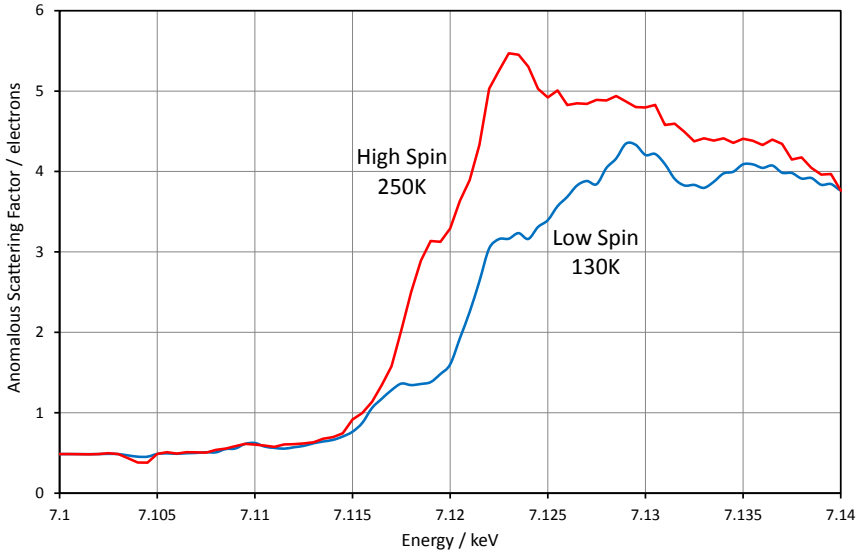


Figure 1.28 Fe K-edge XANES spectrum showing the different absorption edge positions of the low spin (blue) and high spin (red) states of $[Fe(btr)_3](ClO_4)_2$

.

1.5 Thesis Overview

The aim of this project was to explore the applications of tuneable wavelength capabilities of Beamline I19 at Diamond Light Source, Oxford. With reference to variable wavelength experiments undertaken in macromolecular and powder X-ray crystallography, we have explored how these techniques can be exploited in single crystal studies to reveal information about structure that would otherwise be unobtainable. The thesis is split into two main objectives: Firstly we have explored the use of anomalous scattering for discriminating between different states of atoms within molecules; and secondly the application of structure solution techniques previously reserved for macromolecular crystallography to cases of low resolution data in small molecule crystallography.

Chapter 2 focuses on the I19 beamline set-up at Diamond Light Source and the specific pieces of equipment we have used and commissioned. There are also details of other equipment and experimental techniques and synthesis used during the project.

The first part of Chapter 3 provides a description of a program, DetOx, written as part of this research which is used to determine the anomalous scattering factors for different oxidation states in mixed-oxidation state compounds, from absorption fluorescence data. We discuss the application of DetOx to a number of worked examples of mixed-oxidation state compounds. The use of DetOx is then extended to discriminate between spin states in compounds where more than one spin state of a given element is present at any one time.

Chapter 4 is an extension of Chapter 3 for distinguishing elements and their electronic states. It was found that DetOx is limited when applied to cases with large numbers of anomalous scatterers leading to high absorption in the data. We therefore apply a new approach, 'ratio refinement', to overcome the problem. Further worked examples are provided to discriminate elements with near identical scattering and again to mixed-oxidation state compounds.

Chapter 5 describes how we can apply macromolecular techniques for structure solution to the problem of solving 'large' small molecules. In these cases, data quality is generally too low for standard small molecule techniques to be effective and so we explore the incorporation of macromolecular techniques into our solution.

The final chapter is a discussion of the results presented in the thesis. It places the work in context and looks at how variable wavelength X-ray experiments can be applied to future problems in crystallography.

Chapter 2

Experimental Apparatus and Methods

This chapter describes the techniques and equipment used in this research project. All variable wavelength experiments were performed at Diamond Light Source on Beamline I19, with supporting data collected in-house.

2.1 Beamline I19

Beamline I19 is a dedicated small-molecule single-crystal X-ray diffraction beamline at Diamond Light Source, Oxfordshire, UK. Since becoming operational in 2008, I19 has been primarily used for structure determination of small, weakly diffracting crystals. More recently, high pressure experiments, studies of metastable species, variable-temperature crystallography and gas exchange experiments have become more common. [134]

Figure 2.1 shows a schematic of the beamline. The beam enters the beamline from the storage ring on the right hand side of the diagram and first passes into the optics hutch. The wavelength is selected using the cryo-cooled double-crystal monochromator (DCrM) before the beam is focussed by a pair of bimorph mirrors which focus the beam in the horizontal (horizontal focusing mirror, HFM) and vertical (vertical focusing mirror, VFM) planes. Data collection takes place in either experimental hutch 1 (EH1), or experimental hutch 2 (EH2) depending on the nature of the experiment. The beamline also contains a preparation room and control room behind the experimental hutches.

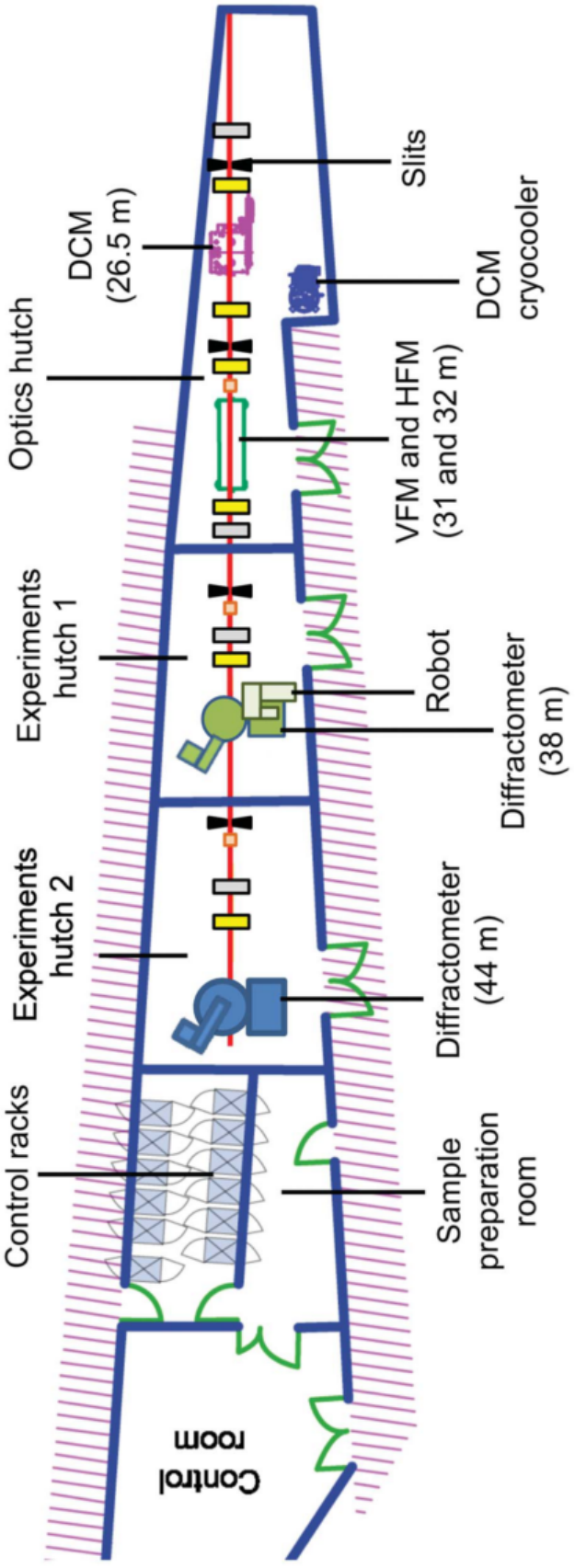


Figure 2.1 Schematic drawing of Beamline I19.

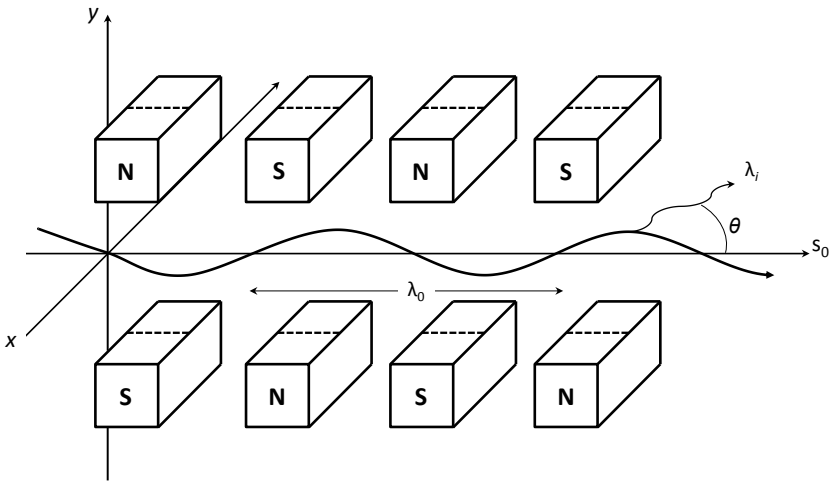


Figure 2.2 Schematic drawing of an undulator.

2.1.1 Undulator

The synchrotron radiation for the beamline is generated by an in-vacuum U21 undulator. An undulator is a insertion device (ID) placed in a straight section of the storage ring and consists of an array of permanent magnets which apply a sinusoidal magnetic field (of wavelength λ_0) transverse to the electron beam (Figure 2.2). The strength of the field can be controlled by altering the gap between the arrays of magnets either side of the beam. In turn the undulator forces individual electrons in the beam to oscillate and emit energy in the form of X-rays. The device is described by the undulator strength parameter, K,

$$K = \frac{eB_0\lambda_0}{2\pi m_e c}, \text{ [135]} \tag{2.1}$$

where B_0 is the magnetic field and m_e is the mass of a electron. For an undulator, $K \ll 1$, and it exploits the constructive interference of radiation emitted from a given electron at points along the device in order to produce very sharp, intense beams with high degrees of coherence at discrete wavelengths. These wavelengths, λ_i, are given by

$$\lambda_i = \frac{\lambda_0}{2\gamma^2 j}\left(1 + \frac{K^2}{2} + \gamma^2\theta^2\right), \text{ [136]} \tag{2.2}$$

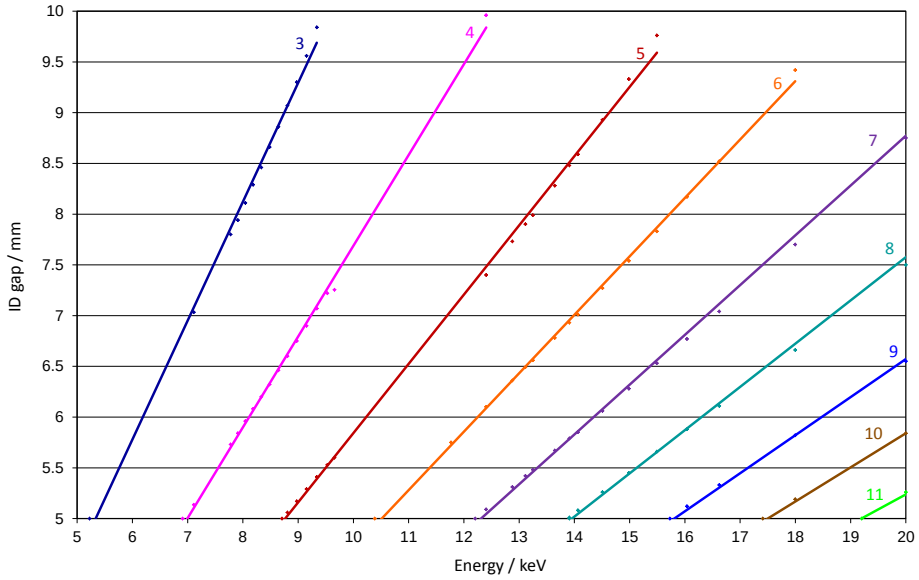


Figure 2.3 Graph to show the maximum flux as a function of energy and ID gap.

where γ is the electron energy, θ is the observation angle of X-rays with respect to the beam's direction of propagation (s_0) and j is the harmonic number (see Figure 2.2). Eqn. 2.2 reveals that these wavelengths are tuneable by varying the value of K. This can be achieved through careful selection of the gap size and therefore the magnetic field strength.

When performing an experiment, since we are regularly changing wavelength of the incident X-ray beam, it is important to ensure the ID gap is adjusted to provide the maximum flux possible. Look-up tables have been compiled for the U21 undulator on I19 showing the maximum flux for a given X-ray energy and ID gap. Figure 2.3 shows this look-up table with best fit lines drawn to indicate the different magnetic harmonics within the energy range. The choice of harmonic to use is therefore a compromise between the required tunability and maximising the flux by choosing the lowest harmonic. For example, when performing mixed-valence studies on manganese complexes, since our energy of interest was around the manganese edge at 6.54 keV, the 3rd harmonic was chosen to maximise flux. Since the minimum ID gap possible is 5 mm, harmonics below the 3rd are not observable (Figure 2.3).

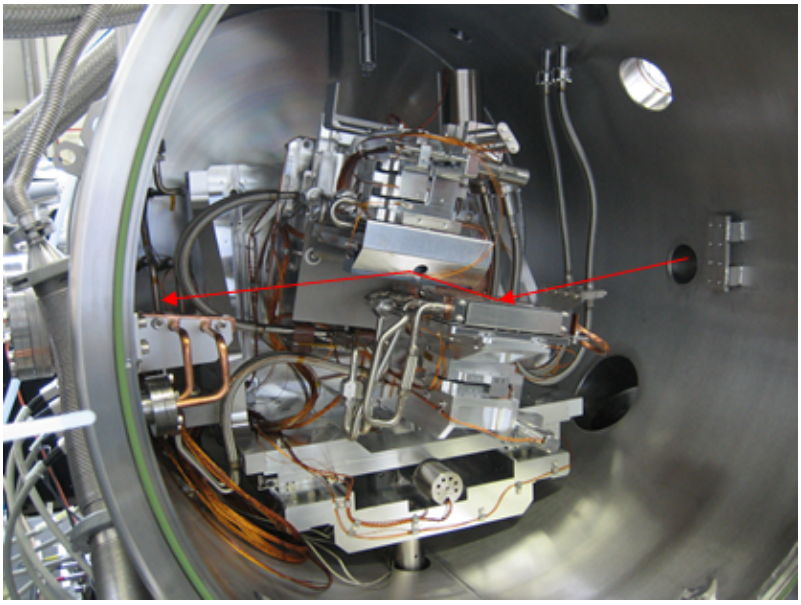


Figure 2.4 Photograph of the DCrM on Beamline I19.

2.1.2 Optics

Monochromator

I19 uses a cryo-cooled DCrM to allow wavelength selection in the range 5 keV to 28 keV (0.44 Å to 2.48 Å). The DCrM consists two Si-111 cut crystals as shown in Figure 2.4.

The two crystals are essentially identical, differing only in their orientation with respect to the beam. They are mounted in a parallel configuration such that the monochromatic exit beam is parallel to the incident 'white' X-ray beam *i.e.* there is no beam offset (as illustrated with red arrows in Figure 2.4). The wavelength is selected by rotating the crystal arrangement in order to satisfy the Bragg diffraction condition for the desired wavelength ($n\lambda = 2\mathrm{d}_{\boldsymbol{h}}\sin\theta$), whilst maintaining their parallelism. Photons of the correct wavelength are diffracted through an angle 2θ. The first crystal must be cryo-cooled to dissipate the heavy heat load from the synchrotron radiation. The whole monochromator is housed in a high vacuum environment.

Stripe	Energy Range / keV
Silicon	5 - 10
Rhodium	10 - 22
Platinum	22 - 28

Table 2.1 Specific energy ranges provided by different stripes on the focusing mirrors.

Focusing Mirrors

The incoming monochromatic radiation is focused in the horizontal and vertical planes independently using the HFM and VFMs (Figure 2.5). The dimensions of the mirrors have been designed to maximise the incidence of the beam across them. The VFM (750 mm) is shorter than the HFM (1000 mm) since the beam divergence is less in the in the vertical plane than the horizontal. Each mirror contains three stripes (silicon, rhodium and platinum), each with different harmonic rejection capabilities. This is required since at the given grazing-incidence of the beam on the mirror (3 μrad), the reflectivity of each stripe is strongly dependent on the energy of the beam, with well-defined cut-off energies. Table 2.1 shows the energy range for which each stripe provides a harmonic-free focused beam.

2.1.3 Experimental Hutch 1

EH1 has been designed for routine, high-throughput crystallography (see Table 2.2). Figure 2.6 shows the sample position in EH1. The Cryostream allows for a sample position temperature range of 30 K to 500 K. Samples can either be mounted manually or by using the robot (Figure 2.7) and the motorised microglide allows samples to be centered from the control room. The 4-circle, κ geometry diffractometer has an air-bearing on the ω axis to allow for rapid scanning and retakes.

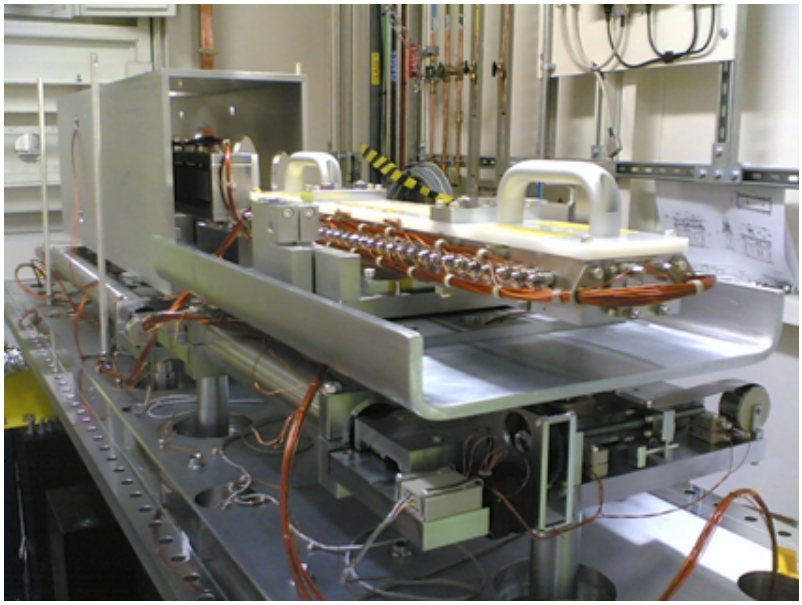


Figure 2.5 Photograph of the focusing mirrors on Beamline I19.

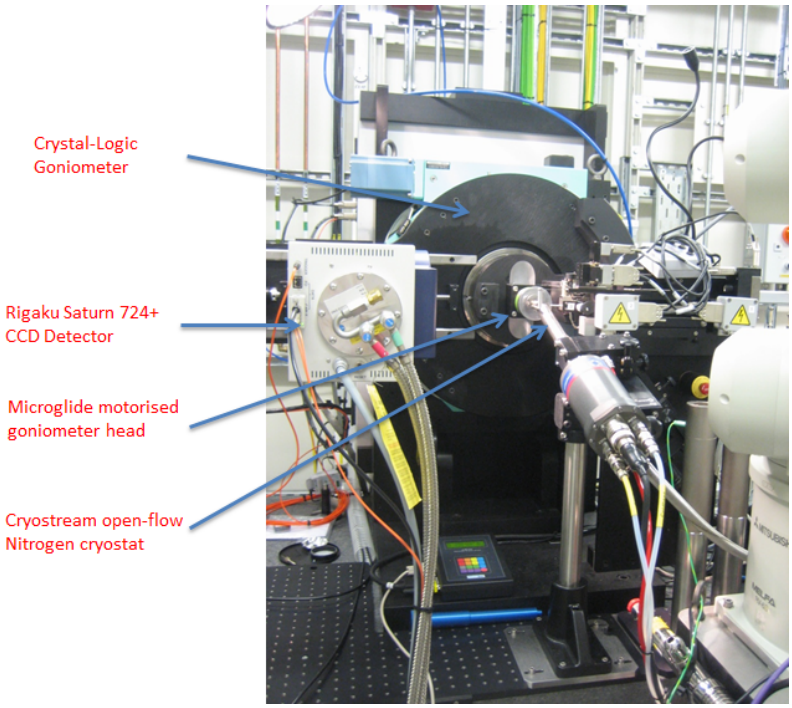


Figure 2.6 Sample position in EH1.

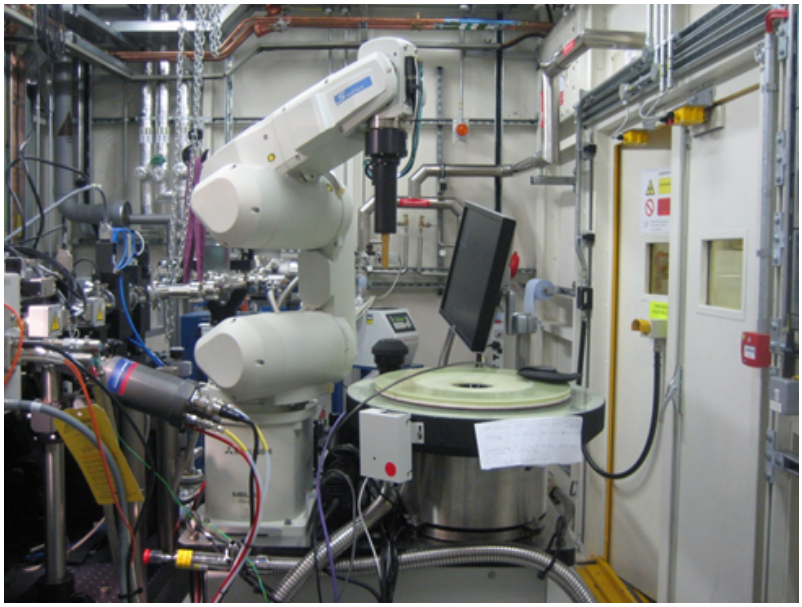


Figure 2.7 Robot in EH1.

Beam Size (uncollimated) / μm	Approximate Gaussian ≈170 (h) by 85 (v) FWHM
Goniometer	Four-circle κ goniometer (ω, κ, ϕ, 2θ)
Temperature Capability / K	≈ 30 to 500
Sample Mounting	Manual or automated
Detector Type	CCD
Detector Model	Rigaku Saturn 724+
Sphere of Confusion / μm	< 10

Table 2.2 Beamline details for EH1.

2.1.4 Experimental Hutch 2

EH2 is designed for less-routine experiments (see Table 2.3). Since the hutch is positioned downstream of EH1, time can be spent setting up for technically difficult experiments whilst EH1 remains operational. Typical experiments performed in EH2 include excited state studies, gas cell studies and variable wavelength studies as in this project. EH2 contains a larger 4-circle goniometer than EH1 allowing for loads of up to 25 kg at the sample position (see Figure 2.8). The Pilatus 300 K detector allows for much faster data collection, as described below.

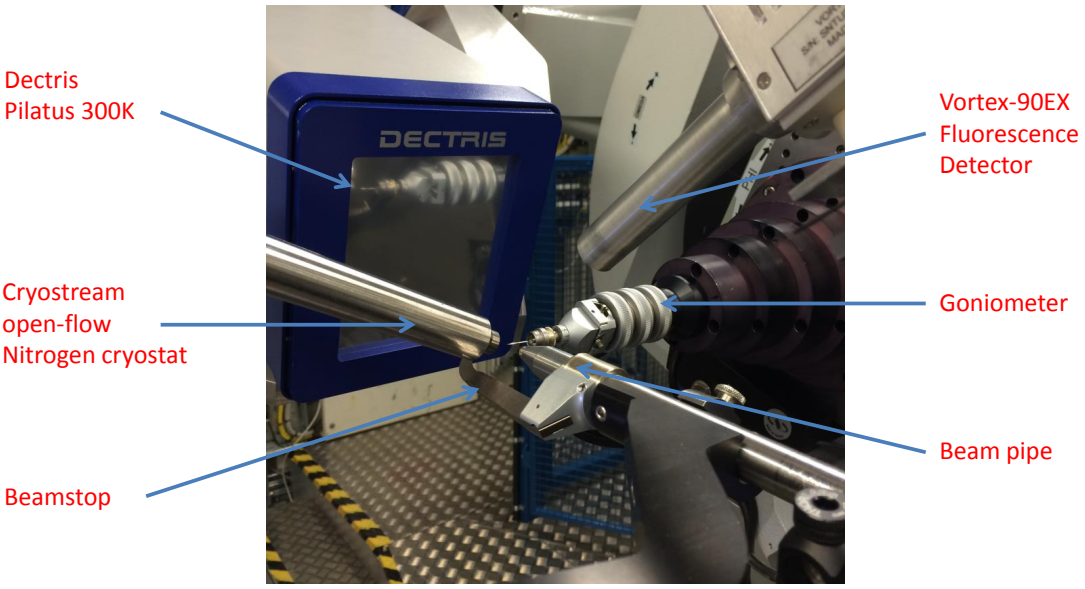


Figure 2.8 Sample position in EH2.

Beam Size (uncollimated) / μm	Approximate Gaussian ≈190 (h) by 130 (v) FWHM
Goniometer	Four-circle κ goniometer (ω, κ, ϕ, 2θ)
Temperature Capability / K	≈ 4 to 1200
Sample Mounting	Manual
Detector Type	Pilatus 300K
Detector Model	Dectris
Sphere of Confusion / μm	< 60

Table 2.3 Beamline details for EH2.

Fluorescence Detector

The Vortex-90EX® silicon drift X-ray detector was commissioned and used as part of this project for X-ray fluorescence studies. The detector has both an excellent energy resolution (<133 eV FWHM at Mn K_α) and a very high input count rate of >1 Mcps. The detector is operational at room temperature and is cooled by a thermoelectric cooler.

The detector is mounted > 90° to the sample (to avoid signal saturation of the detector) and works by measuring the number of photons per second as a function of energy. An initial scan was recorded at an energy above the absorption edge of a chosen element (see Figure 2.9). The initial scan shows all of the emission lines present at and around the incident X-ray energy for atoms present in the sample.

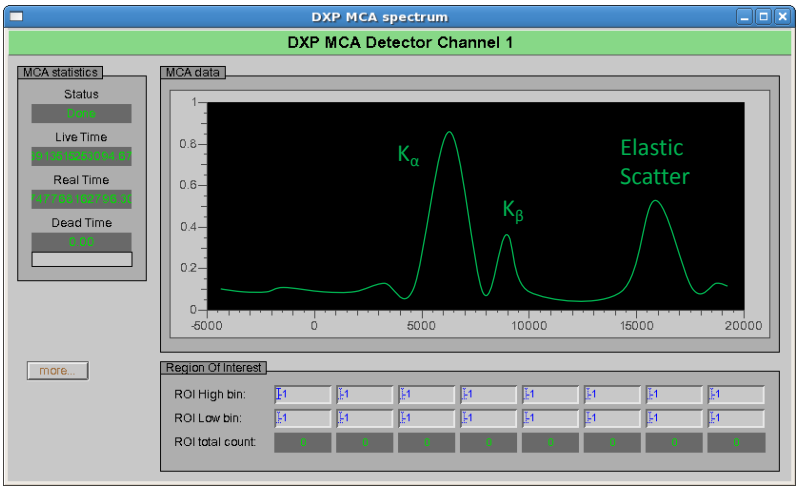


Figure 2.9 Vortex output showing photon counts as a function of energy. Line has been smoothed and labels added for clarity.

The energies of these peaks can be assigned using look-up tables available in the *X-ray Data Booklet*. [137] An additional peak is also present in the spectrum caused by the elastic scattering of the X-ray beam by the sample and air molecules. This peak is found at an energy characteristic of the incident beam – at an energy above that of the emission lines.

Once peaks in the spectrum have been assigned, region of interest (ROI) bins are placed around a selected fluorescence peak *e.g.* the K_α emission. The beamline wavelength was then scanned by changing the angle of the monochromator (and insertion device gap), and at each step in wavelength, the selected peak was integrated to produce a fluorescence spectrum. The spectrum can be viewed in real time as the data collection is in progress using the GDA software – as shown in Figure 2.10. The GDA software was developed (initially at SRS Daresbury but adopted by Diamond Light Source since 2003) as *'an open-source framework for creating customised data acquisition software for science facilities such as neutron and X-ray sources'*.

Pilatus Detector

The Pilatus 300 K detector in EH2 allows for shutterless data collection during an experiment. This means that data collection is much faster, which in turn reduces

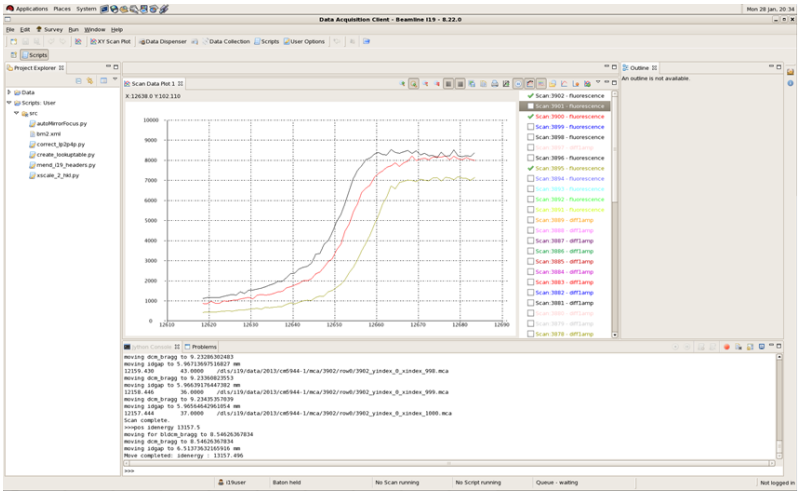


Figure 2.10 GDA output showing the integrated fluorescence peak area as a function of wavelength.

Area	83.8 x 106.5 mm^2
Pixel Size	172 x 172 μm^2
Format	487 x 619 = 310,453 pixels
Dynamic Range	20 bits (1:1,048,576)
Readout Time	7 ms
Point-Spread Function	< 1 pixel
Silicon Sensor Thickness	320 μm
Weight	7.5 kg

Table 2.4 Technical specifications of the Pilatus 300 K detector.

radiation damage to the crystal. Other advantages of the Pilatus over the CCD detector include a higher signal-to-noise ratio (by using energy discrimination thereby eliminating fluorescence) and good resolution of closely spaced peaks since the point spread function is only one pixel. Table 2.4 is an overview of the technical specifications of the Detector.

2.2 In-House Diffractometer

Single crystal X-ray diffraction data collections were performed in-house on a dual-source SuperNova (Figure 2.11). The system includes co-mounted, high intensity X-

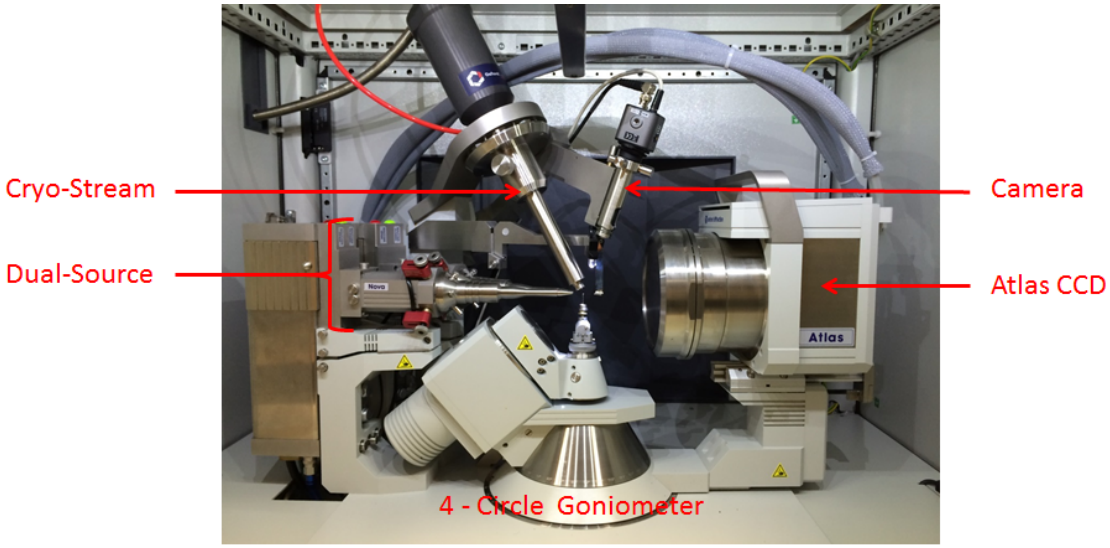


Figure 2.11 Photograph of the in-house SuperNova machine.

ray micro sources with molybdenum and copper K_α wavelengths. The high intensity means that shorter X-ray exposure times are required, and this coupled with the high sensitivity, large area Atlas CCD, dramatically reduces data acquisition times.

2.3 Syntheses

All of the syntheses presented were performed by KJS unless otherwise stated. Supporting crystal structural data can be found in the Appendix A. Routine NMR spectra were collected on a Varian Mercury 300 spectrometer with ^{1}H NMR operating at 300 MHz and $^{13}C\{^{1}H\}$ at 75.5 MHz. Mass spectra were collected on a Bruker microTOF spectrometer. All H_2O used in reactions was de-ionised and microfiltered using a Milli-Q® Millipore machine.

2.3.1 4,4′-bis-1,2,4-triazole - (btr) Ligand

The btr ligand was synthesised as previously described. [138, 139]

N,N-Dimethylformamide Azine Dihydrochloride (**1**): Oxalyl chloride (16 mL, 0.86 mol) was added dropwise to freshly distilled DMF (75 mL) with stirring at 5°C. The solution was then left for 24 hrs. Aqueous hydrazine hydrate (2.5 mL, 0.05 mol) in DMF (10 mL) was added slowly and the solution left to stir for 2 days at r.t. The formed

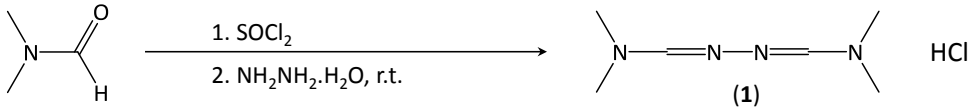


Figure 2.12 Reaction scheme for the synthesis of N,N-Dimethylformamide Azine Dihydrochloride.

white precipitate (dimethylformamide azine dichloride) was collected by filtration and washed with DMF and Et_2O.

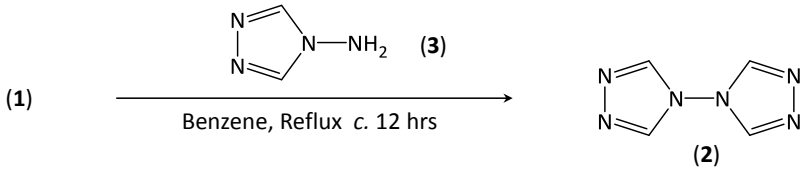


Figure 2.13 Reaction scheme the for synthesis of 4,4′-bis-1,2,4-triazole.

4,4′-bis-1,2,4-triazole (**2**): Setup for this experiment included a dry round bottom flask with a reflux condenser attached. 4-amino-1,2,4-triazole (**3**) (2.12 g, 0.025 mol) was added rapidly to a suspension of (**1**) (4.5 g, 0.021 mol) in 15 mL of benzene with vigorous stirring. The mixture was then heated until the precipitate had dissolved and reappeared as a pasty white mass. The reaction mixture was then refluxed for 12 hrs maintaining vigorous stirring - regularly checking the ^{1}H NMR spectra to ensure the reaction had gone to completion. The precipitate was then collected by filtration and washed with cold EtOH (10 mL) and Et_2O (10 mL). The white powder was recrystallised from MeOH to give colourless crystals - see Appendix A.2.

2.3.2 $Fe(btr)_3(ClO_4)_2$

$Fe(btr)_3(ClO_4)_2$ was synthesised as described previously. [140] 0.40 g (2.94mmol) of the previously prepared btr ligand (**2** from Section 2.3.1) was dissolved in a 5 mL/5 mL mixture of water and methanol, and heated to 60°C. To this was added a second solution containing 0.35 g (0.96 mmol) of $Fe(H_2O)_6(ClO_4)_2$ and 2 mg of ascorbic acid dissolved in 10 mL of water (also heated to 60°C), and stirred for 30 minutes. Colourless single crystals of the compound were grown by slow evaporation from the reaction mixture within 1 week - see Appendix A.3.

2.3.3 $Mn_{12}O_{12}(O_2CCH_3)_{16}(H_2O)_4 \cdot 4\,H_2O \cdot 2\,CH_3CO_2H$

Mn_{12} Acetate was synthesised as described by Lis in 1980. [141] 4 g of $Mn(CH_3COO)_2 \cdot 4\,H_{20}$ was dissolved in a 60% CH_3COOH solution with heating. To a cooled solution was added 1 g of $KMnO_4$ slowly and the mixture heated to 333 K with stirring. The red-brown solution was left at room temperature and pressure for a two days during which reddish-black crystals formed, suitable for X-ray analysis - see Appendix A.4.

2.3.4 $Mn_6(O)_2(O_2CCMe_3)_{10}(C_4H_8O)_4$

The hexanuclear mixed valence compound was synthesised as reported in the literature. [142] The cluster was synthesised by reaction of $MnCl_2 \cdot 4\,H_2O$ (7.75 mmol) with excess $Na(O_2CCMe_3)_{10}$ (16.3 mmol) in MeOH (70 cm^3) and stirred for 24 hrs. The reaction mixture was then dried *in vacuo* to leave the product as a dark brown powder. The powder was then recrystallised from THF to yield dark brown block shaped crystals - see Appendix A.5.

2.3.5 $Co_4^{III}Mn_5^{II}Cl_6(C_8NO_5H_{15})_2(C_8NO_5H_{16})_2(H_2O)(OH)_2$ *

To a stirred pink solution of $[Co^{II}(HOCH_2CH_2)_2NC(CH_2OH)_3Cl]Cl$ (0.36 g, 1.06 mmol) in MeCN (20 mL) and MeOH (10 mL) $NMe_4OH{\cdot}_5H_2O$ (0.36 g, 1.98 mmol) was added. The resulting dark purple solution was stirred for one hr; $MnCl_2{\cdot}4\,H_2O$ (0.40 g, 2.02 mmol) was then added. The solution was stirred for a further 5 hrs, filtered and the purple filtrate stored in a sealed vial. After one week red crystals formed in a 79% yield suitable for X-ray analysis - see Appendix A.7.

2.3.6 $Mn_3O(O_2CPh)_6(pyr)_2(H_2O)[2\,MeCN]$

This trapped-valence, trinuclear oxo-centrered manganese complex was synthesised as previously reported. [143] The synthetic route was a two step process; firstly producing $N-n-Bu_4MnO_4$ followed by the formation of the cluster.

*Sample provided by the Murrie research group in Glasgow.

N−n−Bu$_4$MnO$_4$

The synthesis of this material has been previously outlined in the literature. [144] Aqueous solutions of $KMnO_4$ (5.00 g, 31.6 mmol) and $N{-}n{-}Bu_4MnO_4$ (TBAB) (12.00 g, 37.2 mmol) totalling *c.* 200 mL were combined with vigorous stirring. A purple precipitate formed immediately, which was collected by filtration and washed with water and diethyl ether. The resulting precipitate was then dried *in vacuo*.

Complex Formation

To a solution of $Mn(O_2CMe)\cdot 4\,H_2O$ (2.00 g, 8.15 mmol) and benzoic acid (7.50 g, 61.4 mmol) dissolved in pyridine (3 mL) and EtOH (20 mL), was added $N{-}n{-}Bu_4MnO_4$ (1.14 g, 3.15 mmol) in small portions with stirring. The resulting dark brown homogeneous solution was left to stir for 24 hrs. The grey-green precipitate was then collected by filtration, washed with EtOH and dried *in vacuo*. Recrystallisation was achieved by slowly allowing a saturated MeCN solution to concentrate by evaporation over the course of a few days to produce large, deep green crystals (see Appendix A.6).

2.3.7 TBAB[CO(NH(4 −NO$_2$−C$_6$H$_4$))$_2$]$_2$ - Bromide Urea Salt †

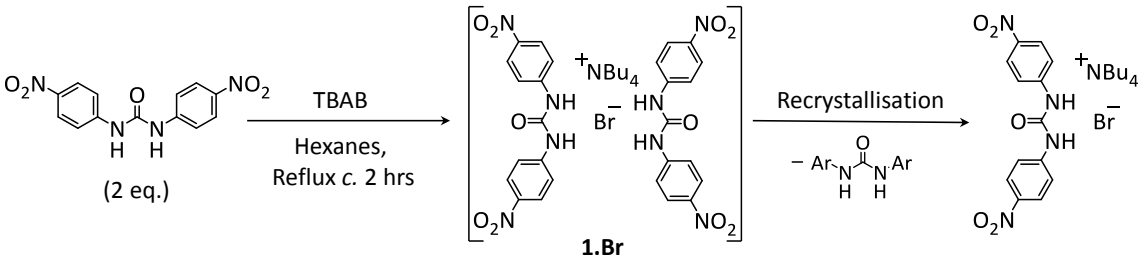


Figure 2.14 Reaction scheme for the synthesis of a bromide urea salt.

TBAB[CO(NH(4 −NO$_2$−C$_6$H$_4$))$_2$]$_2$ was prepared in analogy to a series of TBAF-diarylurea complexes. [145] To an oven-dried 100 mL round bottom flask under a nitrogen atmosphere and equipped with a stir bar and reflux condenser, were added hexanes (50 mL), TBAB (645 mg, 2.0 mmol) and 1,3-bis(4-nitrophenyl)urea (1.21 g, 4.0 mmol).

†Sample provided by the Gouverneur research group in Oxford.

The resulting suspension was allowed to stir in an oil bath at vigorous reflux (80 - 85 °C) for 2 hrs. During the course of the reaction the persistent presence of a dark yellow solid precipitate was observed. The solution was allowed to cool to room temperature, filtered, washed with hexanes (3 x 10 mL), and dried under high vacuum, giving TBAB[CO(NH(4 $-NO_2-C_6H_4$))$_2$]$_2$. Yellow block crystals of X-ray diffraction quality were obtained by vapour diffusion of hexanes into a saturated solution of **1.Br** in EtOAc (which had previously been prepared by dissolving **1.Br** in refluxing EtOAc (approximately 77 °C) until visibly saturated, filtering while hot, and then allowing to stand for approximately 1 minute to cool to near room temperature) - see Appendix A.8.

2.3.8 Bromide Rotaxane ‡

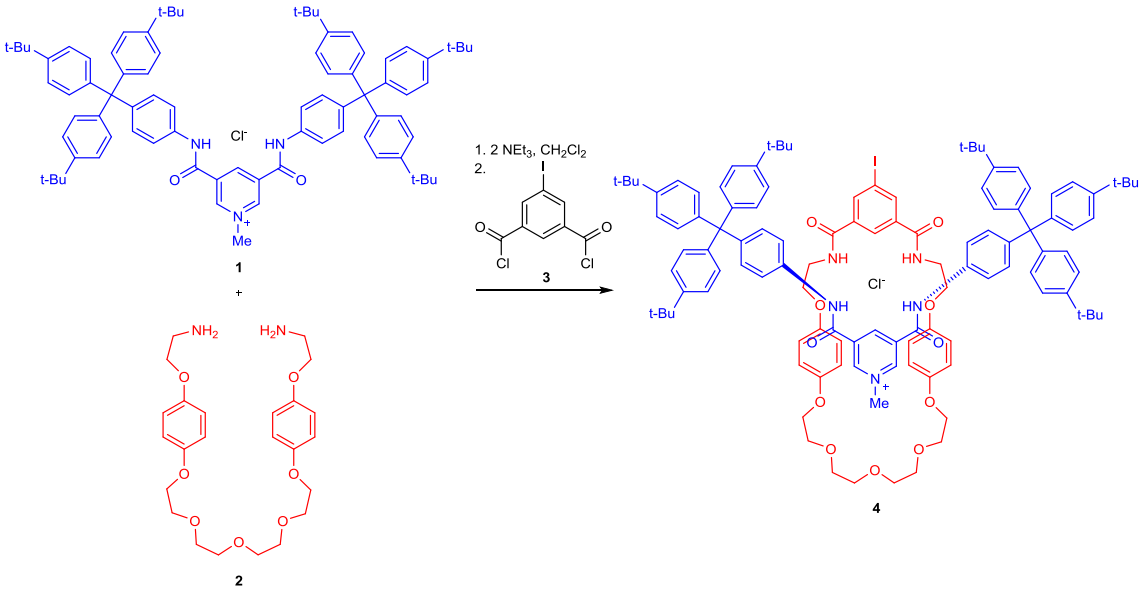


Figure 2.15 Reaction scheme for the synthesis of a bromide rotaxane.

Iodoisophthalic acid, [146] axle **1**, [147] and diamine **2**, [148] were prepared by published literature methods. The pyridinium axle **1** (518 mg, 0.436 mmol) and diamine **2** (202 mg, 0.436 mmol) were dissolved in dichloromethane (25 mL). Triethylamine (151 μL, 1.083 mmol) was added, followed promptly by the dropwise addition of 5-

‡Rotaxane provided by the Beer research group in Oxford. Subsequent anion exchange experiments were carried out by KJS.

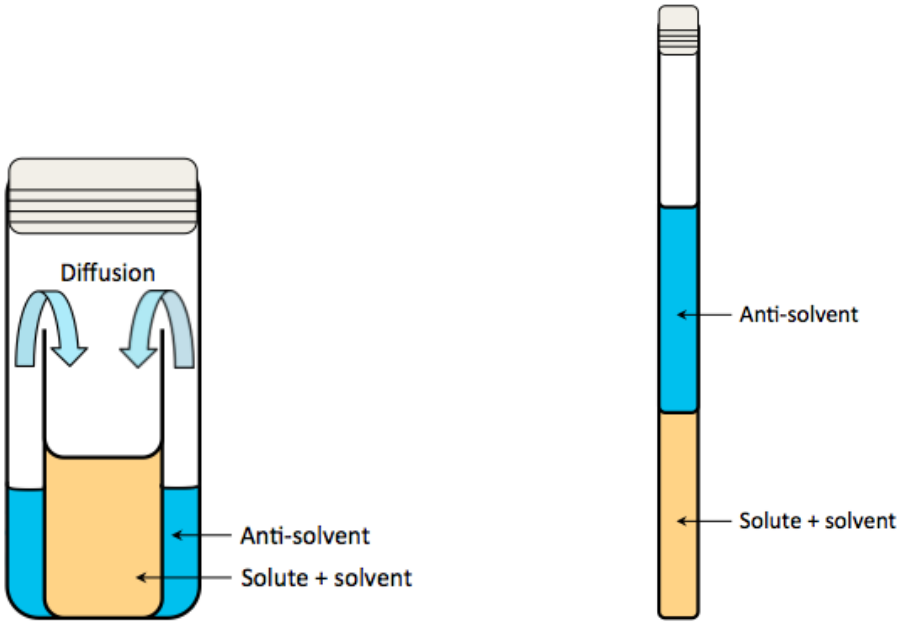
iodoisophthaloyl dichloride (*c.* 0.436 mmol) dissolved in DCM (10 mL) and the reaction left to stir overnight. The volatile components were removed *in vacuo*, the residue dissolved in DCM (25 mL) and washed with 5% sodium carbonate solution (3 x 15 mL), 1 M ammonium chloride (3 x 15 mL) and water (15 mL). The solution was dried ($MgSO_4$), filtered and concentrated *in vacuo*. The pure product was obtained *via* silica gel column chromatography (5% MeOH/DCM) as a yellow solid. To obtain the bromide anion derivative of the structure, an anion exchange reaction was performed. One aliquot (*c.* 10 mL) of **4** was dissolved in DCM (2 mL). A 2 M solution of ammonium bromide (10 mL) was slowly mixed in 2 mL portions with the rotaxane solution, each time removing the aqueous layer. Repeated washing removed the chloride leaving the bromide rotaxane in the DCM layer. The DCM solution was subsequently dried over $MgSO_4$ and the solvent removed *in vacuo*. The remaining yellow powder was recrystallised by vapour diffusion of of $^{i}Pr_2O$ into chloroform to yield yellow needle shaped crystals - see Appendix A.9.

2.4 Crystallisation Techniques

All of the crystallisations performed in the project were either by vapour diffusion or layered diffusion. Although different solvents were used in each case, this section provides a general description of the processes.

2.4.1 Vapour Diffusion

The technique involves an inner vial containing the dissolved compound to be crystallised (the solute) placed in a larger outer vial containing an anti-solvent, which is then sealed (Figure 2.16a). Sometimes the inner vial may have a perforated lid attached in order to slow down the rate of diffusion. Both containers are maintained under the same atmospheric conditions throughout the process. Depending on which solvent is used for the solute, the anti-solvent can be chosen in order to maximise the chance of producing a crystal. Typically around 4 mg of compound is dissolved



(a) Vapour diffusion. **(b)** Layered diffusion.

Figure 2.16 Crystallisation techniques.

in around 0.5 mL of solvent and is surrounded by roughly 2.5 mL of anti-solvent for the best results. If any more compound than this is used, inter-growth of crystals can become problematic. [149] Once set up, the two solutions equilibrate *via* vapour diffusion causing a change in the solubility of the compound, hopefully resulting in the formation of a crystal.

Generally, the anti-solvent is chosen such that it is more volatile than the solute. This means that the anti-solvent slowly diffuses into the solute in order to change its solubility and form a crystal. The anti-solvent must also be miscible with the solute. These properties can be found in solvent miscibility and viscosity tables, as found in *The HPLC Solvent Guide*. [150]

2.4.2 Layered Diffusion

Similar to vapour diffusion, the compound of interest (the solute) is dissolved in a solvent and then an anti-solvent is used to alter the solubility of the compound. Instead of diffusing the anti-solvent into the solvent, this time the anti-solvent is layered with the solute (with the solution of higher density layered beneath the other). The

solutions are then left to mix over time, often in an NMR tube or similar shaped vial to minimise contact area in order to slow down mixing (Figure 2.16b). This process is generally slower than vapour diffusion.

Chapter 3

Discrimination of Elements and their Electronic States I

The energy of an absorption edge for a given element is dependent on the electronic structure of its core. However, its exact energy will be affected by physical properties of the atom such as oxidation state, spin state, and chemical environment. It is this change in energy of the absorption edge that can be used to discriminate elements and their electronic states. Previous work by Attfield has shown the application of resonant powder X-ray diffraction to the valence contrast of Eu^{3+} and Eu^{2+} in Eu_3O_4 by refining the value of f' for a subset of reflections. [124] However, the success of this method is not always applicable in all cases. Firstly, the anomalous signal for Eu is very strong relative to both the element itself ($f' = -22.9(26)$ electrons compared to the atomic number 63 of Europium) and the structure as a whole (since Eu makes such a large proportion of the compound). Not all elements exhibit such large anomalous signals and therefore valence contrast becomes much harder. Secondly the results presented have been obtained from powder data. As the problem is applied to larger structures, solution from powder data becomes a much greater problem due to overlapping peaks.

This chapter describes how we can calculate the different absorption edge energies of an element present in more than one electronic state within a compound. We then go on to show how this can be used to refine the site specific occupancies of the different states within the compound. All of the data are collected on single crystals to provide a more widely applicable technique for electronic state contrast.

3.1 DetOx: Extraction of overlapped absorption edges

The absorption edge energy is dependent upon a number of properties of the atom under observation. During this research I have developed a program, DetOx, [1] to analyse absorption spectra of compounds where more than one oxidation state (or spin state) of the same element is present. In compounds of this type, the observed absorption-fluorescence spectrum is a sum of the spectra from the individual atoms in their different electronic states. The rapid variation of absorption coefficient as a function of energy approaching the absorption edge means that at a certain energy, the anomalous scattering (AS) factors will be significantly different for each oxidation (or spin) state present. The program, written using Python, [151] deconvolutes the absorption-fluorescence spectrum of a mixed-electronic state compound into its component spectra for each state present in the sample, allowing determination of the AS factors, f' and f''.

3.1.1 Description of Program

Fluorescence-absorption data are collected as described in Section 3.2.1. These data are then analysed using DetOx. This section provides a description of how the data are processed to calculate the AS factors for different oxidation and spin states of anomalous scatterers in the structure.

Normalisation

Raw data are firstly corrected for incident beam intensity before being normalised. This is performed by scaling the fluorescence signal at a given energy by the equivalent absorption reading from a diode placed in the beam, with the sample removed. Fluorescence data are input on an arbitrary scale, since there are many factors other than just the nature of the atomic scatterer that can contribute to the signal, including the energy of the incident X-ray beam, volume of the sample, the geometric orientation of the crystal on the diffractometer, the partial concentrations of different

oxidation state heavy-atoms in the sample and the distance of the sample from the detector. Background noise from elastically scattered X-rays and other sources also becomes an issue when making accurate measurements. [23] Therefore, the recorded fluorescence data must undergo a background correction and normalisation. The input spectrum is normalised such that below the edge, the signal tends to zero and above the edge the signal tends to one, leaving only the edge itself and features caused by EXAFS and XANES around the absorption edge.

The normalisation is achieved by fitting a straight line to all input fluorescence data (F_{raw}) more than 50 eV below the theoretical edge (Φ_{below}) and a second to all data from 50 eV above the edge (Φ_{above}) – see Figure 3.1. These lines are extrapolated across the whole range of the data and then used to adjust the raw data to leave the normalised spectrum ($N(E)$),

$$N(E) = [F_{\text{raw}}(E) - \Phi_{\text{below}}(E)] \,/\, [\Phi_{\text{above}}(E) - \Phi_{\text{below}}(E)]\,. \tag{3.1}$$

In order for the program to extrapolate accurately, data should ideally be collected to a range of ±200 eV from the theoretical absorption edge, to give reliable extrapolation of Φ_{below} and Φ_{above}. Away from the absorption edge (± 200 eV), EXAFS and XANES have a negligible effect on the fluorescence signal so it can be assumed that calculated theoretical values are an accurate approximation for the normalised spectrum. [23]

Determination of f''

Theoretical values of f'' above and below the absorption edge are obtained using the program MuCal [152] and linear functions fitted to each (f''_{above} and f''_{below}). These values are approximately correct for most energies except those in the immediate vicinity of an absorption edge. The experimental f'' spectrum is therefore calculated from the normalised data, $N(E)$, by

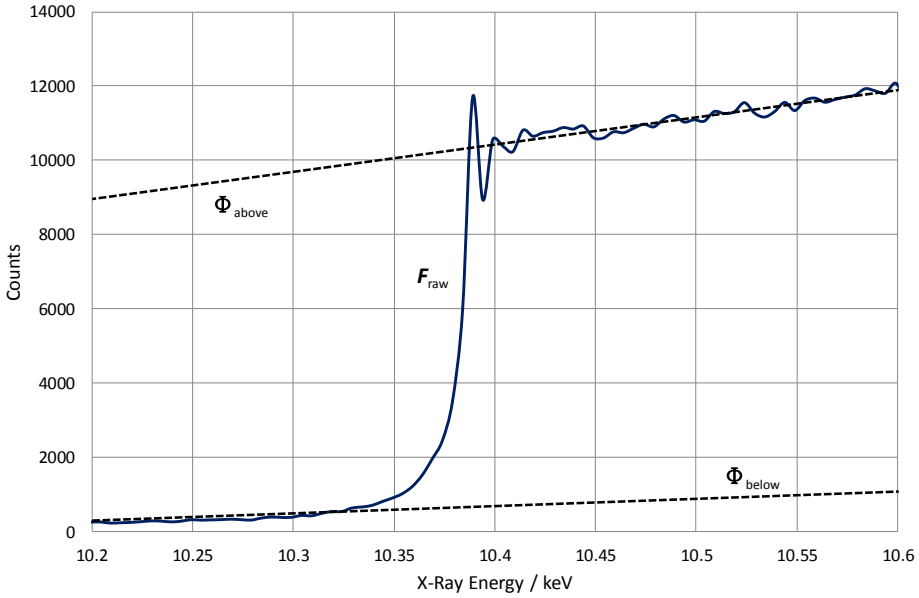


Figure 3.1 Graph to show the raw fluorescence data collected on $GaCl_2$ before normalisation. The dotted black lines represent Φ_{above} and Φ_{below} as described in Eqn. 3.1.

$$f'' = N(E)\left[f''_{\text{above}}(E) - f''_{\text{below}}(E)\right] + f''_{\text{below}}(E). \tag{3.2}$$

Deconvolution

Wilkinson and Cheetham (1992) [123] demonstrated that the $GaCl_2$ spectrum is a sum of the fluorescence spectra for the Ga^{I} and Ga^{III} species in the sample. Since it is not possible to determine exactly how the EXAFS above the edge differ between the two oxidation states, we have made the assumption that the post-absorption-edge EXAFS features do not have an effect on the position of the f' peak. We have tested this assumption by replacing the data points after the absorption edge with theoretical values and determining the related f' values. As shown in Figure 3.2, despite the post-edge EXAFS features being removed from the f'' spectrum, the position of the f' peak remains unchanged. Since we are only interested in the features of spectra before the absorption edge peak, it is not critical to determine the exact contributions to the EXAFS region from individual oxidation state species, as this will not have a significant effect on the AS factors.

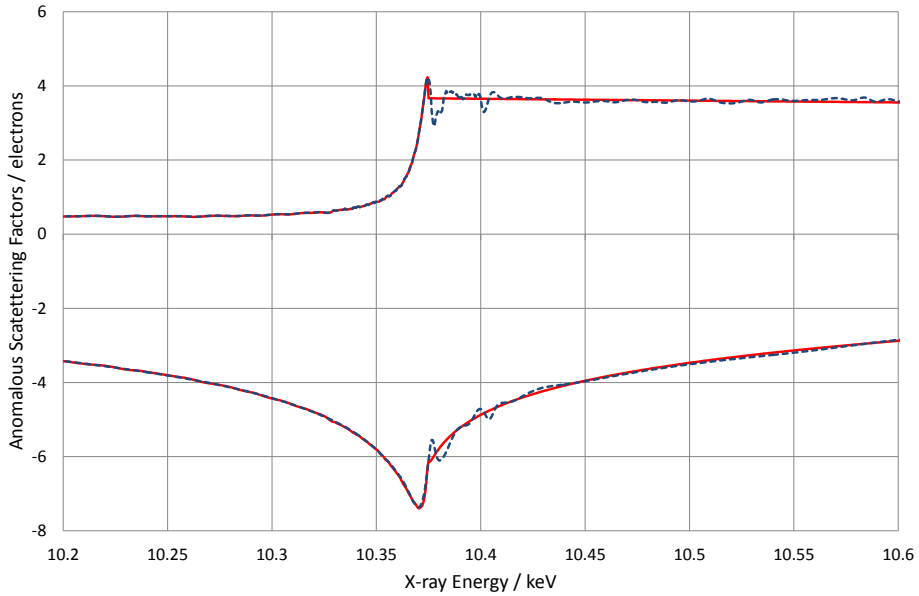


Figure 3.2 The effect on EXAFS on the peak position in the real scattering factors. The dotted blue line shows the DetOx output for the $GaCl_2$ data assuming only one oxidation state is present and the solid red line shows the deconvoluted Ga^{I} oxidation state data (with EXAFS removed).

To overcome the problem of overlapping signals in the spectrum, a Fourier deconvolution is utilised. The measured spectrum, $y(E)$, is defined as a convolution of the ideal signal, $s(E)$, and a response function, $r(E)$, plus a uniform noise function, $n(E)$. If we assume that the signal itself is the sum of two identical underlying fluorescence curves, $u(E)$ and $u(E+\Delta E)$, where one is offset by a shift of ΔE, then the signal, $s(E)$, may be expressed as a convolution $u(E) \otimes d(\Delta E)$, where $d(\Delta E)$ is the sum of two appropriately scaled delta functions separated by ΔE, $m\delta(0) + n\delta(\Delta E)$. The peaks in $d(\Delta E)$ function are scaled such that the ratio of the areas under each, m:n, is proportional to the ratio of the oxidation states present in the sample. In the case of $GaCl_2$, this ratio is simply 1:1.

$$y(E) = u(E) \otimes d(E) \otimes r(E) + n(E) \tag{3.3}$$

In order to determine AS coefficients for each species we must determine the shape of the function $u(E)$. This may be achieved using a Fourier deconvolution.

In the absence of noise, we could recover our unknown data signal by taking the reverse Fourier transform of $\mathcal{F}(y(E))/\mathcal{F}(d(E) \otimes r(E))$. In the presence of noise, $n(E)$, the optimal solution is found by applying a Wiener filter to this expression, the effect of which is to attenuate the influence of terms with a low signal-to-noise ratio in the denominator. [153]

In cases where only one oxidation state exists in the molecule, a single Gaussian delta function can be used for the response function in the above equation. The effect in this case is simply to remove some broadening of the signal caused by X-ray monochromator bandpass, detector response and other random effects. Using an iterative process, we found that the optimum width for the response function, which sharpened the signal without introducing unexpected pre-edge features, was 0.16 eV. This value is expected to be instrument specific. For mixed oxidation state cases, we employ a second iterative process to determine the optimum value of ΔE subject to the conditions that the original signal is reproduced by application of Eqn. 3.3 and no unexpected features (*e.g.* negative values) are introduced to the extracted absorption edge, $u(E)$. The output consists of two f'' curves, separated by a value of ΔE eV, that when added together reproduce the original input absorption-fluorescence signal. These curves are normalised to theoretical f'' values away from the edge to give the imaginary AS factors for the different oxidation states as described above.

Determination of f'

Firstly a Savitzky-Golay filter [154] is applied to the deconvoluted curves in order to smooth the data. The f' values are then calculated by performing the Kramers-Kronig transformation, integrating between 0 and ∞ at each energy, (E_0), in the f'' spectrum. However, the denominator of the integration ($E_0^2 - E^2$), contains a singularity, over which it is not possible to integrate. This can be overcome numerically after a Taylor expansion of the integral as described by Hoyt: [155]

$$
\begin{aligned}
f'(E_0) &= (2/\pi)\int_0^a \left[Ef''(E)/(E_0^2)\right] dE \\
&+ (2/\pi)\int_b^\infty \left[Ef''(E)/(E_0^2)\right] dE \\
&+ (L/\pi)\{\int_a^b \left[Ef''(E)/(E_0^2)\right] dE \\
&- (\ln|b-E_0| - \ln|a-E_0|) - \frac{df''}{dE}|_{E_0}(b-a) \\
&- \sum_{n=2}^{\infty} \frac{1}{(n)n!}\frac{d^n f''^n}{dE^n}(b-E_0)^n|_{E_0} - (a-E_0)^n\}
\end{aligned}
\tag{3.4}
$$

The integration is performed using the Python Quadrature module, from the Numpy and SciPy open source libraries, [156] which computes a definite integral using fixed-tolerance Gaussian quadrature. As implemented, the integration is performed between suggested limits of 0 eV and 50 times (instead of infinity) the theoretical edge energy, [155] to ensure the program runs quickly. Due to the assumption made in the deconvolution that both underlying signals are the same shape, the values of f' will be most reliable below the absorption edge, where there are no significant differences between the shapes of the two curves.

DetOx Output

The default output from DetOx is a graph showing both the f'' and f' calculated from the deconvolved signals. Other outputs from the program include: tabulated AS factors for all oxidation (or spin) states present in the sample at specified energies; a graph of the *effective* f'' spectrum before the deconvolution; the raw data for the deconvoluted spectra in plain text format.

3.1.2 Experimental

The background-corrected data are firstly input into DetOx. Using the program output, appropriate wavelengths can be chosen for data collection so as to maximise

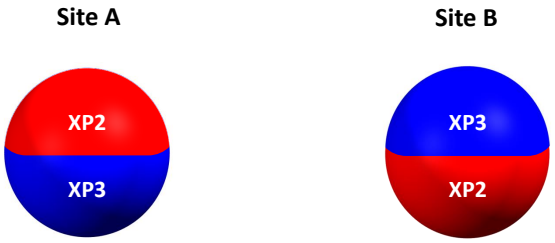


Figure 3.3 Imaginary structure containing two atomic sites each occupied by two atoms.

the differences in f'' and f' between two oxidation or spin states. DetOx then provides the calculated AS factors at the chosen wavelengths. Once the structure has been solved, the relevant AS factors are entered into CRYSTALS [3] and a competitive refinement of the two AS factors for a given site can be undertaken to unambiguously determine the electronic state of the element. We can visualise this competitive refinement by taking an imaginary structure containing two separate atomic sites (A and B) occupied by the same element (X), where the oxidation states are know to be +2 and +3, but the location of each state unknown. After the structure has been solved and refined, two atoms are placed on each ambiguous X atom site and labelled P2 ('plus 2' oxidation) and P3 ('plus 3' oxidation) *i.e.* both the A and B sites now contain an XP2 and an XP3 atom (see Figure 3.3).

We then input the AS factors which have been calculated and output from DetOx such that the XP2 atoms have the correct AS factors pertaining to the +2 oxidation state and XP3 atoms contain those for the +3 state. Before refining only the occupancies at each of these two sites, we must apply restraints on the occupancies such that the combined occupancy of the atoms at a given site is equal to 1.0. The 'list 12' * used by CRYSTALS to apply these restraints is:

*file defining the parameters to be refined in the least-squares calculation, and specifying relationships betweens those parameters

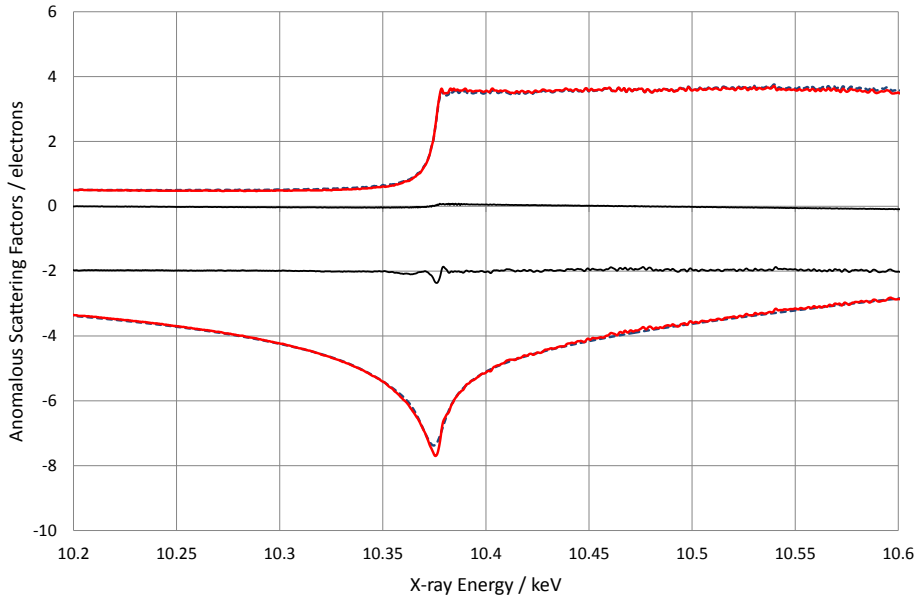
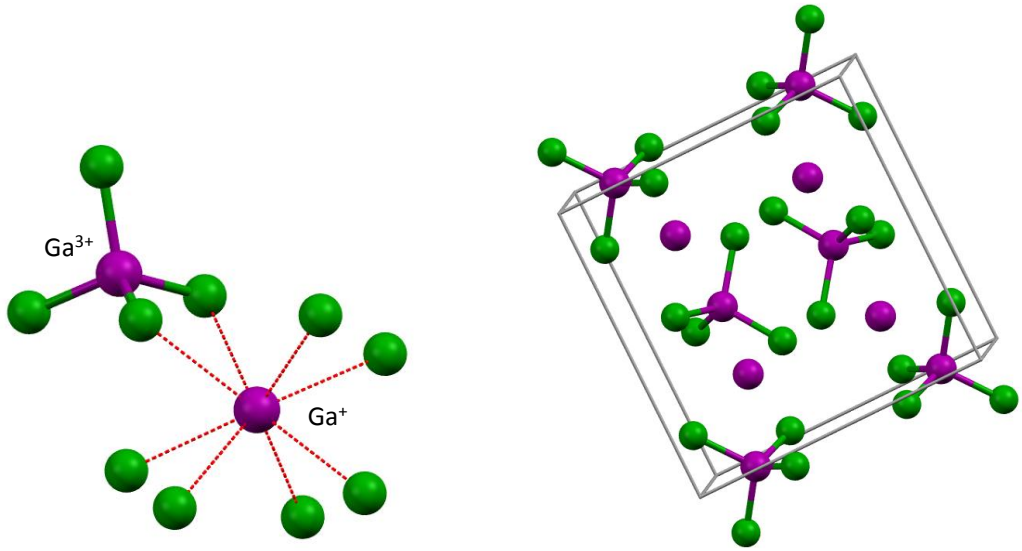


Figure 3.4 Comparisons between the output from DetOx (dotted blue line) and CHOOCH (solid red line) for a sample of $GaCl_3$. The solid black lines represent the difference traces for the f' and f'' curves.

```
#LIST     12
BLOCK SCALE
EQUIV XP2(1,OCC) XP3(1,OCC)
WEIGHT -1 XP3(1,OCC)
EQUIV XP2(2,OCC) XP3(2,OCC)
WEIGHT -1 XP2(2,OCC)
END
```

Note that the commands above restrict the refinement to just two parameters, which define the occupancies of the X atom sites and do not interfere with the rest of the structure which has already been refined. The scale factor is also refined with the occupancies to ensure the most accurate answer.

$GaCl_3$ contains just one oxidation state of Ga^{III}, and was used as a test sample for the normalisation and subsequent Kramers-Kronig transformation in DetOx. The results were verified against those from CHOOCH [23] using the same data (Figure 3.4). CHOOCH is widely used in macromolecular crystallography for calculating AS factors and is therefore a good baseline for testing the accuracy of DetOx. As can be seen from the difference trace in Figure 3.4, there is a negligible deviation in f'' and a maximum difference of 0.2 electrons at the peak in the f' trace. We can therefore



(a) (b)

Figure 3.5 Crystal structure of $GaCl_2$. Supporting information can be found in A.1.

conclude that DetOx works for calculating the AS factors for a single oxidation state atom to the same standard as the existing program CHOOCH.

3.2 $GaCl_2$ - A Mixed Oxidation State Compound

$GaCl_2$ was chosen for initial testing of the DetOx program since it has been extensively studied and the structure has been reported. [157–159] A similar experiment to exploit the anomalous signal in order to determine the oxidation state at each Ga position using powder diffraction data has been previously reported by Wilkinson and Cheetham (1992). [123] The structure consists of tetrahedral $[GaCl_4]^-$ ions and Ga^+ ions (Figure 3.5a). Each Ga^+ ion is surrounded by eight chlorine atoms from neighbouring $GaCl_4$ units in a distorted dodecahedron arrangement. The structure has an orthorhombic unit cell, space group *Pnna*, containing $4\,Ga_2Cl_4$ per unit cell (Figure 3.5b).

Dataset	Ox. State	f''	f'
On Edge	Ga(I)	2.41	−7.36
10368 eV	Ga(III)	1.88	−7.00
Below Edge	Ga(I)	0.59	−4.69
10314 eV	Ga(III)	0.59	−4.62

Table 3.1 Anomalous scattering factors output from DetOx for $GaCl_2$.

3.2.1 Experimental and AS Factor Calculations

Single crystals were obtained from Sigma Aldrich in a sealed vial under nitrogen. The crystals were immersed in dried oil under nitrogen and mounted directly in the cold stream to reduce exposure time to air and moisture.

The theoretical absorption edge of gallium is 10,367 eV. Therefore, absorption-fluorescence data were collected between 10,100 eV and 10,700 eV in steps of 5 eV to locate the absorption edge. After an initial analysis of the output from DetOx, a finer scan was performed between 10,345 eV and 10,415 eV in steps of 0.25 eV. Full X-ray diffraction data collections were then recorded at 10,314 eV (below the absorption edge) and 10,368 eV (on the Ga^I absorption edge). Using the CRYSTALS software, it was possible to determine the oxidation states of the Ga atoms in the structure. Firstly, each Ga site was set to be occupied by two atomic species, Ga^I and Ga^{III} with AS factors calculated in DetOx (Table 3.1), and with occupancies summing to one. The expectation would be that the occupancy of the atom with the Ga^I AS factors refines to one at the Ga^I site and zero at the Ga^{III} site and *vice versa* for the Ga atom with the Ga^{III} AS factors.

3.2.2 Results and Discussion

The output from DetOx is shown in Figure 3.6, and Figure 3.7 shows the deconvoluted f'' superimposed onto the normalised fluorescence spectrum of $GaCl_2$ before deconvolution. Table 3.1 reports the anomalous scattering factors, calculated by DetOx.

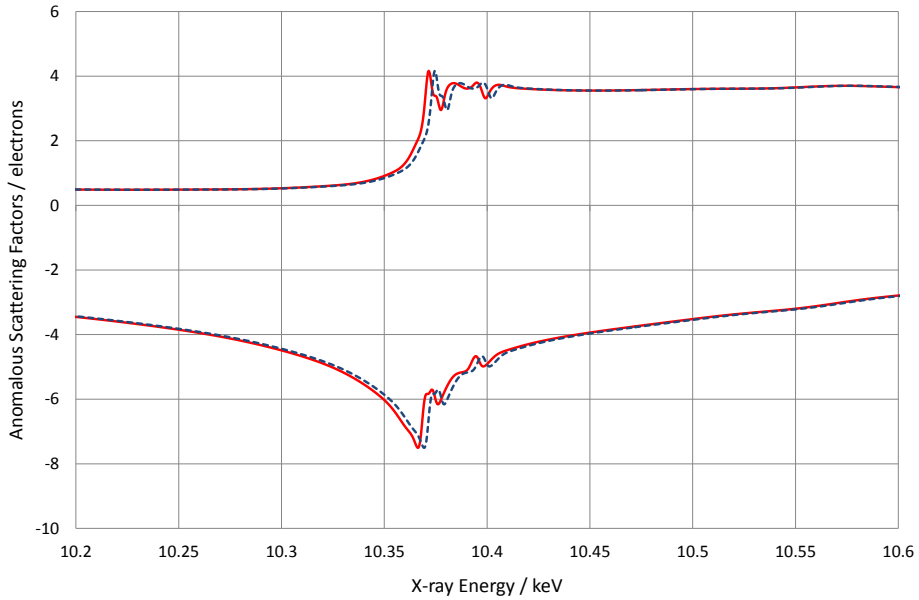


Figure 3.6 Output from DetOx for a sample of $GaCl_2$. The solid red lines show the AS factors for the Ga^{I} ions and those for the Ga^{III} are shown by the dotted blue line.

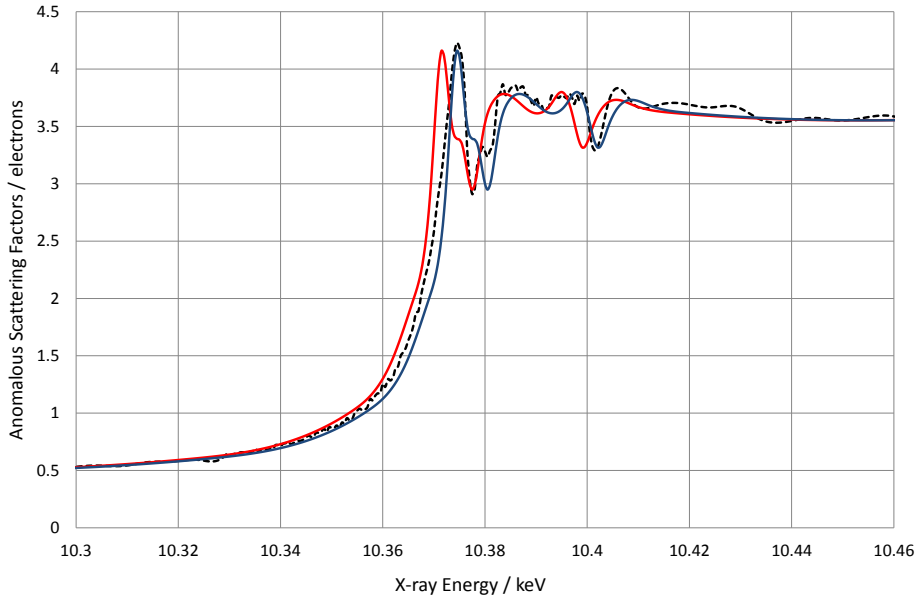


Figure 3.7 Overlay of the two deconvoluted curves (solid red (Ga^{I}) and blue (Ga^{III}) lines) superimposed onto the original spectrum for $GaCl_2$ (dotted black line).

Dataset	Ga Site	Ga(I)	Ga(III)
On Edge	Ga(I)	1.02(20)	−0.02(20)
10368 eV	Ga(III)	−0.02(20)	1.02(20)
Below Edge	Ga(I)	1.3(15)	−0.3(15)
10314 eV	Ga(III)	−0.3(15)	1.3(15)

Table 3.2 Refined $GaCl_2$ occupancies from CRYSTALS.

Structures were solved for both datasets using SIR97 [160] and all refinement and analysis was undertaken using the CRYSTALS crystallographic package. The coordinates for atoms in the crystal structure were consistent with previous single crystal studies by Schmidbaur. [158] A filter on $(\sin\theta/\lambda)^2$ (based on examination of the Wilson Plot) was added to remove the weak, high angle data. The results from the occupancy refinement are shown in Table 3.2.

The DetOx output revealed an optimum energy offset between the Ga^{I} and Ga^{III} absorption edges, consistent with that reported by Wilkinson and Cheetham (1992) from powder data, of 3.3 eV. The values in Table 3.2, show that the occupancies refine as expected, however the reliability of these values must be taken into consideration. Since the errors in occupancy for the on-edge dataset (± 0.20 electrons) are much lower than those below the edge (± 1.5 electrons), it is clear that the values from the below-edge analysis are less reliable. The reason for this is that the below-edge AS values are identical to within ± 0.05 electrons, which results in strong co-linearity for the parameters for the Ga^{I} and Ga^{III} occupancies. However, on the Ga^{I} absorption edge, where f'' varies rapidly with energy, the difference is much larger allowing the occupancies to be determined more reliably.

The accuracy of the f' data has been verified by the test spectrum shown in Figure 3.4 on $GaCl_3$. Therefore, the main source of potential error arises in the deconvolution step. Figure 3.7 shows that the deconvoluted curves align with the appropriate peaks in the original spectrum. Even so, the assumption that the curves have the same form up to the absorption peak (the section of the f'' spectrum we are interested in) is difficult to test. Nevertheless, the results presented in Table 3.2 support the accuracy

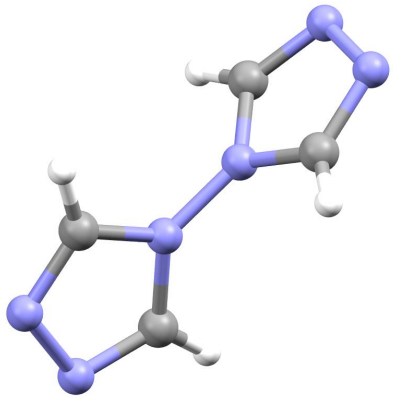


Figure 3.8 Crystal Structure of the btr ligand. Nitrogen and carbon atoms are coloured blue and grey respectively. Supporting information can be found in A.2.

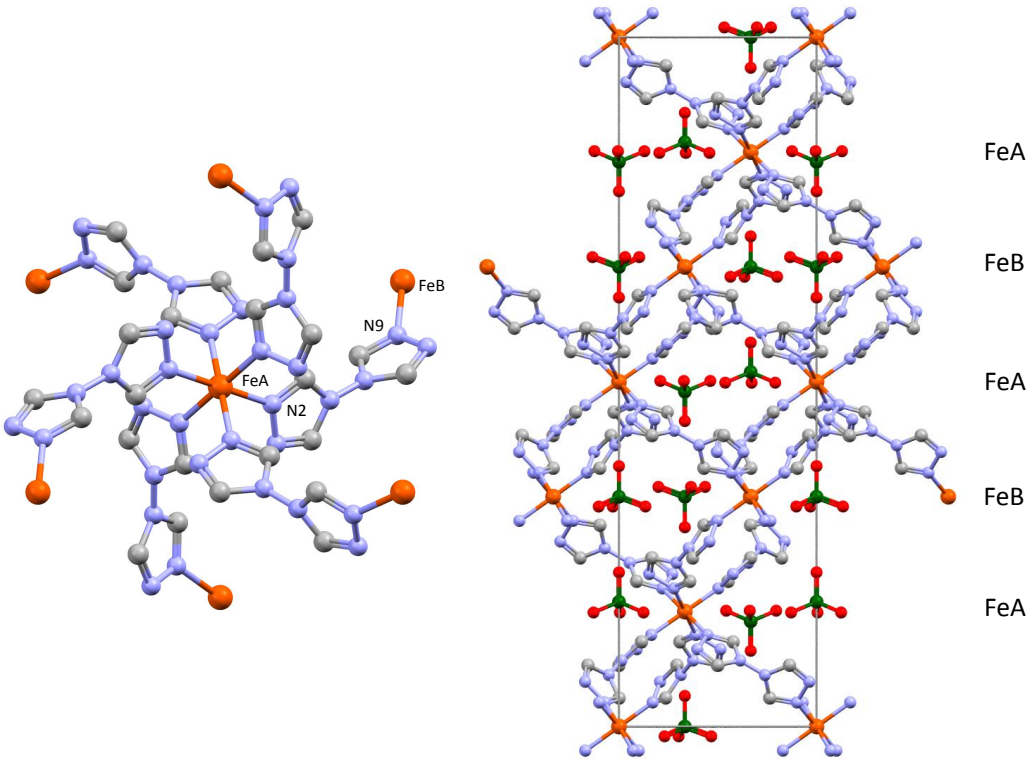
of the calculated AS factors, and indicate that this assumption is appropriate for the function of the program.

3.3 [Fe(btr)$_3$](ClO$_4$)$_2$ - A Spin Crossover Compound

Two-step SCO compounds such as [Fe(btr)$_3$](ClO$_4$)$_2$, [140] present an application of DetOx to the field. At a given temperature, these compounds exhibit bi-stability whereby half of the SCO centres are in a high spin (HS) state with the other 50% low spin (LS). Identification of the different spin states can be obtained using techniques described in section 1.4.4, however as with oxidation state discrimination, information about the location of the different states is not always trivial. This section describes the use of DetOx for the unambiguous positioning of different spin states in two-step SCO systems.

As with GaCl$_2$, [Fe(btr)$_3$](ClO$_4$)$_2$ was chosen for testing of DetOx as applied to SCO materials since it is a well studied system. [140] The compound was the first reported three dimensional SCO material and incorporates the btr ligand (Figure 3.8), which was previously reported as a ligand in other SCO compounds. [161] [162]

Figure 3.9 shows the three dimensional structure of Fe(btr)$_3$(ClO$_4$)$_2$ after reaction of the ligand with Fe(ClO$_4$)$_2$.



(a) Projection along the *c*-axis. **(b)** Projection in the *ac* plane.

Figure 3.9 Crystal structure of $Fe(btr)_3(ClO_4)_2$. Hydrogen atoms have been removed for clarity. Supporting information can be found in A.3.

3.3.1 Two-Step Spin Crossover

The two-step SCO behaviour of $Fe(btr)_3(ClO_4)_2$ can be attributed to the fact that FeA and FeB occupy two similar, but crystallographically distinct sites within the structure (Figure 3.9b). This manifests itself as slightly different iron to perchlorate interactions at the two sites. Mössbauer analysis on the complex has been reported, which reveals the two-step spin transition curve shown in Figure 3.10. The graph clearly shows two LS↔HS transitions at T_1 = 185 K and T_2 = 222 K, which correspond to the spin transitions at the FeA and FeB sites respectively. This assignment of SCO centres to transition temperatures can be shown crystallographically by observing the increase in Fe-N bond length as the centre moves from low to high spin (a consequence of populating the anti-bonding orbitals) - Table 3.3.

However, not all cases provide such a clear transition between the LS-LS, LS-

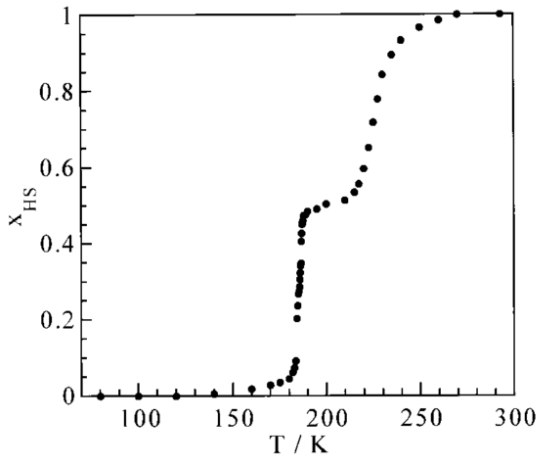


Figure 3.10 Spin Transition Curve for $[Fe(btr)_3](ClO_4)_2$, as derived from the Mössbauer spectra. [140]

	T = 130 K	T = 200 K	T = 250 K
FeA - FeB	8.459	8.566	8.670
FeA - N2	1.998(6)	2.006(6)	2.161(4)
FeB - N9	2.015(6)	2.158(6)	2.165(3)

Table 3.3 FeA-FeB and Fe-N bond lengths (Å) at different temperatures.

HS and HS-HS states. Amoore and co-workers [163] showed that upon desolvation of $[Fe_2(ddpp)_2(NCS)_4] \cdot 4\,CH_2Cl_2$ to $[Fe_2(ddpp)_2(NCS)_4] \cdot CH_2Cl_2$, the two-step SCO of the solvated compound was lost. The two-step SCO behaviour however could be regenerated upon resolvation of the compound in dichloromethane (see Figure 3.11). By measuring the X-ray diffraction between the absorption edges of the desolvated sample (curve b), at various temperatures along the SCO transition, and then refining the occupancies of each iron site as HS or LS, it should be possible to determine if the smooth transition observed is caused by both sites changing from LS to HS equally, or whether one site undergoes the transition before the second site.

3.3.2 Proof-of-Concept

The two-step SCO compound provided an excellent case for testing DetOx since it could firstly be used to test the idea that the absorption-fluorescence spectrum of a sample containing an element in more than one electronic state, is in fact a sum-

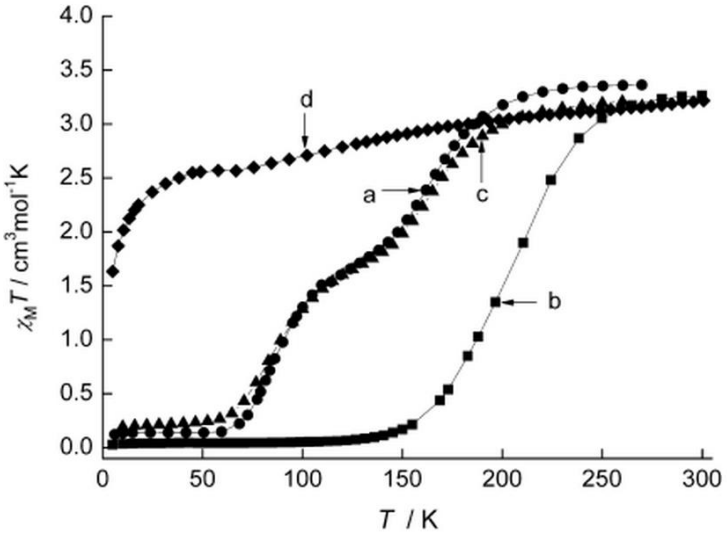


Figure 3.11 A plot of $\chi_M T$, per iron(II), *vs.* temperature for complex $[Fe_2(ddpp)_2(NCS)_4] \cdot x CH_2Cl_2$. where a) $x = 4$, b) $x = 1$, c) $x = 4$ (resolvated) and d) $x = 0$. [163]

mation of the spectra from the individual electronic states. In order to test this we collected fluorescence data from the structure at three temperatures:

1. 130K - all iron sites in the LS configuration,
2. 200K - 50:50 distribution of HS and LS iron sites,
3. 250K - all iron sites in the HS configuration.

The results from the scans (after normalisation of the data with DetOx) are shown in Figure 3.12.

The raw data from the 130 K and 250 K collections were combined and normalised to produce the purple line in Figure 3.13. When overlaid on the 200 K spectrum (dotted red line) we can see from the difference trace that the spectrum has been recreated to within experimental error (maximum difference of 0.23 electrons). We can therefore conclude that our hypothesis holds for the case of more than one spin state of the same element in a compound. This result was used to test DetOx by

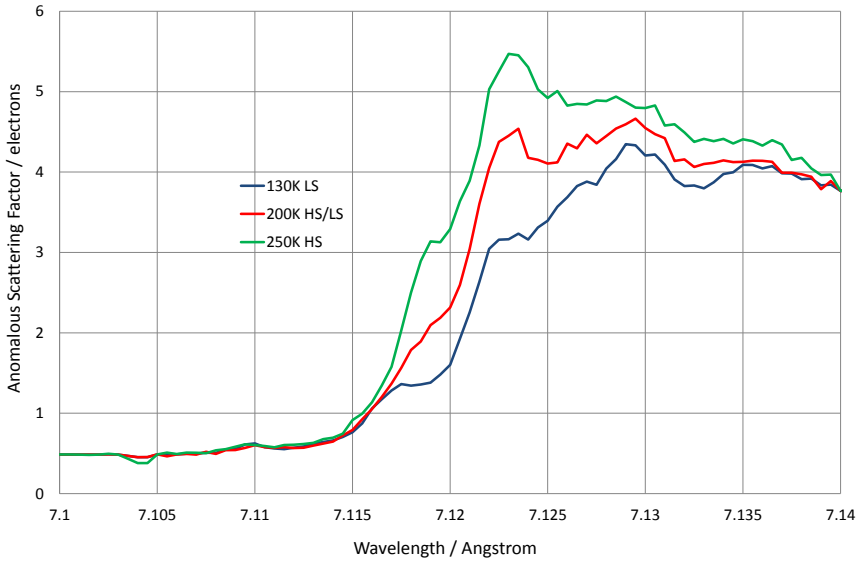


Figure 3.12 Absorption-fluorescence spectrum of $Fe(btr)_3(ClO_4)_2$ collected at 130 K, 200 K and 250 K.

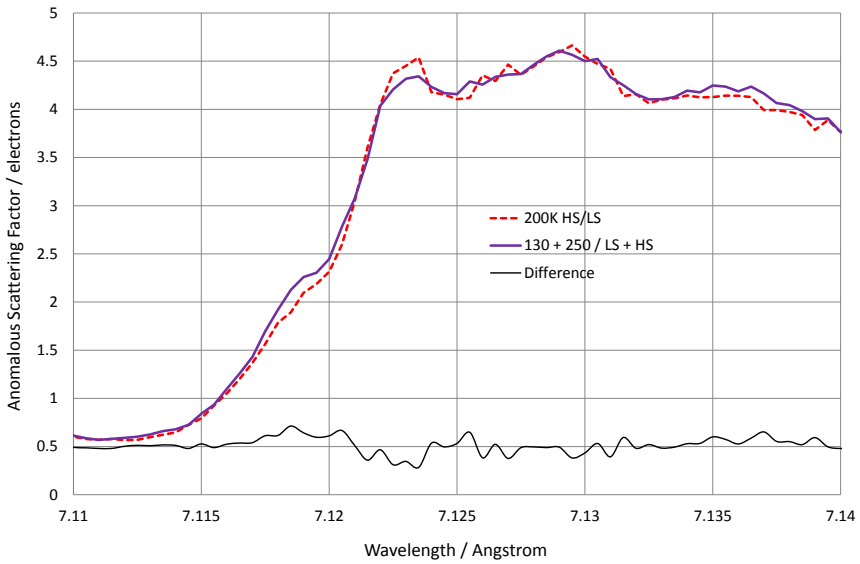


Figure 3.13 Plot of the absorption-fluorescence spectrum of $Fe(btr)_3(ClO_4)_2$ collected at 200 K *vs.* the combined spectrum from the 130 K and 250 K datasets. A difference trace between the two spectra is also shown.

seeing if it could confirm the 50:50 ratio of spin states at 200 K, without any prior knowledge of the shapes of the HS and LS spectra.

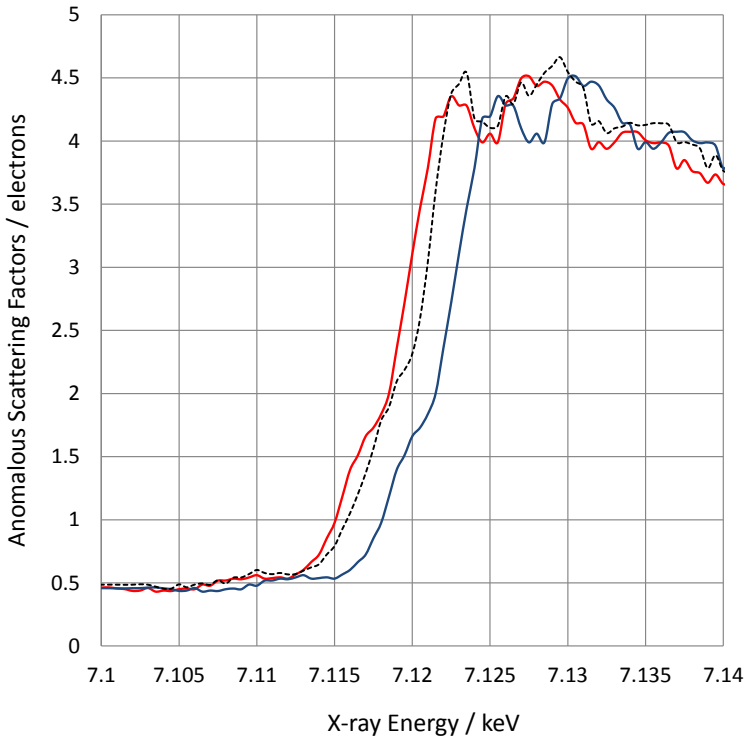


Figure 3.14 Overlay of the two deconvoluted curves (solid red (HS) and blue (LS) lines) superimposed onto the original spectrum for $Fe(btr)_3(ClO_4)_2$ (dotted black line).

3.3.3 Experimental

The synthesis of the btr ligand, and subsequently $Fe(btr)_3(ClO_4)_2$ are described in A.2 and A.3 respectively. Crystals of $Fe(btr)_3(ClO_4)_2$ suitable for X-ray analysis were grown by slow evaporation from the reaction mixture.

The data collections for this experiment were performed at 200 K for the reasons described previously. Since the theoretical absorption edge for iron lies at 7119 eV, the absorption-fluorescence scan was collected between 7105 and 7140 eV in 0.1 eV steps. After processing the data with DetOx (see Figure 3.14), datasets were collected on the edge at 7119 eV (to maximise the difference in f''), and far below the absorption edge at 7105 eV (to avoid contributions from the AS factors for comparison).

As with $GaCl_2$, atoms with the calculated AS factors (Table 3.4) for each spin state were competitively refined at each iron site using CRYSTALS in order to assign the correct spin states. The results from the experiment are discussed below.

Dataset	Spin State	f''	f'
On Edge	HS	2.36	−8.55
7119 eV	LS	1.39	−7.96
Below Edge	HS	0.45	−6.02
7105 eV	LS	0.44	−5.74

Table 3.4 Anomalous scattering factors output from DetOx for $Fe(btr)_3(ClO_4)_2$.

3.3.4 Results and Discussion

Data from the 7119 eV collection were analysed as with the $GaCl_2$ *i.e.* structure solution by direct methods with SIR97 and refinement in CRYSTALS. The optimum energy offset between the edges was found to be 3.5 eV, which is comparable to the 3.3 eV offset in $GaCl_2$.

Since we were able to collect data from the system with all of the iron sites either HS or LS (by varying the temperature), this allowed us to compare our results obtained with DetOx against the 'correct' solution. Figure 3.15 shows the overlay of the calculated HS and LS curves from DetOx onto the experimental absorption-fluorescence data collected at 130 K (LS) and 250 K (HS). The corresponding AS factors from the graph are shown in Table 3.5.

Dataset	Spin State	f''	f'
DetOx	HS	2.36	−8.55
Values	LS	1.39	−7.96
Exptl.	HS - 250 K	3.14	−8.80
Values	LS - 130 K	1.38	−7.36

Table 3.5 Calculated and experimental anomalous scattering factors for $Fe(btr)_3(ClO_4)_2$ at 7119 eV.

The only information available for the deconvolution is the experimental data recorded at 200 K. Therefore the shapes of the deconvoluted spectra do not exactly match the related experimental spectra. This difference is most pronounced at the

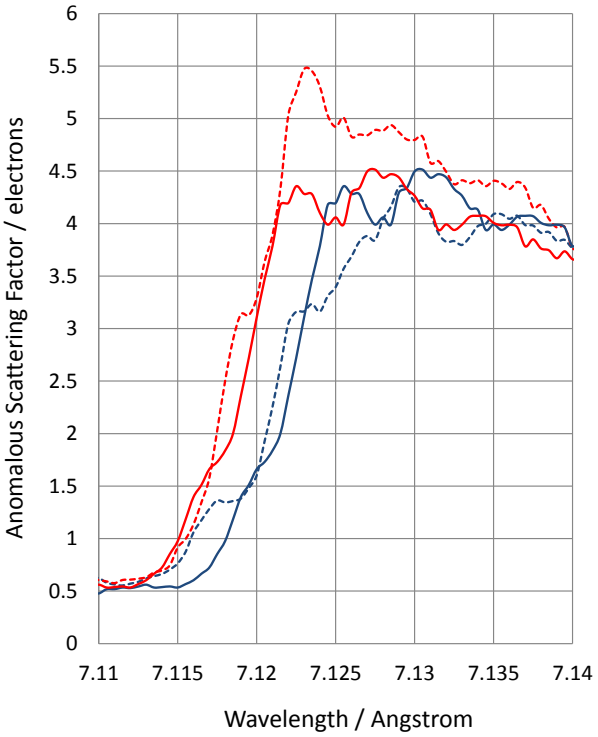


Figure 3.15 Overlay of the two deconvoluted curves (solid lines) superimposed onto the spectra collected at 130 K and 250 K for $Fe(btr)_3(ClO_4)_2$ (dotted lines). Red represents the HS spectra and blue, the LS.

peaks and EXAFS region. However, the area of interest when calculating AS factors for the competitive refinement is that leading up to the peak *i.e.* the shape and position of the absorption edge. As can be seen in Figure 3.15, the curves have much better agreement in this region, with the difference in AS factors ranging from 0.01 electrons to 0.8 electrons at 7119 eV (less than the scattering of a hydrogen atom).

The competitive refinement of the iron site occupancies was carried out as described for $GaCl_2$ by placing two partialy occupied iron atoms at each ambiguous site and refining with the different AS factors. The occupancies from the refinements are shown in Table 3.6 for both the experimental and calculated values.

The results from the competitive refinements, using both the experimental and calculated data, are shown to be in agreement to within experimental error. Table 3.6 shows that at 200 K, almost all of the FeA sites have a LS configuration and almost

Dataset	Fe Site	Fe_{HS} Occ.	Fe_{LS} Occ.
DetOx	FeA	0.07(17)	0.93(17)
Values	FeB	1.01(17)	−0.01(17)
Exptl.	FeA	0.14(15)	0.86(15)
Values	FeB	1.05(15)	−0.05(15)

Table 3.6 Iron site occupancies at 200 K using both calculated and experimental anomalous scattering factors for $Fe(btr)_3(ClO_4)_2$ at 7119 eV.

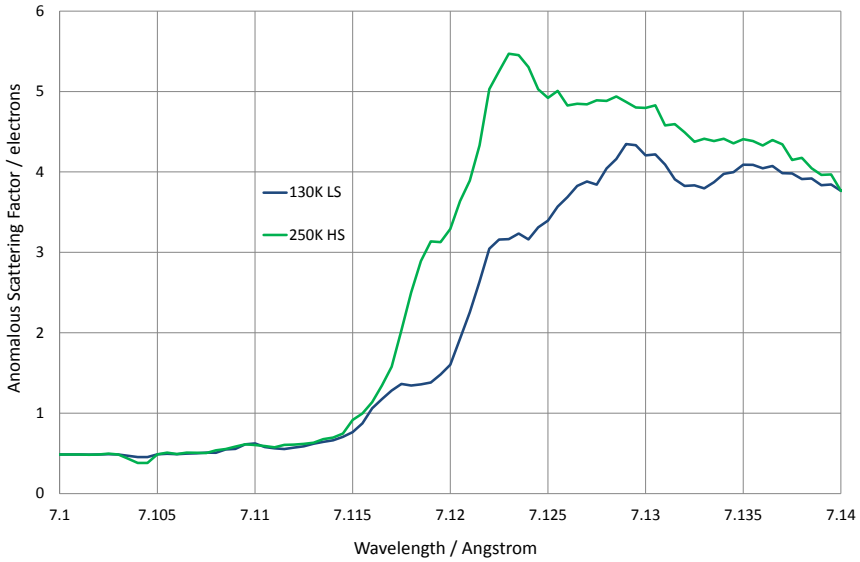


Figure 3.16 Absorption-fluorescence spectrum of $Fe(btr)_3(ClO_4)_2$ collected at 130 K and 250 K.

all of the FeB sites are in a HS state. This result is confirmed by the bond lengths surrounding each iron site. Table 3.3 showed that at 200 K the FeB – N1 is lengthened with respect to the FeA – N6, which is indicative of the FeB site being in a HS state.

The final point to draw from these results is that the ratio of HS to LS states was not calculated to be exactly 50:50. Both sets of results revealed that the FeA site had some HS character when it was expected to be 100% LS *i.e.* there is an overall higher ratio of HS states in the structure. This is reflected the 130 K spectrum (Figure 3.12). At this temperature, *all* of the iron sites are expected to be in a LS state, however the presence of a pre-edge lip in the absorption edge, suggests that there is some HS character present contributing to the spectrum (Figure 3.16).

3.4 Discussion

Sections 3.2 and 3.3 have demonstrated how DetOx can be used to discriminate between more than one electronic state of a given element within a compound. However, the examples given have only included a small number of anomalous scatterers per asymmetric unit. When the problem was applied to samples with a significantly larger number of anomalous scatterers the results proved less reliable. The major reason for this is that the anomalous signal as a fraction of the total scattering from the sample is greatly reduced as the size of the molecule is increased. This can be demonstrated using the same approach in Section 1.4.5 – calculating the ratio of anomalous signal to F_{000}. In each of the following examples, the F_{000} has been divided by Z such that we are only comparing the scattering from a single molecule. Furthermore, since we are using the method for contrasting electronic states of atoms, we are interesting in the *difference* in anomalous scattering between the two states, and therefore we calculate the anomalous contribution, AC, using

$$\mathrm{AC} = \sqrt{(\Delta f')^2 + (\Delta f'')^2}. \tag{3.5}$$

The tables below show the different ACs made by the anomalous differences between states for three compounds. The number of ambiguous sites is included since this is used as a multiplier to the AC. The number of parameters is also included since the AC must subsequently be divided by this value – generally this is equal to the number of sites, however in the case of $GaCl_2$ the two sites' occupancies are linked as a consequence of charge balancing and so the number of parameters is halved.

The three tables reveal that as we move to compounds containing more anomalous scatterers, the difference in anomalous signals as a fraction of the total scattering decreases. With the Mn_6 cluster, although the signal is still greater than that shown in Section 1.4.5 to refine the Flack parameter, we were still unable to refine the occupancies of the different oxidation states. The reason for this is that we have increased the number of parameters being refined six-fold compared to a single Flack parameter.

$GaCl_2$		
Ox. State	f''	f'
Ga(I)	2.41	−7.36
Ga(III)	1.88	−7.00
Difference	0.53	0.36
AC	0.64	
# sites	2	
# param.	1	
F_{000}/Z	130	
AC %	0.98%	

$Fe_2C_{12}H_{12}Cl_2O_8N_{18}$		
Spin State	f''	f'
HS	2.36	−8.55
LS	1.39	−7.96
Difference	0.97	0.59
AC	1.14	
# sites	2	
# param.	2	
F_{000}/Z	334	
AC %	0.34%	

$Mn_6(O)_2(O_2CCMe_3)_{10}(C_4H_8O)_4$		
Ox. State	f''	f'
Mn(II)	3.08	−8.21
Mn(III)	1.55	−8.00
Difference	1.53	0.21
AC	1.54	
# sites	6	
# param.	6	
F_{000}/Z	876	
AC %	0.18%	

This, along with the very small differences, leads to a less stable refinement process. We also run into the problem of increased absorption as the number of anomalous scatterers increases, which introduces noise into the data. When dealing with such small signals, any additional noise in the data will also act to destabilise the refinement.

Chapter 4 discusses a new ratio refinement method used not only to overcome the problem of absorption, but also to try and extract more information from the tiny signals available. We also explore the use of simulated annealing as an alternative to least-squares to avoid any local minima during the refinement process.

Chapter 4

Discrimination of Elements and their Electronic States II

This Chapter describes a new approach to discriminating between elements and their electronic states. The theory remains the same, however the new method uses a ratio refinement to overcome some of the limitations of the procedure described in Chapter 3 associated with noisy data and high absorption.

4.1 Ratio Refinement

Chapter 3 described a method for discriminating between electronic states of an element when more than one was present in a given sample using data collected at only one wavelength. As the problem was extended to systems containing more than two heavy atoms with ambiguous electronic states, the drop in data quality (due to high absorption and lower resolution limit) became an issue leading to high standard uncertainties on occupancy values and an unstable refinement procedure.

4.1.1 Implementation

The ratio refinement method described here uses 3 datasets in order to provide a more robust refinement of the occupancies. The first was a lab dataset collected using Cu K_α X-rays, which provides a good quality model for the refinement. Two further datasets were collected at the synchrotron utilising the tunable wavelength capability. The first of these was collected between the deconvoluted absorption edges as determined using DetOx. The second dataset was collected below both absorption edges where the effects from the anomalous scattering are minimal, but close enough in energy to avoid significant scaling problems between the two datasets.

The idea of taking ratios between datasets has been previously reported by Vorontsov

et al. in the context of time-resolved photocrystallography experiments. [164] In this case, ratios between datasets collected with the sample illuminated (ON) and not illuminated (OFF) at the same wavelength are used, observing the difference in the diffraction pattern between the two conditions. Here we adapt this idea to the case where anomalous scattering differs significantly between diffraction measured at the edge and below the edge (*i.e.* analogous to the ON and OFF states).

Data Integration and Corrections

Once the data had been collected they were integrated using the CrysAlisPro software. For a given dataset three corrections were applied to the integrated data in the software before outputting a .hkl file. The first of these was a frame scale to account for slight changes in beam intensity between frames. The beam intensity from the synchrotron is not completely stable with time and therefore a reference frame is chosen, to which all of the other frames are scaled.

The second data correction is an absorption correction. Although the idea of using a ratio refinement is to minimise the influence of absorption, a correction was still applied to further improve the data quality. The program applies an empirical absorption correction to the data based on the calculated μ value (see Section 3.4).

The final correction to the data was the application of an absolute scale factor. This is required as a consequence of the output format of the .hkl file. The column containing the intensities has a maximum width of 7 characters, so that the maximum value it can take is '999999.'. Therefore, before writing the file, CrysAlisPro took the highest measured intensity value and applied a scale factor to the whole dataset to ensure the maximum value of 999999.0.

Scaling Between Datasets

As described in Section 2.1, as the energy of the incident beam is changed the ID gap of the undulator must also change in order to maximise the flux of the beam. As a

result, the flux at the two wavelengths used for calculating the ratios is not equivalent. We must therefore scale the two datasets *before* calculating the ratios with Eqn. 4.1. This scaling is performed using simulated data calculated from the model in CRYSTALS. Figure 4.1 shows plots of intensities for equivalent reflections at the two different wavelengths. Using a line of best fit, the ratio between the two is calculated as the gradient of the line: 0.9017 for the calculated data and 0.9816 for the observed data. The observed data were then scaled so that the two gradients were equal.

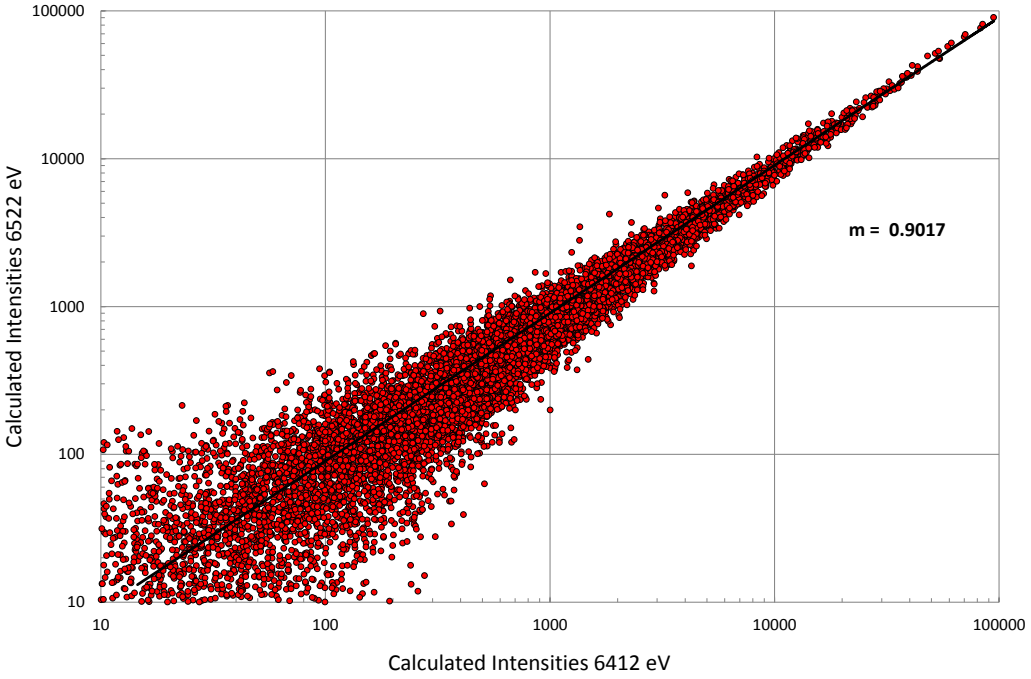
Although the calculated data rely on the exact knowledge of the oxidation state at each ambiguous location in the structure, tests have shown that the changes in intensity caused by the incorrect assignment of states has a negligible effect on the scaling compared to the change in wavelength.

Calculating the Ratios

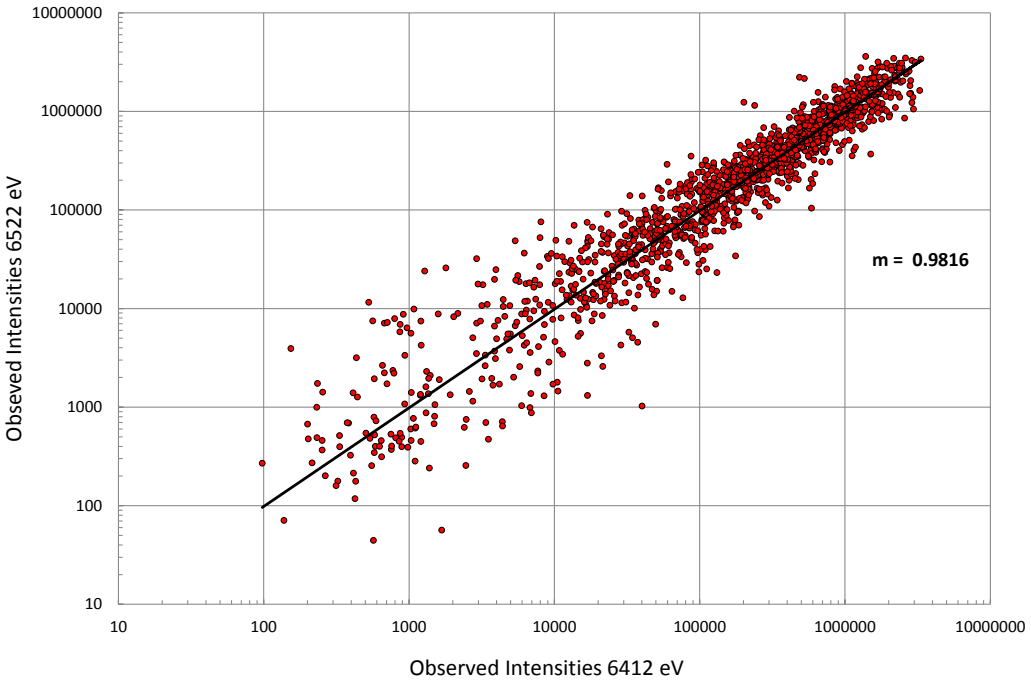
The equation used to take ratios of the difference between two synchrotron datasets (edge and below),

$$\eta_{\text{edge/below}}(\boldsymbol{h}) = \frac{F^2_{\text{edge}}(\boldsymbol{h}) - F^2_{\text{below}}(\boldsymbol{h})}{F^2_{\text{edge}}(\boldsymbol{h}) + F^2_{\text{below}}(\boldsymbol{h})} = \frac{I_{\text{edge}}(\boldsymbol{h}) - I_{\text{below}}(\boldsymbol{h})}{I_{\text{edge}}(\boldsymbol{h}) + I_{\text{below}}(\boldsymbol{h})}, \tag{4.1}$$

is a development of that reported by Vorontsov *et al.*, [164] where F and I are the respective structure factors and intensities of a given reflection. The ratios are calculated using the raw intensity data processed in CRYSALISPRO *i.e.* before merging but after scaling. The reason for doing this is that for a given diffraction spot, the ratio of its intensity below and on the edge should be the same for all other redundant measurements of the same reflection. This helps to overcome the problem of absorption effects and other random noise in the data affecting the ratio for a given reflection. All of the ratios for a given reflection are then merged before use in the refinement process.



(a)



(b)

Figure 4.1 Graphs to show the scaling between datasets using **(a)** calculated data and **(b)** observed data.

Ratio Refinement

Once the structure had been solved and refined using the lab dataset, the occupancies were constrained in the same manner as the experiment described in Section 3.1.2. However, instead of only refining the occupancies, this time the 'list 12' instructions were modified such that the whole structure was refined each time:

```
#LIST     12
BLOCK SCALE X'S  U'S
EQUIV XA(1,OCC) XB(1,OCC)
WEIGHT -1 XB(1,OCC)
EQUIV XA(2,OCC) XB(2,OCC)
WEIGHT -1 XA(2,OCC)
END
```

All of the refinements were performed on F rather than the typical F^2 since this meant that no modification to the code was required in order to produce the partial differentials of the parameters. The major difference between the two techniques however, is the data used during the refinement process. Previously the least-squares error function being minimised contained only the structure factors from the dataset collected between the edges (as described in Section 1.2.3),

$$S_m = \sum_{\boldsymbol{h}} w(\boldsymbol{h}) \mid F_{\text{obs}}(\boldsymbol{h}) - F_{\text{calc}}(\boldsymbol{h}) \mid^2 . \tag{4.2}$$

With the new method however, instead of only refining the occupancies of the unknown atoms against one dataset, we instead refined the whole structure against *both* the lab dataset *and* the newly calculated ratio values of η. The least-squares error function to be minimised can therefore expressed as

$$S_{m+\eta} = \underbrace{\sum_{\boldsymbol{h}} w_\eta(\boldsymbol{h}) \mid \eta_{\text{obs}}(\boldsymbol{h}) - \eta_{\text{calc}}(\boldsymbol{h}) \mid^2}_{\lambda_{\text{edge}},\lambda_{\text{below}}} + \underbrace{\sum_{\boldsymbol{h}} w_F(\boldsymbol{h}) \mid F_{\text{obs}}(\boldsymbol{h}) - F_{\text{calc}}(\boldsymbol{h}) \mid^2}_{\text{Lab}} . \tag{4.3}$$

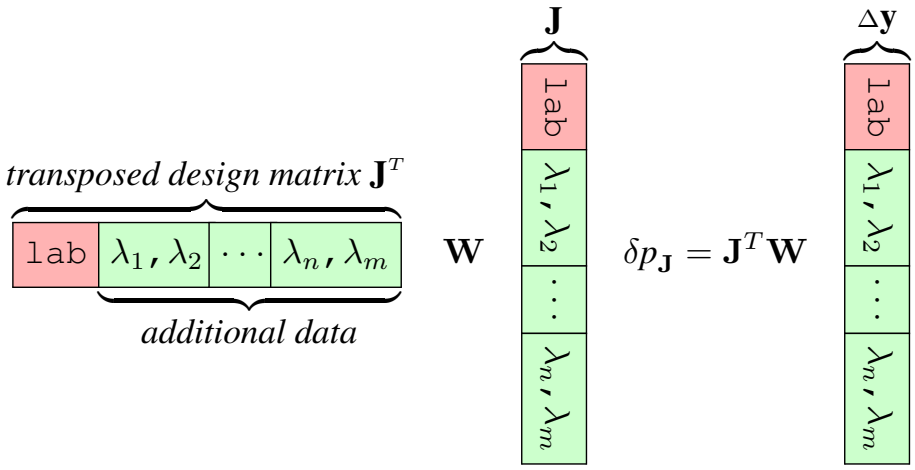


Figure 4.2 Normal equations used during the least-squares refinement.

In terms of implementing this modification into the CRYSTALS code, the additional data collected from the new wavelengths (used to produce the ratio values) are added as extra lines in the matrices in the normal equations – see Eqn. 1.59. Figure 4.2 represents the general addition of data in the normal equations for clarity since more than two additional datasets may be used in the refinement. Additionally, parameters were optimised using a simulated annealing fit to the ratio part of the same minimisation function (Equation 4.3).

4.2 Simulated Annealing

"Simulated annealing is a probabilistic method proposed in Kilpatrick, Gelett and Vecchi (1983) [165] and Cerny (1985) [166] for finding the global minimum of a cost function that may possess several local minima." [167]

The simulated annealing algorithm samples random permutations in the variables and either accepts or rejects a move based on a probability function. This method has advantages over other minimisation techniques since it avoids getting trapped in local minima by occasionally accepting unfavourable moves. The technique is analogous to the physical process of slowly cooling (or annealing) a sample such that the end result is the lowest possible energy configuration. By slowly decreasing the number of 'wrong' moves accepted, the algorithm will converge on the

'best' solution.

We can summarise the algorithm in the following steps:

1. Generate a random solution
2. Calculate its cost, κ, using a pre-defined cost function
3. Generate a neighbouring solution by introducing a random permutation
4. Calculate the cost of the new solution
5. Compare the solutions:
   - If $\kappa_{\text{new}} < \kappa_{\text{old}}$: accept move to the new solution
   - If $\kappa_{\text{new}} > \kappa_{\text{old}}$: accept move based on a probability function
6. Repeat steps 3-5 for a set number of iterations or until an acceptable convergence is reached.

In our model, the variables are the occupancies for different species at ambiguous sites within the structure. For example, in the case of a mixed oxidation state compound, random occupancies are applied to the different states of an atom at a given site in the molecule. The cost is then calculated using the minimisation function described in Equation 4.3. One of the occupancies is then changed randomly and a new cost is calculated. It is worth noting at this point that the sum of occupancies at a given site are constrained to be 1.0. If the new cost is lower than the old value the move is accepted. If, however, the new cost is less favourable, the move is accepted according to the acceptance probability, α, where

$$\alpha = e^{-\frac{\kappa_{\text{new}} - \kappa_{\text{old}}}{\mathrm{T}}}. \tag{4.4}$$

In Equation 4.4, T is the temperature, which follows a cooling regime as a function of the number of iterations. The value of T was lowered from 1 to 0 in two linear stages depicted in Figure 4.3; the second stage for the final 10% of iterations has T = 0.

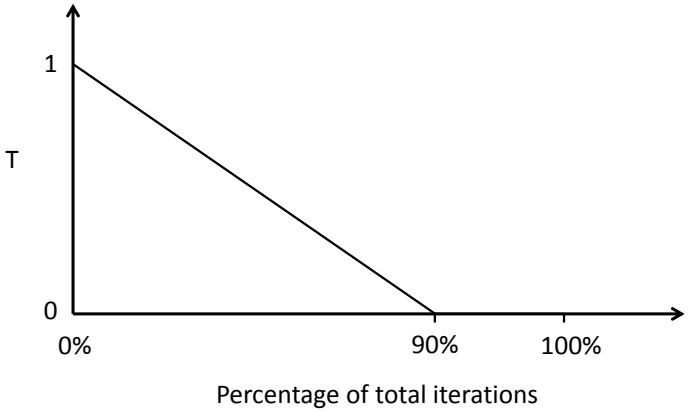


Figure 4.3 Simulated annealing cooling regime.

The form of the cooling regime combined with the acceptance probability means that when T is close to 1, the chance of accepting a bad move is higher than when T→ 0. The second stage of the cooling regime means that towards the end of the algorithm, only 'downhill' moves are accepted. To ensure the best results, it is suggested that the differences in the cost function values should be comparable to T. Since the differences in our data are so small, we apply a scaling factor to the minimisation function to make the difference in κ values comparable to the temperature, T.

4.3 Single Molecule Magnets

Single Molecule Magnets (SMMs) are a class of organic/inorganic materials containing metal clusters that exhibit magnetic properties below a certain blocking temperature. This type of magnetism on a molecular scale was first reported in 1993 on the structure, $Mn_{12}O_{12}(O_2CCH_3)_{16} \cdot 4\,H_2O \cdot 2\,CH_3CO_2H$ (Figure 4.4 – Synthesis can be found in Section 2.3.3 and crystallographic data in A.4), more commonly referred to as Mn_{12}-Ac. [141, 168] Since then, there has been significant interest in this area, due to the potential application to store data at a molecular level. [169]

4.3.1 Origin of Magnetism

The origin of magnetism in SMMs lies in the arrangement of electronic spin at the atomic level. Using high-frequency electron paramagnetic resonance (HFEPR), Mn_{12}-

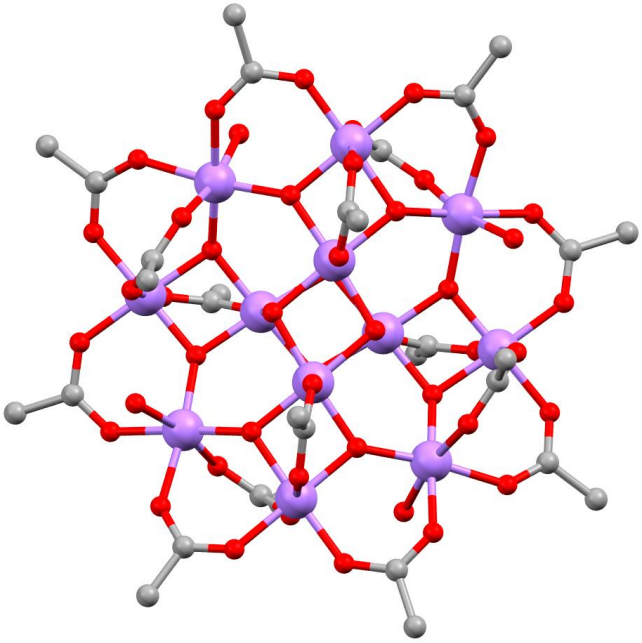


Figure 4.4 Crystal structure of $Mn_{12}O_{12}(O_2CCH_3)_{16}\cdot 4H_2O\cdot 2CH_3CO_2H$. Hydrogen atoms have been removed for clarity. Colour Scheme: Mn purple, O red and C grey.

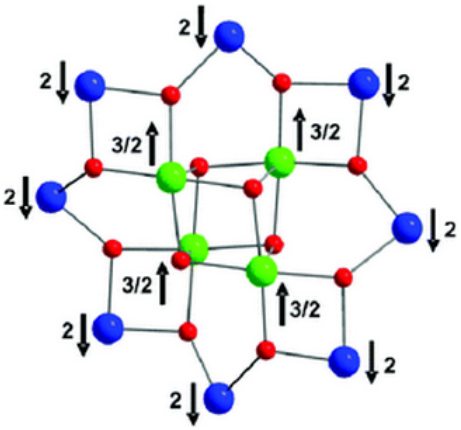


Figure 4.5 Spin alignments in the ground state of Mn_{12} complexes that gives their S = 10 ground states. Colour scheme: Mn^{IV} green, Mn^{III} blue, O red. [171]

Ac has been shown to have a non-zero ground state of S = 10. This large spin ground state is a result of the antiferromagnetic interactions between the Mn^{IV} (S = 3/2) and Mn^{III} (S = 2) ions, [170] *i.e.* the 4 Mn^{IV} ions in the central Mn/O cubane are aligned 'spin down' and the surrounding 8 Mn^{III} are aligned 'spin up' (Figure 4.5).

The single-ion zero-field splitting experienced by the Mn^{III}, leads to magnetic anisotropy of the molecule. This in turn splits the S = 10 grounds state into 21 levels,

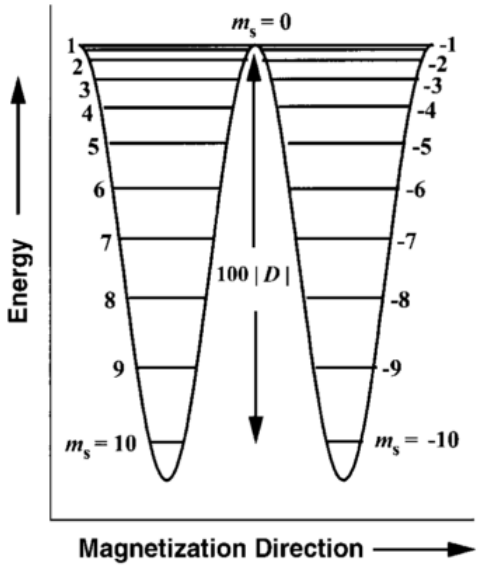


Figure 4.6 Plot of the potential energy *vs.* the magnetisation direction for a single molecule magnet with an S=10 ground state split by axial zero-field splitting. [173]

characterised by a spin quantum number, m_s, where $-S \leq m_s \geq S$. [172] The axial zero-field splitting parameter for Mn_{12}-Ac, D = -50 cm^{-1} ($\equiv -0.70$ K) provides a potential-energy barrier between the $m_s = -10$ ('spin up') and $m_s = 10$ ('spin down') orientations of the spins in an individual Mn_{12} molecule (Figure 4.6). [173]

As a result of this barrier, in order to flip the spin state in the molecule from along the $+z$-axis to along the $-z$-axis, an energy barrier of $E(m_s = 0) - E(m_s = 10) = 100\ D \approx 70\ K$, must be overcome. Herein lies the origin of the magnetism, since the time for reorientation of magnetism scales with the height of the energy barrier. It has been shown that if Mn_{12}-Ac is magnetised using an external magnetic field at 2 K, the relaxation of the spins is such that after 2 months, magnetisation is still at 40%. It is therefore concluded that the magnetism is a property of the individual molecule rather than long range interactions in the bulk sample, thereby meriting the name *single molecule magnet*. The current 'record' for a SMM is $Mn_6O_2(Etsao)_6(O_2CPh(Me)_2)_2(EtOH)_6$ {Et-saoH$_2$ = 2-hydroxyphenylpropanone oxime}, which has a spin ground state of S = 12, and an energy barrier to spin reorientation of ≈ 86 K (blocking temperature ≈ 4.5 K *cf.* ≈ 2.1 K for Mn_{12}-Ac).

4.3.2 Justification

SMMs were chosen as a system to study since the origin of their magnetism, for which they are of such interest, lies in the size of the ground state. This in turn is directly linked to the mixed-oxidation state nature of the structures and the interactions between the spin states of the different centres. By further understanding the absolute distribution of different oxidation states within the molecule, it is possible to design new SMMs to increase the size of the spin ground state and thereby raise the barrier to spin reorganisation. This crucial step is required in order to synthesise an SMM that has applications at or close to room temperature.

Although most SMMs are manganese based, a number mono- and polynuclear metal complexes based on Ni, [174] Co, [169] V [175] or Cr [176] have been reported. Due to the interesting magnetic and electronic properties of these individual ions, the incorporation of two or more into a molecule provides the possibility of control/design over the overall molecular properties. In particular, compounds which couple the large magnetic anisotropy of Co^{II} with the large spin of Mn^{II} have been studied, with the aim to produce larger anisotropy barriers and thereby increase blocking temperatures. In addition, by controlling the synthesis of heterometallic complexes, the possibility of further applications can be explored *e.g.* improvement of functionality *via* the combination of magnetic and optical properties.

In these heterometallic cases, the two different *3d* ions both have similar atomic numbers and do not display vastly different coordination behaviour. Therefore, more advanced techniques, such as multi-wavelength X-ray diffraction across the K-edge absorption resonances of these ions, are required. This unambiguous determination of metal site occupancies is vital for the understanding of how combination of metal ions controls the magnetic properties in these structures.

4.4 Proof of Concept - simulated data

In order to evaluate the reliability of the new method, two test cases were studied. The first was a mixed oxidation state manganese compound and the second, a heterometallic compound containing metals of near identical scattering (manganese and cobalt). Lab datasets collected using Cu K_α radiation were used to provide a model structure for refinement. Using the model, simulated datasets could be calculated for any chosen wavelength in CRYSTALS. The AS factors applied to the anomalous scatterers in the model were calculated with DetOx using experimental data collected on beamline I19. The calculated structure factors, F_{calc}, were then used in Eqn. 4.1 to generate ratio values for each reflection. A new .hkl file could then be read into a modified version of the CRYSTALS software such that the least-squares error function being minimised was that shown in Eqn. 4.3. The results from the simulated ratio refinements are reported below.

4.4.1 Mixed Valence Discrimination

The ratio refinement method was first tested on the hexanuclear mixed-valent compound, $Mn_6(O)_2(O_2CCMe_3)_{10}(C_4H_8O)_4$ (synthesis can be found in Section 2.3.4 and the crystallographic information in Appendix A.5). The structure consists of an edge-sharing bitetrahedal cage bridged by two μ_4-oxides, six 1,3- and four 1,1,3- pivalate ligands. Four further THF ligands bind to the outer Mn atoms in the cage. [142] Bond valence studies and charge balancing reveal that the cage contains a mixture of $4Mn^{II}$ and $2Mn^{III}$ centres (see Figure 4.7).

Absorption fluorescence data were collected at beamline I19 across the Mn edge between 6340 and 6370 eV and analysed with DetOx to calculate the AS factors for each oxidation state. The calculated AS factors from experimental data are shown in Table 4.1.

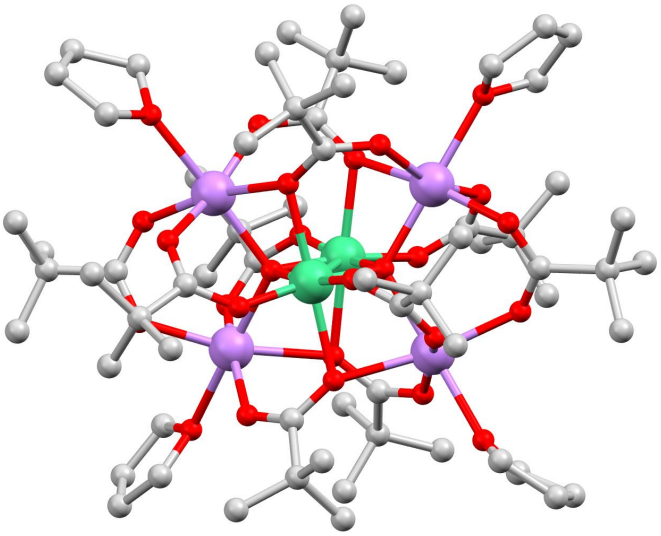


Figure 4.7 Crystal structure of $Mn_6(O)_2(O_2CCMe_3)_{10}(C_4H_8O)_4$. Hydrogen atoms and solvent molecules have been removed for clarity. The different oxidation states of Manganese have been coloured differently for clarity; Light Purple = Mn^{II}, Light Green = Mn^{III}.

Dataset	Ox. State	f'' / electrons	f' / electrons
Between Edges	Mn(II)	3.08	−8.21
6545 eV	Mn(III)	1.55	−8.00
Below Edge	Mn(II)	0.48	−3.89
6450 eV	Mn(III)	0.48	−3.86

Table 4.1 Anomalous scattering factors output from DetOx for Mn_6 cluster.

The occupancies at each manganese site were then competitively refined in CRYSTALS using the calculated AS factors for Mn^{II} and Mn^{III} species. The results from the refinement are reported in Table 4.2, and Figure 4.8 shows the $Mn_6(\mu_4 - O)_2$ core of the molecule with each Mn site labelled. Only four of the six manganese sites are involved in the refinement (and therefore are labelled) since the final two sites are symmetry-related to Mn1 and Mn2 through a centre of inversion.

The results presented in Table 4.2 were shown to be in agreement with the simulated data – calculated using expected valencies as suggested by Murrie. [177] The

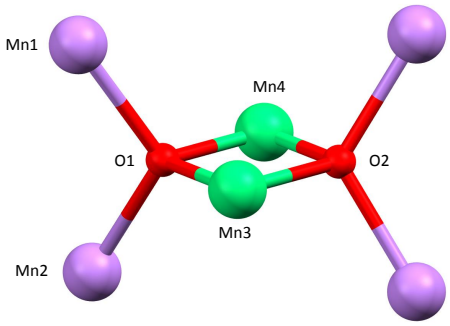


Figure 4.8 Manganese substructure of Mn_6 cluster. The different oxidation states of Manganese have been coloured differently for clarity; Light Purple = Mn^{II}, Light Green = Mn^{III}.

Mn Site	Mn^{II} Occ.	Mn^{III} Occ.
Mn1	1.003(2)	−0.003(2)
Mn2	0.997(2)	0.003(2)
Mn3	0.000(2)	1.000(2)
Mn4	0.002(2)	0.998(2)

Table 4.2 Manganese site occupancies as calculated from a ratio refinement in CRYSTALS using simulated data.

Mn···O bonds that surround the two μ_4-O atoms (O1 and O2 shown in Figure 4.8) have two forms: the shorter Mn···O bond lengths from the central manganese sites (Mn3 and Mn4) are consistent with the higher +3 oxidation state of manganese, whereas the longer Mn···O bond lengths to the outer manganese sites (Mn1 and Mn2) are consistent with the +2 oxidation state. Further evidence for the +3 oxidation states at the vertices of the edge-shared tetrahedra is the presence of a Jahn-Teller distortion perpendicular to the [Mn3, Mn4, O1 and O2] plane.

4.4.2 Near Identical Atom Discrimination

Following the proof of concept for mixed valence species, the method was extended to a heterometallic compound containing atoms of near identical scattering abilities. The compound that was chosen as the test case was the mixed metal cluster, $Co_4{}^{III}Mn_5{}^{II}Cl_6(C_8NO_5H_{15})_2(C_8NO_5H_{16})_2(H_2O)(OH)_2$ (see Figure 4.9 – crystallographic

data are provided in A.7 and synthesis in Section 2.3.5), containing manganese and cobalt atoms (atomic numbers 25 and 27 respectively). This molecule is part of a series of mixed cobalt/manganese species, which are of interest due to their analogy with SMMs. Although such compounds do not strictly exhibit the characteristics of a SMM, a clear understanding of the interactions between heterometallic sites would provide valuable insight for future design and synthesis of new SMMs.

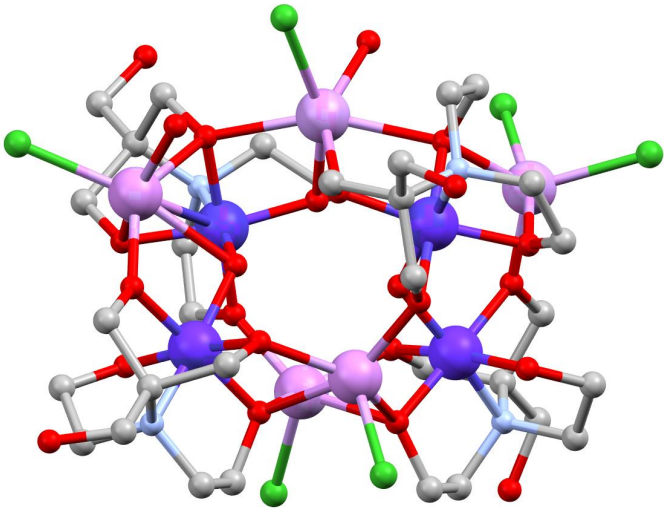


Figure 4.9 Crystal structure of $[Co_4^{III}Mn_5^{II}]$ cluster. Hydrogen atoms and solvent molecules have been removed for clarity. Dark purple = Co, Light purple = Mn and Green = Cl.

Within these structures, the exact locations of the different metals are unknown. In addition to the two metals being indistinguishable by conventional in-house X-ray crystallography experiments, they also adopt very similar coordination geometries. Therefore identification of the atomic species at a given site to date has only be inferred by consideration of charge balance and bond valence sum calculations.

As before, the ratio refinement against simulated diffraction data was performed using AS factors calculated from absorption fluorescence data collected at beamline I19. Two scans were performed, this time across each absorption edge: between 6340 and 6740 eV for the manganese edge and between 7460 and 7960 eV for the cobalt edge. These data were analysed using DetOx and the calculated AS factors are

Dataset	Metal	f'' / electrons	f' / electrons
Below Mn	Mn	0.31	−3.13
Edge 6340 eV	Co	0.68	−1.49
On the Mn	Mn	2.49	−9.28
Edge 6545 eV	Co	0.65	−1.64

Table 4.3 Anomalous scattering factors output from DetOx for $[Co_4^{III}Mn_5^{II}]$ cluster.

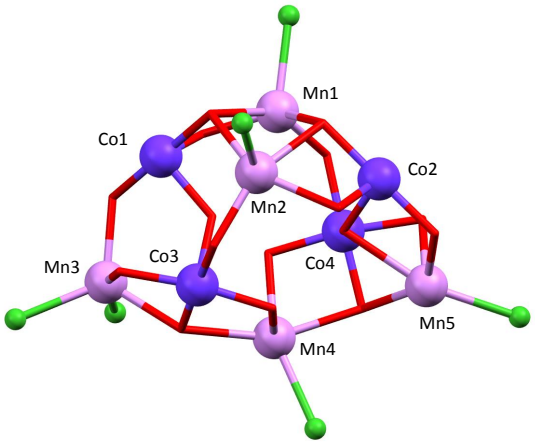


Figure 4.10 Manganese and cobalt substructure of $[Co_4^{III}Mn_5^{II}]$ cluster.

reported in Table 4.3.

At each ambiguous metal site in the structure, the occupancies of manganese and cobalt were competitively refined using the relevant scattering factors shown in Table 4.3. The results from the ratio refinement performed in CRYSTALS are presented in Table 4.4. Figure 4.10 shows a labelled metal oxide core of the molecule corresponding to the metal sites in Table 4.4.

Table 4.4 reports the suggested occupancies from the refinement for each metal centre when using simulated data, thereby proving the method as applied to mixed metal compounds. Within the structure, the manganese sites are all bound to at least one chlorine atom and the cobalt sites are either capped by the $(C_8NO_5H_{15})_2$ or the $(C_8NO_5H_{16})_2$ ligand.

	Ratio Refinement	
Metal Site	Mn Occ.	Co Occ.
Mn1	1.060(2)	−0.060(2)
Mn2	1.056(2)	−0.056(2)
Mn3	1.087(2)	−0.087(2)
Mn4	1.032(2)	−0.032(2)
Mn5	1.022(2)	−0.022(2)
Co1	−0.056(2)	1.056(2)
Co2	−0.046(2)	1.046(2)
Co3	−0.003(2)	1.003(2)
Co4	−0.013(2)	1.013(2)

Table 4.4 Manganese and cobalt site occupancies for the $[Co_4^{III}Mn_5^{II}]$ cluster calculated in CRYSTALS using a ratio refinement on simulated data.

4.4.3 Comparison of Occupancy Refinement Methods

The first comparison to be drawn between the two methods of occupancy refinement (ratio refinement *vs.* single wavelength refinement between the absorption edges described in Chapter 3) is an inspection of the errors. The methods were compared using simulated data generated from the model structure of $Mn_6(O)_2(O_2CCMe_3)_{10}$-$(C_4H_8O)_4$ at 6565 eV (*i.e.* between the absorption edges of the two oxidation states as determined from DetOx) and 6540 eV (below the absorption edges of both oxidation states). The results from the two refinement techniques are shown in Table 4.5.

	Ratio Refinement		Single Wavelength	
Mn Site	Mn^{II} Occ.	Mn^{III} Occ.	Mn^{II} Occ.	Mn^{III} Occ.
Mn1	1.003(2)	−0.003(2)	0.98(5)	0.02(5)
Mn2	0.997(2)	0.003(2)	1.15(5)	−0.15(5)
Mn3	0.000(2)	1.000(2)	−0.27(3)	1.27(3)
Mn4	0.002(2)	0.998(2)	−0.10(3)	1.10(3)

Table 4.5 Comparison of occupancy refinement techniques on Mn_6 using simulated data.

Although both techniques provide the correct answer, as expected the ratio refinement provided a much more reliable and stable refinement than the single-wavelength technique. Given the small differences in AS factors when comparing two oxidation states of the same element, the refinement process was stabilised by refining the whole structure against the model data (collected well below the absorption edge) and using a ratio of data from between the absorption edges and just below the absorption edge for the occupancies.

4.5 Initial Testing

The first experiments using the ratio refinement technique were carried out using the Mn_6 mixed-valence cluster described in Section 4.4 (used for proof-of-concept studies). The results of the experiment are described below although unfortunately the refinement was subject to a number of issues, which are addressed at the end of the section.

4.5.1 $Mn_6(O)_2(O_2CCMe_3)_{10}(C_4H_8O)_4$

The experiment was carried out as described in Section 4.4.1, however this time real data for the ratios were collected at 6545 eV (between edges) and 6450 eV (below the edge). The AS factors for the different oxidation states at each wavelength were calculated in DetOx (Table 4.1). Figure 4.11 shows the output from DetOx with the energies at which data were collected marked with black dotted lines.

Firstly, a least-squares refinement of the ratios was performed in CRYSTALS. The resulting occupancies are presented in Table 4.6, but unfortunately did not agree with the expected values (see Table 4.2). We therefore tried a simulated annealing refinement technique (the results of which are also reported in Table 4.6) to fit the experimental data, however this process too proved unsuccessful. Section 4.5.2 addresses some of the reasons for the failure of this experiment.

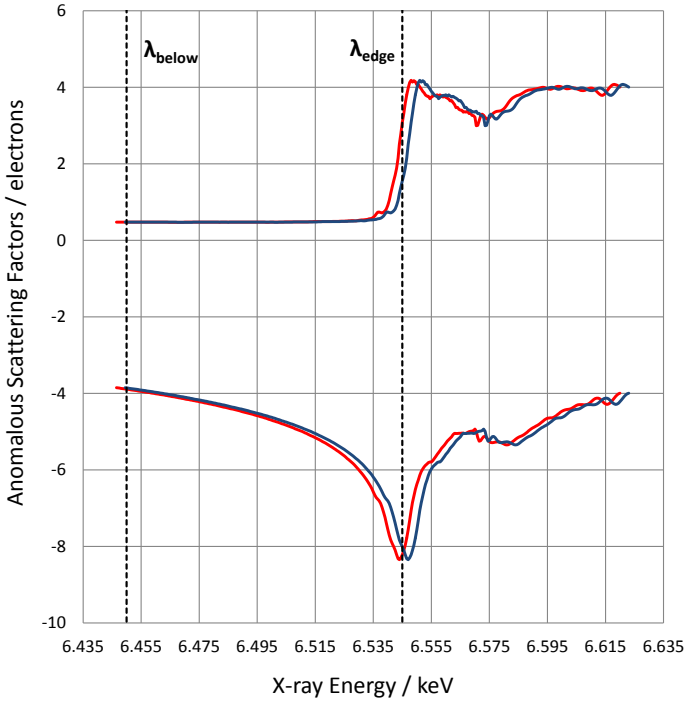


Figure 4.11 Deconvoluted fluorescence graph for Mn_6 cluster showing the energies used for initial data collections. The red line is used for Mn^{II} and the blue line for Mn^{III}.

	Expected Occupancies		Least-Squares	
Mn Site	Mn^{II} Occ.	Mn^{III} Occ.	Mn^{II} Occ.	Mn^{III} Occ.
Mn1	1.0	0.0	1.0(4)	0.0(4)
Mn2	1.0	0.0	−0.2(4)	1.2(4)
Mn3	0.0	1.0	0.3(3)	0.7(3)
Mn4	0.0	1.0	0.8(3)	0.2(3)

Table 4.6 Mn_6 site occupancies as calculated from a least-squares ratio refinement using experimental data. The expected occupancies are presented for comparison.

4.5.2 Limitations

As we can see from the results, the ratio refinement technique was unsuccessful for resolving the oxidation state occupancies in the Mn_6 cluster. The least-squares refinement process was only able to correctly determine the oxidation state at one of the four ambiguous sites. Furthermore, the errors on the occupancies were so high due to the instability of the refinement, that even the 'correct' occupancy at the Mn1 site was an unreliable result. The major reason for this was the high absorption experienced

as we collected data between the absorption edges. The idea that using a ratio of data between and below the edges would eliminate the issue of absorption, since each ratio value is affected equally, did not hold with such a large increase. In fact the ratio between the μ values at 6450 eV and 6545 eV was 4.90 *i.e.* almost a 5× increase in absorption coefficient.

When collecting data at the manganese edge (6.539 keV; 1.896 Å), we are dealing with relatively low energy and therefore long wavelength radiation (*cf.* Mo K_α = 0.7017 Å for a lab source). This has the effect of spreading out the spots in reciprocal space, which for small molecules (with small unit cell dimensions), means that collecting data to a high resolution is very difficult. With the detector position at 120° the maximum resolution possible at the edge of the detector is 0.9 Å *i.e.* only just high enough for direct methods and charge flipping algorithms to work effectively.

Another problem leading to the unreliability of the method was the low completeness of the datasets. Since the crystal has to last for a minimum of two datasets, there is a play-off between the length of time required to collect data of a high completeness, and the time that the crystal can feasibly last in the beam. Given these time constraints we were unable to collect data to a very high completeness and the redundancy was only around 2.0. This low completeness means that the data merging is less accurate since outliers cannot be eliminated statistically. Paired with the low completeness of the dataset, this meant that the reliability and accuracy of the experiment was compromised.

4.6 Method Developments

Section 4.5 described the application of the ratio refinement method to the mixed valence compound $Mn_6(O)_2(O_2CCMe_3)_{10}(C_4H_8O)_4$. However, due to limitations in the method, a number of developments were applied. This section describes such developments and goes on to test the improved method on a number of compounds.

4.6.1 Improvements

One of the leading problems encountered was the large increase in absorption associated with measuring datasets on and below the absorption edge. We therefore decided to collect all of the data below the absorption edge to exploit the changes in f' leading up to the edge. The first benefit of this would be to eliminate the large increase in absorption coefficient between the datasets. Using simulated data (as in Section 4.4) we were able to show that there were in fact enough differences in the AS factors to refine the different occupancies of the oxidation states in the Mn_6 cluster. Figure 4.12 shows the energies of the new data collections below the absorption edge and Table 4.7 shows the associated anomalous scattering factors at the chosen wavelengths.

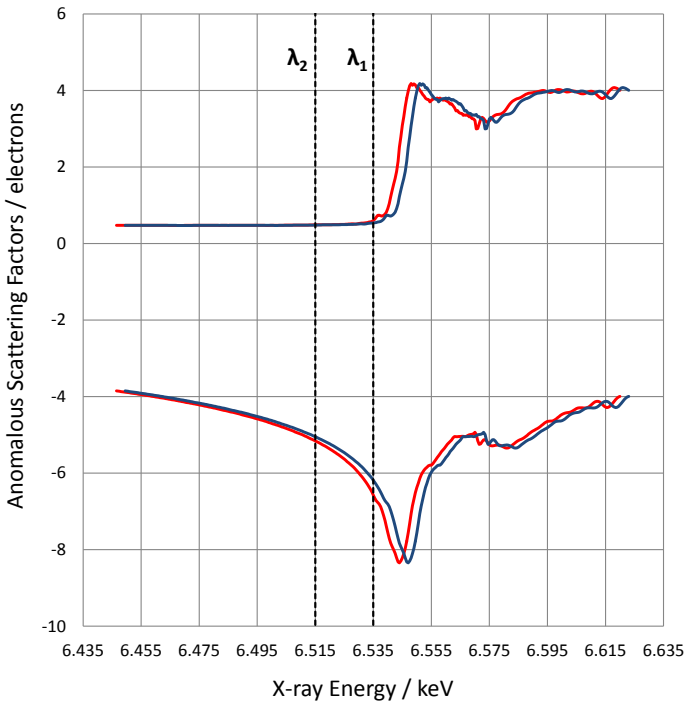


Figure 4.12 Deconvoluted fluorescence graph for Mn_6 cluster showing the energies used for improved-method data collections. The red line is Mn^{II} and the blue line, Mn^{III}.

Dataset	Ox. State	f'' / electrons	f' / electrons
Below Edge, λ_1	Mn(II)	0.60	−6.57
6535 eV	Mn(III)	0.54	−6.19
Below Edge, λ_2	Mn(II)	0.49	−5.16
6515 eV	Mn(III)	0.48	−5.05

Table 4.7 Anomalous scattering factors output from DetOx for Mn_6 cluster.

A second issue that needed to be addressed from the initial experiment on the Mn_6 compound was the distinct lack of both completeness and redundancy in the data. Generally, the anomalous signal only accounts for a fraction of the total intensity and therefore there is great need for accuracy in the data collected. The most reliable way to achieve this is to increase the redundancy. One of the main problems with collecting data at wavelengths as far removed from one another as in the initial experiment was that reflections shifted significantly between frames. Trying to devise a strategy to maximise redundancy is very difficult when there is only time to collect a subset of the full sphere of data. However, by collecting data at wavelengths much closer to one another below the edge, in general we saw that spots remained on the same frames between runs, which allowed us to collect data with a much higher redundancy using a single strategy. As a result, when performing the new refinements with experimental data we only chose reflections with a redundancy of 7 or above. To further improve the quality of data used in the refinement, the highly redundant reflections were also filtered according to an $I_{\text{mean}}/\sigma_{\text{mean}} \leq 3$ rejection criteria.

4.6.2 $Mn_3O(O_2CPh)_6(pyr)_2(H_2O)[2\,MeCN]$

The first compound studied with the improved ratio refinement method was the mixed valence $Mn_3O(O_2CPh)_6(pyr)_2(H_2O)[2\,MeCN]$ compound shown in Figure 4.13 (see Section 2.3.6 and Appendix A.6). This is a well studied compound and known to have two distinct oxidation states present. The molecule is non-centrosymmetric and so the different oxidation states occupy crystallographically distinct locations. The oxidation states are further confirmed by the longer Mn···O bonds associated with the

Mn^{II} centre (bound to the terminal H_2O group) and two shorter Mn···O bonds from the Mn^{III} (bound to the pyridine ligands). [143] The compound was chosen over the Mn_6 cluster for testing, firstly because there were fewer Mn sites to determine, and secondly, the crystals were of a very high quality and would provide the best data for refinement.

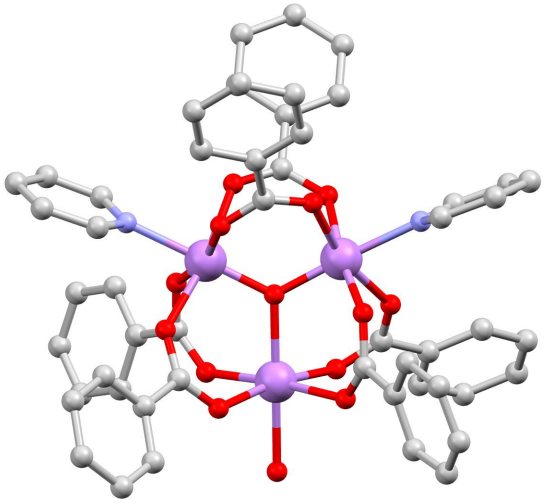


Figure 4.13 Crystal structure of Mn_3 cluster. Hydrogen atoms and solvent molecules have been removed for clarity.

Experimental

Absorption data were collected on the sample across the Mn edge between 6340 eV and 6740 eV and analysed in DetOx. Using the output deconvoluted fluorescence graph, 6 wavelengths were chosen at which to collect data. The first reason for this large number of datasets is that it allowed us to take ratios between a number of different datasets and iteratively determine the best wavelength difference to use in future refinements. Secondly, as described in Section 4.1, more than one set of ratio values can be used in the refinement by adding additional lines in the normal equations (see Figure 4.2). The wavelengths were chosen by finding f' values in steps of approximately 1.0 electrons from its lowest value in the spectrum (as shown in

Figure 4.14). The AS factors at these chosen wavelengths are reported in Table 4.8.

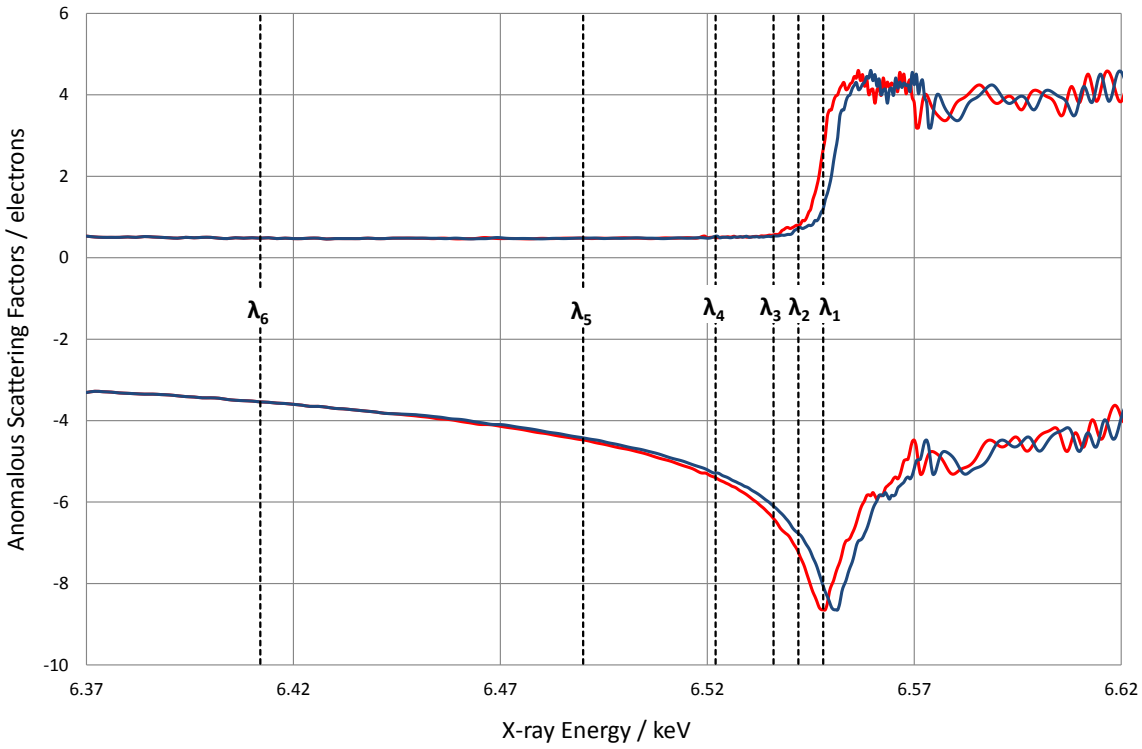


Figure 4.14 Deconvoluted fluorescence graph for Mn_3 cluster showing the energies used for data collections. The red line is Mn^{II} and the blue line, Mn^{III}.

Dataset	Ox. State	f'' / electrons	f' / electrons
λ_1	Mn(II)	2.663	−8.644
6548 eV	Mn(III)	1.200	−8.060
λ_2	Mn(II)	0.789	−7.223
6542 eV	Mn(III)	0.732	−6.773
λ_3	Mn(II)	0.561	−6.405
6536 eV	Mn(III)	0.536	−6.101
λ_4	Mn(II)	0.498	−5.403
6522 eV	Mn(III)	0.530	−5.288
λ_5	Mn(II)	0.483	−4.474
6490 eV	Mn(III)	0.480	−4.423
λ_6	Mn(II)	0.485	−3.540
6412 eV	Mn(III)	0.485	−3.540

Table 4.8 Anomalous scattering factors output from DetOx for Mn_3 cluster.

Results

Following a number of ratio refinements using different pairs of datasets, we found that the best result was obtained using λ_3 and λ_6 (6542 eV and 6412 eV respectively). These two wavelengths provide enough contrast between the two oxidation states whilst avoiding absorption effects from the increasing f'' experienced near to the manganese absorption edge. For the refinement, only reflections with a redundancy of 7 or higher following outlier filtering were used, thereby only using the most reliable data. The results of the refinement are presented in Table 4.9: the manganese sites refer to those labelled in Figure 4.15 as described by Vincent *et al.* [143]

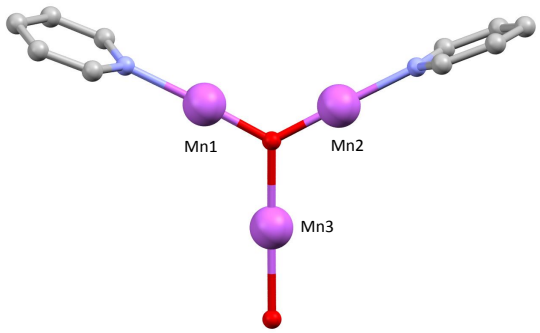


Figure 4.15 Manganese substructure of Mn_3 cluster. Mn1 and Mn2 are in the Mn^{III} oxidation state while Mn3 is in the Mn^{II} state.

	Expected Occupancies		Least-Squares		Simulated Annealing	
Mn Site	Mn^{II} Occ.	Mn^{III} Occ.	Mn^{II} Occ.	Mn^{III} Occ.	Mn^{II} Occ.	Mn^{III} Occ.
Mn1	0.0	1.0	0.23(8)	0.73(8)	−0.06	1.06
Mn2	0.0	1.0	0.45(11)	0.55(11)	−0.06	1.06
Mn3	1.0	0.0	0.33(11)	0.67(11)	1.00	0.00

Table 4.9 Mn_3 site occupancies as calculated from both a least-squares and simulated annealing ratio refinement using experimental data. The expected occupancies are presented for comparison.

Table 4.9 reveals that where the least-squares refinement was unsuccessful in accurately assigning the oxidation states at each Mn site in the structure, the simulated

annealing technique proved successful. The results of the simulated annealing do not have any associated standard uncertainties since these cannot be computed directly from the algorithm. Instead, the refinement was run multiple times from random starting points and we could consistently recover the results presented to a high degree of accuracy. The merits of simulated annealing over the least-squares for the ratio refinement are discussed in Section 4.6.4.

4.6.3 $Co_4^{III}Mn_5^{II}Cl_6(C_8NO_5H_{15})_2(C_8NO_5H_{16})_2(H_2O)(OH)_2$

This molecule was used previously for proof of concept testing and is described in Section 4.4.2. It was proposed that the compound contained 5 Mn^{II} ions bound to chlorine atoms and 4 Co^{II} ions capped by ligands. The crystals provided were good quality and therefore a good candidate for collecting data at multiple wavelengths.

Experimental

Firstly fluorescence data were collected on the crystal across both the Mn (6340 eV to 6740 eV) and Co (7460 eV to 7960 eV) edges and analysed with DetOx. As before, the energies at which to collect data were decided by stepping 1.0 electrons in f' from the minimum value of the manganese trace. Data were only collected below the manganese absorption edge (the lower energy edge of the two metals present) to reduce absorption problems. A total of 6 datasets were collected to again allow for optimisation of the ratio refinement technique. The position of these datasets are shown in Figure 4.16 and the corresponding AS factors reported in Table 4.10.

Results

It was found that the best refinement was achieved using λ_3 and λ_6 *i.e.* 6539 eV and 6490 eV respectively. The difference of 2.59 electrons in f' of the manganese provided

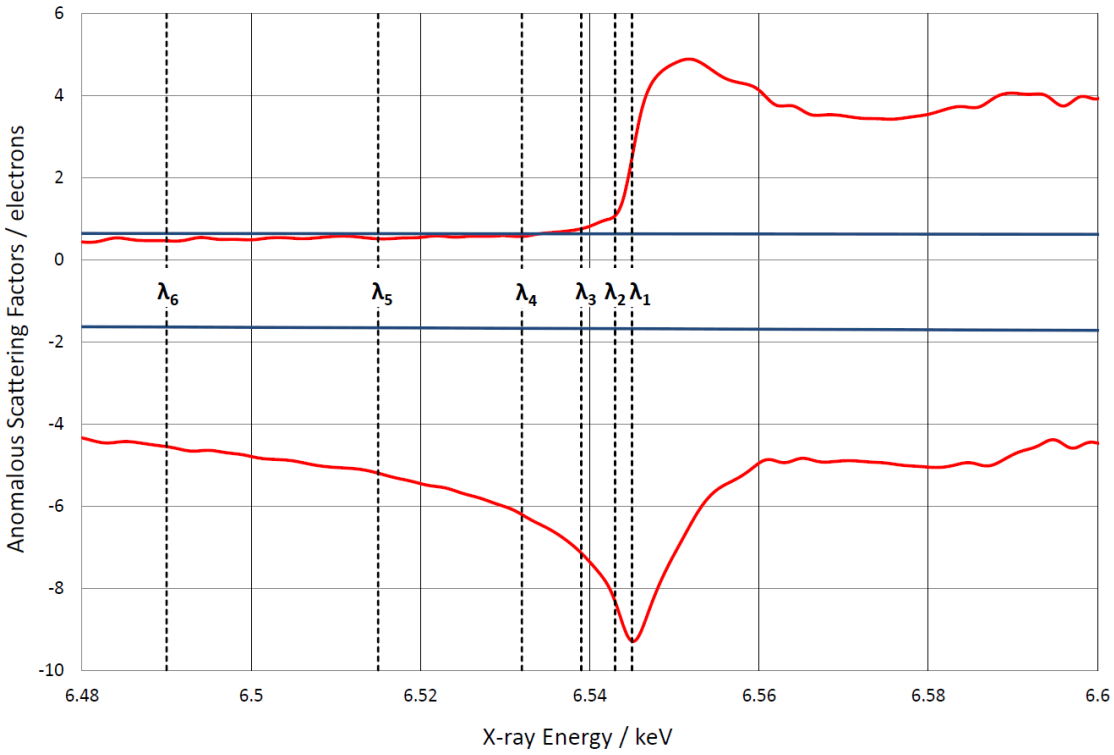


Figure 4.16 Deconvoluted fluorescence graph for $[Co_4^{III}Mn_5^{II}]$ cluster showing the energies used for data collections. The red line is used for Mn and the blue line for Co.

Dataset	Ox. State	f'' / electrons	f' / electrons
λ_1	Mn	2.493	−9.280
6545 eV	Co	0.646	−1.637
λ_2	Mn	1.078	−8.312
6543 eV	Co	0.646	−1.637
λ_3	Mn	0.762	−7.134
6539 eV	Co	0.646	−1.637
λ_4	Mn	0.574	−6.200
6532 eV	Co	0.646	−1.637
λ_5	Mn	0.518	−5.189
6515 eV	Co	0.646	−1.637
λ_6	Mn	0.470	−4.543
6490 eV	Co	0.646	−1.637

Table 4.10 Anomalous scattering factors output from DetOx for $[Co_4^{III}Mn_5^{II}]$ cluster.

	Simulated Data		Experimental Data	
Metal Site	Mn Occ.	Co Occ.	Mn Occ.	Co Occ.
Mn1	1.060(2)	−0.060(2)	0.197(9)	0.803(9)
Mn2	1.056(2)	−0.056(2)	0.193(7)	0.807(7)
Mn3	1.087(2)	−0.087(2)	0.148(9)	0.852(9)
Mn4	1.032(2)	−0.032(2)	0.190(9)	0.810(9)
Mn5	1.022(2)	−0.022(2)	0.199(9)	0.801(9)
Co1	−0.056(2)	1.056(2)	0.220(8)	0.780(8)
Co2	−0.046(2)	1.046(2)	0.153(9)	0.847(9)
Co3	−0.003(2)	1.003(2)	0.194(9)	0.806(9)
Co4	−0.013(2)	1.013(2)	0.045(8)	0.955(8)

Table 4.11 Simulated *vs.* experimental manganese and cobalt site occupancies for the $[Co_4^{III}Mn_5^{II}]$ cluster.

enough contrast between the elements to resolve their occupancies within the structure. Using only reflections with a redundancy greater than 7 and $I_{mean}/\sigma_{mean} \leq 3$, 217 reflections remained for inclusion in the least-squares ratio refinement in CRYSTALS – the results of which are presented in Table 4.11.

The results from the ratio refinement on simulated data are also included in Table 4.11 for comparison *i.e.* datasets for calculating the ratios were generated using a model structure, which assumed the metal positions to be exactly as shown in Figure 4.10. The refined model fitting the experimental data did not agree with the proposed model. Instead, we now propose that the structure mainly consists of cobalt with some manganese disorder across each of the sites (roughly 20%). We investigated further in order to support this new hypothesis.

Firstly, we checked the data being used in the refinement to see if noise was causing the occupancies to differ from the expected values. Figure 4.17 shows a plot of the ratios calculated from the experimental data (used in the refinement) *vs.* the simulated ratios for both the predicted 'discrete-metal-site' model (red circles) and the

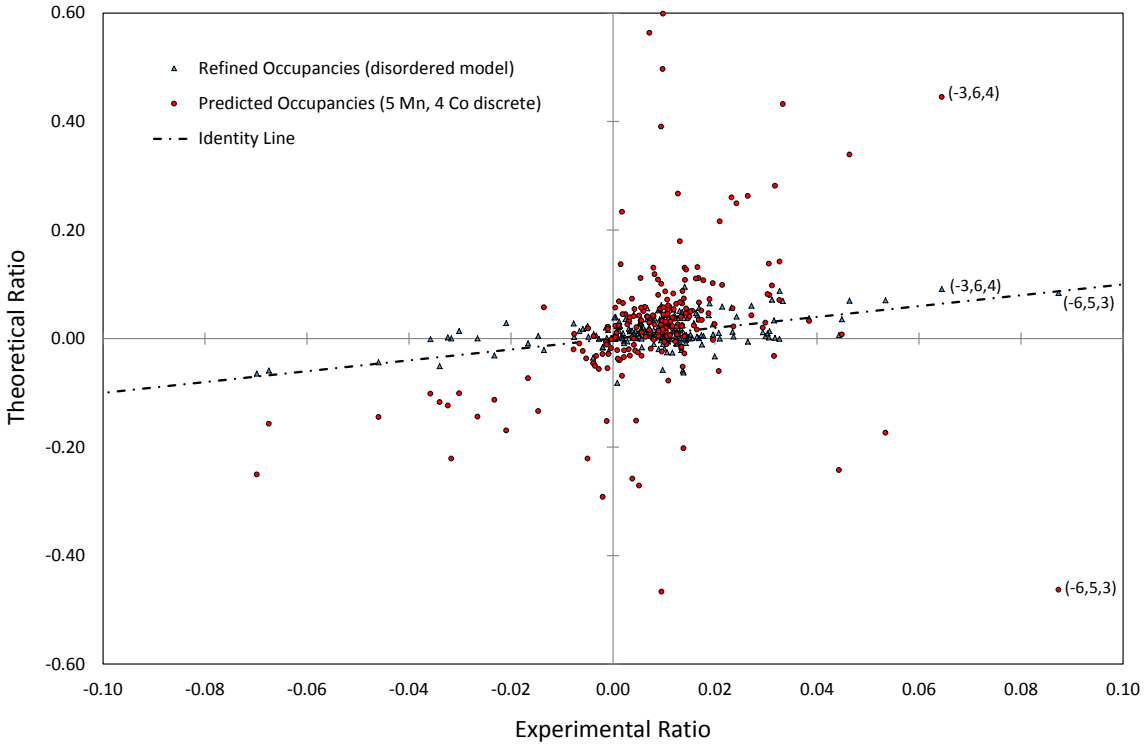
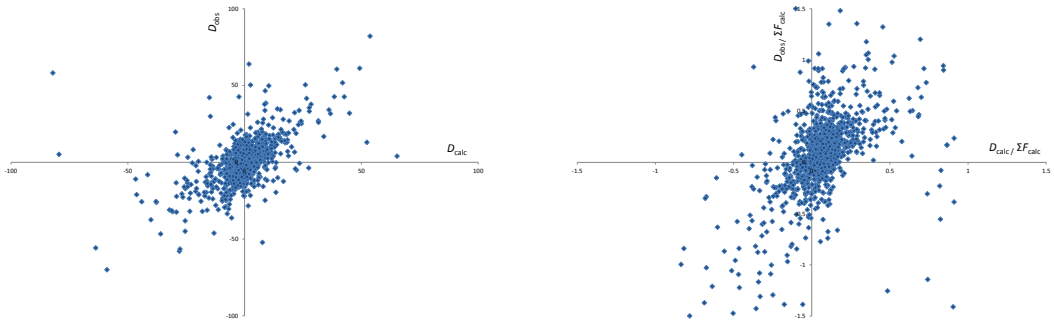


Figure 4.17 Comparison of calculated and simulated occupancies for $[Co_4^{III}Mn_5^{II}]$ cluster at different occupancies.

newly proposed 'disordered-metal-site' model (blue triangles). Allowing for experimental error, the more correct the model, the more closely the ratios should follow the identity line (dotted black line). As demonstrated by two reflections marked on the graph (the (-3,6,4) and the (-6,5,3)), the experimental ratios are much closer in magnitude to the simulated ratios calculated from the 'disordered-metal-site' model. This provides strong evidence that the structure does not consist of 5Mn and 4Co atoms in discrete positions within the structure.

Secondly, the same refinement was carried out, but this time using the simulated annealing technique to fit the data rather than least-squares. Again, the results were consistent with the occupancies in Table 4.11, suggesting a disorder of around 20% Mn across each site with Co making up the rest.

Finally, in the synthesis of this compound (see Section 2.3.5), $MnCl_2 \cdot 4H_2O$ is added in the final step of the synthesis. Therefore, a cobalt-only cluster could potentially have formed first and upon addition of the manganese chloride, the sites



(a) **(b)**

Figure 4.18 Comparison of the calculated *vs.* observed signals available for **(a)** resolving the Flack parameter using differences in Friedel pairs and **(b)** resolving oxidation states in the Mn_3 cluster using normalised differences between like reflection collected at different wavelengths. Graph **(a)** has been reproduced from Flack *et al.* [131]

become randomly disordered across the crystal. This could be tested by performing the synthesis again without the addition of manganese chloride to see if an isomorphous structure can be formed – thereby ensuring zero manganese contamination in the structure.

4.6.4 Discussion

This chapter has demonstrated that by using a ratio refinement technique we are able to discriminate between oxidation states and neighbouring elements in favourable cases. However, due to such small signals and a greater number of ambiguous sites compared to Chapter 3, this proved difficult. In the case of the Mn_3 cluster, the difference in anomalous signal was only 0.12% of F_{000}. Figure 4.18 puts this value in context by comparing to the size of the signal to that used for refining the Flack parameter. We were initially confident that enough signal would be available to discriminate the oxidation states, however the refinement was not so straightforward for a number of reasons.

Firstly, we found that when the datasets used for the ratio refinement were collected between and below the absorption edges, the increase in absorption associated with the edge introduced systematic errors into the data. As a result, scaling the datasets was more difficult with the consequence that the ratio refinement was unsuccessful (using either least-squares or simulated annealing to fit the data). This led us to modify the data collection method and analysis. We chose to collect both datasets for the ratio calculation below the absorption edge (thereby avoiding errors due to absorption effects as a result of the increasing f''). We also filtered the data such that only reflections with a redundancy of 7 or greater and $I_{\text{mean}}/\sigma_{\text{mean}} \leq 3$ were used for the refinement *i.e.* the most reliable reflections.

Despite proving the technique on simulated data, the least-squares ratio-refinement was still unable to discriminate the oxidation states when applied to experimental data. We believe this was due to the combination of subtle signals and noise in the data and a greater number of parameters being refined, leading to the solution getting trapped in local minima. The refinement was shown to be unstable and consistently provided the incorrect occupancies for structures where the site occupancy is know *a priori*. As a result we explored a different technique for fitting the data. The simulated annealing algorithm provides a random approach to sampling the fit of the parameters to the data, and so is less prone to getting trapped in false minima. Using this method we were able to fit a signal from hundreds of pairs of reflections to consistently confirm the expected oxidation states in the Mn_3 cluster, thereby proving the technique's viability for use in the ratio refinement process.

When studying the mixed-metal Co_4Mn_5 cluster, although the available signal was much greater (0.93% of F_{000}) and comparable to that of $GaCl_2$, the ratio refinement technique was still required due to the greater number of ambiguous sites in the structure. Least-squares refinement using data from a single wavelength was unstable and did not provide an accurate result. However, due to the atoms' larger anomalous difference signal in this system the least-squares approach was stable for ratio re-

finement. The result obtained did not agree with the expected 'discrete-site' model consisting of 5Mn and 4Co sites, but instead a disordered-site model was proposed. We were able to verify this result using the simulated annealing technique, which we had already proven to work with smaller signals in Mn_3.

There is still room to improve the method for both data collection and analysis. For example, the Mn_6 cluster was used to test the simulated ratio refinement, however, the fit to the signal in the experimental data was poor, and neither least-squares nor simulated annealing could recover meaningful site occupancies. The first improvement would be to perform a leverage analysis to determine the most important reflections in order to increase the accuracy of the refinement. We have previously shown that it is possible to calculate this set of reflections, however, a new strategy must be devised in order to collect them with a high redundancy. [2] With the current collision model on EH2, new strategies must be tested manually one step at a time which is very time consuming. It was therefore not viable to create and test a new strategy for each sample. Once the model has been improved and software is available to test a given strategy quickly and precisely, the leverage analysis will improve the accuracy of the refinement by reducing the standard uncertainty of influential measurements.

Another improvement to the current technique would be to further address the absorption problems. Although collecting data below the edge alleviated much of the absorption problems, it did so at the expense of signal and there was still noise in the data (which was potentially one of the main reasons that the least-squares approach was unsuccessful for valence contrast). The current set-up on EH2 does not allow for face-indexing of the crystals, however the addition of this feature would lead to a much better signal-to-noise in the data, thereby improving the refinement technique and reducing the chance of getting trapped in false minima. Face-indexing would also be beneficial for the single-wavelength technique described in Chapter 3, but it is less of a requirement. In that case, where we only collect at one wavelength, a full

sphere can be rapidly obtained to give high redundancy in the data, which leads to more accurate empirical absorption corrections in CRYSALISPRO.

Chapter 5

Structure Solution of 'Large' Small Molecules

Sections 1.2 and 1.3 described a number of current techniques that are regularly applied in X-ray crystallography for solving small molecule and macromolecular structures respectively. The trend in small molecule crystallography is to study larger and larger structures, and the quality of the data collected decreases as a result. This is a common problem encountered in macromolecular crystallography and this chapter describes how we have applied structure solution techniques that have traditionally been reserved for non-centrosymmetric macromolecular structures, to the field of small molecule crystallography.

5.1 Bromide Urea Salt

Bromo-*p*-NO_2-diphenyl-urea was chosen for a proof-of-concept experiment to test whether it was possible to solve the structure using the MAD phasing technique. The structure of this centrosymmetric small molecule has been previously reported (spacegroup $P2_1/n$) and was therefore an ideal choice for these initial studies. [145] Since we were able to grow good quality crystals, the data collected were very high resolution and therefore the structure could, in fact be easily solved by standard techniques such as direct methods or charge flipping (Figure 5.1). Therefore, in order to mimic a poor quality dataset such as would be expected when solving the structure of a very large molecule (hundreds of non-hydrogen atoms in the asymmetric unit), the data collected were limited to a resolution of 2 Å. At this resolution we found that the structure could no longer be solved by conventional methods. This section describes the MAD experiment that was carried out in EH1 on beamline I19 and how we were able to solve the structure at 2 Å resolution using data from multiple wavelengths.

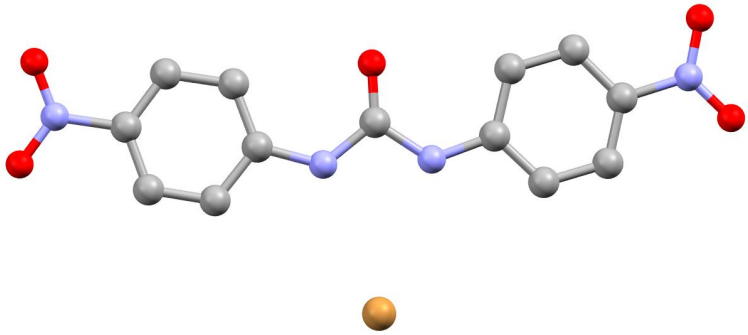


Figure 5.1 Crystal structure of TBAB[CO(NH(4 –NO_2–C_6H_4))$_2$]$_2$. The ethyl acetate and tetrabutylammonium counter ion have been removed for clarity. Red = O, Blue = N, Grey = C and Orange = Br.

5.1.1 Experimental

Suitable crystals for analysis were grown by vapour diffusion of hexane into a hot ethyl acetate solution containing the sample. Since this proved to be a very stable compound with respect to radiation damage, the initial studies were carried in EH1 (Section 2.1.3) using a Rigaku Saturn 724+ CCD detector at 100 K.

The first part of the experiment involved collecting an absorption-fluorescence scan across the bromine edge using the Vortex detector. An initial scan was collected between 13,220 eV and 13,720 eV in steps of 1 eV *i.e.* ±250 eV around the theoretical absorption edge for bromine, at 13,473.7 eV. A second finer scan was collected between 13,450 eV to 13,500 eV in steps of 0.5 eV, since the spectrum in this region is most sensitive as a function of energy. Using the normalised spectrum from these scans (Figure 5.2), wavelengths for data collection were chosen so as to maximise the information available from the anomalous signal. Table 5.1 shows the chosen energies for data collections and the associated AS factors output from DetOx.

At each chosen wavelength a full sphere dataset was collected and subsequently integrated using the Agilent CrysAlisPro software. [178] In order to cut the resolution of the data to 2 Å the data were read into Crystals and filtered.

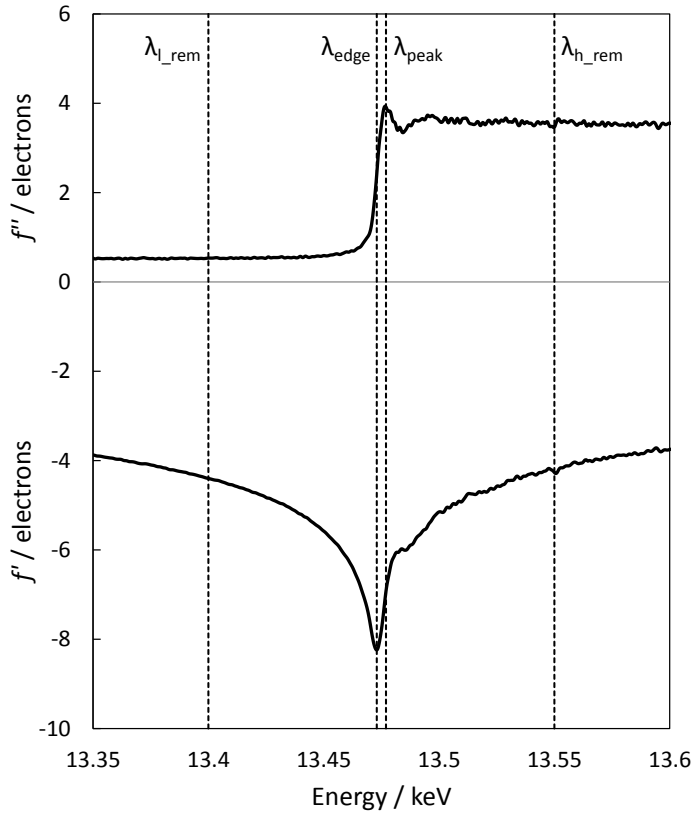


Figure 5.2 Normalised absorption-fluorescence spectrum for bromide urea salt. Dotted black lines represent the energies at which datasets were collected.

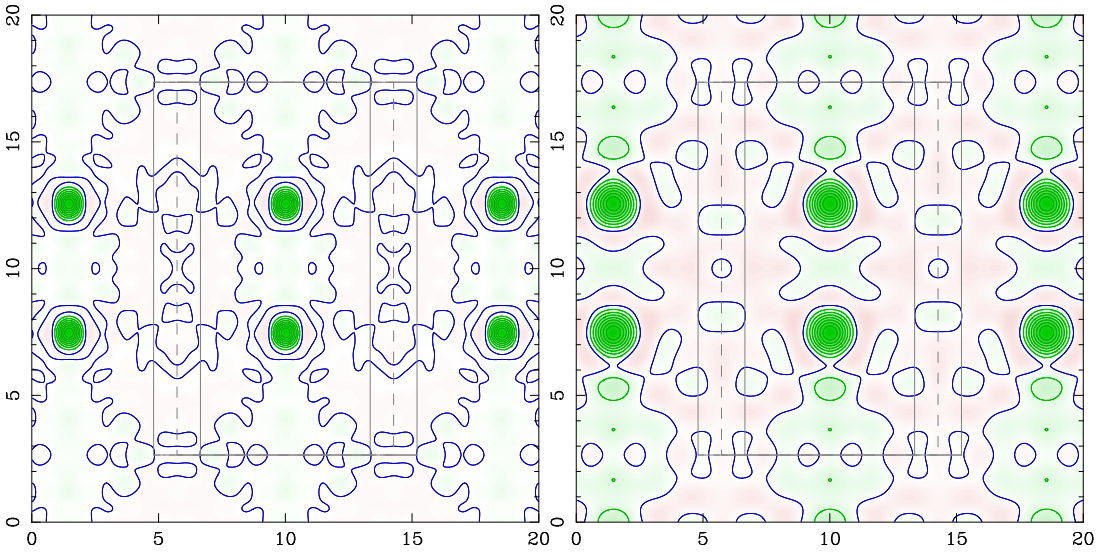
Dataset	Energy / keV	f'' / electrons	f' / electrons
Low Remote	13.400	0.533	−4.402
Inflection	13.473	2.414	−8.242
Peak	13.477	3.916	−6.924
High Remote	13.550	3.463	−4.268

Table 5.1 Anomalous scattering factors for bromide urea salt at selected wavelengths.

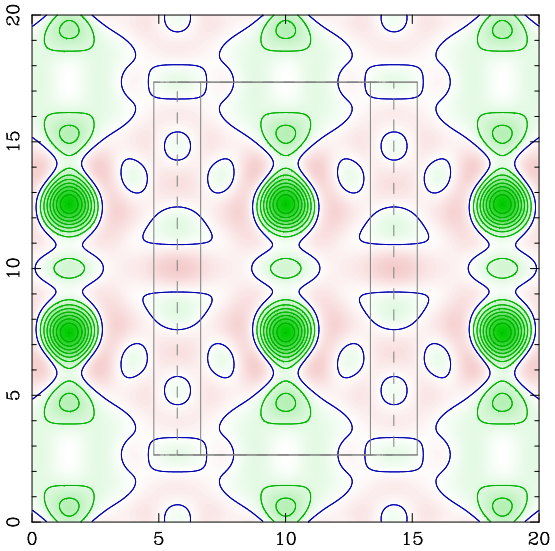
5.1.2 Substructure Determination

The substructure was solved using the Bayesian estimate technique previously described in Section 1.3.6 to calculate values of F_A, F_O and $\Delta\varphi$. The merits of this technique over a Bijovet difference Patterson or the anomalous difference Patterson were also described in Section 1.3.6 using the bromide urea salt as an example (*cf.* Figure 1.14 *vs.* Figure 1.16). Figure 5.3 shows three calculated Patterson maps using experimental data at different resolutions.

Figure 5.3 clearly shows that as the resolution of the data is lowered, the precision



(a) Full Resolution **(b)** 1.5 Å



(c) 2.0 Å

Figure 5.3 Bromine urea salt Bayesian Patterson maps calculated with **(a)** full resolution, **(b)** 1.5 Å and **(c)** 2.0 Å, at the $w = 1/2$ Harker Section. All maps are viewed down the c-axis and further details can be found in Figure 1.9.

Dataset	x	y	z
Ref. Model	0.863	0.087	0.576
Full Res.	0.861(4)	0.089(2)	0.576(1)
1.5 Å	0.865(2)	0.088(1)	0.576(1)
2.0 Å	0.859(3)	0.089(2)	0.576(1)

Table 5.2 Bromine atom coordinates as refined using data at different resolutions. Exact coordinates from the model structure are included for comparison (for which the errors are negligible to 3 d.p.).

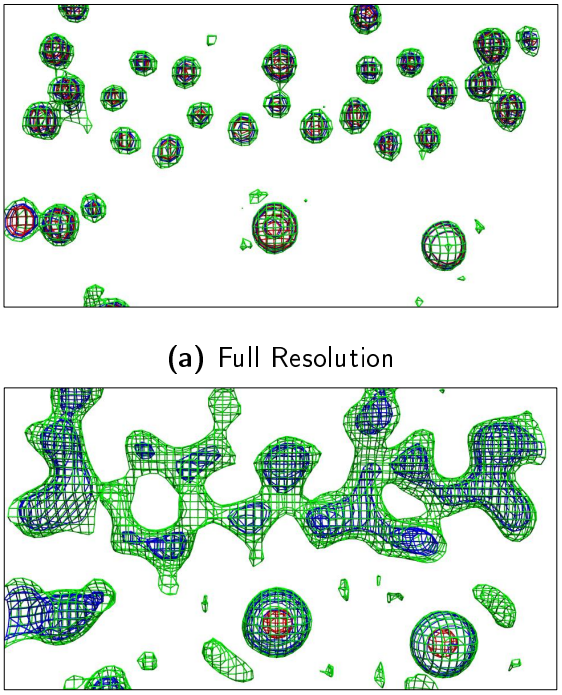
of the Patterson map decreases. The points of high density corresponding to the Br–Br vectors become less well defined and the background noise also increases with decreasing resolution. However, we have shown that even at 2.0 Å resolution, the Patterson map is still sufficient to determine the Br atom position, and to the same degree of accuracy as at other resolutions following refinement of the substructure (see Table 5.2).

5.1.3 Complete Structure

The final step is to use calculated phases from the anomalous substructure, φ_A^{calc}, to provide phase estimates for the complete structure. Substructure phases were calculated using CRYSTALS and output as a .pch file:

```
#PUNCH 6 C
END
```

Using Eqn. 1.83 ($\varphi_T = \Delta\varphi^{\text{obs}} + \varphi_A^{\text{calc}}$) we can now calculate values of φ_T using our calculated estimates of $\Delta\varphi^{\text{obs}}$ and φ_A^{calc}. The new phases were combined with the experimental structure factors, F_{obs}, to create a new .pch file. Structure factors from the low remote dataset were used since of the four datasets they are least susceptible to absorption effects. The new .pch file was then read back into CRYSTALS and an F_{obs}



(a) Full Resolution

(b) 2.0 Å

Figure 5.4 Electron density map of bromine urea salt at **(a)** full resolution and **(b)** 2.0 Å. Both images are produced using the same contour levels.

Fourier map (electron density) calculated. The electron density maps for the full and 2 Å resolution datasets are shown in Figure 5.4.

As expected, the electron density maps in Figure 5.4 show that at lower resolution the atomic positions are much less well defined than with full resolution data. However, the map is still interpretable and key features in the structure such as the benzene rings and nitrate groups are apparent. The final step therefore is to place atoms at peaks in the electron density map and refine the structure. After several rounds of searching for atoms in Fourier difference maps, refining the model each time, we were able to solve the structure from 2 Å datasets (*n.b.* this was not possible using standard small molecule crystallographic solution techniques). Figure 5.5 shows the structure solved using 2 Å resolution data overlaid onto the model structure presented in Figure 5.1.

Figure 5.5 shows that the structure solved using the MAD phasing technique from low resolution data is in very good agreement with the model structure; there is only

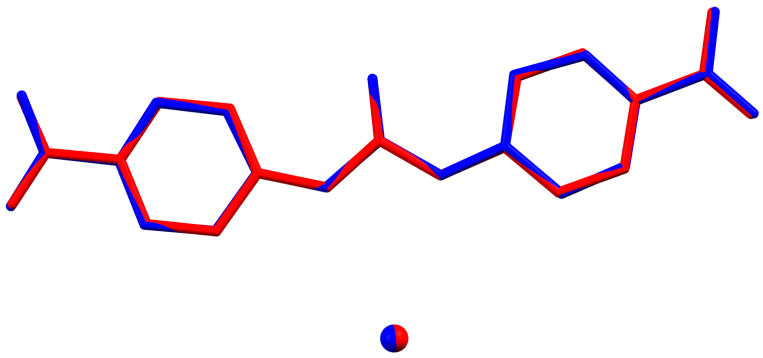


Figure 5.5 Overlay of bromide urea salt solved by; MAD phasing using 2Å data (red) and Superflip using full resolution data (blue).

a very minor discrepancy in the relative angles of the benzene rings. We can therefore conclude that the implementation of the phasing method is correct and that we can indeed apply these techniques to centrosymmetric small molecules.

5.1.4 EH1 Limitations - Bromide Rotaxane

So far this section has described how we were able to successfully solve the structure of a bromide urea salt by exploiting the anomalous signal from the bromine atom at wavelengths on and around the absorption edge. The next experiment was therefore to test the technique with a larger structure where the quality of data available is much lower due to problems such as poor crystallinity and high solvent content. For this experiment we chose a bromide-bound rotaxane structure (Figure 5.6). The crystal structure has been previously reported, [145] but with a larger R-factor (14.38 %) than the bromide urea salt (5.80 %).

The experiment was performed as described above. Firstly, an absorption-fluorescence scan was recorded between 13,220 eV and 13,720 eV and the data normalised. Data were collected at the same four wavelengths as before and the calculated AS factors are shown in Table 5.3.

The data collection however was not as successful as with the bromide urea salt

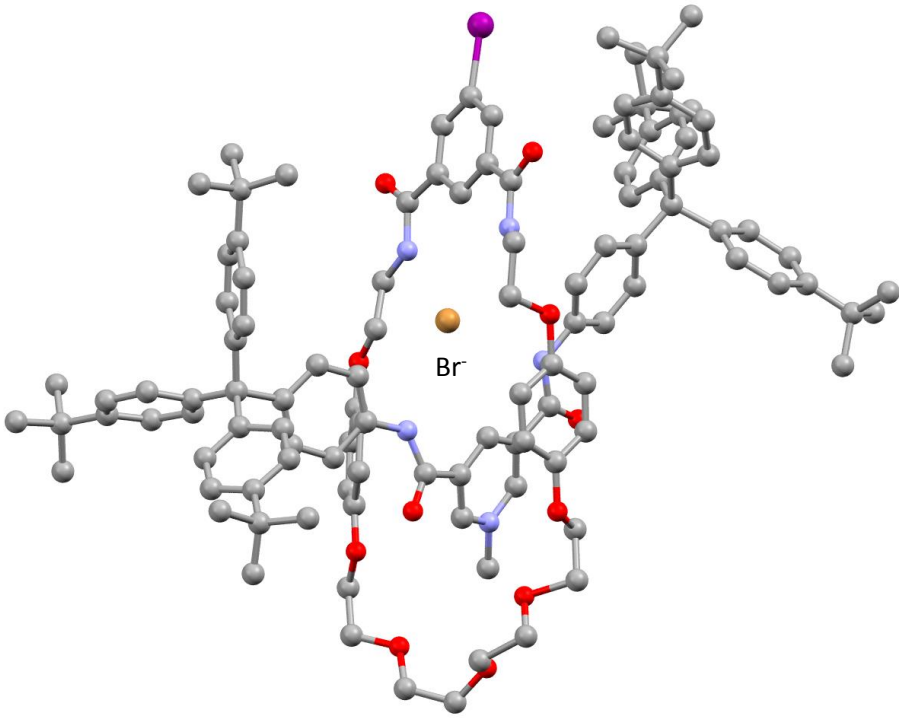


Figure 5.6 Bromide rotaxane crystal structure. Bromide atom is orange and the iodine is purple.

Dataset	Energy / keV	f'' / electrons	f' /electrons
Low Remote	13.400	0.51	−4.44
Inflection	13.473	2.04	−8.74
Peak	13.477	4.73	−7.61
High Remote	13.550	3.68	−3.98

Table 5.3 Anomalous scattering factors for bromide rotaxane at selected wavelengths.

since the rotaxane was much more susceptible to radiation damage. Each data set took around 1.5 hrs to collect a full sphere of data and therefore by the end of the fourth data set, much of the resolution of the data had been lost as a result of sample degradation. Figure 5.7 shows the calculated Bayesian-Patterson map using the first dataset collected *i.e.* the best available dataset in terms of resolution and completeness.

Due to the poor data resolution, the peaks in the Patterson map are not very well defined. Figure 5.7 shows a sharp peak at the origin (as expected) and two weak peaks corresponding to the Br–Br vector either side of the origin (since the Patterson map is

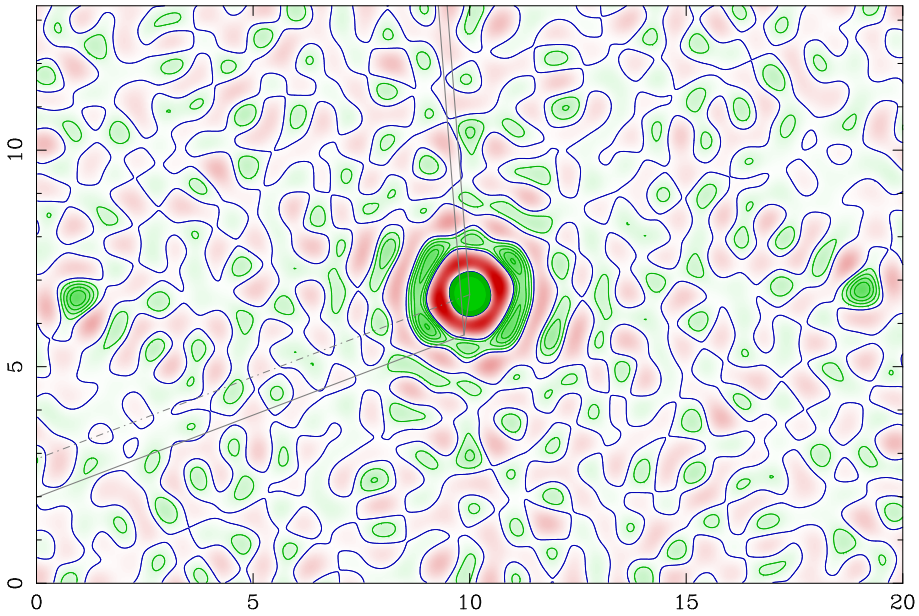


Figure 5.7 Bromide rotaxane Bayesian Patterson map at full resolution. Map viewed at an oblique angle; further details can be found in Figure 1.9.

Dataset	x	y	z
Ref. Model	0.075	0.242	0.070
Full Res.	0.069(20)	0.241(13)	0.070(10)

Table 5.4 Refined bromide coordinates for the bromide rotaxane. Errors in the model coordinates are negligible to 3 d.p.

centrosymmetric). We also observe Fourier ripples in the Patterson map around the origin peak. This is indicative of missing reflections, which can be caused by radiation damage or disorder in the structure. [179]

Despite the weakness of the peaks in the Patterson map, the substructure could still be determined although the errors in the coordinates are significantly higher than previously (Table 5.4).

The bromine substructure was then used to calculate estimates for φ_A^{calc}. However, when the new phases for the low remote data set were input into CRYSTALS, the complete structure could not be solved. Parts of the structure *e.g.* benzene rings, were visible in the electron density map however this did not provide enough ad-

ditional phase information to complete or refine the structure. In order to try and overcome this problem, subsequent experiments were carried out in EH2 using the Pilatus detector to minimise radiation damage on the sample.

5.2 *o*-Xylylenebis(triphenylphosphonium Bromide)

o-Xylylenebis(triphenylphosphonium Bromide) (*o*-Xtpb) was used as a test sample for a MAD phasing experiment in EH2 on beamline I19. The major advantage of using EH2 over EH1 is the availability of the Pilatus detector. As described in Section 2.1.4, the shutterless capabilities of the Pilatus means that a data collection takes a fraction of the time taken using the CCD detector in EH1. The upshot of this was to reduce the effects of radiation damage on the sample (as experienced with the bromide rotaxane) and therefore improve data quality.

5.2.1 Experimental

Crystals of *o*-Xtpb were used straight out of the bottle from Sigma-Aldrich and the structure had been previously reported. [180] As with the urea salt, the crystals were of high quality and the structure could be solved by standard small-molecule methods using data from the in-house diffractometer (Figure 5.8). Therefore the sample was a good candidate for commissioning EH2 for MAD phasing experiments.

As before, an absorption-fluorescence scan was performed and normalised between 13,220 eV and 13,720 eV to decide at which wavelengths to collect data. The selected wavelengths and the corresponding calculated AS factors from DetOx are shown in Table 5.5

Although the Pilatus detector collects images at a much faster rate than the CCD detector in EH1, the crystal was still subject to enough radiation damage that after the first dataset had been collected (low remote) the crystal could not be used again.

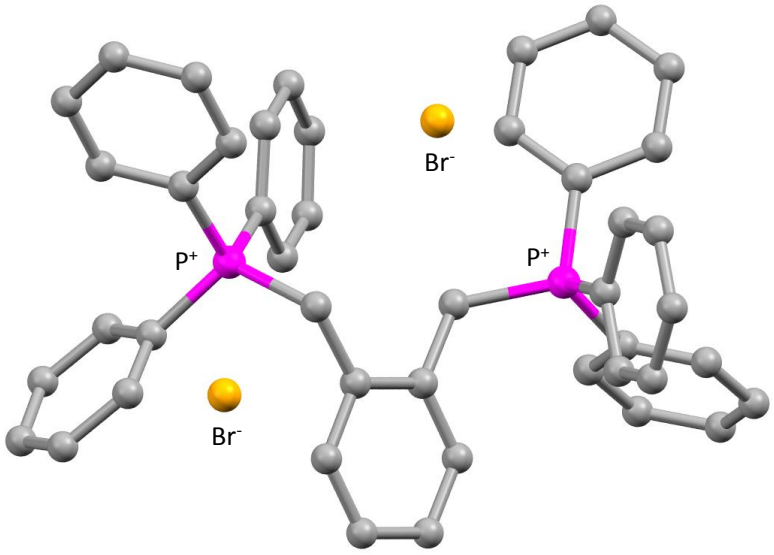


Figure 5.8 Crystal structure of *o*-Xylylenebis(triphenylphosphonium Bromide). Hydrogen atoms have been removed for clarity.

Dataset	Energy / keV	f'' / electrons	f' / electrons
Low Remote	13.350	0.518	−3.866
Inflection	13.476	4.162	−8.114
Peak	13.480	4.576	−6.866
High Remote	13.700	3.696	−3.330

Table 5.5 Anomalous scattering factors for bromide urea salt at selected wavelengths.

Therefore the experiment was adjusted so that data were only collected at two detector positions (2θ = 20° and 80°). The maximum resolution data that can be collected at a given wavelength at these detector positions is shown in Table 5.6.

		Detector Position	
Dataset	Energy / keV	20°	80°
Low Remote	13.350	1.047	0.576
Inflection	13.476	1.037	0.570
Peak	13.480	1.036	0.569
High Remote	13.700	1.020	0.560

Table 5.6 Maximum resolution (in Å) and completeness of datasets collected at different detector positions and wavelengths.

Table 5.6 shows that data collected at 20° is of high enough resolution to solve the structure by direct methods or charge flipping. However, the quality of data was limited at this resolution due to the poor completeness of the data. As a result the structure could not be solved using conventional techniques using only the 20° data, and the solution using both the 20° and 80° data was not of an acceptable quality for a model structure. Instead, data collected from the first extended-strategy dataset (before the crystal was subject to radiation damage) were used to provide a model structure for comparison with later MAD phasing solutions.

5.2.2 Data Collection

The images collected from the Pilatus detector were not compatible with the CRYSALISPRO software for data integration. Therefore, before analysing the data, the frames firstly had to be converted to a readable format. This was achieved using a program developed in the group called DIAMOND2CRYSALIS, [181] which converted the frames from a .cbf to .esperanto file format. Since the Pilatus can collect up to 20 frames per second, data are often collected with 'fine ϕ-slicing' to improve quality by removing overloads and reducing background *i.e.* data are collected in oscillation ranges less than half of the estimated mosaicity of the crystal - typically 0.1°. The downside of collecting such small ranges of data is that spot intensities are reduced. Therefore, using DIAMOND2CRYSALIS, frames were stacked in sets of 10 (by merging 10 images of 0.1°, the oscillation range becomes one degree) so that they could be integrated more accurately*. Stacking frames also dramatically reduces the data integration time by a factor of 10. Each dataset was integrated twice – firstly using only data collected at 20° (average completeness of 54.3%) and secondly using data collected at both 20° and 80° (average completeness of 87.0%).

*accuracy based on the R_{int} of integration in CRYSALISPRO. No further investigations have been undertaken.

Dataset	Atom	x	y	z
80° & 20°	Br1	0.599	0.195	0.156
	Br2	0.221	0.407	−0.011
20°	Br1	0.593(6)	0.194(4)	0.161(6)
	Br2	0.210(6)	0.411(4)	−0.013(5)

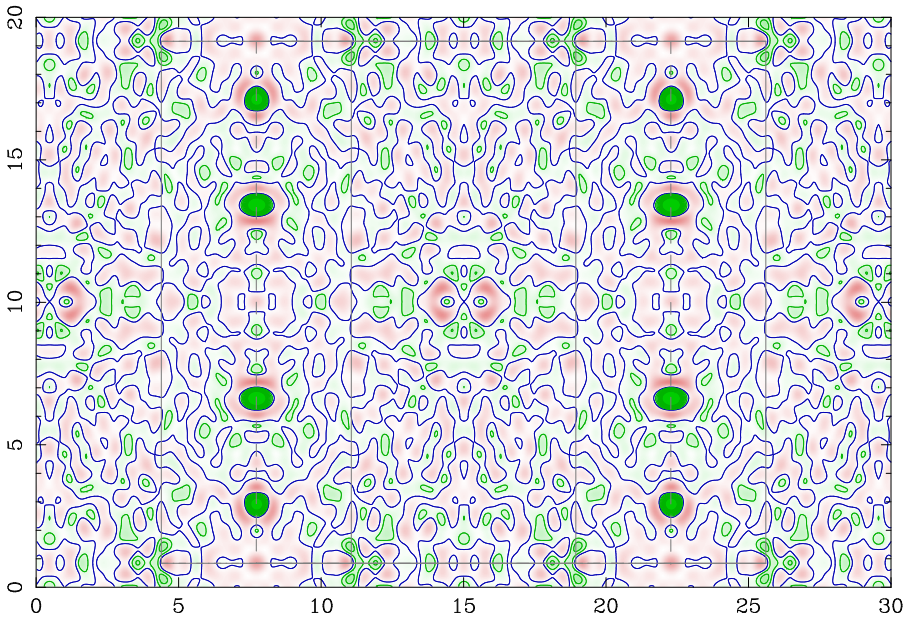
Table 5.7 Xylylene bromine substructure coordinates as refined using data at different resolutions. Errors in the 80° & 20° dataset coordinates are negligible to 3 d.p.

5.2.3 Substructure Determination

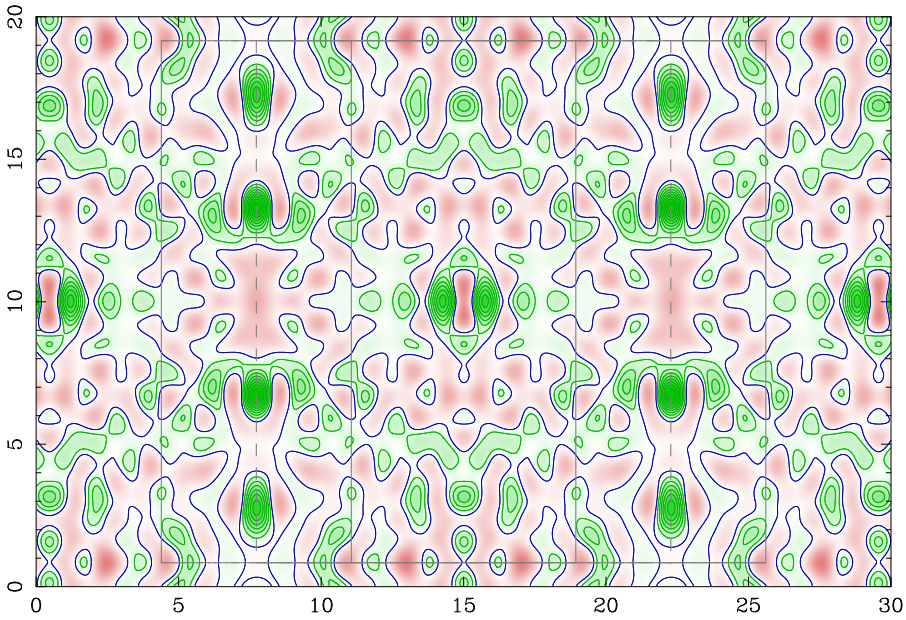
The substructure was solved using a Bayesian F_A Patterson map. Figure 5.9 shows the calculated Patterson maps using the data from **(a)** both the 20° and 80° detector positions and **(b)** only the 20° detector position.

These maps reveal similar features to previous Patterson maps using data at different levels of resolution. When using the full resolution data available (87% completeness), the Patterson shows 8 sharply defined intense peaks in the $w = 1/2$ Harker plane. However, upon removing the higher resolution data collected at $2\theta = 80°$, the same 8 peaks become less well defined and the level of background noise increases. Since the completeness is comparatively very low (54.3% completeness) when only using data at $2\theta = 20°$, the missing reflections cause Fourier ripples in the Patterson map. This can be seen by the alternating concentric rings of green (positive) and red (negative) density propagating from the centre of the map. It is worth noting that the area of high positive density in the centre of Figure 5.9b is not a split peak but an artefact of the Fourier ripples and the symmetry of the map.

The calculated positions of the bromine substructure from the two Patterson maps are shown in Table 5.7.



(a) 20° and 80°



(b) 20°

Figure 5.9 *o*-Xtpb Bayesian Patterson maps calculated with **(a)** data collected at both 20° and 80°, **(b)** data collected only at 20°, at the $w = 1/2$ Harker Section. All maps are viewed down the *c*-axis. Further details can be found in Figure 1.9.

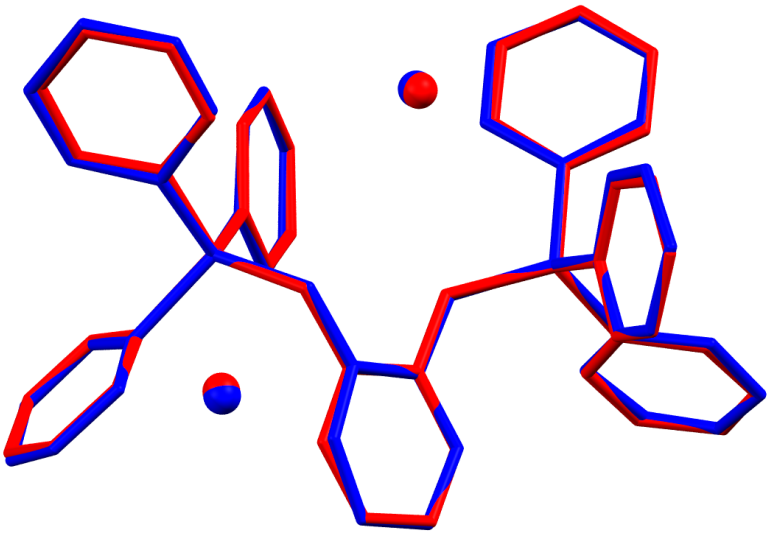


Figure 5.10 Overlay of *o*-Xtpb solved by; MAD phasing (red) and Superflip (blue) using full resolution data available.

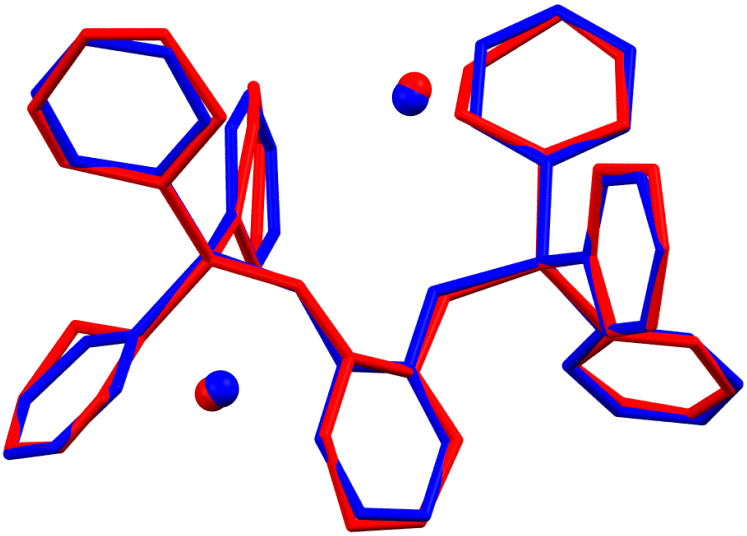


Figure 5.11 Overlay of *o*-Xtpb solved by; MAD phasing (red) using only data collected at $2\theta = 20$ ° and Superflip (blue) using full resolution data available.

5.2.4 Complete Structure

The complete structure was solved as described in Section 5.1.3 for the urea salt. Calculated phases from the anomalous substructure, φ_A^{calc} were combined with the Bayesian estimates for $\Delta\varphi^{\text{obs}}$ to give values for φ_T. These new phases were applied to the structure factors of the low remote dataset to produce a direct F_{obs} electron density map in CRYSTALS.

Using the full resolution low remote dataset, the structure was solved using Superflip to give an R-factor of 7.02%. This provided a model structure against which we could compare the results from solving the structure by MAD phasing techniques at the different resolutions. Figures 5.10 and 5.11 show the overlay of the MAD solved structure onto the Superflip solution using the different datasets. As expected, the fit when using data from both 80° and 20° is much better than when only using data collected at 20°. However, despite the large lack of completeness in the 20° data, the structure could still be solved.

As a measure of how well the MAD structure solution fitted against the Superflip solution, a RMSD (root mean square deviation) on the atomic positions was calculated for each overlay *i.e.* the average distance between like atoms when superimposed. For the structure solved using the 80° and 20° data, the RMSD was 0.0878 Å, whereas that for the 20° structure was 0.2265 Å. However, the major discrepancies between the two structures in Figure 5.11 arose because the phenyl rings in the MAD structure were not a regular shape *i.e.* the bond lengths and angles were not all equal. If we regularise all of the phenyl groups such that they are chemically correct, the RMSD is reduced from 0.2265 Å to 0.1309 Å. The general correction applied to 6 atoms (C1 – C6) in a ring is:

```
#REGULARISE REPLACE
GROUP 6
TARGET  C(1) C(2) C(3) C(4) C(5) C(6)
PHENYL
END
```

Figure 5.12 shows the improved MAD structure superimposed onto the Superflip structure. Although the overlap of the two structures is still not perfect, the improved result was sufficient for the purpose of the experiment.

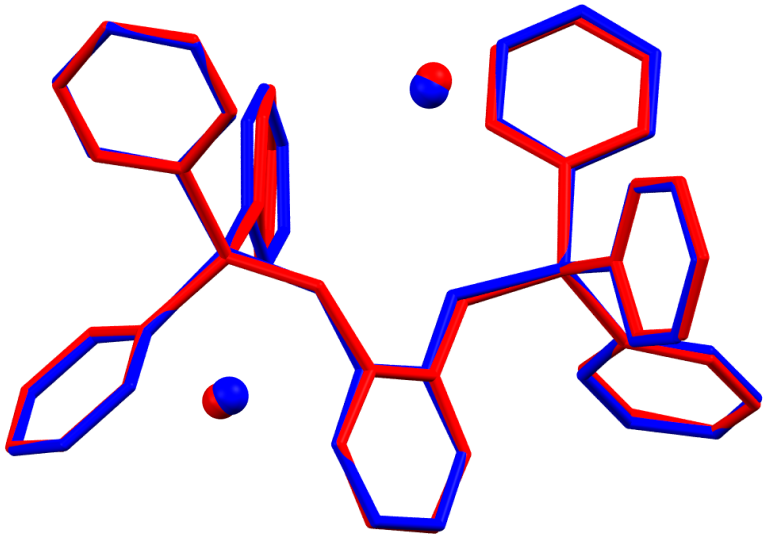


Figure 5.12 Overlay of *o*-Xtpb solved by MAD phasing (red) using only data collected at $2\theta = 20\,^{\circ}$ with restraints of the phenyl rings, and Superflip (blue) using full resolution data available.

5.3 Discussion

Current macromolecular techniques for structure solution from multiple wavelength experiments exploit the non-centrosymmetric nature of the molecules. As a consequence, the break-down of Friedel's Law means that twice as much data pertaining to the anomalous scatterers is present in each dataset. Moreover, since proteins are made up of a combination of just 20 amino acids, much information about the structure is already known before data have even been collected. In this chapter we have investigated the application of MAD phasing to the case of centrosymmetric structures containing relatively few heavy atoms. The size and centrosymmetric nature of such compounds meant that much less data about the anomalous scatterers could be obtained during the experiment and the MAD phasing procedure had to be modified accordingly.

This chapter has shown that we were successful in solving the structures of two small, centrosymmetric molecules using multiple wavelength techniques without assuming any prior knowledge of the structure. Although we were only able to apply the technique to these relatively small molecules during the course of this project,

the results have proven very promising. We were able to demonstrate that enough information about the anomalous scatterers could be obtained from four datasets collected on and around the absorption edge of bromine in order to solve the anomalous substructures. Using calculated phases from the substructure the rest of the structure could then be solved in both cases.

As we plan to apply the technique to larger molecules, the first foreseeable problem we are likely to encounter is over-complication of the Patterson map. As the number of anomalous scatterers, n, is increased, the number of peaks in the map increases as n^2. Therefore the map quickly becomes uninterpretable. However, it is has been shown that more complex substructures can be solved using a charge flipping algorithm as described in Section 1.2.1 (from data resolution as poor as 5-6 Å). [32]

Another problem of increasing the size of the molecule is the associated problem of radiation damage. Section 5.1.4 described the need for faster data collection; the long exposure times of the bromide rotaxane to the X-ray beam in EH1 destroyed the crystal (despite collecting at 100 K and placing a thin film of aluminium in the beam path to reduce the beam intensity). This problem was alleviated by collecting data using the shutterless detector in EH2, however, the *o*-Xtpb was still unable to survive four full-sphere data collections, and so the data collection strategy had to be shortened. The high intensity of the X-ray beam is mainly due to the beamline using an undulator to produce the radiation (see Section 2.1.1). Ideally, these experiments would be carried out on a beamline using a bending magnet insertion device since the intensity of radiation produced is much lower. Another benefit of the bending magnet is that tuning the energy is much simpler, since there is no need to change the ID gap in order to match the harmonics to the monochromated energy. As a result, the crystals should withstand irradiation for longer, thereby allowing more data to be collected on the sample. Figure 5.13 shows the relative intensity of radiation produced by an undulator and a bending magnet as a function of X-ray energy. Such

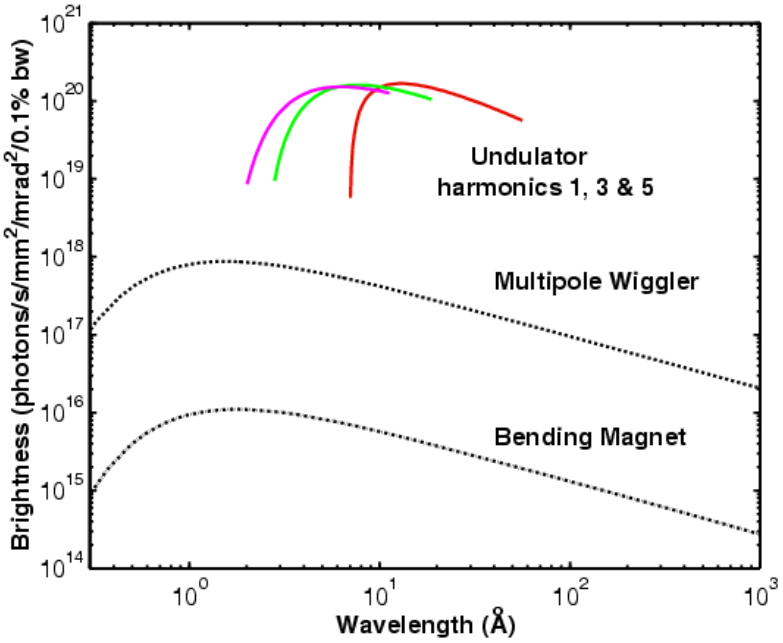


Figure 5.13 Graph of radiation intensity *vs.* energy for different insertion devices.[†]

an experiment however would require data collection on a different beamline – an option which we were unable to explore during this project.

The final point of discussion from the chapter concerns the completeness of data. We have shown that a solution is possible with as low as 54.3% completeness. Completeness is a major limitation for current small molecule solution techniques (more so for charge flipping than direct methods) and was the reason why a good solution was not possible using the data collected at 20° and 80°. The ability to solve the structure with such a low completeness is a major advantage since this is likely to become more of a problem as the size of the molecule being studied is increased (with consequent reduction of data resolution). Macromolecular crystallography often employs techniques to theoretically enhance the resolution (and completeness) of a dataset such as: Patterson maximum entropy methods; [182] electron-density peak deconvolution; [183] or a combination of density-modification and other artificial methods. [184, 185] Application of such techniques to 'large' small molecules would be the next step if completeness were to become more of an issue during the solution.

[†]image courtesy of http://pd.chem.ucl.ac.uk/pdnn/pdindex.htm#inst2

Chapter 6

Conclusions and future work

6.1 Conclusions

The work described in this thesis demonstrates that resonant X-ray diffraction enables the study of properties of materials that would otherwise have to be inferred from a variety of other techniques. All of the experiments described were carried out at Diamond Light Source to exploit the tunable wavelength capabilities available on Beamline I19.

Firstly, we commissioned the use of the Vortex detector on the beamline. This was subsequently used to collect absorption-fluorescence data on a variety of compounds containing one or more electronic states (oxidation and spin) of the same element. The Python program, DetOx, [1] determines anomalous scattering factors of mixed electronic-state species from these absorption-fluorescence data. Testing on the mixed-oxidation state compound $GaCl_2$ proved successful, showing that carefully selecting the X-ray energy in order to maximise differences in the anomalous scattering (between the different species), allows us to unambiguously identify the oxidation state at each Ga site within the crystal structure. The technique was also applied to the two-step SCO compound, $Fe(btr)_3(ClO_4)_2$, which contains one HS and one LS site at 200 K. Again, we showed that the occupancies of the different spin states could be refined, revealing that the HS site occupancy was entirely localised but the LS site exhibited some HS character. However, for samples containing more anomalous scatterers, a new technique was required.

We showed that a ratio refinement technique can be used to overcome the problem of high absorption in larger structures. We firstly confirmed the Mn^{II} and Mn^{III} sites within the mixed-valence Mn_3 cluster. However, due to the difference in anoma-

lous scattering between states being so small compared to the F_{000} (0.12%), a simulated annealing algorithm was used to determine the occupancies rather than the least-squares technique currently used in CRYSTALS. We next showed that the experimental data did not match the proposed structure for the heterometallic cluster $[Co_4{}^{III}Mn_5{}^{II}]$. With a larger difference in anomalous signal between the Mn and Co atoms, this same result was obtained using both a least-squares and simulated annealing ratio refinement, thereby providing strong evidence for a disordered-site model rather than the expected discrete Co_4Mn_5 model. Extension of the method to the more complex six-site, mixed-valence Mn_6 cluster revealed a limitation of the technique. The signal-to-noise in the data was too great to extract enough information to resolve the site occupancies and this has led us to consider improvements in data collection and analysis for future experiments.

Finally, we studied the application of macromolecular structure solution techniques to centrosymmetric, 'large' small molecules. These were the first multiple wavelength experiments to be carried out in EH2 of beamline I19 at Diamond, and therefore commissioning of the Pilatus detector was required. Optimisation of the data collection strategy and development of data-handling software such as DIAMOND2CRYSALIS were required before useful data could be collected. This has allowed for the successful solution of two bromide containing molecules (a bromide urea salt and *o*-Xylylenebis(triphenyl phosphonium bromide)) to within an RMSD of 0.1309 Å using 2 Å data – at this resolution, conventional small molecule structure solution techniques proved unsuccessful.

6.2 Future work

The next step for the mixed-atomic state discrimination work is for its integration into the CRYSTALS software. A full integration would allow other users to test the method on new samples and push the limits of its application. From the initial fluorescence-absorption scan, DetOx would suggest the ideal energies for data collection and sub-

sequently, following a multiple-dataset input into CRYSTALS, the occupancies at selected, ambiguous sites in the structure would be refined. Currently, only least-squares refinement of data can be performed; addition of the simulated annealing algorithm would permit for valence-contrast studies such as that of the Mn_3 cluster described.

The structure solution technique is still in the testing stage since it has only been applied so far to relatively small molecules as a proof of concept. The future direction for this part of the project would be to test larger and larger molecules, modifying the technique as necessary. Eventually, this too could be integrated into the CRYSTALS software as a new low resolution structure solution technique for small molecule crystallographers.

References

[1] K. J. Sutton, S. A. Barnett, K. E. Christensen, H. Nowell, A. L. Thompson, D. R. Allan, R. I. Cooper, *Journal of Synchrotron Radiation* **2013**, *20*, 200–4.

[2] P. Parois, K. J. Sutton, R. I. Cooper, *Journal of Chemical Crystallography* **2015**, in press.

[3] P. W. Betteridge, J. R. Carruthers, R. I. Cooper, K. Prout, D. J. Watkin, *Journal of Applied Crystallography* **2003**, *36*, 1487.

[4] M. J. Cooper, P. F. Mijnarends, N. Shiotani, N. Sakai, A. Bansil, *X-Ray Compton Scattering*, Oxford University Press, **2004**, p. 392.

[5] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, 3rd, Springer, **2006**.

[6] C. Giacovazzo, H. L. Monaco, G. Artioli, D. Viterbo, G. Ferraris, G. Gilli, G. Zanotti, M. Catti, *Fundamentals of Crystallography*, (Ed.: C. Giacovazzo), Oxford University Press, **2002**.

[7] B. Warren, *X-ray Diffraction*, Addison-Wesley Publishing Co., **1969**.

[8] R. W. Grosse-Kunstleve, P. D. Adams, *Journal of Applied Crystallography* **2002**, *35*, 477–480.

[9] S. Hassani, English in *Mathematical Methods SE - 5*, Springer New York, **2009**, pp. 139–170.

[10] W. L. Bragg, *Nature* **1912**, *90*, 410.

[11] H. Hönl, *Annalen der Physik* **1933**, *18*, 625.

[12] H. Hönl, *Zeitschrift für Physik* **1933**, *84*, 11.

[13] L. G. Parratt, C. F. Hempstead, *Physical Review* **1954**, *94*, 1593–1600.

[14] D. T. Cromer, D. A. Liberman, *Acta Crystallographica Section A* **1981**, *37*, 267–268.

[15] H. K. Wagenfeld, *Zeitschrift für Physik B - Condensed Matter* **1987**, *65*, 437–441.

[16] W. A. Hendrickson, C. M. Ogata, *Methods in Enzymology* **1997**, *276*, 494–523.

[17] W. A. Hendrickson, J. L. Smith, S Sheriff, *Methods in Enzymology* **1985**, *115*, 41–55.

[18] W. Hendrickson, *Science* **1991**, *254*, 51–58.

[19] T. C. Terwilliger, J Berendzen, *Acta Crystallographica. Section D Biological Crystallography* **1996**, *52*, 749–57.

[20] R. W. James, *The Optical Principles of the Diffraction of X-rays*, G. Bell and Sons Ltd, London, **1969**.

[21] H. Fujiwara, *Spectroscopic Ellipsometry: Principles and Applications*, John Wiley and Sons, **2007**.

[22] D. R. Hartree, d. L. Kronig, H. Petersen, *Physica* **1934**, *1*, 895–924.

[23] G. Evans, R. F. Pettifer, *Journal of Applied Crystallography* **2001**, *34*, 82–86.

[24] D. T. Cromer, *Journal of Applied Crystallography* **1983**, *16*, 437.

[25] H. Lipson, W. Cochran, *The Determination of Crystal Structures*, G. Bell and Sons Ltd, London, London, **1953**.

[26] C. Giacovazzo, *Phasing in Crystallography: A Modern Perspective*, Oxford University Press, **2013**, p. 432.

[27] A. J. Blake, W. Clegg, J. M. Cole, J. S. O. Evans, P. Main, S. Parson, D. J. Watkin, *Crystal Structure Analysis: Principles and Practice*, 2nd, (Ed.: W. Clegg), Oxford University Press, **2009**.

[28] L. Ooi, *Principles of X-ray Crystallography*, Oxford University Press, **2010**.

[29] R. Y. Rubinstein, D. P. Kroese, *Simulation and the Monte Carlo Method*, 2nd, John Wiley and Sons, **2008**.

[30] L. Palatinus, *Acta Crystallographica Section B Structural Science Crystal Engineering and Materials* **2013**, *69*, 1–16.

[31] G. Oszlányi, A. Süto, *Acta Crystallographica. Section A Foundations of crystallography* **2008**, *64*, 123–34.

[32] C. Dumas, A. van der Lee, *Acta Crystallographica. Section D Biological Crystallography* **2008**, *D64*, 864–73.

[33] L. Palatinus, A. van der Lee, *Journal of Applied Crystallography* **2008**, *41*, 975–984.

[34] E. F. Garman, *Acta Crystallographica. Section D Biological Crystallography* **2010**, *66*, 339–51.

[35] J. Kmetko, M. Warkentin, U. Englich, R. E. Thorne, *Acta Crystallographica. Section D Biological Crystallography* **2011**, *67*, 881–93.

[36] G. N. Ramachandran, R. Srinivasan, *Fourier Methods in Crystallography*, John Wiley and Sons, **1970**, p. 259.

[37] M. G. Rossmann, *Acta Crystallographica* **1960**, *13*, 221–226.

[38] E. Garman, J. W. Murray, *Acta Crystallographica Section D Biological Crystallography* **2003**, *59*, 1903–1913.

[39] M. F. Perutz, *Acta Crystallographica* **1956**, *9*, 867–873.

[40] J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, D. C. Phillips, *Nature* **1958**, *181*, 662–666.

[41] G. H. Stout, L. H. Jensen, *X-Ray Structure Determination: A Practical Guide*, 2nd, John Wiley and Sons, **1989**.

[42] G. Scapin, *Acta Crystallographica. Section D Biological Crystallography* **2013**, *69*, 2266–75.

[43] M. G. Rossmann, D. M. Blow, *Acta Crystallographica* **1962**, *15*, 24–31.

[44] A. Vagin, A. Teplyakov, *Acta Crystallographica. Section D Biological Crystallography* **2010**, *66*, 22–5.

[45] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *Journal of Applied Crystallography* **2007**, *40*, 658–674.

[46] E. Prince, P. T. Boggs in *International Tables for Crystallography*, **2006**, Chapter 8.1, pp. 678–688.

[47] J. L. Sussman, S. R. Holbrook, G. M. Church, S. Kim, *Acta Crystallographica Section A* **1977**, 580–584.

[48] B. E. Prince, D. M. Collins in *International Tables for Crystallography, Vol. C*, **2006**, Chapter 8.2, pp. 689–692.

[49] E. Nishibori, T. Nakamura, M. Arimoto, S. Aoyagi, H. Ago, M. Miyano, T. Ebisuzaki, M. Sakata, *Acta Crystallographica. Section D Biological Crystallography* **2008**, *64*, 237–47.

[50] J. Kuriyan, A. T. Brünger, M. Karplus, W. A. Hendrickson, *Acta Crystallographica Section A Foundations of Crystallography* **1989**, *45*, 396–409.

[51] W. C. Hamilton, *Acta Crystallographica* **1965**, *18*, 502–510.

[52] C. Canuto, A. Tabacco, English in *Mathematical Analysis I SE - 7*, Universitext, Springer Milan, **2008**, pp. 223–255.

[53] D. W. Bennett, *Understanding Single-Crystal X-Ray Crystallography*, Wiley, **2010**.

[54] C. Hunter, *Nature* **2011**, *469*, 39–41.

[55] C. J. Gilmore, *Acta Crystallographica Section D Biological Crystallography* **2000**, *56*, 1205–1214.

[56] P. Müller, *Crystallography Reviews* **2009**, *15*, 57–83.

[57] N. K. S. Davis, A. L. Thompson, H. L. Anderson, *Journal of the American Chemical Society* **2011**, *133*, 30–1.

[58] J. K. Sprafke, D. V. Kondratuk, M. Wykes, A. L. Thompson, M. Hoffmann, R. Drevinskas, W.-H. Chen, C. K. Yong, J. Karnbratt, J. E. Bullock, M. Malfois, M. R. Wasielewski, B. Albinsson, L. M. Herz, D. Zigmantas, D. Beljonne, H. L. Anderson, *Journal of the American Chemical Society* **2011**, *133*, 17262–17273.

[59] M. J. Langton, J. D. Matichak, A. L. Thompson, H. L. Anderson, *Chemical Science* **2011**, *2*, 1897.

[60] M. Nowostawska, S. A. Corr, S. J. Byrne, J. Conroy, Y. Volkov, Y. K. Gun'ko, *Journal of Nanobiotechnology* **2011**, *9*, 13.

[61] Y. Chen, K. Wang, *Photochemical & Photobiological Sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology* **2013**, *12*, 2001–7.

[62] M. G. Walter, A. B. Rudine, C. C. Wamser, *Journal of Porphyrins and Phthalocyanines* **Sept. 2010**, *14*, 759–792.

[63] W. A. Hendrickson, *Journal of Synchrotron Radiation* **1999**, *6*, 845–851.

[64] W. A. Hendrickson, M. M. Teeter, *Nature* **1981**, *290*, 107–113.

[65] G. L. Taylor, *Acta Crystallographica. Section D Biological Crystallography* **2010**, *66*, 325–38.

[66] J. Karle, *International Journal of Quantum Chemistry* **1980**, *18*, 357–367.

[67] J. Karle, *Acta Crystallographica Section A Foundations of Crystallography* **1989**, *45*, 765–781.

[68] J. L. Smith, W. A. Hendrickson, T. C Terwilliger, J. Berendzen, *International Tables for Crystallography, Vol. F*, Kluwer Academic Publishers, **2006**, Chapter 14.2, pp. 299–309.

[69] D. T. Cromer, *The Journal of Chemical Physics* **1970**, *53*, 1891.

[70] R. L. Kingston, *Acta Crystallographica Section D Biological Crystallography* **2001**, *57*, 101–107.

[71] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, **1982**, p. 395.

[72] G. D. Smith, B. Nagar, J. M. Rini, H. A. Hauptman, R. H. Blessing, *Acta Crystallographica Section D Biological Crystallography* **1998**, *54*, 799–804.

[73] A. Pähler, J. L. Smith, W. A. Hendrickson, *Acta Crystallographica Section A Foundations of Crystallography* **1990**, *46*, 537–540.

[74] W. Yang, W. A. Hendrickson, R. J. Crouch, Y. Satow, *Science* **1990**, *249*, 1398–1405.

[75] T. C. Terwilliger, *Acta Crystallographica. Section D Biological Crystallography* **1994**, *50*, 11–6.

[76] T. C. Terwilliger, J Berendzen, *Acta Crystallographica. Section D Biological Crystallography* **1997**, *53*, 571–9.

[77] W. C. Hamilton, *Statistics in Physical Science: estimation, hypothesis testing, and least squares*, Ronald Press Co., **1964**, p. 230.

[78] C. W. Carter, R. M. Sweet, *Macromolecular Crystallography*, Academic Press, **1997**.

[79] A. J. C. Wilson, *Acta Crystallographica* **1949**, *2*, 318–321.

[80] T. C. Terwilliger, D. Eisenberg, *Acta Crystallographica* **1987**, 6–13.

[81] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, *Acta crystallographica. Section D Biological crystallography* **2011**, *67*, 235–42.

[82] Z. Otwinowski, W. Minor in *International Tables for Crystallography Volume F: Crystallography of biological macromolecules SE - 24*, (Eds.: M. G. Rossmann, E Arnold), International Tables for Crystallography, Springer Netherlands, **2001**, pp. 226–235.

[83] CCP4, *Acta crystallographica. Section D Biological crystallography* **1994**, *50*, 760–3.

[84] S. French, K. Wilson, *Acta Crystallographica* **1978**, *A34*, 517–525.

[85] Z. Otwinowski in Daresbury Study Weekend proceedings, **1991**.

[86] G. M. Sheldrick, *Acta crystallographica. Section D Biological crystallography* **2010**, *66*, 479–85.

[87] P. D. Adams, R. W. Grosse-Kunstleve, L.-W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter, T. C. Terwilliger, *Acta Crystallographica Section D Biological Crystallography* **2002**, *58*, 1948–1954.

[88] P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallographica Section D* **2010**, *66*, 213–221.

[89] A. D. McNaught, A. Wilkinson., *Compendium of Chemical Terminology, 2nd ed. (the "Gold Book")*, 2nd, Blackwell Scientific Publications, Oxford, **1997**.

[90] H. Vahrenkamp, A. Geiß, G. N. Richardson, *J. Chem. Soc. Dalton Trans.* **1997**, 3643–3651.

[91] N. S. Hush, *Prog. Inorg. Chem.* **1967**, *8*, 391–444.

[92] M. B Robin, P. Day, *Adv. Inorg. Chem. Radiochem* **1967**, *10*, 247–422.

[93] P. Day, N. S. Hush, R. J. H. Clark, *Philosophical Transactions. Series A Mathematical Physical and Engineering Sciences* **2008**, *366*, 5–14.

[94] R. Burns, D. A. Nolet, K. M. Parkin, C. A. McCammon, K. B. Schwartz in *Mixed-Valence Compounds SE - 10, Vol. 58*, (Ed.: D. Brown), Springer Netherlands, **1980**, pp. 295–336.

[95] L. Cambi, L. Szego, *Ber. Drsch. Chem. Ges.* **1931**, *64*, 259.

[96] A. Bousseksou, G. Molnár, L. Salmon, W. Nicolazzi, *Chemical Society Reviews* **2011**, *40*, 3313–35.

[97] P. Gütlich, A. Hauser, H. Spiering, *Angewandte Chemie (International Ed. in English)* **1994**, *33*, 2024–2054.

[98] G. Molnár, L. Salmon, W. Nicolazzi, F. Terki, A. Bousseksou, *Journal of Materials Chemistry C* **2014**, *2*, 1360.

[99] M. Cavallini, *Physical Chemistry Chemical Physics* **Sept. 2012**, *14*, 11867–76.

[100] M. A. Halcrow, *Spin-Crossover Materials, Properties and Applications*, (Ed.: M. A. Halcrow), Wiley, **2013**, p. 564.

[101] M. A. Halcrow, *Chemical Communications* **2013**, *49*, 10890–2.

[102] T. Matsumoto, G. N. Newton, T. Shiga, S. Hayami, Y. Matsui, H. Okamoto, R. Kumai, Y. Murakami, H. Oshio, *Nature Communications* **2014**, *5*, 3865.

[103] P. Gütlich, H. A. Goodwin, *Spin Crossover in Transition Metal Compounds I*, Springer, **2004**, p. 341.

[104] P. Gütlich, H. A. Goodwin, *Topics in Current Chemistry* **2004**, *233*, 1–47.

[105] O. Kahn, J. Kröber, C. Jay, *Advanced Materials* **1992**, *4*, 718–728.

[106] P. Gütlich, H. Andreas, *Coordination Chemistry Reviews* **1990**, *97*, 1–22.

[107] O. Kahn, C. J. Martinez, *Science* **1998**, *279*, 44–48.

[108] G. Ritter, E. Konig, W. Irler, H. A. Goodwin, *Inorganic Chemistry* **1978**, *17*, 224–228.

[109] D. T. Cromer, *The Journal of Chemical Physics* **1967**, *47*, 1892.

[110] M. Ladd, R. Palmer, *Structure Determination by X-ray Crystallography: Analysis by X-rays and Neutrons*, 5th, Springer, **2013**, p. 756.

[111] E. Shoko, M. Smith, R. McKenzie, *Physical Review B* **2009**, *79*, 134108.

[112] J. K. Burdett, T. J. McLarnan, *American Mineralogist* **1984**, *69*, 601–621.

[113] R. M. Wood, G. J. Palenik, *Inorganic Chemistry* **1998**, *37*, 4149–4151.

[114] P. Marcon, K. Ostanina, *Progress in Electromagnetic Research Symposium* **2012**, 420–424.

[115] O. Kahn, *Molecular Magnetism*, Wiley, **1993**.

[116] F. Palacio in *Molecular Magnetism: From Molecular Assemblies to the Devices SE - 2, Vol. 321*, Springer Netherlands, **1996**, pp. 5–63.

[117] H. W. Fulbright, *Journal of the American Chemical Society* **1963**, *85*, 838.

[118] R. H. Herber, *Journal of Chemical Education* **1965**, *42*, 180–187.

[119] M. Newville, Fundamentals of XAFS, **2004**.

[120] C. Fadley, *Journal of Electron Spectroscopy and Related Phenomena* **2010**, *178-179*, 2–32.

[121] C. R. Brundle, A. D. Baker, *Electron Spectroscopy: Theory, Techniques and Applications*, Academic Press, **1978**, p. 145.

[122] S. M. Kanan, N. Abdo, M. Khalil, X. Li, I. A. Abu-Yousef, F. Barilrobert, H. H. Patterson, *Applied Catalysis B: Environmental* **2011**, *106*, 350–358.

[123] A. P. Wilkinson, A. K. Cheetham, *Journal of Applied Crystallography* **1992**, *25*, 654–657.

[124] J. P. Attfield, *Nature* **1990**, *343*.

[125] S. Sasaki, T. Toyoda, K. Yamawaki, K. Ohkubo, *Journal of synchrotron radiation* **1998**, *5*, 920–2.

[126] J. K. Warner, A. K. Cheetham, D. E. Cox, R. B. V. Dreele, *Journal of the American Chemical Society* **1992**, *114*, 6074–6080.

[127] G. H. Kwei, R. B. V. Dreele, S.-W. Cheong, Z. Fisk, J. D. Thompson, *Physical Review B* **1990**, *41*, 1889–1893.

[128] H. Palancher, S. Bos, J. F. Bérar, I. Margiolaki, J. L. Hodeau, *The European Physical Journal Special Topics* **2012**, *208*, 275–289.

[129] Y. Zhang, A. P. Wilkinson, P. L. Lee, S. D. Shastri, D. Shu, D.-Y. Chung, M. G. Kanatzidis, *Journal of Applied Crystallography* **2005**, *38*, 433–441.

[130] H. Ohara, S. Sasaki, Y. Konoike, T. Toyoda, K. Yamawaki, M. Tanaka, *Physica B: Condensed Matter* **2004**, *350*, 353–365.

[131] H. D. Flack, M. Sadki, A. L. Thompson, D. J. Watkin, *Acta crystallographica. Section A Foundations of crystallography* **2011**, *67*, 21–34.

[132] H. D. Flack, G. Bernardinelli, *Chirality* **2008**, *20*, 681–690.

[133] D. F. Swinehart, *Journal of Chemical Education* **1962**, *39*, 333.

[134] H. Nowell, S. A. Barnett, K. E. Christensen, S. J. Teat, D. R. Allan, *Journal of Synchrotron Radiation* **2012**, *19*, 435–41.

[135] S. Corde, K. Ta Phuoc, G. Lambert, R. Fitour, V. Malka, A. Rousse, A. Beck, E. Lefebvre, *Reviews of Modern Physics* **2013**, *85*, 1–48.

[136] G. Margaritondo, *Introduction to Synchrotron Radiation*, Oxford University Press, **1988**, p. 280.

[137] A. Thompson, D. Attwood, E. Gillikson, M. Howells, K. Kim, J. Kirz, J. Kortright, I. Lindau, P. Pianetta, A. Robinson, J. Scofield, J. Underwood, D. Vaughan, G. Williams, H. Winick, *X-ray Data Booklet*, Second, (Eds.: A. C. Thompson, D Vaughan), Lawrence Berkeley National Laboratory, University of California, **Jan. 2001**.

[138] A. Naik, J. Marchand-Brynaert, Y. Garcia, *Synthesis* **2008**, *2008*, 149–154.

[139] J. G. Haasnoot, W. L. Groeneveld, *Z. Naturforsch* **1979**, *34b*, 15000.

[140] Y. Garcia, O. Kahn, L. Rabardel, B. Chansou, L. Salmon, J. P. Tuchagues, *Inorganic Chemistry* **1999**, 4663–4670.

[141] T. Lis, *Acta Crystallographica* **1980**, *B36*, 2042–2046.

[142] M. Murrie, S. Parsons, R. E. P. Winpenny, *J. Chem. Soc. Dalton Trans.* **1998**, *3*, 1423–1424.

[143] J. B. Vincent, H.-R. Chang, K. Folting, J. C. Huffman, G. Christou, D. N. Hendrickson, *Journal of the American Chemical Society* **1987**, *109*, 5703–5711.

[144] T. Sala, M. V. Sargent, *J. Chem. Soc. Chem. Commun* **1978**, 253–254.

[145] K. M. Engle, G. W. Pidgeon, L. Pfeifer, A. L. Thompson, H. A. Sparkes, R. S. Paton, J. M. Brown, V. Gouverneur, "Unpublished Results", **2014**.

[146] A. Kraft, *Liebigs Annalen-Recueil* **1997**, 1463–1471.

[147] A. J. McConnell, C. J. Serpell, A. L. Thompson, D. R. Allan, P. D. Beer, *Chemistry: A European Journal* **2010**, *16*, 1256–64.

[148] L. M. Hancock, L. C. Gilday, S. Carvalho, P. J. Costa, V. Félix, C. J. Serpell, N. L. Kilah, P. D. Beer, *Chemistry: A European Journal* **2010**, *16*, 13082–94.

[149] B. Spingler, S. Schnidrig, T. Todorova, F. Wild, *Cryst Eng Comm* **2012**, *14*, 751.

[150] P. C. Sadek, *The HPLC Solvent Guide*, 2nd, John Wiley and Sons, **2002**, p. 664.

[151] G. van Rossum, F. L. Drake, Python Reference Manual, Virginia, USA, **2001**.

[152] P. Bandyopadhyay, C. U. Segre, *http://www.csrri.iit.edu/mucal.html*.

[153] D. Miller, W. Scott, *Deconvolution with Inverse and Weiner Filters*, OpenStax CNX, **2006**.

[154] P. A. Gorry, *Analytical Chemistry* **1990**, *62*, 570–573.

[155] J. J. Hoyt, D Defontaine, W. K. Warburton, *Journal of Applied Crystallography* **1984**, *17*, 344–351.

[156] D. Ascher, P. Dubois, K. Hinsen, J. Hugunin, T. Oliphant, Numerical Python, Lawrence Livermore National Laboratory, California, USA, **2001**.

[157] G. Garton, H. Powell, *Journal of Inorganic and Nuclear Chemistry* **1957**, *4*, 84–89.

[158] H. Schmidbaur, R. Nowak, W. Bublak, P. Burkert, B. Huber, G. Muller, *Zeitschrift für Naturforschung. Teil b Anorganische Chemie Organische Chemie* **1987**, *42*, 553–556.

[159] A. P. Wilkinson, A. K. Cheetham, D. E. Cox, *Acta Crystallographica Section B Structural Science* **1991**, *47*, 155–161.

[160] A. Altomare, M. C. Burla, M. Camalli, G. L. Cascarano, C. Giacovazzo, A. Guagliardi, A. G. Moliterni, G. Polidori, R. Spagna, *Journal of Applied Crystallography* **1999**, *32*, 115–119.

[161] W. Vreugdenhil, J. H. Van Diemen, R. A. G. De Graaff, J. G. Haasnoott, J. Reedijk, *Polyhedron* **1990**, *9*, 2971–2979.

[162] A. Ozarowski, Y. Shunzhong, B. R. McGarvey, A. Mislankar, J. E. Drake, *Inorganic Chemistry* **1991**, 3167–3174.

[163] J. J. M. Amoore, C. J. Kepert, J. D. Cashion, B. Moubaraki, S. M. Neville, K. S. Murray, *Chemistry: A European Journal* **2006**, *12*, 8220–7.

[164] I. Vorontsov, S. Pillet, R. Kaminski, M. S. Schmokel, P. Coppens, *Journal of Applied Crystallography* **2010**, *43*, 1129–1130.

[165] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, *Science* **1983**, *220*, 671–680.

[166] V. Cerny, *Journal of Optimization Theory and Applications* **1985**, *45*, 41–51.

[167] D. Bertsimas, J. Tsitsiklis, *Statistical Science* **1993**, *8*, 10–15.

[168] R. Sessoli, H. Tsai, A. R. Schake, S. Wang, J. B. Vincent, K. Folting, D. Gatteschi, G. Christou, D. N. Hendrickson, *Journal of the American Chemical Society* **1993**, 1804–1816.

[169] M. Murrie, *Chemical Society Reviews* **2010**, *39*, 1986–95.

[170] K. M. Mertes, Y. Suzuki, M. P. Sarachik, Y. Myasoedov, H. Shtrikman, E. Zeldov, E. M. Rumberger, D. N. Hendrickson, G. Christou, *Solid State Communications* **2003**, *127*, 131–139.

[171] R. Bagai, G. Christou, *Chemical Society Reviews* **2009**, *38*, 1011–26.

[172] G. Christou, D. Gatteschi, D. N. Hendrickson, R. Sessoli, *MRS Bulletin* **2011**, *25*, 66–71.

[173] D. Ruiz-Molina, G. Christou, D. N. Hendrickson, *Mol. Cryst. and Liq. Cryst.* **2000**, *343*, 17–27.

[174] C. Cadiou, M. Murrie, C. Paulsen, V. Villar, W. Wernsdorfer, R. E. P. Winpenny, *Chemical Communications* **2001**, 2666–2667.

[175] S. L. Castro, Z. Sun, C. M. Grant, J. C. Bollinger, D. N. Hendrickson, G. Christou, S. Diego, L. Jolla, *Journal of the American Chemical Society* **1998**, *7863*, 2365–2375.

[176] S. K. Langley, D. P. Wielechowski, V. Vieru, N. F. Chilton, B. Moubaraki, B. F. Abrahams, L. F. Chibotaru, K. S. Murray, *Angewandte Chemie (International ed. in English)* **2013**, 12014–12019.

[177] M. Murrie, *Personal Communication* **2014**.

[178] Agilent, *CrysAlisPro* **2012**.

[179] P. Müller, R. Herbst-Irmer, A. Spek, T. Schneider, M. Sawaya, *Crystal Structure Refinement: A Crystallographer's Guide to SHELXL*, International Union of Crystallography, **2006**.

[180] M. Hanif, P. Lu, H. Xu, Z. Tian, B. Yang, Z.-m. Wang, L. Tian, Y. Xu, Z. Xie, Y. Ma, *Chemical Research in Chinese Universities* **2009**, *25*, 754–759.

[181] P. Parois, K. Sutton, M. Warren, R. I. Cooper, "Unpublished Results", **2014**.

[182] L. Palatinus, W. Steurer, G. Chapuis, *Journal of Applied Crystallography* **2007**, *40*, 456–462.

[183] A. Altomare, C. Cuocci, C. Giacovazzo, G. S. Kamel, A. Moliterni, R. Rizzi, *Acta Crystallographica. Section A Foundations of Crystallography* **2008**, *64*, 326–36.

[184] Y. Jia-xing, M. M. Woolfson, K. S. Wilson, E. J. Dodson, *Acta Crystallographica. Section D Biological Crystallography* **2005**, *61*, 1465–75.

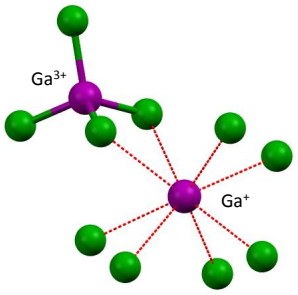
[185] R. Caliandro, B. Carrozzini, G. L. Cascarano, L. De Caro, C. Giacovazzo, D. Siliqi, *Acta Crystallographica. Section D Biological Crystallography* **2005**, *61*, 556–65.

Appendices

Appendix A

Crystallographic Information

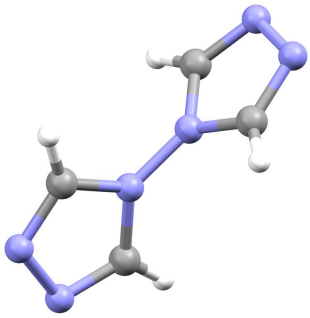
A.1 $GaCl_2$



Synchrotron Data	Below
Empirical Formula	Ga Cl2
Molecular Weight	281.26
Temperature	100 K
Wavelength	1.2011 Å
Crystal system, space group	Orthorhombic, $Pnna$
Unit cell dimensions	a = 7.1811(3) Å α = 90
	b = 9.5714(2) Å β = 90
	c = 9.4712(4) Å γ = 90
Volume	650.98(3) Å^3
Z	4
Density	2.870 Mg/m^3
Absorption coefficient	9.757 mm^{-1}
F(000)	520
Crystal size	0.10 x 0.10 x 0.20 mm
Theta range for data collection	5.12 to 31.94°
Index ranges	$-5 \leq h \geq 6$
	$-8 \leq k \geq 8$
	$-8 \leq l \geq 8$
Reflections collected	1253
Independent reflection	222
Completeness	99.5% to 26.83°
Absorption correction	Integration
Refinement method	Full-matrix least squares on F^2
Data / restraints / parameters	222 / 0 / 30
Goodness-of-fit on F^2	0.873
Final R indicies [I>2sigma(I)]	R_1 = 0.039, wR_2 = 0.124
R indices (all data)	R_1 = 0.061, wR_2 = 0.127
Extinction coefficient	None

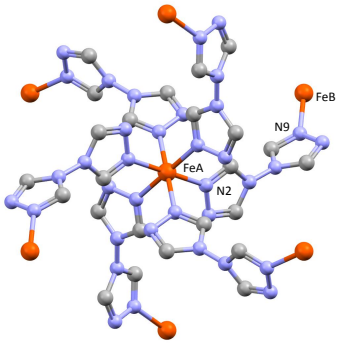
Synchrotron Data	On Edge
Empirical Formula	Ga Cl2
Molecular Weight	281.26
Temperature	100 K
Wavelength	1.1949 Å
Crystal system, space group	Orthorhombic, $Pnna$
Unit cell dimensions	a = 7.1858(3) Å α = 90
	b = 9.5784(3) Å β = 90
	c = 9.4775(5) Å γ = 90
Volume	652.32(3) Å^3
Z	4
Density	2.864 Mg/m^3
Absorption coefficient	9.757 mm^{-1}
F(000)	520
Crystal size	0.10 x 0.10 x 0.20 mm
Theta range for data collection	5.09 to 31.74°
Index ranges	$-5 \leq h \geq 6$
	$-8 \leq k \geq 8$
	$-8 \leq l \geq 8$
Reflections collected	1265
Independent reflection	227
Completeness	99.5% to 26.67°
Absorption correction	Integration
Refinement method	Full-matrix least squares on F^2
Data / restraints / parameters	227 / 0 / 30
Goodness-of-fit on F^2	1.083
Final R indicies [I>2sigma(I)]	R_1 = 0.039, wR_2 = 0.100
R indices (all data)	R_1 = 0.044, wR_2 = 0.106
Extinction coefficient	None

A.2 Btr Ligand



Lab Data	
Empirical Formula	C2 N3 H2
Molecular Weight	136
Temperature	150 K
Wavelength	Mo K_α 0.7017 Å
Crystal system, space group	Orthorhombic, $Pnna$
Unit cell dimensions	a = 6.9260(18) Å α = 90
	b = 7.2685(19) Å β = 90
	c = 11.097(3) Å γ = 90
Volume	558.65(16) Å^3
Z	4
Density	1.567 Mg/m^3
Absorption coefficient	0.119 mm^{-1}
F(000)	280
Crystal size	0.40 x 0.40 x 0.20 mm
Theta range for data collection	5.47 to 27.57°
Index ranges	$-9 \leq h \geq 9$
	$-7 \leq k \geq 8$
	$-14 \leq l \geq 14$
Reflections collected	3381
Independent reflection	532
Completeness	94.5% to 25.1°
Absorption correction	Integration
Refinement method	Full-matrix least squares on F^2
Data / restraints / parameters	532 / 4 / 57
Goodness-of-fit on F^2	0.9281
Final R indicies [I>2sigma(I)]	R_1 = 0.051, wR_2 = 0.097
R indices (all data)	R_1 = 0.075, wR_2 = 0.112
Extinction coefficient	None

A.3 $Fe(btr)_3(ClO_4)_2$



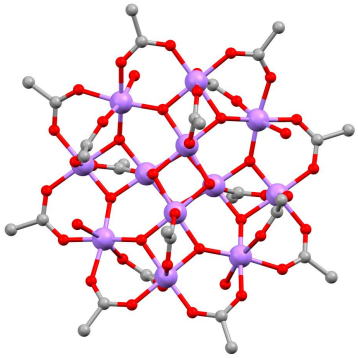
Lab Data	130 K
Empirical Formula	Fe C12 H12 Cl2 O8 N18
Molecular Weight	663
Temperature	130 K
Wavelength	Mo K_α 0.7017 Å
Crystal system, space group	Trigonal, $R\bar{3}$
Unit cell dimensions	a = 11.0422(14) Å α = 90
	b = 11.0422(14) Å β = 90
	c = 33.356(5) Å γ = 120
Volume	3522.3(5) $Å^3$
Z	6
Density	1.876 Mg/m^3
Absorption coefficient	0.955 mm^{-1}
F(000)	2004.0
Crystal size	0.10 x 0.10 x 0.10 mm
Theta range for data collection	5.25 to 25.39°
Index ranges	$-13 \leq h \geq 13$
	$-10 \leq k \geq 11$
	$-37 \leq l \geq 39$
Reflections collected	1135
Independent reflection	937
Completeness	64.4% to 25.39°
Absorption correction	Integration
Refinement method	Full-matrix least squares on F^2
Data / restraints / parameters	937 / 0 / 126
Goodness-of-fit on F^2	0.943
Final R indicies [I>2sigma(I)]	R_1 = 0.067, wR_2 = 0.151
R indices (all data)	R_1 = 0.096, wR_2 = 0.168
Extinction coefficient	None

Lab Data	200 K
Empirical Formula	Fe C12 H12 Cl2 O8 N18
Molecular Weight	663
Temperature	250 K
Wavelength	Cu K_α 1.541 Å
Crystal system, space group	Trigonal, $R\bar{3}$
Unit cell dimensions	a = 10.9611(4) Å α = 90
	b = 10.9611(4) Å β = 90
	c = 34.6338(16) Å γ = 120
Volume	3603.6(17) $Å^3$
Z	6
Density	1.833 Mg/m^3
Absorption coefficient	7.844 mm^{-1}
F(000)	2004.0
Crystal size	0.10 x 0.10 x 0.10 mm
Theta range for data collection	3.83 to 74.67°
Index ranges	$-13 \leq h \geq 11$
	$-9 \leq k \geq 13$
	$-43 \leq l \geq 40$
Reflections collected	8909
Independent reflection	1642
Completeness	99.5% to 74.7°
Absorption correction	Integration
Refinement method	Full-matrix least squares on F^2
Data / restraints / parameters	1642 / 0 / 137
Goodness-of-fit on F^2	0.852
Final R indicies [I>2sigma(I)]	R_1 = 0.048, wR_2 = 0.099
R indices (all data)	R_1 = 0.050, wR_2 = 0.101
Extinction coefficient	None

Lab Data	250 K
Empirical Formula	Fe C12 H12 Cl2 O8 N18
Molecular Weight	663
Temperature	250 K
Wavelength	Cu K_α 1.541 Å
Crystal system, space group	Trigonal, $R\bar{3}$
Unit cell dimensions	a = 11.0152(3) Å α = 90
	b = 11.0152(3) Å β = 90
	c = 35.5705(11) Å γ = 120
Volume	3737.73(11) $Å^3$
Z	6
Density	1.767 Mg/m^3
Absorption coefficient	7.563 mm^{-1}
F(000)	2004.0
Crystal size	0.10 x 0.10 x 0.10 mm
Theta range for data collection	3.73 to 74.66°
Index ranges	$-11 \leq h \geq 13$
	$-13 \leq k \geq 11$
	$-44 \leq l \geq 43$
Reflections collected	3043
Independent reflection	1654
Completeness	99.5% to 66.4°
Absorption correction	Integration
Refinement method	Full-matrix least squares on F^2
Data / restraints / parameters	1654 / 0 / 124
Goodness-of-fit on F^2	0.871
Final R indicies [I>2sigma(I)]	R_1 = 0.072, wR_2 = 0.134
R indices (all data)	R_1 = 0.079, wR_2 = 0.141
Extinction coefficient	None

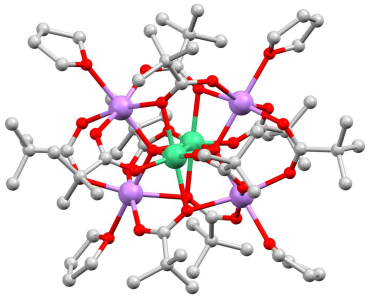
Synchrotron Data	Below - 7105 eV	On Edge - 7119 eV
Temperature	250 K	250 K
Wavelength	1.7450 Å	1.7416 Å
Crystal system, space group	Trigonal, $R\bar{3}$	Trigonal, $R\bar{3}$
Unit cell dimensions	a = 10.9939(6) Å	a = 10.9948(5) Å
	b = 10.9939(6) Å	b = 10.9948(5) Å
	c = 35.645(2) Å	c = 35.6485(18) Å
	α = 90	α = 90
	β = 90	β = 90
	γ = 120	γ = 120
Volume	3731.0(4) $Å^3$	3732.0(3) $Å^3$
Z	6	6
Absorption coefficient	44.73 mm^{-1}	44.49 mm^{-1}
Reflections collected	4697	4754
Completeness	94.7% to 73.34°	94.7% to 74.05°

A.4 $Mn_{12}O_{12}(O_2CCH_3)_{16} \cdot 4\,H_2O \cdot 2\,CH_3CO_2H$



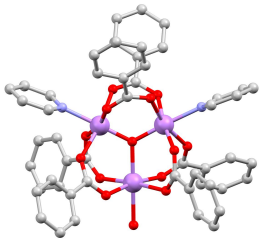
Lab Data	
Empirical Formula	Mn3 C9 H18 O14
Molecular Weight	2060.18
Temperature	150 K
Wavelength	Mo K_α 0.7017 Å
Crystal system, space group	Tetragonal, $I\bar{4}$
Unit cell dimensions	a = 17.2376(2) Å α = 90
	b = 17.2376(2) Å β = 90
	c = 12.11280(10) Å γ = 90
Volume	3599.14(3) Å^3
Z	2
Density	1.775 Mg/m^3
Absorption coefficient	2.131 mm^{-1}
F(000)	1864
Crystal size	0.30 x 0.20 x 0.40 mm
Theta range for data collection	5.16 to 27.48°
Index ranges	$-22 \leq h \geq 22$
	$-15 \leq k \geq 15$
	$-15 \leq l \geq 15$
Reflections collected	4132
Independent reflection	4129
Completeness	99.1% to 27.5°
Absorption correction	Integration
Refinement method	Full-matrix least squares on F^2
Data / restraints / parameters	4129 / 0 / 226
Goodness-of-fit on F^2	0.873
Final R indicies [I>2sigma(I)]	R_1 = 0.046, wR_2 = 0.104
R indices (all data)	R_1 = 0.048, wR_2 = 0.105
Extinction coefficient	None

A.5 $Mn_6(O)_2(O_2CCMe_3)_{10}(C_4H_8O)_4$



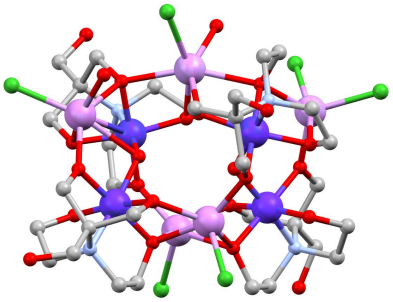
Lab Data	
Empirical Formula	Mn6 C70 H130 O27
Molecular Weight	1733.39
Temperature	150 K
Wavelength	Mo K_α 0.7017 Å
Crystal system, space group	Othorhombic, $Pbca$
Unit cell dimensions	a = 13.9732(3) Å α = 90
	b = 22.8225(4) Å β = 90
	c = 27.2470(5) Å γ = 90
Volume	8689.15(17) Å^3
Z	4
Density	2.13 Mg/m^3
Absorption coefficient	7.435 mm^{-1}
F(000)	3504
Crystal size	0.10 x 0.10 x 0.30 mm
Theta range for data collection	3.71 to 75.20°
Index ranges	$-17 \leq h \geq 16$
	$-20 \leq k \geq 28$
	$-32 \leq l \geq 33$
Reflections collected	50333
Independent reflection	8885
Completeness	99.3% to 73.69°
Absorption correction	Integration
Refinement method	Full-matrix least squares on F^2
Data / restraints / parameters	8885 / 0 / 444
Goodness-of-fit on F^2	0.9622
Final R indicies [I>2sigma(I)]	R_1 = 0.064, wR_2 = 0.184
R indices (all data)	R_1 = 0.068, wR_2 = 0.191
Extinction coefficient	None

A.6 $Mn_3O(O_2CPh)_6(pyr)_2(H_2O)[2\,MeCN]$



Lab Data	
Empirical Formula	Mn3 C56 H48 O14 N4
Molecular Weight	1165.81
Temperature	150 K
Wavelength	Cu K_α 1.541 Å
Crystal system, space group	Monoclinic, $P2_1/n$
Unit cell dimensions	a = 15.2317(2) Å α = 90
	b = 23.4787(4) Å β = 91.8480(12)
	c = 15.2716(2) Å γ = 90
Volume	5458.62(8) $Å^3$
Z	2
Density	1.42 Mg/m^3
Absorption coefficient	6.142 mm^{-1}
F(000)	2396
Crystal size	0.20 x 0.20 x 0.20 mm
Theta range for data collection	3.45 to 74.12°
Index ranges	$-18 \leq h \geq 17$
	$-27 \leq k \geq 29$
	$-13 \leq l \geq 19$
Reflections collected	37846
Independent reflection	10884
Completeness	99.5% to 71.90°
Absorption correction	Integration
Refinement method	Full-matrix least squares on F^2
Data / restraints / parameters	10884 / 0 / 694
Goodness-of-fit on F^2	0.9805
Final R indicies [I>2sigma(I)]	R_1 = 0.029, wR_2 = 0.069
R indices (all data)	R_1 = 0.035, wR_2 = 0.072
Extinction coefficient	None

A.7 $Co_4^{III}Mn_5^{II}Cl_6(C_8NO_5H_{15})_2(C_8NO_5H_{16})_2(H_2O)(OH)_2$



Lab Data	
Empirical Formula	Co4 Mn5 Cl6 C33 N4 O24 H70
Molecular Weight	1630.06
Temperature	150 K
Wavelength	Cu K_α 1.541 Å
Crystal system, space group	Monoclinic, $P2_1/c$
Unit cell dimensions	a = 17.9405(6) Å α = 90
	b = 13.2148(3) Å β = 98.444(3)
	c = 24.3802(8) Å γ = 90
Volume	5717.41(19) Å^3
Z	4
Density	1.962 Mg/m^3
Absorption coefficient	1.935 mm^{-1}
F(000)	3389.496
Crystal size	0.20 x 0.20 x 0.20 mm
Theta range for data collection	3.66 to 74.77°
Index ranges	$-22 \leq h \geq 22$
	$0 \leq k \geq 16$
	$0 \leq l \geq 30$
Reflections collected	41479
Independent reflection	11479
Completeness	99.7% to 70.28°
Absorption correction	Integration
Refinement method	Full-matrix least squares on F^2
Data / restraints / parameters	11479 / 4 / 733
Goodness-of-fit on F^2	1.074
Final R indicies [I>2sigma(I)]	R_1 = 0.082, wR_2 = 0.163
R indices (all data)	R_1 = 0.090, wR_2 = 0.169
Extinction coefficient	None

A.8 TBAB[CO(NH(4 −NO_2−C_6H_4))$_2$]$_2$

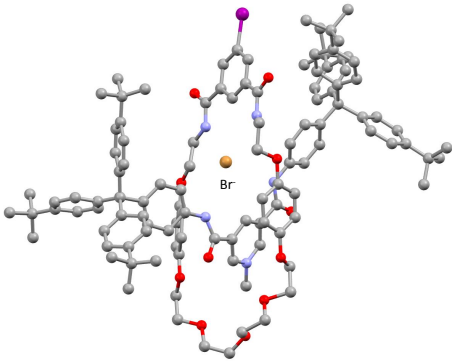
Tuneable wavelength data

- All data collected at 100 K

	Low Remote - 13400 eV	Inflection - 13473 eV
Empirical formula	Br1 C29 N5 O5 H37	Br1 C29 N5 O5 H37
Wavelength	0.9259 Å	0.9209 Å
Crystal system, space group	Monoclinic, $P2_1/n$	Monoclinic, $P2_1/n$
Unit cell dimensions	a = 8.5403(3) Å	a = 8.5391(2) Å
	b = 14.7030(4) Å	b = 14.4998(4) Å
	c = 27.5241(5) Å	c = 27.5220(5) Å
	α = 90	α = 90
	β = 93.8718(17)	β = 92.8292(18)
	γ = 90	γ = 90
Volume	3448.39(13) Å^3	3446.95(14) Å^3
Z	4	4
Absorption coefficient	2.29 cm^{-1}	7.83 cm^{-1}
Reflections collected	32250	32764
Completeness	71.9% to 0.80 Å	73.2% to 0.80 Å
R_{int}	0.058	0.068

	Peak - 13477 eV	High Remote - 13550 eV
Empirical formula	Br1 C29 N5 O5 H37	Br1 C29 N5 O5 H37
Wavelength	0.9206 Å	0.9156 Å
Crystal system, space group	Monoclinic, $P2_1/n$	Monoclinic, $P2_1/n$
Unit cell dimensions	a = 8.5414(1) Å	a = 8.5334(1) Å
	b = 14.7051(4) Å	b = 14.6918(4) Å
	c = 27.5302(5) Å	c = 27.5010(5) Å
	α = 90	α = 90
	β = 93.8364(16)	β = 93.8671(16)
	γ = 90	γ = 90
Volume	3450.13(12) Å^3	3439.97(12) Å^3
Z	4	4
Absorption coefficient	13.40 cm^{-1}	13.20 cm^{-1}
Reflections collected	32793	33337
Completeness	73.2% to 0.80 Å	74.6% to 0.80 Å
R_{int}	0.072	0.058

A.9 Bromide Rotaxane



Empirical Formula	C114 H140 N5 O11 I1 Br1 Cl3
Molecular Weight	2041.35
Temperature	150 K
Wavelength	0.6889 Å
Crystal system, space group	Triclinic, $P\bar{1}$
Unit cell dimensions	a = 11.882(11) Å α = 106.223(16)
	b = 19.633(19) Å β = 94.444(3)
	c = 24.57(2) Å γ = 98.652(15)
Volume	5397(9) Å^3
Z	2
Density	1.256 Mg/m^3
Absorption coefficient	0.782 mm^{-1}
F(000)	2135.599
Crystal size	0.10 x 0.30 x 0.30 mm
Theta range for data collection	1.15 to 27.45°
Index ranges	$-15 \leq h \geq 15$
	$-25 \leq k \geq 24$
	$0 \leq l \geq 32$
Reflections collected	49991
Independent reflection	14535
Completeness	88.9% to 25.3°
Absorption correction	Integration
Refinement method	Full-matrix least squares on F^2
Data / restraints / parameters	14535 / 1130 / 1446
Goodness-of-fit on F^2	0.9166
Final R indicies [I>2sigma(I)]	R_1 = 0.1206, wR_2 = 0.2508
R indices (all data)	R_1 = 0.1439, wR_2 = 0.2692
Extinction coefficient	None

A.10 o-Xylylenebis(triphenylphosphonium bromide)

Tuneable wavelength data

- All data collected at 100 K

	Low Remote - 13350 eV	Inflection - 13476 eV
Empirical formula	Br1 P1 C22 H19	Br1 P1 C22 H19
Molecular Weight	788.53	788.53
Wavelength	0.9287 Å	0.9200 Å
Crystal system, space group	Monoclinic, $P2_1/c$	Monoclinic, $P2_1/c$
Unit cell dimensions	a = 14.4898(14) Å	a = 14.4970(15) Å
	b = 18.3381(14) Å	b = 18.3524(18) Å
	c = 15.2248(14) Å	c = 15.2531(15) Å
	α = 90	α = 90
	β = 115.521(11)	β = 115.549(12)
	γ = 90	γ = 90
Volume	3650.8(5) Å^3	3661.3(6) Å^3
Z	4	4
Absorption coefficient	4.49 cm^{-1}	24.34 cm^{-1}
Reflections collected	3642	3711
Completeness	32.3% to 0.80 Å	61.9% to 0.99 Å
R_{int}	0.060	0.065

	Peak - 13480 eV	High Remote - 13700 eV
Empirical formula	Br1 P1 C22 H19	Br1 P1 C22 H19
Molecular Weight	788.53	788.53
Wavelength	0.9198 Å	0.9050 Å
Crystal system, space group	Monoclinic, $P2_1/c$	Monoclinic, $P2_1/c$
Unit cell dimensions	a = 14.512(3) Å	a = 14.504(2) Å
	b = 18.361(3) Å	b = 18.3245(18) Å
	c = 15.32492 Å	c = 15.3323(19) Å
	α = 90	α = 90
	β = 115.55(2)	β = 115.442(16)
	γ = 90	γ = 90
Volume	3684(1) Å^3	3679.8(8) Å^3
Z	4	4
Absorption coefficient	24.33 cm^{-1}	23.17 cm^{-1}
Completeness	61.8% to 0.99 Å	61.6% to 0.97 Å
R_{int}	0.064	0.047