Developments in Structure Solution for Electron Crystallography

By

Andrew A. Stewart

Theoretical Crystallography

Department of Chemistry

University of Glasgow

2003

"Change a Persons perception and you will change their views"

-Edward De Bono 1964

ABSTRACT

The concept of electron crystallography is by no means new, although the majority of crystallographers will never have contemplated trying the technique, to solve a structural problem.

Electron crystallography is fraught with problems from sample preparation, and data collection to structure solution, all have unique and challenging problems when being used in the context of electron crystallography. New techniques and methodologies for structure solution using electron crystallographic data are presented with the aim of making structure solution a more routine part of the procedure for electron crystallography. The problems of phase extension into the missing cone region, structure solution and refinement of the structures are all dealt with here.

Acknowledgements

I would first of all like to thank Prof. Chris Gilmore for giving me the opportunity of doing a PhD. Also for all his generosity both academically, personally and always with a good pinch of humour.

Stuart McKay for his support and encouragement and for being a friend. The endless banter and sarcasm, with out you I think the entire group would have gone insane.

Gordon Barr and Graeme Tate for making the Gilmore group the palace to be. Thanks guys for all the help support and encouragement, during the highs and lows of the PhD adventure.

Elizabeth Maclean for keeping us all in check, and having a wicked sense of humour.

I would also like to thank Drs Adrian Lapthorn, Klas Andersson, Wei Dong, Douglas Dorset, Kelvin Tyler, and Kenneth Shankland for useful discussions and helpful advice.

A special thank you goes to Jennifer Henderson, for helping me improving my written English. Thank you Jen, without your help Chris may have hired a hitman to solve the problem of my appalling written English.

And to the rest of my PhD compatriots through out the department who have made my undergraduate and postgraduate studies a more bearable experience with all the Alchemist events and spontaneous nights out. A special thanks should go to the organic research office who have endured more than a few of my lunch time rants on the bug bare topic of the day. Thank you! With special mentions for Calum 'Stoon' McCleod, Claire Martin, David Robins, Gordon McKernian,, John 'the' Hand. Paddy ,Bernie, Gillian, and Christine.

A big thank you to Andrea Docherty for all those big chat late night discussions and for being an amazing friend and flatmate.

To all the Carluke guys, Blair, Ian, Colin, Jamie, Benyon, Lucas, Graeme, Sean, Gary, Guy. For having faith in me when I had almost given up on myself.

To all my fencers, especially the Sabre group and Lisa. Thanks for making fencing fun again! Now remember, hand before foot!

Finally to my family, Mum, Dad, and Elaine, thank you for all the love, support, and encouragement. Especially during the difficult times, when I was told that due to Dyslexia that I would never be able to achieve the goals I desired. Thank you for making me believe in my own ability. Thank you for holding the door of opportunity ajar for me.

1 NEW TECHNIQUES FOR ELECTRON CRYSTALLOGRAPHY 10

1.1 INTRODUCING ELECTRON CRYSTALLOGRAPHY 10
1.2 SCATTERING AND DIFFRACTION IN AN ELECTRON DIFFRACTION EXPERIMENT 12
1.3 THE EFFECT OF TILTING A CRYSTAL 14

2 THE PHASE PROBLEM AND CLASSICAL DIRECT METHODS 16

2.1 THE PHASE PROBLEM 16
2.2 DIRECT METHODS 16
2.3 MAXIMUM ENTROPY AND BAYES THEOREM 18
2.3.1 Maximum entropy principle 18
2.3.2 Bayes theorem and entropy 18
2.3.3 Why use maximum entropy? 19
2.4 MAXIMUM ENTROPY AS APPLIED TO CRYSTALLOGRAPHY 19
2.4.1 Normalisation 19
2.4.2 Data partitioning 20
2.4.3 Constrained entropy maximisation 20
2.4.4 The line search 23
2.4.5 Plane search 24
2.4.6 Technical details of calculation 27
2.5 THE MICE ALGORITHM IN ACTION 30
2.5.1 Maps 34
2.6 P(_Q) FUNCTIONS 35
2.7 ERROR-CORRECTING CODES (ECCS) 35
2.8 FORUIER METHODS 37
2.9 REFINEMENT 37

3 AB INITIO STRUCTURE SOLUTION FROM ELECTRON DIFFRACTION DATA 43

3.1 AB-INITIO PROCEDURE 43
3.2 AB INITIO CALCULATIONS 44
3.3 BRUCITE 45
3.3.1 Solution 45
3.4 ISOTACTIC POLY(1-BUTENE) FORM III 47
3.4.1 Solution 48
3.5 MANNAN I 49
3.5.1 Projection solution 51
3.5.2 Mannan I 3-dimensional structure 52
3.6 POLY (1,4-TRANS - CYCLOHEXANEDIYL DIMETHYLENE SUCCINATE) (T-CDS) 53
3.6.1 Structure Solution of Poly (1,4-trans - cyclohexanediyl dimethylene succinate) (T-cds) 54
3.7 BUCKMINSTERFULLERENE (C_{60}) 55
3.7.1 Structure solution of Buckminsterfullerene 56
3.8 DISCUSSION OF *AB INITIO METHODS IN ELECTRON DIFFRACTION* 58
3.9 CONCLUSIONS 59

4 THE MISSING CONE PROBLEM 61

4.1 EXPERIMENTAL TECHNIQUE 62
4.2 RESULTS 64
4.2.1 Basic Copper Chloride 64
30° tilt series Copper Chloride results 66
45° tilt series Copper Chloride results 67
60° tilt series Copper Chloride results 69
4.2.5 Isotactic Poly(1-butene) form III 70
30° tilt series Isotactic Poly(1-butene) results 72
45° tilt series Isotactic Poly(1-butene) results 74
4.2.8 Data from Poly(1,4,trans-cyclohexanediyl dimethylene succinate) (T-cds) 75
4.2.9 Poly (1,4-trans –cyclohexanediyl dimethylene succinate) ±30° tilt results 76
4.2.10 Poly (1,4-trans –cyclohexanediyl dimethylene succinate) ±45° tilt results 78

4.2.11 Beyond the observed hkl range ... *79*
4.2.12 Discussion and conclusions ... *88*

5 LEAST SQUARES REFINEMENT OF ATOMIC MODELS ... **91**

5.1 STRUCTURE REFINEMENT ... 91
5.2 USING REFINEMENT PACKAGES WITH ELECTRON DIFFRACTION DATA ... 97
5.3 REFINEMENT EXPERIMENTAL ... 100
5.4 RESULTS ... 102
5.4.1 Table of all results ... *102*
5.4.2 Al_mFe ... *102*
5.4.3 Brucite ($Mg(OH)_2$) ... *104*
5.4.4 10-Cyano-9,9'-Bianthryl (CNBA) ... *106*
5.4.5 Basic Copper Chloride ... *108*
... *110*
5.4.6 2,6-bis[4-(dimethylamino)benzylidene]cyclohexanone DMABC ... *110*
5.4.7 4-dimethylamino-3-cyanobiphenyl (DMACB) ... *112*
5.4.8 Mannan I ... *114*
5.4.9 Copper Perchlorophthalocyanine ... *116*
5.4.10 Isotactic Poly(1-butene) Form III ... *119*
5.4.11 Poly(_-caprolactone) ... *120*
5.4.12 Polyethylene ... *122*
5.4.13 Poly(1,4,trans-cyclohexanediyl dimethylene succinate) (T-cds) ... *123*
5.5 REFINEMENT DISCUSSION ... 125
5.6 REFINEMENT CONCLUSIONS ... 126

6 THE $P(\delta q)$ FUNCTION ... **129**

6.1 THE $P(\delta q)$ FUNCTION ... 129
6.2 EXPERIMENTAL ... 130
6.3 RESULTS ... 132
6.3.1 Al_mFe Aluminium alloy ... *132*
6.3.2 10-Cyano-9,9'-Bianthryl (CNBA) ... *135*
6.3.3 2,6-bis[4-(dimethylamino)benzylidene]cyclohexanone (DMABC) ... *137*
6.3.4 4-dimethylamino-3-cyanobiphenyl (DMACB) ... *138*
6.3.5 Copper Perchlorophthalocyanine ... *140*
6.3.6 Isotactic Poly(1-butene) Form III ... *142*
6.4 CONCLUSIONS OF $P(\delta q)$ RESULTS ... 144

7 CONCLUSIONS ... **147**

8 APPENDIX A: CIF FILES FROM DATABASE ... **150**

8.1 AL_MFE CIF ... 152
8.2 BRUCITE CIF ... 167
8.3 10-CYANO-9,9',-BIANTHRYL ... 172
8.4 BASIC COPPER CHLORIDE ... 177
8.5 2,6-BIS[4(DIMETHYLAMINO)BENZYLIDENE]CYCLOHEXANONE ... 182
8.6 4-DIMETHYLAMINO-3-CYANOBIPHENYL ... 186
8.7 MANNAN I ... 190
8.8 COPPER PERCHLOROPHTHALOCYANINE ... 194
8.9 ISOTACTIC POLY(1-BUTENE) FORM III ... 199
8.10 POLY (ε-CAPROLACTONE) ... 203
8.11 POLYETHYLENE ... 207
8.12 POLY(1,4,TRANS-CYCLOHEXANEDIYL DIMETHYLENE SUCCINATE) ... 211

TABLE OF FIGURES

FIGURE 1-1 THE DIFFERENT WORKING MODES OF AN ELECTRON MICROSCOPE 11
FIGURE 3-1 FLOW DIAGRAM OF THE *AB INITIO* CALCULATION PROCESS 44
FIGURE 3-2 CENTROID MAP FOR NODE 141: BRUCITE 47
FIGURE 3-3 FINAL CENTROID MAP FOR ISOTACTIC POLY(1-BUTENE) FORM III 49
FIGURE 3-4 STRUCTURE SOLUTION OF MANNAN I IN PROJECTION CENTROID MAP OF NODE 22 52
FIGURE 3-5 THREE-DIMENSIONAL ELECTRON DENSITY CENTROID MAP OF MANNAN I , NODE 163 53
FIGURE 3-6 CENTROID MAP OF T-CDS, NODE 2725 55
FIGURE 3-7 STRUCTURE SOLUTION OF BUCKMINSTERFULLERENE 58
FIGURE 4-1 EXAMPLE SCRIPT FOR MISSING CONE CALCULATIONS 63
FIGURE 4-2 COMPARISON OF ±30° TILT SERIES MAPS BEFORE AND AFTER THE CALCULATION 67
FIGURE 4-3 COMPARISON OF ±45° TILT SERIES MAPS BEFORE AND AFTER THE CALCULATION 69
FIGURE 4-4 COMPARISON OF ±60° TILT SERIES MAPS BEFORE AND AFTER THE CALCULATION 70
FIGURE 4-5 COMPARISON OF ±30° TILT SERIES MAPS BEFORE AND AFTER THE CALCULATION 73
FIGURE 4-6 COMPARISON OF ±45° TILT SERIES MAPS BEFORE AND AFTER THE CALCULATION 74
FIGURE 4-7 COMPARISON OF ±30° TILT SERIES MAPS BEFORE AND AFTER THE CALCULATION 77
FIGURE 4-8 COMPARISON OF ±45° TILT SERIES MAPS BEFORE AND AFTER THE CALCULATION 78
FIGURE 4-9 BASIC COPPER CHLORIDE SUPER RESOLUTION Q^{ME} MAP 80
FIGURE 4-10 ISOTACTIC POLY(1-BUTENE) FORM III SUPER RESOLUTION Q^{ME} 83
FIGURE 4-11 THE POLY(1,4,TRANS-CYCLOHEXANEDIYL DIMETHYLENE SUCCINATE) Q^{ME} MAP 88
FIGURE 5-1 FLOW DIAGRAM OF THE REFINEMENT PROCESS 101
FIGURE 5-2 REFINED BRUCITE MODEL 105
FIGURE 5-3 REFINED 10-CYANO-9,9'-BIANTHRYL MODEL 108
FIGURE 5-4 REFINED BASIC COPPER CHLORIDE MODEL 110
FIGURE 5-5 REFINED 2,6-BIS[4-(DIMETHYLAMINO)BENZYLIDENE]CYCLOHEXANONE MODEL 112
FIGURE 5-6 REFINED 4-DIMETHYLAMINO-3-CYANOBIPHENYL MODEL 114
FIGURE 5-7 REFINED MANNAN I MODEL 116
FIGURE 5-8 REFINED COPPER PERCHLOROPHTHALOCYANINE MODEL 118
FIGURE 5-9 REFINED ISOTACTIC POLY(1-BUTENE) FORM III MODEL 120
FIGURE 5-10 REFINED POLY(E-CAPROLACTONE) MODEL 121
FIGURE 5-11 REFINED POLYETHYLENE MODEL 123
FIGURE 5-12 REFINED POLY(1,4,TRANS-CYCLOHEXANEDIYL DIMETHYLENE SUCCINATE) MODEL 124
FIGURE 6-1 AL_MFE CENTROID MAP OF NODE 235 135
FIGURE 6-2 COPPER PERCHLOROPHTHALOCYANINE CENTROID MAP OF NODE 272 142
FIGURE 6-3ISOTACTIC POLY(1-BUTENE) FORM III CENTROID MAP OF NODE 73 144

TABLE OF TABLES

TABLE 3-1 BRUCITE, TOP 8 NODES FROM ANALYSIS 46
TABLE 3-2 POLYBUT-1-ENE, ROUTE TO FINAL SOLUTION 48
TABLE 3-3 POLYBUT-1-ENE, ANALYSIS RESULTS FROM FINAL SET OF NODES 48
TABLE 3-4 MANNAN I ANALYSIS RESULTS IN PROJECTION 51
TABLE 3-5 MANNAN I 3D, ANALYSIS OF RESULTS 52
TABLE 3-6 T-CDS PATH TO FINAL STRUCTURE 54
TABLE 3-7 T-CDS FINAL ANALYSIS OF RESULTS 54
TABLE 3-8 FINAL NODE FOR BUCKMINSTERFULLERENE 57
TABLE 3-9 ANALYSIS OF BUCKMINSTERFULLERENE 57
TABLE 4-1 NUMBER OF KNOWN PHASES AND REFLECTIONS FOR EACH TILT SERIES AND COMPOUND 63
TABLE 4-2 BASIC COPPER CHLORIDE REFLECTIONS AND PHASES 65
TABLE 4-3 EXTRAPOLATED MAGNITUDES AND PHASES AND THEIR OBSERVED VALUES, FOR ±30° TILT SERIES. 66
TABLE 4-4 EXTRAPOLATED MAGNITUDES AND PHASES AND THEIR OBSERVED VALUES, FOR ±45° TILT SERIES 68
TABLE 4-5 EXTRAPOLATED MAGNITUDES AND PHASES AND THEIR OBSERVED VALUES, FOR ±60° TILT SERIES 69
TABLE 4-6 ISOTACTIC POLY(1-BUTENE) FORM III REFLECTIONS AND PHASES 71
TABLE 4-7 EXTRAPOLATED MAGNITUDES AND PHASES AND THEIR OBSERVED VALUES, FOR ±30° TILT SERIES 72
TABLE 4-8 EXTRAPOLATED MAGNITUDES AND PHASES AND THEIR OBSERVED VALUES, FOR ±45° TILT SERIES 74
TABLE 4-9 POLY(1,4,TRANS-CYCLOHEXANEDIYL DIMETHYLENE SUCCINATE) (T-CDS) MAGNITUDES AND PHASES 75
TABLE 4-10 EXTRAPOLATED MAGNITUDES AND PHASES AND THEIR OBSERVED VALUES, FOR ±30° TILT SERIES 77
TABLE 4-11 EXTRAPOLATED MAGNITUDES AND PHASES AND THEIR OBSERVED VALUES, FOR ±45° TILT SERIES 78
TABLE 4-12 BASIC COPPER CHLORIDE SUPER RESOLUTION PREDICTED REFLECTIONS 79
TABLE 4-13 THE ISOTACTIC POLY(1-BUTENE) FORM III RESULTS FOR SUPER RESOLUTION 81
TABLE 4-14 THE POLY(1,4,TRANS-CYCLOHEXANEDIYL DIMETHYLENE SUCCINATE) (T-CDS) RESULTS FOR SUPER RESOLUTION 84
TABLE 5-1 SUMMARY OF ALL REFINEMENT CALCULATIONS 102
TABLE 5-2 ALUMINIUM IRON ATOM POSITIONS 103
TABLE 5-3 ALUMINIUM IRON BOND LENGTHS (Å) 104
TABLE 5-4 ALUMINIUM IRON BOND ANGLES 104
TABLE 5-5 BRUCITE ATOM POSITIONS 105
TABLE 5-6 BRUCITE BOND LENGTHS (Å) 105
TABLE 5-7 BRUCITE BOND ANGLES 105
TABLE 5-8 CNBA ATOM POSITIONS 106
TABLE 5-9 CNBA BOND LENGTHS (Å) 107
TABLE 5-10 CNBA BOND ANGLES 107
TABLE 5-11 BASIC COPPER CHLORIDE ATOM POSITIONS 109
TABLE 5-12 BASIC COPPER CHLORIDE BOND LENGTHS(Å) 109
TABLE 5-13 BASIC COPPER CHLORIDE BOND ANGLES 109
TABLE 5-14 DMABC ATOMIC POSITIONS 111
TABLE 5-15 DMABC BOND LENGTHS (Å) 111
TABLE 5-16 DMABC BOND ANGLES 111
TABLE 5-17 DMACB ATOMIC POSITIONS 112
TABLE 5-18 DMACB BOND LENGTHS (Å) 113
TABLE 5-19 DMACB BOND ANGLES 113
TABLE 5-20 MANNAN I ATOM POSITIONS 115
TABLE 5-21MANNAN I BOND LENGTHS (Å) 115
TABLE 5-22 MANNAN I BOND ANGLES 115
TABLE 5-23 COPPER PERCHLOROPHTHALOCYANINE ATOM POSITIONS 117
TABLE 5-24 COPPER PERCHLOROPHTHALOCYANINE BOND LENGTHS (Å) 117
TABLE 5-25 COPPER PERCHLOROPHTHALOCYANINE BOND ANGLES 118

TABLE 5-26 POLY(1-BUTENE) FORM III ATOM POSITIONS 119
TABLE 5-27 POLY(1-BUTENE) FORM III BOND LENGTHS (Å) 119
TABLE 5-28 POLY(1-BUTENE) FORM III BOND ANGLES 119
TABLE 5-29 POLY(_-CAPROLACTONE)ATOM POSITIONS 121
TABLE 5-30 POLY(E-CAPROLACTONE) BOND LENGTHS (Å) 121
TABLE 5-31 POLY(E-CAPROLACTONE) BOND ANGLES 121
TABLE 5-32 POLYETHYLENE ATOM POSITIONS 122
TABLE 5-33 POLYETHYLENE BOND LENGTHS (Å) 122
TABLE 5-34 POLYETHYLENE BOND ANGLES 123
TABLE 5-35 T-CDS ATOM POSITIONS 124
TABLE 5-36 T-CDS BOND LENGTHS (Å) 124
TABLE 5-37 T-CDS BOND ANGLES 124
TABLE 6-1 AL_MFE ORIGIN REFLECTION 133
TABLE 6-2 AL_MFE REFLECTIONS SELECTED BY NEXT COMMAND FOR PERMUTATION BY P(_Q) FUNCTION 133
TABLE 6-3 AL_MFE RESULTS FROM ANALYSIS OF ENTROPY MAXIMISED NODES 134
TABLE 6-4 10-CYANO-9,9'-BIANTHRYL ORIGIN REFLECTIONS 136
TABLE 6-5 10-CYANO-9,9'-BIANTHRYL REFLECTIONS SELECTED BY NEXT COMMAND FOR PERMUTATION BY P(_Q) FUNCTION 136
TABLE 6-6 10-CYANO-9,9'-BIANTHRYL RESULTS FROM ANALYSIS OF ENTROPY MAXIMISED NODES 136
TABLE 6-7 2,6-BIS[4-(DIMETHYLAMINO)BENZYLIDENE]CYCLOHEXANONE ORIGIN REFLECTIONS 137
TABLE 6-8 2,6-BIS[4-(DIMETHYLAMINO)BENZYLIDENE]CYCLOHEXANONE REFLECTIONS SELECTED BY NEXT COMMAND FOR PERMUTATION BY P(_Q) FUNCTION 138
TABLE 6-9 2,6-BIS[4-(DIMETHYLAMINO)BENZYLIDENE]CYCLOHEXANONE RESULTS FROM ANALYSIS OF ENTROPY MAXIMISED NODES 138
TABLE 6-10 4-DIMETHYLAMINO-3-CYANOBIPHENYL ORIGIN REFLECTION 139
TABLE 6-11 4-DIMETHYLAMINO-3-CYANOBIPHENYL REFLECTIONS SELECTED BY NEXT COMMAND FOR PERMUTATION BY P(_Q) FUNCTION 139
TABLE 6-12 4-DIMETHYLAMINO-3-CYANOBIPHENYL RESULTS FROM ANALYSIS OF ENTROPY MAXIMISED NODES 139
TABLE 6-13 COPPER PERCHLOROPHTHALOCYANINE ORIGIN REFLECTION 140
TABLE 6-14 COPPER PERCHLOROPHTHALOCYANINE REFLECTIONS SELECTED BY NEXT COMMAND FOR PERMUTATION BY P(_Q) FUNCTION 140
TABLE 6-15 COPPER PERCHLOROPHTHALOCYANINE RESULTS FROM ANALYSIS OF ENTROPY MAXIMISED NODES 141
TABLE 6-16 ISOTACTIC POLY(1-BUTENE) FORM III ORIGIN REFLECTIONS 143
TABLE 6-17 ISOTACTIC POLY(1-BUTENE) FORM III REFLECTIONS SELECTED BY NEXT COMMAND FOR PERMUTATION BY P(_Q) FUNCTION 143
TABLE 6-18 ISOTACTIC POLY(1-BUTENE) FORM III RESULTS FROM ANALYSIS OF ENTROPY MAXIMISED NODES 143

1 NEW TECHNIQUES FOR ELECTRON CRYSTALLOGRAPHY

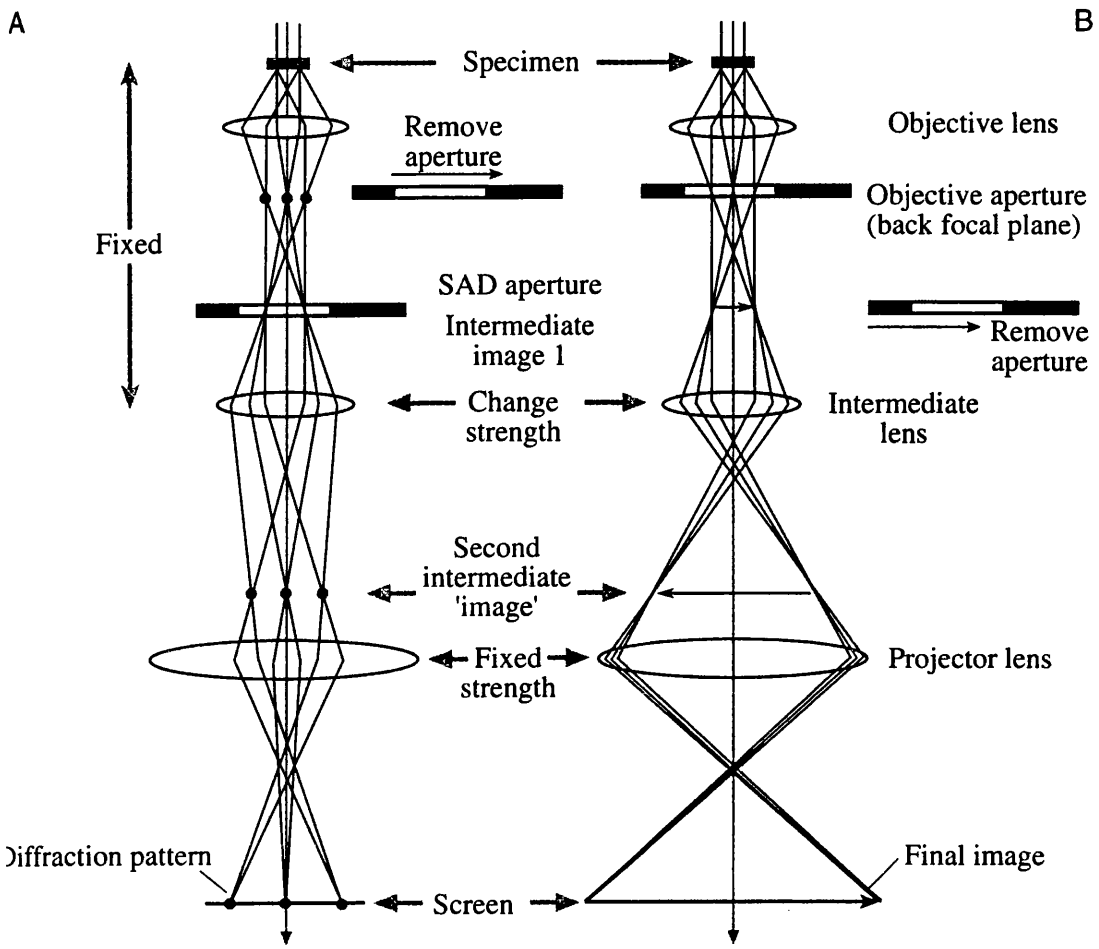
1.1 Introducing Electron Crystallography

The small but vitally important area of electron crystallography emerged in the late 1920s in Russia, mainly as a method for locating hydrogens in metal structures[1]. Simultaneously, they realised that this methodology could be used to determine structures where it was found difficult or impossible to grow crystals for single crystal X-ray diffraction experiments. While the usefulness of electron diffraction for finding hydrogen atoms in metal complexes is somewhat debatable, its main use today is unquestionably in areas where it is difficult to grow single crystals of the appropriate quality for single crystal experiments. From ribosomes and membrane proteins through to polymers and minerals[2], the technique is applied and slowly produces the desired results.

The process of electron crystallography is a complex and technical one; at no stage does the method yield a simple or quick step. Every step presents possible pitfalls, from obtaining a suitable crystal of the correct thickness, through to collecting the data, or getting structure solution from sparse data sets. All these steps have their own unique problems and difficulties, which need to be overcome.

The electron microscope has two modes that are utilised in electron crystallography, diffraction mode and imaging mode, something that is unique for a crystallography experiment. The microscope has electromagnets that are used to focus the electron beam and produce an image, comparable to a conventional camera[3].

Figure 1-1 The different working modes of an electron microscope



The question has to be asked why can electrons be focused, and not X-rays, even though they are of similar wavelengths? The answer lies in the interaction of different radiation types with matter. X-rays in general, pass straight through a given sample, hence a beam stop is required, but electrons almost always interact with the sample. The reason for electron diffraction, even over small thicknesses of crystal is that electrons are charged. The interaction between the transmitting electron beam and the electron potential is dominated by the electron-electron repulsion.

The strength of an electron interacting with a sample is 10^6 times stronger, whereas, neutrons are 10^{-2} times weaker than X-rays[1-3]. The strength of the electron interaction with the sample is mainly due to the magnetic properties of electrons, which allows the electromagnets to focus the beams once they have passed through the sample, and form an image. The same properties that allow electromagnets to focus the beam into an image are the same properties which cause the problem of strong sample interaction.

Due to the affinity that the electron has for the sample it is possible that it may scatter more than once on its path through the material. This is a major disadvantage because all the known methods for structure solution are based upon the kinematical approximation, which assumes the beam only diffracts once on its path through the sample. There are practical methods for overcoming these problems, such as having a sample with no heavy atoms, or making the sample less than

100Å thick, so that the sample will not be thick enough for the beam to diffract more than once on its path through the sample[4-6].

Another key difference between X-rays and electron diffraction is what the radiation "sees" when it probes the sample. X-rays are diffracted from the electron density, whereas the electrons are diffracted from the electrostatic potential, which is a combination of electron density and nuclear charge. The electron potential charge drops off at a lower gradient, as it move away from the centre of an atom, making it useful for locating small atoms in the presence of larger ones.

1.2 Scattering and diffraction in an electron diffraction experiment

In an electron crystallographic experiment, the data produced needs to be kinematic (or approximately so) in nature. This means that the data must be akin to that found in an X-ray experiment. It is generally obvious when dynamical scattering has occurred because the strength of the diffraction spots are almost as strong as, or a high percentage of, the main beam. This is particularly obvious when you go to higher θ where it is know that the scattering should be weaker and the measured intensities small.

The experimenter can help to make the data kinematical and useable by making the crystal thin and the electron potential of the beam of the appropriate strength i.e. higher accelerating voltage. Usually the higher the potential difference the better the data will be. Occasionally the dynamical scattering can be hard to detect and multi-slice calculations are used to check the quality of the data, but this is not always possible, since it is essentially necessary to know the structure before you can perform these calculations. The multi-slice calculations are essentially modelling the thickness of the crystal and the effect it has on the scattering of electrons.

In electron crystallography and in microscopy in general it is essential to understand, there are several different types of scattering, and the differences between them, dynamical: secondary: elastic: and inelastic scattering[3].

Dynamical scattering occurs when the electron beam interacts intensely with the crystal. A beam which has diffracted in perfect Bragg diffraction, may interact again and be diffracted back into the main beam, by the same plane of atoms. The thicker the sample the higher the probability of this occurring becomes. This is somewhat of an over simplification of dynamical scattering theory, but is sufficient for the purposes of this volume of work. See Cowley [7,8], for a more detailed explanation. Dynamical scattering often manifests itself in the appearance of forbidden reflections, or reflections that have unusually high intensity for their θ angle.

Secondary scattering is one specific form of dynamical diffraction, which leads to the appearance of forbidden reflections. The appearance of these forbidden occurs because, the radiation wavelength is small (~0.05Å) making the radius of the Ewald sphere very large, which means that more than one lattice point a at a time will be illuminated. This leads to a situation where one of the beams acts as an incident beam and is diffracted again, causing the appearance of the forbidden reflection.

Elastic scattering is the process of diffraction that occurs during Bragg diffraction, where the electrons only diffract once, on the way through the sample and generally from the outer edge of the atoms without energy loss. Inelastic scattering on the other hand, occurs when the electron penetrates further into the core of the atom and is then scattered with energy loss. This deep penetration of the electron shells causes the electron to bounce off of other electrons and not to diffract in the normal manner. This can lead to other useful processes such as electron energy loss spectrometry (EELS) and its accompanying X-ray emission lines, which are of use in analytical microscopy, but are considered a hindrance in electron crystallography. Inelastic scattering is also associated with many of the problematic aspects of using a microscope for these experiments, particularly radiation damage, which takes on two forms: radiolysis and knock-on damage.

Radiolysis occurs when the inelastic scattering breaks the chemical bonds by ionisation. Polymers and organic compounds such as alkali halides are particularly susceptible to this form of damage.

Knock-on damage is directly related to the beam energy, and occurs when the energy of the beam is transferred to an atom. This energy then causes the atom to vibrate so violently that it breaks all its bonds, and therefore free itself from the crystal. The size of an atom and how strongly it is bonded to its neighbours all affect how, when and if an atom will be displaced in this manner. Also the higher the beam's potential difference, the more likely the occurrence of knock-on damage.

The size and shape of a crystal affect the way in which an experiment will behave; the most influential effect on a diffraction experiment is the amount of crystal bending. The size of area that must be illuminated to produce a diffraction pattern is critical to how much crystal bending will influence the experiment. The larger the area which needs to be illuminated the higher the probability of bending effects influencing the data. Bending effects are unavoidable, due to the crystals being on such a small scale. On the nanometer and micrometer scale the shape of the individual molecule affects the way in which the surface of the crystals looks, and bends[9,10].

Crystal bending affects the mean free path of the beam through the sample and therefore leads to different path lengths and absorbency effects, which ultimately affect the measured intensities.

Once an approximately kinematical data set of reflections has been collected it should be noted that these reflections are not the same as X-ray reflections, because they interact with the matter differently. The scattering factor for each element is different, and electron scattering factors must be used in any calculations to or from real space. The electron interacts with the electron potential of an atom or molecule rather than the electron density. Its volume is larger than the electron density, and falls off at a gentler gradient. The electron potential is a combination of the positive nuclei and the electron density, making the peaks in the maps corresponding to the atomic nuclei. The different shape of the electron potential compared to the electron density makes it essential that electron scattering factors are used in calculations to allow for the different shape of the scattering curves that each element has with respect to the different radiation types[1].

1.3 The effect of tilting a crystal

During electron crystallographic experiments, if three-dimensional data is to be collected, it is necessary to tilt the sample around the vertical axis. This allows a

larger sample of reciprocal space to be covered and avoids issues that occur in projection, which can lead to misinterpretation of a structure. Other advantages of tilting the sample are finding the three-dimensional symmetry, and more accurate unit cell dimensions.

The reconstruction of the three-dimensional diffraction pattern has to be assembled from a series of two-dimensional patterns. The patterns are all assembled into a three-dimensional data set. The number of patterns needed to adequately sample reciprocal space varies between space groups and orientation of the sample. In general the higher the symmetry, the fewer diffraction patterns required, although the more tilts collected and added into the sampling of reciprocal space, the easier structure solution will become.

There are of course problems associated with doing this. Tilting the sample increases the mean free path through the sample, and leads to all the associated problems, discussed above. The probability of dynamical diffraction, inelastic and secondary scattering all increase the sample is inclined, increasing the chances of obtaining an unusable diffraction pattern[2,11].

Even if an adequate number of diffraction patterns have been collected, it is physically impossible to cover all of reciprocal space. The microscope itself restricts the tilting of the sample between 45° and 60° (This varies between individual microscopes depending on the sample holders available and the goniometer stage) this restricted sampling of reciprocal space is known as the dead zone or missing cone problem. The missing data can lead to problems in determining the symmetry and leaves a relatively poorly sampled axis, which can lead to problems in solving structures, and in particular determining atom position accurately on the axis where the data is sparse.

2 THE PHASE PROBLEM AND CLASSICAL DIRECT METHODS

2.1 The phase problem

In a crystallographic experiment, finding the spatial arrangement for the unit cell contents is hindered by the fact that the phase information is lost in the diffraction experiment. The relationship between real and reciprocal space, of a given unit cell, works in both directions when the phases and the structure factor moduli are present. As already stated the phases are lost during a crystallographic experiment. Therefore the arrangement of atoms can only be, constructed by somehow estimating, deducing the missing phases, because the experiment only records the moduli of the data, and not the phase information. This is known as the crystallographic phase problem.[12]

2.2 Direct methods

Traditionally the approaches for solving structures have all been based upon the triplet phase relationship. This states that the phases from three related structure factors should be approximately sum to zero.

Direct methods of this type also require, that the data is of at least 1.2Å resolution or above, and complete. They assume that the atoms are randomly distributed within the unit cell, which of course is not always true. There are many programs that will solve small molecule structures by these methods; some of the more popular ones are Shake and Bake[13], DirDiff[13,14], Shelxs[13,15], Sir[16], and Mithril[17], all of which use the same basic concepts.

The first real application of direct methods in crystallography was the Sayre equation[18]

$$F_{\underline{h}} = \left(\frac{\theta_{\underline{h}}}{V}\right)\sum_{\underline{k}} F_{\underline{k}} F_{\underline{h}-\underline{k}} \,, \quad \textbf{(2-1)}$$

where phases could be determined directly from the structure factors. $F_{\underline{h}}$ is the phased structure factor, V is the volume of the unit cell , and $\theta_{\underline{h}}$ is a function of the atomic scattering factors. The Sayre equation works for equal atom structures and usually needs some type of basis or apriori to start the calculation working efficiently.

Structure invariants are phase relationships that do not vary with respect to the origin. The structure invariant allows the triplet phase relationship to work. Within the triplets the three individual phases will change but the sum of the three will be zero, no matter where the origin is placed. $\varnothing_{\underline{h}} + \varnothing_{\underline{k}} + \varnothing_{-\underline{h}-\underline{k}} \approx 0$[19]. Where $\varnothing$ is the phase of a given reflection.

The observed structure factors $\left|F_{\underline{h}}\right|$ need a correction factor to account for the falloff in scattering at high angle, and which makes the atoms point-like scatterers,

$$\left|E_{\underline{h}}\right|^2 = kI_{\underline{h}}^{obs} \Big/ \varepsilon \sum_{i=1}^{N} f_i'^2 \,. \text{ (2-2)}$$

Where ε compensates for reflections restricted by symmetry, f_i is the atomic scattering factor of the ith atom, N is the number of atoms in the unit cell, and k is a scale factor.

When the triplets are combined the Tangent formula is produced[19].

$$\tan\phi_{\underline{h}} = \frac{\sum_{\underline{h}'} K(\underline{h},\underline{k})\sin\left(\phi_{\underline{k}} + \phi_{\underline{h}-\underline{k}}\right)}{\sum_{\underline{h}'} K(\underline{h},\underline{k})\cos\left(\phi_{\underline{k}} + \phi_{\underline{h}-\underline{k}}\right)}, \text{ (2-3)}$$

Where

$$K(\underline{h},\underline{k}) = 2N^{-\frac{1}{2}}\left|E_{\underline{h}}, E_{\underline{k}}, E_{\underline{h}-\underline{k}}\right| \quad , \text{ (2-4)}$$

if all the atoms are equal then N is the number of atoms in the unit cell. The tangent formula is the basis for all direct methods programs.

A new approach was suggested in the early 1980s by Bricogne[20], which would allow the limits of direct methods to be pushed back, *i.e.* solve structures where the data set is incomplete and the resolution is above 1.2 Å. The method of choice for achieving this goal was maximum entropy, a statistical approach using Bayesian statistics. In a departure from the traditional method, new types of statistics were employed in a different area. Rather than work in reciprocal space and work with phase relationships, this method was in real space, and has its origins in information theory.

2.3 Maximum entropy and Bayes theorem

2.3.1 Maximum entropy principle

The maximum entropy principle, in the context of crystallography, is to produce an image or map which is maximally non-committal, or minimally biased with respect to missing data, and maximises the entropy of the map subject to the constraint that the map produced, must contain the data which generated the map, to within experimental error. In other words, the input map must be recovered from the calculation unchanged, and the calculation must not be prejudiced toward predicting the missing data. The calculation uses the entropy of the input map to help it make choices of how to maximise the entropy of the output map[21].

2.3.2 Bayes theorem and entropy

The roots of Maximum entropy lie in Bayesian theory, where the use of currently known information is used to help assist the statistics and make more informed decisions about the posterior[22].

$$\Pr ob(image|data) \propto \Pr ob(image) \times \Pr ob(data|image) \quad (2\text{-}5)$$

Posterior distribution Prior distribution Likelihood

If the initial data is biased by some factor, then the likelihood should indicate any problems here, by showing a low probability of the data given the image.

Bayes theorem not only looks for a probability of a given situation, it can update a situation given the current data and knowledge of a given situation. In a crystallographic environment the entropy of a map is defined as

$$S = -\sum_{i=1}^{n} p_i \log p_i \quad (2\text{-}6)$$

where p_i is a normalised probability of finding an electron in a particular volume of the unit cell

$$p_i = p_i \Big/ \sum_{i=1}^{n} p_i \,. \quad (2\text{-}7)$$

Where n is the number of pixels in a map or image.

The maximum entropy method helps remove the bias that a prior can introduce to the statistics.

In the entropy calculation itself the logarithm can be to any base but for the purposes of our calculations the base is always *e* the natural logarithm.

The main advantage of this approach in crystallography is that it allows the incorporation of known data into the results, concepts such as non-crystallographic symmetry (NCS), envelopes and known fragments from molecular replacement methods[21,23,24].

2.3.3 Why use maximum entropy?

The driving force behind using maximum entropy was initially to move away from the requirements of atomic resolution data, the main limiting factor of direct methods. Maximum entropy is stable regardless of data resolution. Another limiting factor of traditional methods was the use of Wilson statistics, where the distribution of atoms within the unit cell is assumed to be random. The use of maximum entropy allows new probability distributions to be accessed. In addition to allowing some of the traditional barriers to be direct methods calculations to be removed, maximum entropy allows the user to bring all their phasing knowledge to help solve the problem, such as, non-crystallographic symmetry, molecular envelopes, and known partial fragments, which can all be used as sources of phasing information.

2.4 Maximum entropy as applied to Crystallography

2.4.1 Normalisation

The majority of direct methods use normalised structure factors, but the MICE algorithm uses unitary structure factors of the form;

$$\left|U_{\underline{h}}\right|^{obs} = k\frac{\left|F_{\underline{h}}\right|^{obs}}{\sum_{j=1}^{N} f_j \exp\left(BSin\theta/\lambda\right)} \quad \textbf{(2-8)}$$

where $0 \leq U_{\underline{h}} \leq 1.0$. k is a scale factor, and B is an overall temperature factor.

Both the k and B factors can be found from a Wilson plot[25,26].

The $\sigma\left|U_{\underline{h}}\right|$ values are calculated as follows:

$$\sigma^2\left|U_{\underline{h}}\right| = \left|U_{\underline{h}}^2\right|\left\{\frac{\sigma^2\left|F_{\underline{h}}\right|}{\left|F_{\underline{h}}^2\right|} + \frac{\sigma^2 k}{k^2} + s^4\sigma^2 B + \frac{2s^2 k}{\quad} r(k,B)\sigma k\sigma B\right\} \quad \text{(2-9)}$$

Where the k and B factors and their errors can be found from a Wilson plot. The B factor is found from the gradient of the line and k is found at the intercept.

The use of unitary structure factors allows the U_0 to be equal to one. The use of N, the number of atoms in the unit cell has dubious meaning when the data are not at atomic resolution, and is a source of debate and research for many direct methods groups.

2.4.2 Data partitioning

After the data have been normalised, it is then split into 2 sets for the calculation stage. Group one is known as the set {H} and contains the structure factors with phases, these form the basis set and are used to form the constraints of the maximum entropy calculation. The second set is {K} which contains the structure factors with known magnitude but no known phases. These form the set {K} which will measure the log likelihood gain of a particular map, by comparing the magnitudes of the observed to those magnitudes predicted by the calculation. In electron crystallography a third set {U} can be used, it contains the unmeasured reflections, this set is disjoint from {H} and {K}.

2.4.3 Constrained entropy maximisation

The constrained entropy maximisation algorithm as applied in the MICE program is a complex but stable method for calculating the maximum entropy of any given node. The calculation can be deemed slow by other direct methods standards, but can cope with data that is unusable by conventional direct methods, so there is a trade off of between efficiency and data quality.

When using optimisation techniques, such as the Newton-Raphson method, they will not always converge to a solution, if the equations they are ill conditioned. An alternative method to this is to use Lagrangian multipliers, they are nearly always stable and converge to a solution. The Lagrangian multipliers turn a constrained problem into an unconstrained problem, when applied to a Newton or quasi-Newton problem.

The initial stage of the calculation requires initial Lagrangian multipliers to be estimated there is one Lagrangian multiplier per structure factor. In this context the Lagrangian multipliers are finding an approximation to the Hessian matrix.

The basis set of U magnitudes and phases are used to calculate an electron density along with any other prior knowledge such as an envelope or non-crystallographic symmetry.

$$\rho(\underline{x}) = \frac{1}{V}\left[1 + \sum_{\underline{h}\in H} U_{\underline{h}} e^{-2\pi i \underline{h}\bullet\underline{x}}\right] \cdot m(\underline{x}). \quad (2\text{-}10)$$

Where $m(\underline{x})$ is the prior knowledge and 1 is the U_0 value.

This map is truncated so that all the negatives are removed from the electron density, because maximum entropy requires the electron density to be positive everywhere. This is now as the $\rho'(\underline{x})$ electron density.

Next the natural logarithm of the $\rho'(\underline{x})$ map is taken so as to normalise the map and have everything on the suitable scale for what follows. This is known as the ω map.

$$\varpi(\underline{x}) = \ln \rho'(\underline{x}) \quad (2\text{-}11)$$

An inverse Fourier transform of the omega map gives the initial starting zetas or Lagrangian multipliers ζ, for all reflections, including those for unmeasured reflections belonging to the set {U} which consists of the unmeasured, unphased reflections, and unphased reflections belonging to the set {K}.

$$\varpi(\underline{x}) \xrightarrow{IFT} \xi_{\underline{h}\in(H,K,U)} \quad (2\text{-}12)$$

These zetas are truncated to give only the set , {H}, which give the initial starting set for calculation to begin. This is spectrum truncation.

$$\xi_{\underline{h}\in(H,K,U)} \xrightarrow{\text{Spectrum truncation}} \xi^{(0)}_{\underline{h}\in H} \quad (2\text{-}13)$$

To maximise the entropy of the map we must maximise

$$S_m = -\int \rho(\underline{x}) \log \frac{\rho(\underline{x})}{m(\underline{x})} d\underline{x}^3 . \quad (2\text{-}14)$$

subject to the constraints that $\rho(\underline{x})$ reproduces $U_{\underline{h}\in H}$ within experimental error.

Note that an additional prior term has been added to the usual definition of maximum entropy, and the summation has been replaced by integration.

The approach taken in MICE is to treat the problem like a Newton-Raphson refinement, which is allowed because the Lagrangian multipliers turn the problem from a constrained one into an unconstrained problem, thus allowing the calculation to be essentially a Newton-Raphson refinement.

From the initial starting zetas, the set is expanded into space group P1 and then a Fourier transform is taken to produce an initial ω map. The next stage is to exponentiate the map, suppressing regions of density below exp(-4.5) e^- $Å^{-3}$ (This figure is born out of experience, and has no real reason for using this particular cut off, other than it has been found to work consistently). The map is then normalised by summing over all the pixels, and prior knowledge is included.

$$\underset{\underline{h}\in H}{\zeta^{(initial)}} \xrightarrow{\text{Expand to P1}} \xrightarrow{FT} \omega^{(initial)} \xrightarrow{\exp} e^{\omega^{initial}(\underline{x})} \xrightarrow{Normalise} Z = \sum_{\underline{x}} e^{\omega^{initial}(\underline{x})} \cdot m(\underline{x}) \quad (2\text{-}15)$$

The resultant normalisation allows the initial entropy maximisation map to be produced.

$$q^{(initial)}(\underline{x}) = \frac{1}{Z} e^{\omega^{initial}(\underline{x})} \cdot m(\underline{x}) \xrightarrow{FT} U^{ME}_{\underline{h}\in(H,K,L)} \xrightarrow{\text{Spectrum truncation}} U^{ME}_{\underline{h}\in H} \quad (2\text{-}16)$$

At this stage we calculate the gradient $\delta q^{initial}(\underline{x})$. The gradient is calculated by

$$U^{obs}_{\underline{h}\in H} - U^{calc}_{\underline{h}\in H} = \Delta U^{initial}_{\underline{h}\in H} \xrightarrow{\text{FT + spectrum truncation}} \delta q(\underline{x}) \quad (2\text{-}17)$$

A Fourier transform of these $\Delta U^{initial}_{\underline{h}\in H}$ will produce a gradient map. Now we have a gradient and a Hessian matrix approximation[20] in the form of, $Hessian \approx \frac{1}{q^{ME}(\underline{x})}$.

$$x^{i+1} = x^i - gradient/Hessian. \quad (2\text{-}18)$$

The calculation as it stands would be too unstable and not converge, damping is used to move in the suggested direction by a percentage of the original suggestion. This helps the calculation converge in a more stable manner.

The next stage is finding how much to allow the calculation should move, this is known as the $\delta\zeta_{\underline{h}\in H}$ term.

$$\delta\xi_{\underline{h}} = FT - \frac{\delta q(\underline{x})}{q^{ME}(\underline{x})} \quad \textbf{(2-19)}$$

As with many optimisation techniques it is somewhat reckless to move the full amount that is suggested by the calculation. Therefore other factors are used to help suggest a more prudent amount to move at each step. Combinations of line and plane searches are used to find a more conservative value.

$$\zeta_{\underline{h}}^{new} = \zeta_{\underline{h}}^{current} + t\delta\zeta_{\underline{h}}^{current} - s\zeta_{\underline{h}}^{current} \quad \textbf{(2-20)}$$

2.4.4 The line search

The line search is used to find the most suitable distance to move.

$$\xi_{\underline{h}}^{(i+1)} = \xi_{\underline{h}}^{(i)} + t\delta\xi_{\underline{h}}^{(i)} \quad \textbf{(2-21)}$$

The t value is calculated such that a constraint function is a minimum,

$$C = \frac{1}{2}\sum_{\underline{h}\in H}\frac{1}{\sigma_{\underline{h}}^2}\left(\left|U_{\underline{h}}\right|^{obs} - \left|U_{\underline{h}}\right|^{ME}\right)^2$$

The constraint function C is essentially a χ^2 function. This is applied to make sure that the observed and calculated magnitudes are within experimental error. When C is modelled as a cubic function of t, there are four unknowns with four equations to be solved

$$C = a_0 + a_1 t + a_2 t^2 + a_3 t^3. \quad \textbf{(2-22)}$$

The four constants are found by finding the values of t that gives C a minimum. The values are computed as follows:

$$C|_{t=0} = a_0 \quad (2\text{-}23)$$

$$\frac{dC}{dt}|_{t=0} = a_1 \quad (2\text{-}24)$$

$$C|_{t=\alpha} = a_0 + a_1\alpha + a_2\alpha^2 + a_3\alpha^3 \quad (2\text{-}25)$$

$$\frac{dC}{dt}|_{t=\alpha} = a_1 + 2a_2\alpha + 3a_3\alpha^2 \quad (2\text{-}26)$$

Where α is a user-defined parameter that is normally 0.2

2.4.5 Plane search

The plane search uses another variable s as well as t to add more stability to the calculation.

$$\xi_{\underline{h}}^{(i+1)} = (1-s)\xi_{\underline{h}}^{(i)} + t\delta\xi_{\underline{h}}^{(i)} \quad (2\text{-}27)$$

The new variable s is a pullback, from the suggested value of ζ. The line search then takes a percentage of the suggested ζ and the plane search then takes a small amount from this value again to ensure that the newly predicted ζ are not moving too far. This time the constraints are modelled as a bi-cubic, in s and t, rather than as a cubic in just t.

$$C[s,t] = C_{00} + C_1 s + C_2 t + C_{11}s^2 + C_{12}st + C_{21}t^2 + C_{111}s^3 + C_{112}s^2 t + C_{222}t^3 \quad (2\text{-}28)$$

Once the constraints have been determined for C_{min}, we now move to C_{aim} which is a half of the proposed distance from the current point. Then map the constraint function on to the entropy space at the point C_{aim} also modelled in s and t. Now find where the function $S[s,t]$ is a maximum. From which the ω and q maps can be calculated.

The plane search is somewhat more complicated than the line search to calculate. In the line search the cycle of events occurs as follows:

$$\xi_{\underline{h}}^{current} \rightarrow \omega(\underline{x}) \rightarrow q(\underline{x}) \xrightarrow{\text{Spectrum truncate}} U_{\underline{h}\in H}^{ME} \rightarrow \delta U_{\underline{h}\in H} \rightarrow \delta\xi_{\underline{h}} \quad (2\text{-}29)$$

The new $\delta\zeta_{\underline{h}}$ are added to the current position, then the whole process starts again. Whereas in the plane search, the $\omega(\underline{x})$and $q(\underline{x})$ maps need to be calculated for each of the 4 trial positions, (0,0), (s_{trial},0), (0,t_{trial}) and (s_{trial},t_{trial}). Then for each of the 4 positions the constraint $C[s,t]$ is modelled as a bi-cubic function. Once the constraint function has been minimised, move half the proposed distance, then map the current

position onto an entropy function $S[s,t]$ also modelled as a bi-cubic in s and t. Then a similar process is used to solve the equations to maximise the entropy S in s and t.

$$S[s,t] = S_{00} + S_1 s + S_2 t + S_{11}s^2 + S_{12}st + S_{21}t^2 + S_{111}s^3 + S_{112}s^2 t + S_{122}st^3 + S_{222}t^3 \quad \textbf{(2-30)}$$

The 4 maps are calculated by starting from the current point, this means that the current _ map starts each of the plane search calculations[28].

$$\omega_0(\underline{x}) = \omega[0,0](\underline{x}) \quad \textbf{(2-31)}$$

The other ω maps are calculated at the four positions by the following equation, with s and t being either 0, s_{trial}, t_{trial}.

$$\omega[s,t](\underline{x}) = \omega_0(\underline{x}) + s\Delta_s\omega(\underline{x}) + t\Delta_t\omega(\underline{x}) \quad \textbf{(2-32)}$$

Where

$$\Delta_s\omega(\underline{x}) = -\omega(\underline{x}) \quad \textbf{(2-33)},$$

and

$$\Delta_t\omega(\underline{x}) = FT\left[\frac{\delta q(\underline{x})}{q(\underline{x})}\right]_{\underline{h}\in H} \quad \textbf{(2-34)}$$

For each of the 4 maps the maximum entropy map q^{ME} is calculated.

$$q[s,t] = \frac{e^{w[s,t](\underline{x})}}{Z[s,t]}. \quad \textbf{(2-35)}$$

Where

$$Z[s,t] = \sum_i^n N_i, \quad \textbf{(2-36)}$$

is a normalising constant, summing all the pixels N_i of the map.

The constraint function is then minimised,

$$\frac{dC[s,t]}{ds} = \left\langle \nabla_q[s,t]C, q[s,t], \Delta_s\omega \right\rangle - \left\langle q[s,t], \nabla_q[s,t]C \right\rangle \left\langle q[s,t], \Delta_s\omega \right\rangle \quad \textbf{(2-37)}$$

$$\frac{dC[s,t]}{dt} = \left\langle \nabla_q[s,t]C, q[s,t], \Delta_t\omega \right\rangle - \left\langle q[s,t], \nabla_q[s,t]C \right\rangle \left\langle q[s,t], \Delta_t\omega \right\rangle \quad \textbf{(2-38)}$$

Where

$$(\nabla_q[s,t]S)(\underline{x}) = -\omega[s,t](\underline{x}), \quad (2\text{-}39)$$

and

$$(\nabla_q[s,t]C)(\underline{x}) = -\delta q(\underline{x}) \quad (2\text{-}40)$$

The values for each equation are then computed at each point giving the 10 equations with 10 unknowns that can be minimised to give satisfy the equation $C[s,t]$.

$$C[0,0], C[0,t_{trial}], C[s_{trial},0], C[s_{trial},t_{trial}],$$
$$\frac{dC[0,0]}{dt}, \frac{dC[0,t_{trial}]}{dt}, \frac{dC[s_{trial},0]}{dt}, \frac{dC[0,0]}{ds}, \frac{dC[0,t_{trial}]}{ds}, \frac{dC[s_{trial},0]}{ds}$$

Once the 10 equations with 10 unknowns have been solved and a minimum found then,

$$C_{aim} = \frac{1}{2}C_{\min} + \frac{1}{2}C[0,0] \quad (2\text{-}41)$$

Contour the function $C[s,t]$ at contour C_{aim}. Now repeat the process of computing 10 equations and 10 unknowns to find the maximum value of $S[s,t]$,

$$S[0,0], S[0,t_{trial}], S[s_{trial},0], S[s_{trial},t_{trial}],$$
$$\frac{dS[0,0]}{dt}, \frac{dS[0,t_{trial}]}{dt}, \frac{dS[s_{trial},0]}{dt}, \frac{dS[0,0]}{ds}, \frac{dS[0,t_{trial}]}{ds}, \frac{dS[s_{trial},0]}{ds}$$

Now repeat the process of computing 10 equations and 10 unknowns to find the maximum value of $S[s,t]$.

When using maximum entropy there is often a problem with the high angle reflections build density too fast, so a damping factor is introduced to allow the low resolution structure to build before the high resolution features. The damping factor takes the form

$$\delta\zeta_{\underline{h}} = \delta\zeta_{\underline{h}} e^{-B\sin^2\frac{\theta}{\lambda^2}}, \quad (2\text{-}42)$$

Where B is a temperature factor, damping also involves three other factors; r is a user parameter usually set to 0.04, which sets a maximum move in entropy space; s_{trial} and

t_{trial} are more distance restraints and restrict the search by making the steps smaller as more features appear in the map

$$S_{trial} = \min\left[\frac{s_{trial}}{\max(\omega(\underline{x}),|\omega(\underline{x})|)}\right], s_{trial}, \text{(2-43)}$$

as the maps increase in detail, they become sharper, this in turn damps the s parameter more. t_{trial} is chosen on the basis of the smallest move between trial and trial divided by the new map produced by the delta ζ,

$$t_{trial} = \min\left[\frac{t_{trial}}{\max\left(FT\left[\delta\xi_{\underline{h}\in H}\right](\underline{x}),\left|FT\left[\delta\xi_{\underline{h}\in H}\right](\underline{x})\right|\right), t_{trial}}\right] \text{(2-44)}$$

The t parameter stops the entropy moving too far when the ζ values are increasing, which occurs when the maps become more detailed.

The above process occurs until the calculation converges or is stopped by reaching a preferred figure of merit. Essentially the whole process can be summarised as follows.

2.4.6 Technical details of calculation

From the above explanation it may not be clear why certain equations and figures of merit are being implemented. The following should help to explain why and how these equations are used. The χ^2 statistic is used to guarantee that the observed and calculated U-magnitudes are fitted within experimental error as long as the $\sigma_{\underline{h}}$ are known. Ideally the χ^2 value should be 1.0 or as close to as possible when the calculation finishes, although the calculation can be terminated by a log likelihood decrease before the ideal χ^2 value has been reached. The χ^2 calculation takes the form,

$$\chi^2 = (2n_n + n_c)^{-1} \sum S_{\underline{h}}^{-2} \left| |U_{\underline{h}}|^{obs} - |U_{\underline{h}}|^{ME} \right|^2 \text{(2-45)}$$

Where n_a is the number of acentric reflections, n_c is the number of centric reflections, and $s_{\underline{h}}$ the variance is,

$$s_{\underline{h}} = \sigma_{\underline{h}}^2 + p\varepsilon_{\underline{h}} \sum \text{(2-46)}$$

Where p is a mixing parameter ideally 1. $\varepsilon_{\underline{h}}$ is a statistical weight and $\sum$ is the sum of the scattering factors from the unit cell. The χ^2 function is used at the end of each

cycle of entropy maximisation, which allows the user to check that the calculation is converging in a well-behaved manner and that the calculated and observed U-magnitudes are close enough to make the calculations plausable.

The R factor is also calculated at the end of each entropy maximisation cycle without the addition of weighting factors to give an indication of how well the calculated and observed magnitudes are behaving. It takes the form,

$$R_{\underline{h}} = \frac{\sum_{\underline{h}\in H}\left|\left|U_{\underline{h}}\right|^{obs} - \left|U_{\underline{h}}\right|^{ME}\right|}{\sum_{\underline{h}\in H}\left|U_{\underline{h}}\right|^{obs}} \quad \textbf{(2-47)}$$

Reasonable values are considered between, 0.3 and 0.2, this indicating that the fit between $\left|U_{\underline{h}}\right|^{obs}$ and $\left|U_{\underline{h}}\right|^{ME}$ is optimal.

The final set of equations that are calculated at the end of each cycle of refinement are the constraints. The constraints are then used in the next cycle of entropy maximisation. They constraints take the form,

$$C = \frac{1}{2}\sum_{\underline{h}\in H}\frac{1}{\sigma_{\underline{h}}^{2}}\left(\left|U_{\underline{h}}\right|^{obs} - \left|U_{\underline{h}}\right|^{ME}\right)^{2} . \quad \textbf{(2-48)}$$

The other figure of merit that has been mentioned but not yet defined is the log likelihood gain (LLG). The LLG is used to measure how well the basis set is predicting the set of reflections {K}, the thinking being that the more accurately the set of reflections {K} can be predicted, compared to the observed reflections in the set {K}, the more accurate the prediction of the phases for the set {H}, and its accompanying entropy maximisation. In the following equations the reflections are treated independently within the equations.

For each extrapolated reflection $\underline{k}$ defines

$$\Lambda_{\underline{k}} = \frac{\left|U_{\underline{k}}\right|^{obs}}{\varepsilon_{\underline{k}}\sum + \sigma_{\underline{k}}^2} \exp\left\{-\frac{1}{2}\frac{\left(\left|U_{\underline{k}}\right|^{obs}\right)^2 + \left(\left|U_{\underline{k}}\right|^{ME}\right)^2}{\varepsilon_{\underline{k}}\sum + \sigma_{\underline{k}}^2}\right\} I_0\left(\frac{\left|U_{\underline{k}}\right|^{obs}\left|U_{\underline{k}}\right|^{ME}}{\varepsilon_{\underline{k}}\sum + \sigma_{\underline{k}}^2}\right) \quad \textbf{(2-49)}$$

acentric case

$$\Lambda_{\underline{k}} = \frac{\left|U_{\underline{k}}\right|^{obs}}{\pi\left(2\varepsilon_{\underline{k}}\sum + \sigma_{\underline{k}}^2\right)} \exp\left\{-\frac{1}{2}\frac{\left(\left|U_{\underline{k}}\right|^{obs}\right)^2 + \left(\left|U_{\underline{k}}\right|^{ME}\right)^2}{2\varepsilon_{\underline{k}}\sum + \sigma_{\underline{k}}^2}\right\} \cosh\left(\frac{\left|U_{\underline{k}}\right|^{obs}\left|U_{\underline{k}}\right|^{ME}}{2\varepsilon_{\underline{k}}\sum + \sigma_{\underline{k}}^2}\right) \quad \textbf{(2-50)}$$

centric case

This equation asks the question how well do the observed and calculated maximum entropy magnitudes agree. As with any probabilistic question a hypothesis is needed as a reference point, in this case we use the classic Wilson statistics as our comparison. In Wilson statistics the atoms are assumed to be random within the unit cell, which is known to be wrong but still works effectively for small molecules. In terms of the Rice functions used, the $\left|U_{\underline{k}}\right|^{ME}$ terms are set to zero, the equation now becomes,

$$\Lambda_{\underline{k}}^{0} = \frac{\left|U_{\underline{k}}\right|^{obs}}{\varepsilon_{\underline{k}}\sum + \sigma_{\underline{k}}^2} \exp\left\{-\frac{1}{2}\frac{\left(\left|U_{\underline{k}}\right|^{obs}\right)^2}{\varepsilon_{\underline{k}}\sum + \sigma_{\underline{k}}^2}\right\}. \quad \textbf{(2-51)}$$

The log of the ratio is taken for each reflection,

$$L_{\underline{k}} = \log\frac{\Lambda_{\underline{k}}}{\Lambda_{\underline{k}}^{0}}, \quad \textbf{(2-52)}$$

This is then summed for all reflections in the set {K} to obtain the log-likelihood gain (LLG),

$$LLG = \sum_{\underline{k}} L_{\underline{k}} \quad \textbf{(2-53)}$$

As the LLG is used as a measure of phasing power, and not mean phase error; it will only indicate that there is the possibility of having a low mean phase error. Likelihood only indicates how well the predicted reflections in the set {K} have been calculated, and is therefore not a direct indicator of phase error, because likelihood only uses amplitudes and not phases.

2.5 The MICE algorithm in action

MICE is the FORTRAN 77 implementation of the maximum entropy principle as applied to crystallography by Bricogne in his 1984 paper. MICE stands for Maximum entropy In a Crystallographic Environment. The procedure has two main parts, RALF and MICE, RALF is the normalisation program which calculates the U-magnitudes from the raw data files, whereas MICE is the segment of the program which deals with the structure solution *via* maximum entropy[26-32].

For *ab initio* calculations the user first of defines the origin, the procedure involves assembling a list of the large U-magnitudes, there phases are assumed to be known. Each reflection is in turn eliminated from the list on the basis of which reflection is most unfavourably optimising the second neighbourhood. This reflection is then removed and the procedure repeated. Once all the reflections have been eliminated, a reverse list of the most suitable reflections has been created. The origin defining reflections will be selected from the most suitable phased reflections from the top of the revised list of reflections, with respect to the symmetry.

Once the origin has been defined, the user then selects the KNOW command to tell the program to use that particular origin, they may enter this as reflection number followed by phase angle, in the KNOW command.

After origin definition the next stage is to permute some reflections. The reason for this is that the maximum entropy method cannot build phase relationships like the traditional direct methods can so it needs to acquire phase information from elsewhere to help start the calculation. This is achieved by permuting a set of reflections, the idea being that if a large portion of phase space is covered efficiently then the probability of having a set of reflections with phase close to their true value will be high. There are three stages to the process, the first of which is picking the method of phase permutation. This can be done by magic integers (MI)[33], error-correcting codes (ECCs)[31], full factorial designs (FFDs) or incomplete factorial designs (IFDs). The most common method in MICE is to use ECCs, because they tend to be the most efficient way of covering phase space with a reasonable number of nodes. Magic integers produce accurate phase predictions but the permutations tend to cluster rather than evenly sample phase space, IFDs also suffer from the same problem, whereas FFDs cover too much of phase space and therefore become impractical due to the

number of calculations they produce. Nonetheless all four options are available. Once a method of phase permutation has been decided the user must then select the reflections to be permuted, which is done *via* the NEXT command. The NEXT command selects reflections on the basis of the resolution range choosing the largest U-magnitudes and optimising the second neighbourhood of the basis set. The final stage of this process is now to permute the reflections *via* the PERMUTE command. This takes the origin (the current set {H}) and adds the chosen set of reflections with a set of phases, to create a set of new basis sets {H}, which are known as nodes. This process is repeated until all the permutations have been exhausted. The user now has large collection of basis sets all with the same reflections but different phases, and each individual basis set must now be maximised.

The reflections to be added from {K} to {H} are selected on the basis of optimising the second neighbourhood. For a given set {H} containing *n* reflections, the second neighbourhood of {H} consists of the N_2 symmetry unique set of reflections distinct from the basis set,

$$k = \underline{h}_1 \pm {}^t\underline{R}_g \underline{h}_2 \text{ for } \underline{h}_1 \underline{h}_2 \in H \quad \textbf{(2-54)}$$

Where ${}^t\underline{R}_g$ is the transpose of the rotation matrix for the space group. Reflection k can also be defined *via* a second writing of the form

$$\underline{k} = \underline{h}_3 \pm {}^t\underline{R}_g \underline{h}_4 \text{ for } \underline{h}_3 \underline{h}_4 \in \{H\}. \quad \textbf{(2-55)}$$

Reflections are then transferred from {K} to {H} such that the quantity:

$$\sum_{\underline{k}} \left| E_{\underline{k}}^2 - 1 \right\| U_{\underline{h}1} \left\| U_{\underline{h}2} \right\| U_{\underline{h}3} \left\| U_{\underline{h}4} \right| \quad \textbf{(2-56)}$$

is maximised. It can now be stated that origin defining reflections are preferably chosen in a similar fashion: one starts with a complete set of strong reflections and eliminates them stepwise using the equation[19]

$$\sum_{\underline{k}} \left| E_{\underline{k}}^2 - 1 \right\| U_{\underline{h}1} \left\| U_{\underline{h}2} \right\| U_{\underline{h}3} \left\| U_{\underline{h}4} \right| \quad \textbf{(2-57)}$$

as a measure of strength.

Once the reflections have been chosen they must be phased before they can be added to the basis set {H}. To do this, the PERM command is invoked, where a variety of

phase permutation techniques can be chosen: error-correcting codes, magic integers, incomplete factorial designs, or full factorial designs. Usually error-correcting codes are used, in which case the phase assignment uses quadrant fixing for acentric reflections, $\pm\pi/4$, $\pm 3\pi/4$, and centric reflections are assigned one of the two possible phase choices, *e.g.* 0, π or $\pm\ \pi/2$. A node is created for each of the possible phase combinations assigned by the error correcting code. Each node forms a branch on one level of a phasing tree: the root node, on the first level, is defined by the origin reflections; the second level consists of the nodes generated by the NEXT and PERM commands, during phase permutation. Additional levels can be built by choosing a particular node and carrying out further phase permutation.

Before proceeding with entropy maximisation the calculations must all occur on the same scale. To achieve this they use a common p-factor. The p-factor is a fit between Wilson statistics and the errors on the unitary structure factors. This stops the calculation from over fitting the U-magnitudes in the set

$$\{\mathbf{H}\},\ s_{\underline{h}}^2 = \sigma_{\underline{h}}^2 + p\sum\ ,\ \text{where}\ \sum\ = \sum_{j=1}^{N} f_j^2 \quad (2\text{-}58)$$

The p-factor is calculated so that the χ^2 starts at ≈ 8.5. The p-factor is calculated using Brent's optimisation method, this uses a bracketing method which picks to points a and b these are your boundary conditions, three more points are picked, u, v, w, within the boundaries a and b. The points u and v are the brackets of the minimum points found thus far, w is a new point being evaluated, if it is lower than u and v then the bracket u or v is moved to restrict the evaluation of the next point w. Eventually the point w will make little or no change the minimum of the bound function is considered to be found. This is considered to be a fast method of minimisation, and it does not require any derivatives to be calculated.

Once a p-factor has been determined the entropy maximisation can start. Rather than maximise each node individually, MICE has an accompanying auxiliary program mega-MICE (MM). Mega-MICE automatically does each calculation sequentially, and to speed up the process, it sends each node out across a network of 20 computers. When all the calculations have been completed, MICE collates all the data back together in a usable form. When a calculation reaches this stage a user is faced with potentially thousands of basis sets to choose from so some pruning of the phase tree

must occur. Initially the pruning was done on the basis of the highest log-likelihood values. This method works to a limited extent, but then a more sophisticated method was developed. The new method, known as analysis, uses Student t-tests to analyse each node with respect to the singlet, doublet and triplet phase relationships of the reflections with usually a 2% level of significance, although the level of significance can be altered depending on the quality of the data. The significance level is often lowered to 10% when codes are employed.

The analysis process exists because there are random fluctuations in the probability distributions of both the hypothesis and the null hypothesis that lead to the possibility of the null hypothesis being larger than the hypothesis when the basis set being used is correct. So we therefore analyse the statistical relationships within the distribution to ensure that they are consistent with the observation that the hypothesis is larger than the null hypothesis or *vice versa*.

The analysis process looks to see if any of the permuted phases have any single, double or triplet relationships between them. If they do then this suggests that they have a significant influence on the probability distribution, and that the LLG has a high probability of overcoming the random fluctuations in H_1. The simplest case is detection of the main effect on a single centric reflection. The main effect is the sign associated with a particular reflection. This is calculated by μ^+ the LLG average and its variance V^+ for the sets in which the sign is +. Then the converse is produced and the μ^- and associated V^- are computed for the sets where the sign is -. The t-test is then:

$$t = \frac{\left|\mu^+ - \mu^-\right|}{\sqrt{V^+ - V^-}}.$$

The t-test then allows the sign to be allocated with an associated significance level. The default is ≤2%. The test is done for every single-phase indication, and is then repeated for any doublet and triplet phase relationships, where permitted. (Some codes like the Nordström-Robinson code have to sparse a coverage of phase space to enable triplet phase relationships)

For each of the m phase relationships, i, a weight is calculated indicating how reliable the phase relationship is believed to be. The weight

$$w_i = \left(1 - \frac{I_1(s_i)}{I_0(s_i)}\right) \quad \textbf{(2-59)}$$

Is determined by the significance, s_i of the Student t-test result for reflection i I_0 and I_1 are Bessel functions of the zeroth and first order. The Bessel functions take the form of $I_n \approx \frac{1}{n!}\left(\frac{x}{2}\right)^n$ For each node, n, an overall score is calculated by summing all the weights where there is an agreement between the basis set {H} and the t-test. This is then multiplied by the LLG:

$$score_n = LLG_n \sum_{j=1}^{m} w_j . \quad \textbf{(2-60)}$$

The scores and associated nodes are then ranked and only the top 8-16 are kept.

If any of the nodes look promising *via* the score statistics and visual inspection of the maps then they are kept, more reflections are then permuted. Each set of the permuted reflections and phases generated by the permutation are added to the basis set {H}, creating the next level in the phasing tree. Entropy maximisation is applied to each node followed by analysis, the procedure is repeated until a solution is found.

2.5.1 Maps

There are 2 types of maps produced by the MICE program to allow the user to view their results;

1 The $q^{ME}(\underline{x})$ map, which is the current maximum entropy map.

2 A centroid map, which is a Sim weighted $q^{ME}(\underline{x})$ map where all the extrapolated reflections in the set {K}, which now have predicted phases, are given a weight. The weights are designed to minimise the mean square error of the electron density in the maps and minimise the uncertainty of the peak positions. Reflections with a weight ≤ 0.1 are usually excluded from the maps.

$q^{ME}(\underline{x})$maps are not maps in the traditional sense but are probability distributions. The peaks in the map correspond to atom positions, just like the traditional maps. The centroid maps use Fourier coefficients calculated as follows:

$$|U_k|^{obs}\left[\frac{I_1(X_k)}{I_0(X_k)}\right]\exp\left(i\varphi_k^{ME}\right) \quad (2\text{-}61)$$

where:

$$X_k = \left(\frac{2N}{\varepsilon_k}\right)|U_k|^{obs}|U_k|^{ME} \quad (2\text{-}62)$$

The reflections {H} are also included with $|U_{\underline{h}}|^{obs} + \phi_{\underline{h}}^{ME}$ *i.e.* with unit weight.

The centroid map type can be produced as either $|U|$ or $|F|$ maps. U-maps produce more peaky maps with more defined atom positions, whereas F-maps show more bonding type feature, due to the different contribution of the scattering factor in the calculation.

2.6 P(δq) functions

The maximum entropy method as described above is a long and computationally expensive process. Attempts have been made to shorten the time of the calculation, without degrading the quality of the results. One such method is the P(δq) function[28], which was used in the early maximum entropy papers and has since fallen out of favour, because of fluctuations in its effective use, as a source of phasing power. The associated figures of merit are also a source of concern due to their lack of discrimination between nodes. Nonetheless the function still remains, and is a possible source of improvement for the calculations. $P(\delta q)$ is defined *via*:

$$P(\delta q) = -\int_V \frac{\delta q(\underline{x})^2}{q^{ME}(\underline{x})} d^3\underline{x} \quad (2\text{-}63)$$

This function is considerably quicker to calculate than a whole entropy maximisation procedure. As it only requires one Fourier transform, the point of this function is to quickly produce features in the map where q^{ME} is large, making it potentially very useful near the start of an *ab initio* calculation, or when a low resolution envelope has been determined and we wish to build density at a higher resolution on it. A more detailed description is given in chapter 6.

2.7 Error-correcting codes (ECCs)

Within MICE, error-correcting codes are one of the most efficient methods of phase permutation, which Bricogne originally implemented in the BUSTER program[34]. The

reason for using error-correcting codes over other methods, such as quadrant permutation or magic integers, is the very efficient gain in sampling of phase space. The Golay code[35] for instance allows you to permute 24 degrees of freedom, and only generate 4096 nodes, or phase sets, instead of 2^{24}=16777216, and one of the 4096 nodes is guaranteed to have a maximum of 4 wrong phase choices. The origin of error-correcting codes lies in the same arena as that of maximum entropy theory; they both originate from information theory, which is mainly concerned with transmitting messages across noisy communication channels such as telephone wires or beaming back information from satellites to earth. The reason it is a useful tool in crystallography is because of the way some codes are designed. The design of codes is such that, if error occurs, then that error can be detected and corrected. The spatial arrangement of the points of the code are such that the nearest point to the error occurring should be the correct answer, although that depends on the individual code properties, making only some codes useful as experimental designs.

A more practical explanation of how a code works is explained below. First of all some definitions from coding theory are required[36]. First of all there is the alphabet, which defines the symbols chosen to write the code in. All the examples used in MICE are binary codes using 0,1 as their alphabet. F_q is the alphabet, of q elements, so in the case of MICE it is an F_2={0,1} alphabet. The code C is made up of M codewords $(F_q)^n$ where each codeword is of length n. Codewords are a sequence of symbols from the given alphabet.

The detection of the error is possible because of the Hamming distance, d, between the various codewords. The Hamming distance is the shortest distance between any 2 codewords, and is the number of places in which any 2 codewords differ. Codes are usually described as, (n, M, d), where n is the length of each codeword, M is the number of codewords and d is the minimum distance between any two codewords. The three codes used in MICE are Hamming (16, 16, 4), Nordström-Robinson (16, 256, 6), and Golay (24, 4096, 8). An ECCs can detect $d-1$ errors and can correct $(d-1)/2$ errors, by assigning the closest codeword[37]. The easier it is for a code to detect and correct an error the more useful it is as a potential source of phase permutation. The fewer codewords that are needed to detect and correct errors are the best codes from a phase permutation perspective, because this means that the code

covers phase space very efficiently in the fewest permutations. Golay codes are the most efficient codes for this type of phase space coverage. The error that is being detected is not a phase angle. Instead it is the proximity to one of the phase choices being permuted.

Error-correcting codes in MICE are used for phase permutation in the following manner. Centric reflections are represented by one binary digit 1 or 0 for their phase choice. 0 represents the phase choice 0°, and 1 the alternative of π.

Acentric reflections on the other hand need two bits to assign their phase quadrant. One bit is used to assign the real part and the other on the imaginary, *i.e.* $0,0 = \pi/4; 1,0 = 3\pi/4; 1,1 = 5\pi/4; 0,1 = 7\pi/4$[31].

2.8 Foruier methods

Fourier methods are used to locate missing atoms within a structure, often known as Fourier synthesis recycling[41]. Starting from a fragment, the electron density is calculated using the calculated phases from the fragment. The map should contain the original fragment and reveal the positions of other atoms within the structure. The new atoms along with the original fragment are then used to obtain phases and recalculate the electron density. This procedure is continued until the structure is complete.

The methods works by splitting the structure factor into known and unknown portions.

The Fourier synthesis will give peaks at the sites of unknown atom positions, some will be incorrect and peak picking should be done by taking into account the size of the peak and does it make chemical sense for the peak to be in that position. Alternatively $2|F_{obs}| - |F_{calc}|$ can be used to improve accuracy of the calculation.

The $2|F_{obs}| - |F_{calc}|$ Fourier synthesis gives an improved weighting scheme and for the unknown peaks compared to the background, it also helps to suppress wrongly included atoms in the structure.

2.9 Refinement

The concept of least squares atomic parameter refinement has gone hand in hand with crystallography ever since its inception, and especially as a method for validating the

structural model proposed by direct methods, Patterson methods, molecular replacement solutions, or any other way of deriving a structure. For small molecule crystallographers the dominant method is least-squares. In recent times biological crystallographers have started to use maximum likelihood refinement as a variant of this.

Once a structure is believed to be complete in the sense that all the atoms have been located, then the process of least-squares refinement is used to find a more accurate fit between the observed and calculated data. In doing so the refinement process can show a possible error in the proposed model, such as an incorrectly placed molecule, or the wrong spacegroup. As the least squares starts to converge it is possible to check the atomic displacement parameters on the atoms in the structure. There is a lot of information available to the crystallographer, that allows the atom assignment to be checked by allowing the atoms to be modelled by an additional six parameters rather than four[38]. If the thermal parameters become too small or big, this implies that there is a problem with the atom assignment.

The refinement process is heavily reliant on the R factor

$$R = \frac{\sum \left| F_{\underline{h}}^{obs} \right| - \left| F_{\underline{h}}^{calc} \right|}{\sum \left| F_{\underline{h}}^{obs} \right|} \quad \textbf{(2-64)}$$

or a weighted form which is currently more used than (2-64)

$$R_w = \frac{\sum_i^n w_i \Delta F_i^2}{\sum_i^n w_i \left| F_i^o \right|^2} \quad \textbf{(2-65)}$$

This figure of merit is very simple and can be misleading, and great caution must be taken when using it as a figure of merit. In general the following can be used as guides to how well a structure determination is progressing for small molecules, R = $0.45 \leq 0.35$ the model has some merit; R = $0.35 \leq 0.25$, the model should refine to true positions; R = $0.25 \leq 0.15$, Within 0.1Å of their true positions[2].

Singular value decomposition (SVD)[39], is different approach to the least squares method which inverts the matrix in a somewhat different and more robust manner. Maximum likelihood (ML)[31,40] uses the errors on the observed magnitudes and phases

to help fit the calculated to the observed data. This can lead to dramatic improvements when dealing with poor data parameter ratios, such as those produced in protein samples, and is the only alternative to least squares method which has been widely implemented.

References

(1) Vainshtein, B.K. *Structural Analysis by Electron Diffraction*; The MacMillan Company: New York, **1964**.

(2) Dorset, D.L. *Structural Electron Crystallography*; Plenum Publishing Coorporation: New York, **1995**.

(3) Williams, D.B.; Carter, C.B. *Transmition Electron Microscopy*; Plenum Press: New York and London, **1996**.

(4) Voigt-Martin, I.G.; Yan, D.H.; Gilmore, C.J.; Shankland, K.; Bricogne, G. *Ultramicroscopy* **1994**, *271*, 271-288.

(5) Voigt-Martin, I.G.; Yan, D.H.; Yakimansky, A.; Schollmeyer, D.; Gilmore, C.J.; Bricogne, G. *Acta Crystallographica* **1995**, *A51*, 849-868.

(6) Voigt-Martin, I.G.; Zhang, Z.X.; Kolb, U.; Gilmore, C.J. *Ultramicroscopy* **1997**, *68*, 43-59.

(7) Cowley, J. M. editor. Electron diffraction techniques. **1992**, 1.

(8) Cowley, J. M. editor. Electron Diffraction Techniques. **1993**, 2. 0-19-855733-7(v.2).

(9) Dorset, D.L.; Hu, H.; Jäger, J. *Acta Crystallographica* **1991**, *A47*, 543-549.

(10) Dorset, D.L. *Micron* **1994**, *25*, 423-430.

(11) Amos, L.A.; Henderson, R.; Unwin, P.N.T. *Progress in Biophysical Molecular Biology* **1982**, *39*, 183-231.

(12) Woolfson, M.M. *An Introduction to X-ray Crystallography*; Cambridge University Press: **1997**

(13) Miller, R.; Gallo, S.M.; Khalak, H.G.; Weeks, C.M. *Journal of Applied Crystallography* **1994**, *27*, 613-621.

(14) DIRDIF. Beurskens, P.T.; Admiraal, G.; Beurskens, G.; Bosman, W.P.; Garcia-Granda, S.; Gould, R.O.; Smits, J.M.M.; Smykalla, C. **1980**.

(15) Shelx97. Program for the refinement of crystal structures, University of Göttingen, Germany. Sheldrick, G.M. **1997**

(16) Altomare, A.; Burla, M.C.; Camali, M.; Cascarano, G.L.; Giacovazzo, C.; Guagliardi, A.; Moliterni, A.G.G.; Polidori, G.; Spagana, R. *Journal of Applied Crystallography* **1999**, *32*, 115-119.

(17) Gilmore, C.J. *Journal of Applied Crystallography* **1984**, *17*, 42-46.

(18) Sayre, D. *Acta Crystallographica* **1952**, *5*, 60-65.

(19) Hauptman, H.; Karle, J. Solution of the phase problem: I. The centrosymmetric crystal. **1953**, Monograph 3.

(20) Bricogne, G. *Acta Crystallographica* **1984**, *A40*, 410-445.

(21) Gilmore, C.J. *Acta Crystallographica* **1996**, *A52*, 561-589.

(22) Bayes, R.T. *Philosophy Transactions Royal Society London* **1763**, *53*, 370-418.

(23) Bricogne, G. *Acta Crystallographica* **1988**, *A44*, 517-545.

(24) Gilmore, C.J.; Shankland, K.; Fryer, J.R. *Ultramicroscopy* **1993**, *49*, 132-146.

(25) Wilson, A.J.C. *Acta Crystallographica* **1949**, *2*, 318-321.

(26) Gilmore, C.J.; Bricogne, G. *Methods in Enzmology* **1997**, *277*, 65-78.

(27) Dong, W.; Gilmore, C.J. *Acta Crystallographica* **2001**, *A54*, 438-446.

(28) Gilmore, C.J.; Bricogne, G. *Acta Crystallographica* **1990**, *A46*, 297-308.

(29) Gilmore, C.J.; Henderson, K.; Bricogne, G. *Acta Crystallographica* **1991**, *A47*, 830-841.

(30) Gilmore, C.J.; Henderson, A.N.; Bricogne, G. *Acta Crystallographica* **1991**, *A47*, 842-846.

(31) Gilmore, C.J.; Dong, W.; Bricogne, G. *Acta Crystallographica* **1999**, *A55*, 70-83.

(32) Shankland, K.; Gilmore, C.J.; Bricogne, G.; Hasizume, H. *Acta Crystallographica* **1993**, *A49*, 493-501.

(33) Main, P. *Acta Crystallographica* **1977**, *A33*, 750-757.

(34) Bricogne, G. *Acta Crystallographica* **1993**, *D49*, 37-60.

(35) Golay, M.J.E. *Procedings IRE (IEEE)* **1949**, *37*, 23-28.

(36) Pretzel, O. *Error-correcting codes and finite fields*; Oxford University Press: New York, **1992**.

(37) Hill, R. *A first Course in Coding Theory*; Oxford University Press: **1986**.

(38) Watkin, D. *Acta Crystallographica* **1994**, *A50*, 411-437.

(39) Press, W.H.; Teukolsky, S.A.; Vetterling, W.T.; Flannery, B.P. *Numerical Recipices in C*; Cambridge University Press: **1996**

(40) Pannu, N.S.; Read, R. *Acta Crystallographica* **1996**, *A52*, 659-668.

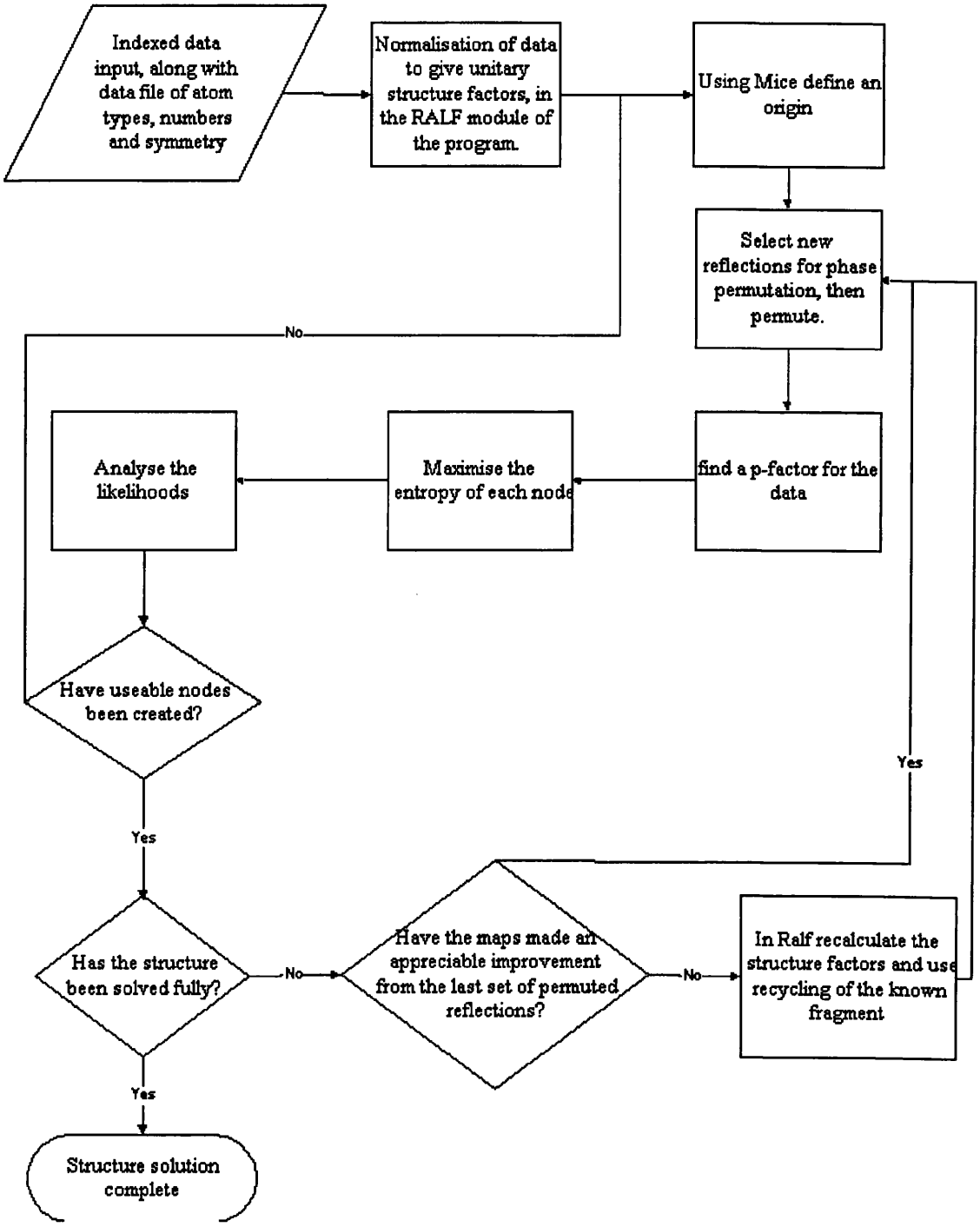
(41) Karle, J. *Partial Structures and the Tangent Formula*. In: *Crystallographic Computing*; Ahmed, F.R. Ed. **1976**.

3 AB INITIO STRUCTURE SOLUTION FROM ELECTRON DIFFRACTION DATA

3.1 *Ab-initio* procedure

Structure solution by the maximum entropy program MICE is a flexible and varied process, as each solution can vary quite dramatically in terms of how the final solution is obtained. The flowchart in Figure 3-1 shows a summary of how solutions can be obtained from the program.

Figure 3-1 Flow diagram of the *ab initio* calculation process



3.2 *Ab initio* calculations

Several molecules have been solved from electron diffraction data using maximum entropy methods. This is an area where direct methods normally fail, because of the sparse nature, and quality, of electron diffraction data sets. The incomplete nature of the data due to the restrictions of the microscope (the missing cone problem, to be discussed later), and its error prone origins arising from dynamical scattering. The unknown thickness and/or bending within a given crystal sample means that the data

has a high probability of producing some dynamical scattering. This implies that at best the data produced is only approximately of a kinematical nature. These factors make electron diffraction data more suitable for the robust technique of maximum entropy rather than traditional direct methods. The statistical framework allows for many of the foibles that electron diffraction data can possess, whereas many of the other statistical alternatives cannot tolerate these problems.

The field of direct methods as, applied to electron crystallography is relatively new compared to X-ray crystallography. The reasons for this being that the electron crystallographic community believed that their data was not complete enough in terms of having very over determined data at atomic resolution and the quality of the data was dubious, in that data would not all be close enough to the kinematic approximation to allow the direct methods to work, as it does in the X-ray case. This was proven to be a wrong assumption in the 1991 by Fan and co-workers, who applied the Sayre equation to phase a extend from 2Å to 1Å for the Copper perchlorophthalocyanine. Since then Dorset has lead the field in applying direct methods and to various types of electron diffraction data sets, from small molecules to polymers and even proteins. Dorset has used Fourier synthesis. Symbolic addition and the Tangent formula to solve electron diffraction data sets[8]

The following examples are of varying difficulty for the maximum entropy algorithm to solve. All but the Brucite structure fail with routine direct methods packages, but even here the power of the technique with electron diffraction data is highlighted with its ease of solution, and its ability to locate hydrogen atoms.

3.3 Brucite

The Brucite structure was determined by Boris B. Zvyagin and co. workers in 1997 [1]. The cell parameters are, a = 3.149, b = 3.149, c = 4.769, in the space group $P\bar{3}m1$, and Z = 1 Brucite has the molecular formula $Mg(OH)_2$; there are a total of 70 unique reflections.

3.3.1 Solution

First the origin was chosen and defined using the ORIGIN and KNOW commands in MICE. The top 30 reflections between 0.7 and 50Å resolution were examined for the

most appropriate origin, and this was then defined as : 1 0 3 with phase angle 360°. The NEXT command is then used to select reflections for permutation; the number of reflections for permutation is dictated by the error-correcting code being used. Full factorial designs can be used, but these are often avoided due to their computationally intensive nature. The PERM or PERMUTATION command is then used to create new data sets to be entropy maximised. This means that all reflections selected by the NEXT command are given a permuted phase. Each combination of phases given to the reflections defines a basis set. Each basis set was then entropy maximised under a common p-factor.

Each basis set or node has its entropy maximised in turn. Once every node has had its entropy maximised the information is collated and can be looked at in two different modes: (i) One simply looks for the highest log-likelihood from the complete set of nodes that have been entropy maximised. This method is crude and could be misleading in the sense that LLGs assume that there is no relationship between the reflections which is clearly wrong otherwise traditional direct methods would not work, although LLGs are an excellent method for distinguishing between phase sets given only a few phased reflections. (ii) To analyse the nodes looking for meaningful phase relationships in conjunction with the log-likelihood results. This results in a score, which is used to rank the results. (see pages 30-31)

The Brucite solution was achieved after origin definition and one set of phase permutation with a Nordström-Robinson error-correcting code, followed by analysis. The solution was obvious upon visual inspection of the maps, and was within the top 8 solutions suggested by the analysis.

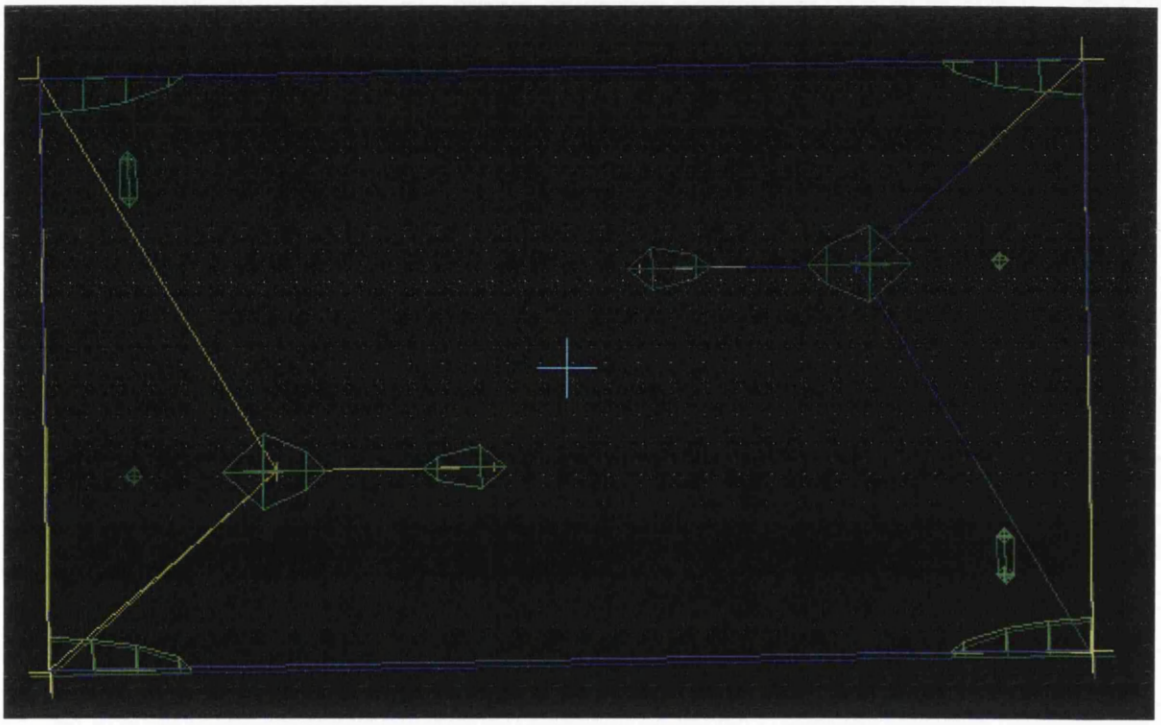
Table 3-1 Brucite, top 8 nodes from analysis

Set No.	Log Likelihood Gain	Score
234	12.930	1.000
10	12.930	1.000
58	10.078	0.779
150	5.507	0.313
22	5.507	0.313
141	5.337	0.303*
230	4.154	0.236
7	5.233	0.206

* Best solution.

Nodes 234, 10, 5, all have higher scores than the final solution but their maps upon visual inspection show only one significant peak and show no real sign of an atomic structure. The other nodes 230, 150, 7, and 22 all have what looks like noise in the map with no significant peaks that could represent an atomic structure. Whereas node 141 clearly has several peaks of density at sensible distances which chemically make sense as the Brucite $Mg(OH)_2$ structure.

Figure 3-2 Centroid map for node 141: Brucite



3.4 Isotactic Poly(1-butene) Form III

The data set for Isotactic Poly(1-butene) Form III was obtained by electron diffraction by Dorset and co-workers in 1994[2]. They obtained the data set from solution and epitaxially grown crystals. A tilt series was used collect as much three-dimensional data as possible. The orthorombic space group is $P2_12_12_1$ with unit cell dimensions a = 12.38, b = 8.88, c = 7.56Å, Z = 4. The data set has 125 unique reflections. The original solution was achieved by the tangent formula in the QTAN program[3].The figure of merit used was the NQUEST figure of merit, which has been found, in general, to be poor figure of merit for electron diffraction work due to the poor sampling of reciprocal space. Because of this poor figure of merit [4], 20 of the phases were determined separately by symbolic addition. These phases were then used as a comparison for each permutation produced by the QTAN program.

3.4.1 Solution

The final model for polybut-1-ene was arrived at *via* several stages of phase permutation and extension of the basis set. After each stage of permutation, a node was selected for further permutation and enlargement of the basis set. The process of selecting which node was suitable was a mixture of using the scores and log-likelihoods from the analysis process and visual inspection of the top few suggested maps to look for the most suitable candidate. Suitable candidates were chosen on the premise of which electron densities visually looked most likely to contain a solution, with density peaks being at sensible carbon-carbon distances. The following nodes were found the appropriate route to the final solution. 1->789->4623 ->8466. Node1, the root node, was defined:

h	k	l	phase
4	5	3	90
9	6	3	360
2	2	0	90

Table 3-2 Polybut-1-ene, route to final solution

Node	Maximum Likelihood	Chi Squared
789	13.05	1.20511
4623	13.47	2.43038
8466	15.10	5.37407

Table 3-3 Polybut-1-ene, analysis results from final set of nodes

Set No.	Log Likelihood Gain	Score
8474	15.531	0.918
8466	15.095	0.892*
8730	16.926	0.840
8722	15.029	0.746
8478	14.805	0.735
8470	14.051	0.697
8506	13.804	0.685
8498	13.650	0.678

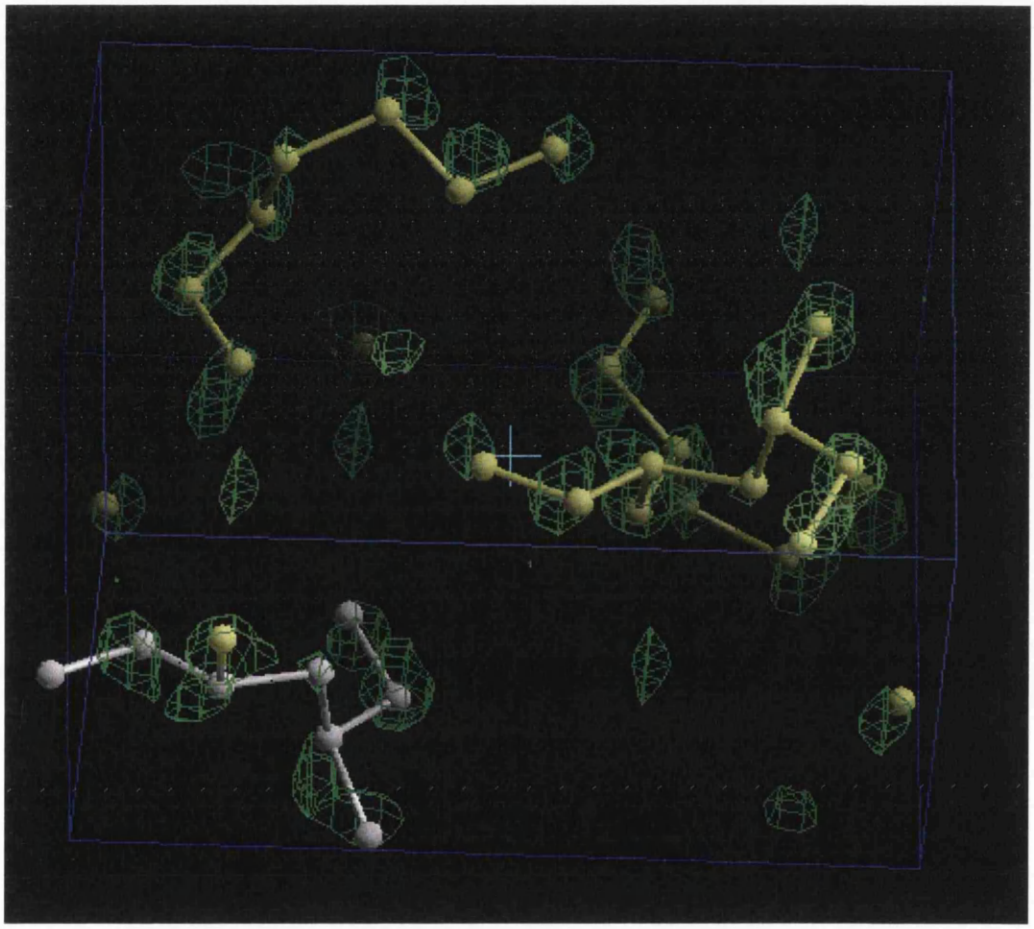
* Best solution

As would be expected the solution of isotatic poly(1-butene) form III has an increasing LLG value at each branch of the tree. The χ^2 value increases at each stage this may seem alarming but is an artefact of the calculations, the more reflections in the basis set {H} there are then the more difficult it can become for the maximum

entropy equations to be satisfied completely, which leads to the fit between observed and calculated U's becoming more inaccurate, hence larger χ^2 values at each stage.

The best solutions was chosen on the grounds of LLG analysis score and the interperetability of the maps in real space. All the maps were of comparable quality in the a-b plane and so the map with the best spatial resolution in the c-axis was accepted as the final solution.

Figure 3-3 Final centroid map for isotactic Poly(1-butene) Form III



3.5 Mannan I

The three-dimensional electron diffraction data set was published by Chanzy [5]. The orthorhombic space group is $P2_12_12_1$ with unit cell dimensions a = 8.92, b = 7.21, c = 10.27Å. The data set has 58 unique reflections and 12 non-hydrogen atoms in the asymmetric unit Z = 1 and chemical formula $C_6O_6H_6$.

There have been two attempts to solve Mannan I: in projection and in three dimensions. The three dimensional solution was attempted and failed to find the

whole structure. So a projection solution was tried in a similar vein to the original determination was achieved by fitting a model to a three dimensional electron density.

3.5.1 Projection solution

The origin was defined as:

h	k	l	phase
1	0	1	90
1	2	0	90
1	1	0	90

via the ORIGIN and KNOW commands. The NEXT command was used to search for suitable reflections for phase permutation with the parameters as so, 0.7 50 16. A Nordström-Robinson code was then used to permute the reflections. Thereafter a p-factor was obtained and MICE was used to maximise the entropy of all the newly generated nodes.

After maximisation analysis was used to find the most likely solution.

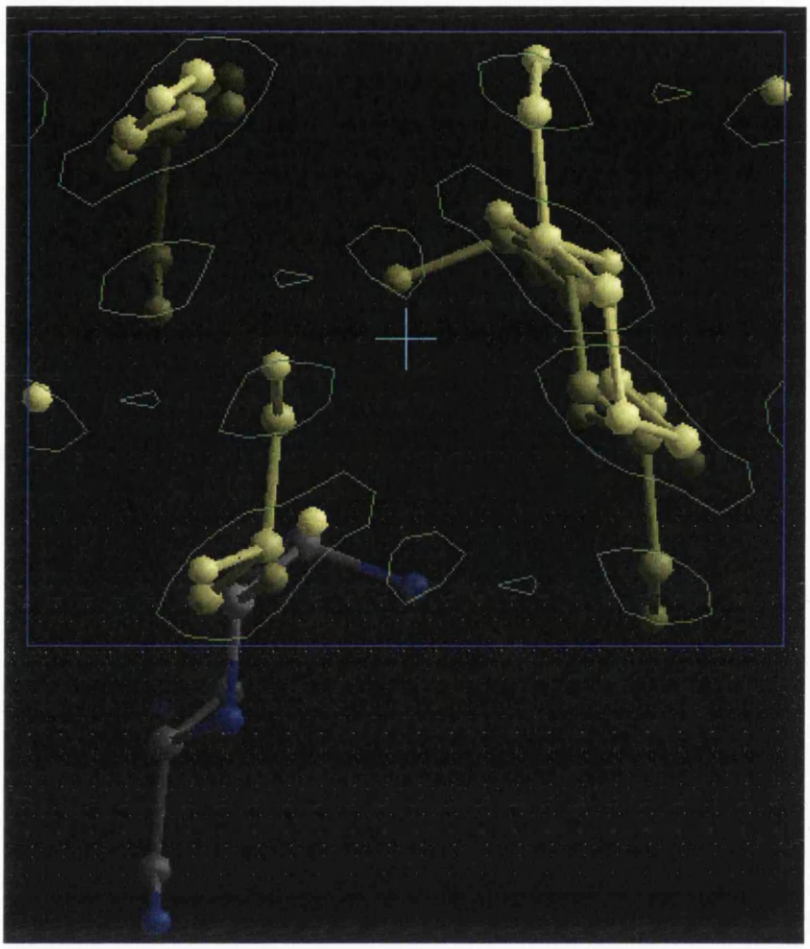
Table 3-4 Mannan I analysis results in projection

Set No.	Log Likelihood Gain	Score
78	3.494	1.000
90	3.494	1.000
42	3.494	1.000
242	3.494	1.000
230	3.494	1.000
22	3.494	1.000*
150	3.494	1.000
170	3.494	1.000

* Best solution

The solutions all converged to the same point in phase space with little or no differences in the LLG and scores of the different solutions. The electron densities all looked similar, with the differences occurring in the contouring level of the final maps. Node 22 was considered the favoured solution because the map was still interpretable at very high contouring levels, unlike some of the other nodes.

Figure 3-4 Structure solution of mannan I in projection centroid map of node 22



3.5.2 Mannan I 3-dimensional structure

Several different approaches have been attempted for solving this structure: three different error-correcting codes, and four different origin definitions, all with no success. So a small selection of reflections was chosen using the NEXT command, and a full factorial design rather than an EEC was implemented.

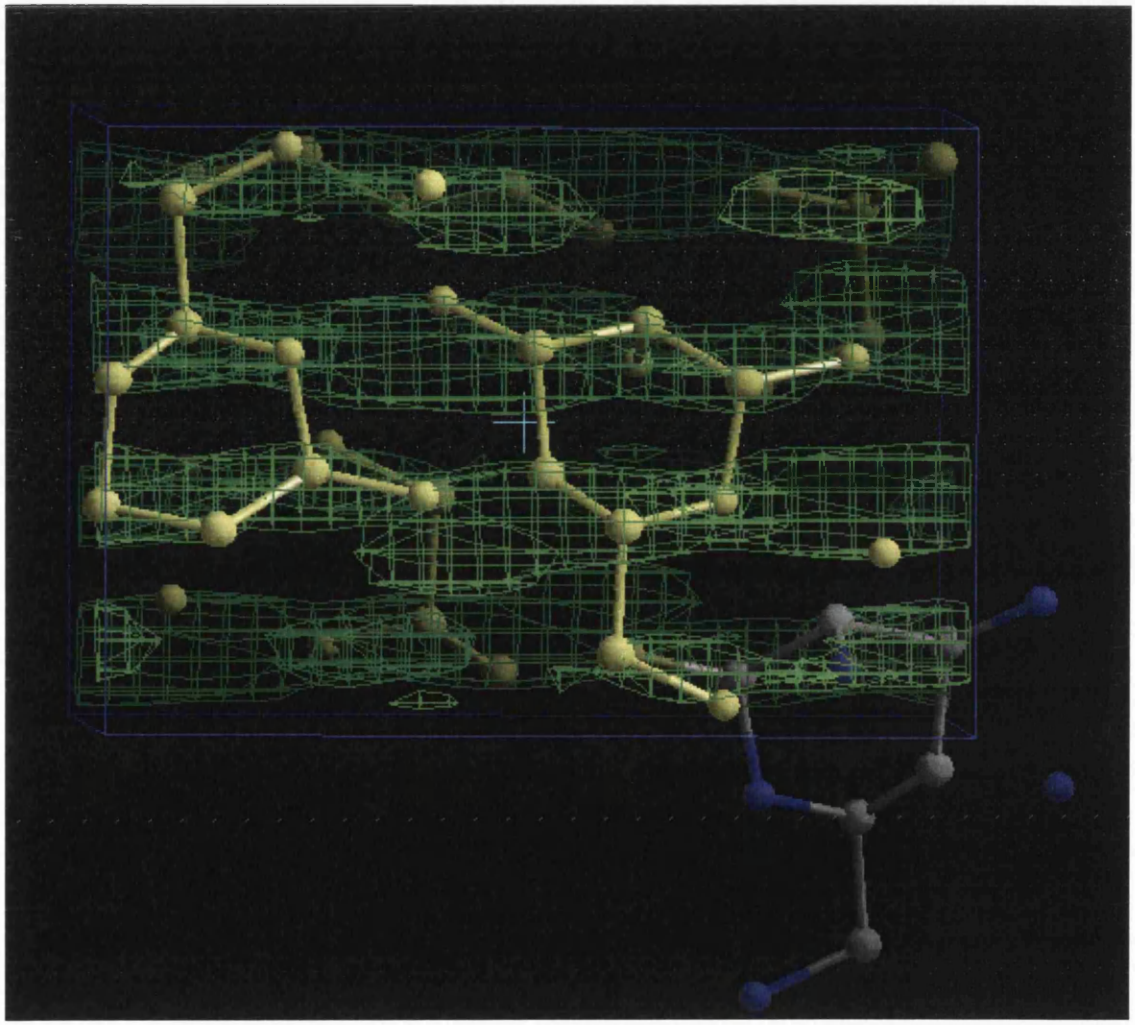
Table 3-5 Mannan I 3D, analysis of results

Set No.	LLG	Score
163	4.251	0.360*
212	3.674	0.311
17	1.832	0.276
61	1.229	0.185
108	1.773	0.150
151	1.518	0.129
242	0.839	0.126
27	0.832	0.125

* Best structure

The three dimensional determination of Mannan I was unsuccessful in terms of direct methods and finding atom positions of the structure. This is possibly unsurprising in that the original solution was achieved using molecular replacement, suggesting that the phase relationships were too weak to even determine a small fragment of the structure, or that there was not enough of reciprocal space sampled to solve the structure even by more robust methods like maximum entropy.

Figure 3-5 Three-dimensional electron density centroid map of mannan I , node 163



3.6 Poly (1,4-trans - cyclohexanediyl dimethylene succinate) (T-cds)

The three-dimensional electron diffraction data set was published by Brisse and co-workers in 1984[5]. The monoclinic space group is $P2_1/n$ with unit cell dimensions a = 6.48, b = 9.48, c = 13.51Å, $\beta = 45.9°$ and Z = 2. The data set has 87 observed reflections. The original determination was achieved by model fitting to the electron density. Because the b-axis was used for tilt purposes, only half the unique intensity data was collected, and a Wilson plot showed that an average temperature factor of B

= 6.1Å^2 was achieved for the collected data. T-cds has the molecular formula $C_{12}O_4H_{12}$.

3.6.1 Structure Solution of Poly (1,4-trans - cyclohexanediyl dimethylene succinate) (T-cds)

The solution of T-cds was obtained by origin definition and two rounds of Nordström-Robinson error-correcting codes. The path to the solution was as follows: origin definition for node one:

h	k	l	phase
2	9	2	0
1	1	0	0
1	10	1	180

This was followed by a Nordström-Robinson code; analysis of the results gave possible nodes with potential solutions. For nodes 898 and 906, which both had positive LLGs and high scores from the analysis. Upon visual inspection they both appeared to have sensible electron densities which could lead to a sensible solution with peaks at reasonable C-C and C-O distances. For each a further Nordström-Robinson codes was applied and the solutions analysed. The final structure solution node was 2725.

Table 3-6 T-cds path to final structure

Node	Maximum Likelihood	Chi Squared
898	1.84	0.99770
906	3.80	0.98244
2725	7.36	0.99202

Table 3-7 T-cds final analysis of results

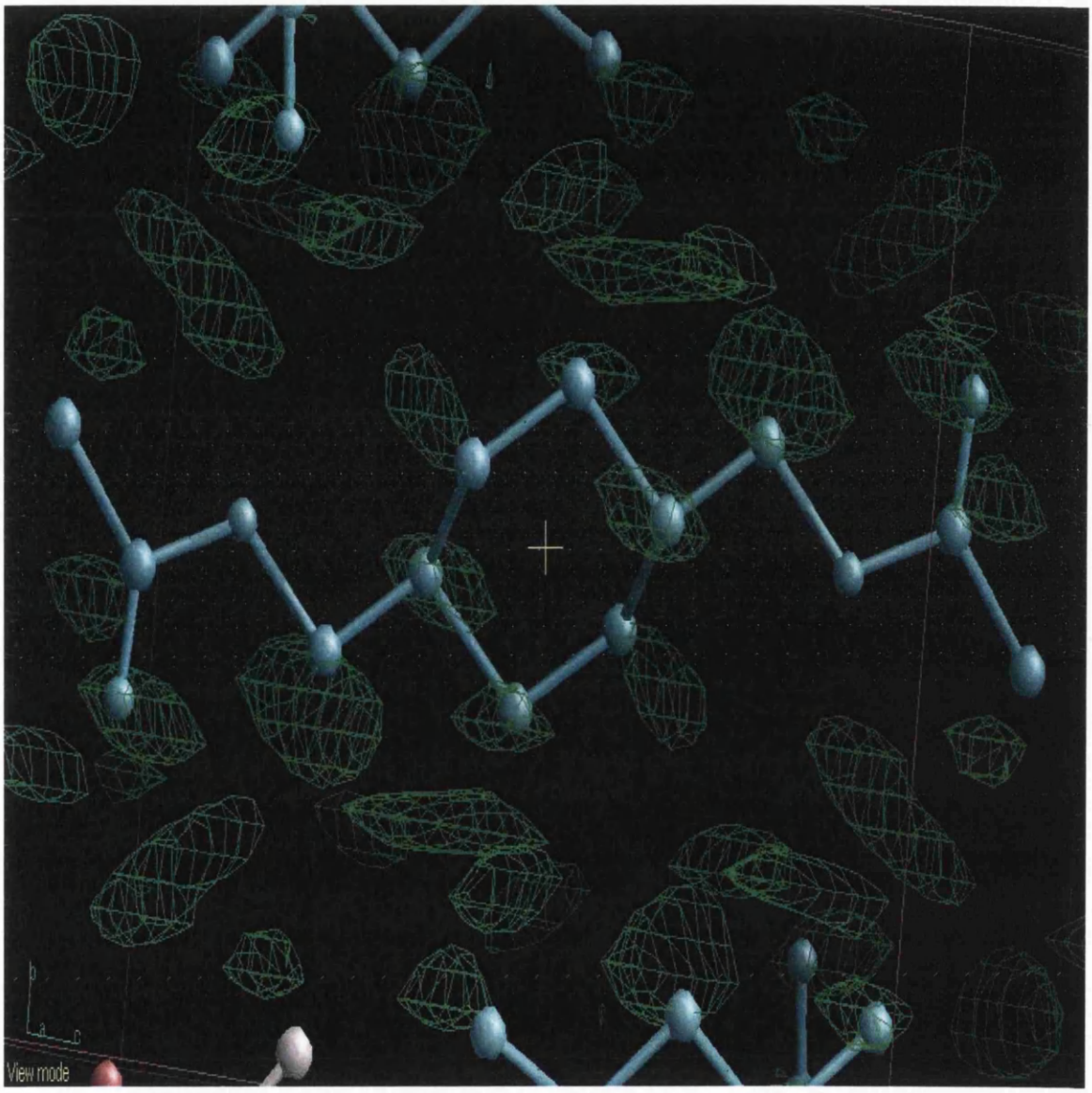
Node	Log Likelihood Gain	Score
2745	7.672	0.741
2741	7.612	0.735
2985	7.576	0.732
2981	7.567	0.731
2729	7.511	0.725
2725	7.358	0.710*
3001	7.284	0.703
2616	7.202	0.695

* Best solution

Node 2725 was chosen because it had all the atoms present in the molecule whereas the other nodes had more well defined atomic positions, but were missing some of the atoms, which were not found upon subsequent calculations. The lack of resolution on

some atomic positions is not at all surprising if only half the data set is present in the calculation.

Figure 3-6 Centroid map of t-cds, node 2725



3.7 Buckminsterfullerene (C_{60})

The three-dimensional electron diffraction data set was published by Dorset and co-workers [6]. The cubic space group is Fm3m with unit cell dimensions a = 14.26Å. The data set has 42 unique reflections and 2 atoms in the asymmetric unit. The original solution was achieved by calculating a series of triplet and positive quartets, although these gave surprising results until a multislice calculation was performed showing that the h00 intensities were too strong and are in fact weak reflections and not strong ones as observed. This evidence was also backed up by disappearance of the strong

reflections as the sample was tilted, again evidence of secondary scattering effects in the h00 reflections.

3.7.1 Structure solution of Buckminsterfullerene

The solution of buckminsterfullerene was achieved relatively easily, with origin definition and one level of Nordström-Robinson code based phase permutation. Upon analysis of the LLGs the final solution was found to be just outside the top eight. This made the solution unusual because, generally, only nodes in the top eight of an analysis are considered credible. Asking a user to look at all the maps generated by a calculation is unrealistic, and so only the top eight are normally examined. There are exceptions to this, especially with electron diffraction data sets. The main reasons for these exceptions are the quality of the data. For example, if dynamical diffraction is suspected to have distorted the data slightly or that the estimated standard deviations are very poor. Then the first thing to do is look slightly outside the top eight suggested nodes. If this is fruitless then rerun the analysis with lower confidence limit, for example 5% rather than the normal 2%.

Table 3-8 Final node for buckminsterfullerene

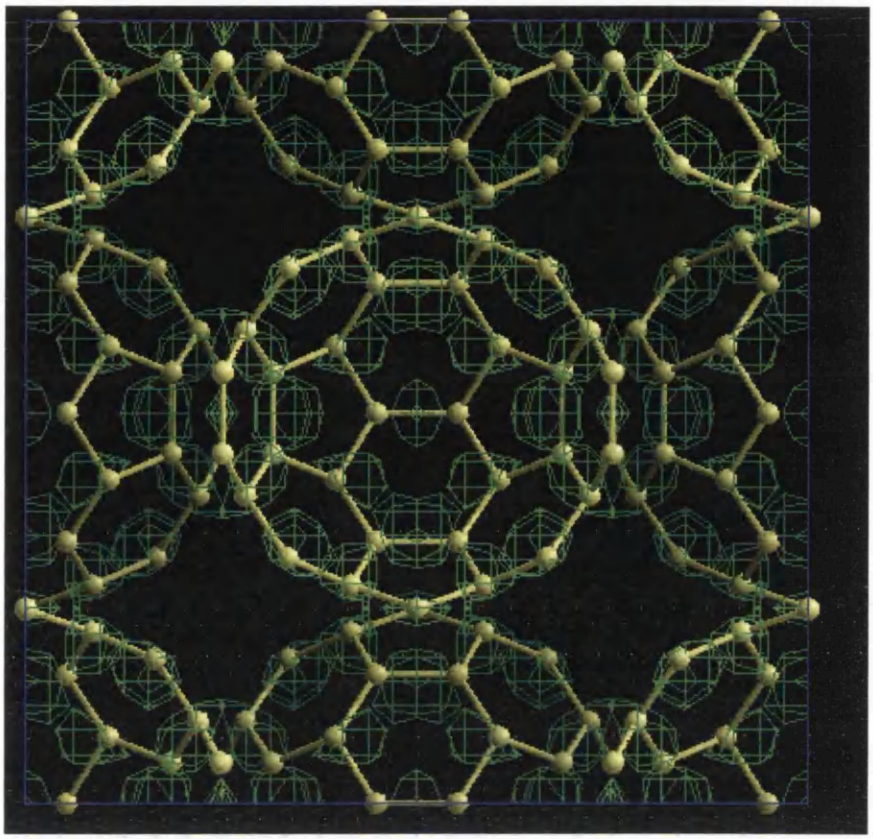
Node	Maximum Likelihood	Chi Squared
65	0.03	0.97625

Table 3-9 Analysis of Buckminsterfullerene

Set No.	LLG	Score
257	12.908	0.444
57	3.249	0.252
18	0.143	0.011
190	0.299	0.010
152	0.130	0.010
63	0.056	0.002
155	0.052	0.002
5	0.052	0.002

Maximum entropy solutions from MICE suffer from the same problems found in the direct methods structure solution, that secondary scattering in the original data set has skewed the results. Although the result produced here does have difficulty locating the atom positions, which indicates a structure having free rotation disorder at room temperature, this is concurrent with the room temperature X-ray solution, even though the traditional method finds a static solution at room temperature due to the large h00 reflections. It would appear that the maximum entropy solution manages to overcome the inaccurate intensities with its more robust statistical approach. Previously there has been another solution of the buckminsterfullerene using maximum entropy[7], but omitting the dynamical reflections.

Figure 3-7 Structure solution of buckminsterfullerene



3.8 Discussion of *ab initio methods in electron diffraction*

From the five structures tested, four have reasonable solutions and the one with no three dimensional solution proved to be easily solvable in projection, and can so be labelled as lacking in sufficient data to be solved in three dimensions.

Brucite, poly (1,4 -trans-cyclohexanediyl dimethylene succinate) and poly(1-butene) From III are all solved with relative ease, requiring no deviation from the standard routine, even though the data sets are far from ideal, like the poly (1,4 -trans-cyclohexanediyl dimethylene succinate) case where only half of the reflections were recorded. The buckminsterfullerene required a relaxation in the analysis process to allow for the sparse data set with a small amount of secondary scattering. The calculation managed to avoid the problems of the original structure solution and compensate for the secondary scattering, leading to a correct solution of a rotationally disordered structure, and not a static structure at room temperature as predicted by conventional direct methods. This demonstrates the robustness of the Bayesian statistical approach that even data sets with a small amount of error within a data set can be tolerated and the correct solution found. This cannot be said for any other form of direct methods.

The most difficult case was the attempt to solve Mannan I in three dimensions, the data set has only 58 reflections with more than half belonging to one projection, the hk0. This explains why the structure proved to be so difficult to solve in three dimensions, there is too little data and therefore detail to work with in the third dimension.

3.9 Conclusions

The maximum entropy direct methods approach as implemented in MICE has been shown to be a robust method for structure solution of electron diffraction data sets. Unlike many of its alternatives, MICE seems able to solve the majority of structures from the given data sets, through a routine methodology. In the literature there is no set method for trying to solve electron diffraction data: potentially MICE could fill this void and become a routine method for solving electron diffraction data sets *ab initio*. Although a better LLG function could potentially be developed to help make the discrimination between the sets more easily distinguishable.

References

(1) Zhukhlistov, A.P.; Zvyagin, B.B.; Avilov, A.S.; Ferraris, D.; Poltnikov, V.P. *Crystallography Reports* **1997**, *42*, 774-777.

(2) Dorset, D.L.; McCourt, M.P.; Kopp, S.; Wittmann, J.C.; Lotz, B. *Acta Crystallographica* **1994**, *B50*, 201-208.

(3) Langs, D.A.; DeTitta, G.T. *Acta Crystallographica* **1975**, *A31*, S16

(4) DeTitta, G.T.; Edmonds, J.W.; Langs, D.A.; Hauptman, H. *Acta Crystallographica* **1975** , *A31*, 472-479.

(5) Chanzy, H.; Pérez, S.; Miller, D.P.; Paradossi, G.; Winter, W.T. *Macromolecules* **1987**, *20*, 2407-2413.

(6)Dorset, D.L.; McCourt, M.P. *Acta Crystallographica* **1994**, *A50*, 344-351.

(7) Gilmore, C.J.; Dong, W.; Bricogne, G. *Acta Crystallographica* **1999**, *A55*, 70-83.

(8) Dorset, D.L. *Structural Electron Crystallography*; Plenum Publishing Coorporation: New York, **1995**.

4 THE MISSING CONE PROBLEM

The missing cone problem arises from the experimental situation in which a full sampling of reciprocal space is not possible. This is caused by the lens pole gap restricting the visibility of the electron beam. This is a serious problem for the electron crystallographer, and can cause a variety of problems in differing situations. Difficulties can arise in space group and unit cell determination. More significantly missing reflections can lead to incorrect structure determination, or ill defined regions of electron density[1], all of which makes the solution of the missing cone an important problem.

One approach to overcoming the difficulties of the missing cone is to grow epitaxial crystals. These crystals are grown on a inorganic or organic substrate with a suitable lattice spacing that causes the crystal to grow in a particular projection. The new crystal projection allows further sampling of reciprocal space. It is not always possible to produce epitaxial crystals because it is a difficult technique to use, and can only be applied to small molecules when the desired lattice parameters match[2].

An alternative approach could be to apply direct techniques such as maximum entropy, which has the advantage of being more stable than other direct methods approaches when using the sparse data sets produced by electron diffraction. Maximum entropy also naturally extrapolates data. Although it is quite different to the methodology of the Sayre equation[3] the basic principles are the same.

The maximum entropy approach to the missing cone problem works by taking a basis set of phased reflections and asking what other electron density can be accurately predicted. The procedure predicts electron density in real space, so a Fourier transform is applied to the entropy maximised map, new reflections and their phases are shown by the calculation, as well as recovering the original reflections and phases.

Some work on this area has been attempted by Dorset and co-workers[4,5] on model data sets generated by simulating electron diffraction data at various tilt angles from the model. They have used the Sayre equation to extrapolate the magnitudes and phases of the missing reflections. The work found that some unusual behaviour in the relationship between the tilt angle and the accuracy of the phase and magnitude prediction. A ±45° the phase prediction was more accurate that from the ±60°, but the missing amplitudes were worse. This is somewhat counter intuitive because the higher

the tilt angle the more of reciprocal space can be sampled and therefore the more reflections can be collected the fewer reflections that have to be predicted. Intuitively this calculation should become more accurate the more information that is available, because there is more information and because there is therefore fewer reflections and phases to predict.

They were also able to confirm the long standing belief that accurate phases are more important than accurate magnitudes, for producing interpretable electron density maps[6].

Although there is a more important need for accurate phases than amplitudes, the problem arises from the fact that we only have the magnitudes to work with. So the magnitudes need to be of a reasonable accuracy to predict accurate phases. Although when phase extending, this is not the same as that of an *ab initio* determination, because the magnitudes used in the phase extensions are not randomly sampled. This implies that there is a high correlation between the high resolution limits of the basis set and the Fourier transform of the electron density.

4.1 Experimental technique

The first step is to generate a set of sigma values for the intensities $\sigma(|F_{\underline{h}}|)$. The normal procedure was to take 10% of the structure factor *i.e.* $\sigma(|F_{\underline{h}}|) = 0.1|F_{\underline{h}}|$, but this could lead to problems in the calculation, because the sigma values were too large. This meant the predicted intensity values allowed in the maximum entropy calculation would vary too much from the observed value. This problem can be fixed by a command in RALF, called SCALE. SCALE multiplies the $\sigma(|F_{\underline{h}}|)$ by a constant, so that the observed and calculated values of the structure factors can be constrained to be more favourably comparable.

The basis set of reflections was normalised in the RALF program and passed forward to the MICE module. In MICE the program was told all the known amplitudes and phases from micrographs to a known tilt angle ±30°, ±45° or ±60°. Next a maximum entropy calculation was performed with these reflections in the basis set and the results were analysed. The analysis process involved looking at how well the extrapolated results compared to the observed intensities and the phases.

Table 4-1 Number of known phases and reflections for each tilt series and compound

Compound	Number of reflections used with known phases		
	±30°	±45°	±60°
Basic Copper Chloride	97	110	114
Polybut-1-ene Form III	114	130	140
Poly (1,4-trans - cyclohexanediyl dimethylene succinate)	67	76	85

The following script is an example of how the program is used to carry out the following calculation.

Figure 4-1 Example script for missing cone calculations

```
#!/bin/csh –f
mice –emolecule.e –zmolecule.z –kknown.reflections –ooutput.mice <<EOF
know f
getp
exec
setp {to value obtained above}
node 1
target 1.0
damp 0.05,0.05,0.04,5
sele 1
debug
level 2
exec
EOF
```

The script first of all obtains a p-factor, the p-factor is calculated, using Brent's method of optimisation. The p-factor is adjusted to find the value of 8.5 for the χ^2 value.This has been found by experienced users of the program, as the value where entropy maximisation should begin. This sets the U^{obs} -values at a reasonable distance from the $|U|^{ME}$ values before entropy maximisation begin, so as to allow the calculation to run for a few cycles before convergence. Next the root node is selected and the χ^2 TARGET is set to 1.0, which then means that the calculation will stop once convergence between the calculated and observed values for the U's are on average within 1_ of each other. The DAMP command damps the entropy maximisation, stopping it from moving too far ahead at any one cycle in the process. The DAMP command has 4 values s, t, r, B, s and t are the maximum trial values

allowed in the line and plane searches, *c.f.* chapter 2.4; r is a user parameter to restrict the distance moved in ω space, usually set to 0.04, this is also an empirical parameter, and B is the temperature factor. SELE selects the node of calculation and the DEBUG and LEVEL 2 commands are to output data in verbose mode, thus allowing comparison of the calculated new reflections and the actual observed reflections. EXEC executes the commands in the preceding block to start the calculation.

The calculations use subsets of real data to simulate a smaller tilting angle than the original data set, thus allowing comparison of the predicted magnitudes and phases with real data.

4.2 Results

4.2.1 Basic Copper Chloride

Basic copper chloride structure was solved by electron diffraction in 1958 by Vainstein and co-workers[7]. The structure crystallised in space group $P2_1/n$ with cell dimensions a = 5.73 b = 6.1, c = 5.63Å, β = 93.75°. The data set contains 120 unique reflections. Z= 4

No $\sigma(|F_{\underline{h}}|)$ values were published with the structure so 10% of the $|F_{\underline{h}}|^{obs}$ was used. In RALF all the following reflections were normalised with their phases to allow the comparison of predicted and known phases later in the calculation. The $\sigma(|F_{\underline{h}}|)$ values were also scaled so as to produce sensible U-magnitudes and associated $\sigma(|U|)$ values.

Table 4-2 Basic copper chloride reflections and phases

H	K	L	F	_F	_	H	K	L	F	_F	_
0	0	1	1.96	0.20	0	1	6	1	3.12	0.31	0
0	0	4	3.73	0.37	0	1	6	3	2.42	0.24	0
0	0	5	0.88	0.09	180	2	0	-3	2.84	0.28	0
0	0	6	3.04	0.30	0	2	0	-2	7.35	0.74	0
0	1	1	3.17	0.32	180	2	0	-1	3.93	0.39	180
0	1	2	1.13	0.11	180	2	0	1	2.64	0.26	0
0	1	3	2.75	0.28	0	2	0	2	3.94	0.39	0
0	2	0	2.74	0.27	180	2	0	3	1.86	0.19	180
0	2	1	6.38	0.64	0	2	0	4	3.30	0.33	0
0	2	2	1.44	0.14	0	2	1	-3	1.10	0.11	180
0	2	3	7.80	0.78	0	2	1	-1	1.45	0.15	0
0	2	5	2.02	0.20	0	2	1	1	1.35	0.14	0
0	2	6	1.27	0.13	0	2	1	2	3.27	0.33	180
0	2	7	1.86	0.19	0	2	1	3	1.65	0.17	180
0	3	1	1.59	0.16	0	2	2	-4	2.06	0.21	180
0	3	2	1.00	0.10	0	2	2	-3	2.02	0.20	0
0	3	3	2.16	0.22	180	2	2	-2	2.05	0.20	0
0	3	4	0.68	0.07	180	2	2	-1	6.35	0.63	0
0	4	0	8.00	0.80	0	2	2	1	5.63	0.56	0
0	4	1	1.47	0.15	0	2	2	2	1.77	0.18	180
0	4	2	2.53	0.25	0	2	2	3	3.09	0.31	0
0	4	3	0.71	0.07	180	2	2	4	1.60	0.16	0
0	4	4	2.50	0.25	0	2	3	-2	1.28	0.13	0
0	4	6	2.32	0.23	0	2	3	-1	2.13	0.21	180
0	5	1	1.49	0.15	180	2	3	0	2.06	0.21	180
0	6	1	2.54	0.25	0	2	3	2	1.62	0.16	0
0	6	3	4.07	0.41	0	2	4	-4	2.51	0.25	0
0	6	5	2.04	0.20	0	2	4	-1	2.09	0.21	180
0	8	1	0.29	0.03	0	2	4	0	2.23	0.22	0
1	0	-6	1.86	0.19	0	2	6	-1	2.32	0.23	0
1	0	-2	10.10	1.01	0	2	6	1	2.32	0.23	0
1	0	1	3.40	0.34	180	2	6	3	1.74	0.17	0
1	0	2	3.73	0.37	0	3	0	-4	2.79	0.28	0
1	0	4	4.65	0.47	0	3	0	-2	6.80	0.68	0
1	1	-5	1.40	0.14	180	3	0	2	2.91	0.29	0
1	1	-3	1.10	0.11	0	3	1	-3	1.34	0.13	0
1	1	-1	3.07	0.31	180	3	1	2	1.04	0.10	180
1	1	0	1.78	0.18	180	3	2	-5	2.04	0.20	0
1	1	1	3.24	0.32	0	3	2	-3	2.47	0.25	0
1	1	5	1.04	0.10	180	3	2	-2	1.77	0.18	0
1	2	-6	0.36	0.04	0	3	2	-1	4.00	0.40	0
1	2	-5	3.04	0.30	0	3	2	1	6.85	0.69	0
1	2	-2	3.66	0.37	180	3	2	3	2.02	0.20	0
1	2	-1	2.46	0.25	0	3	3	-3	1.25	0.12	180
1	2	0	3.76	0.38	0	3	3	-2	1.15	0.12	180

1	2	1	8.51	0.85	0	3	3	-1	1.51	0.15	180
1	2	2	2.17	0.22	180	3	4	-4	1.77	0.18	0
1	2	4	1.31	0.13	0	3	4	-2	2.20	0.22	0
1	2	5	2.03	0.20	0	3	4	0	2.07	0.21	0
1	2	6	0.44	0.04	0	3	6	1	2.32	0.23	0
1	3	1	2.92	0.29	180	4	0	-2	3.11	0.31	0
1	3	2	1.62	0.16	180	4	0	2	4.98	0.50	0
1	3	4	1.38	0.14	0	4	0	4	2.79	0.28	0
1	4	-3	2.14	0.21	0	4	2	-5	1.76	0.18	0
1	4	-2	5.32	0.53	0	4	2	-1	4.50	0.45	0
1	4	0	4.77	0.48	0	4	2	1	3.34	0.33	0
1	4	2	1.27	0.13	0	4	3	3	1.98	0.20	0
1	4	4	3.07	0.31	0	4	4	-4	1.77	0.18	0
1	6	-3	2.67	0.27	0	4	4	2	1.80	0.18	0
1	6	-2	0.75	0.08	180						

The basic copper chloride extrapolated results are shown below. The following tables contain the missing reflections that have been predicted. The tables show the newly calculated maximum entropy values and compare them to the observed values that were actually measured at the higher tilt angle, along with a comparison of the phase values.

4.2.2 ±30° tilt series Copper Chloride results

The following table lists the results from the ±30° tilt set of reflections. Each reflection listed has been predicted by maximum entropy, and was not part of the basis set

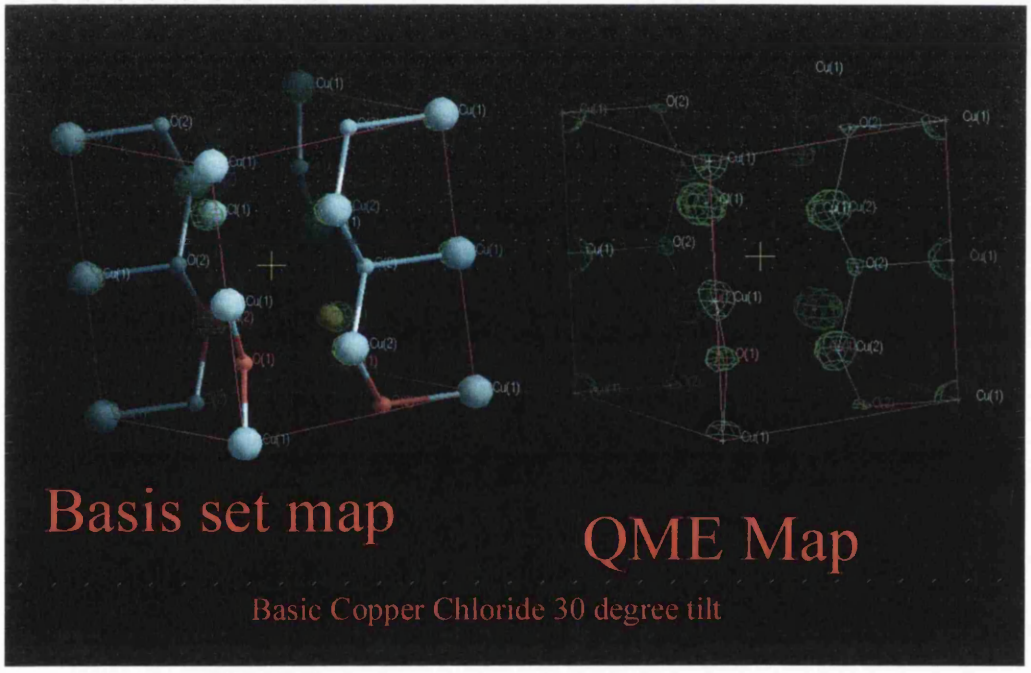
Table 4-3 Extrapolated magnitudes and phases and their observed values, for ±30° tilt series.

H	K	L	U(Obs)	U(ME)	PHASE (Uobs)	PHASE (U(ME))
-2	–2	4	0.132	0.114	180.00	-180.00
-1	-4	3	0.100	0.039	0.00	0.00
-1	-2	5	0.156	0.204	0.00	0.00
-1	-2	6	0.026	0.075	0.00	0.00
-1	-1	3	0.083	0.052	0.00	0.00
-1	-1	5	0.152	0.026	180.00	-180.00
-1	0	2	0.350	0.420	0.00	0.00
-1	0	6	0.134	0.173	0.00	0.00
0	1	1	0.144	0.053	180.00	180.00
0	2	2	0.061	0.053	0.00	0.00
0	3	2	0.083	0.013	0.00	180.00
0	3	3	0.185	0.093	180.00	-180.00
0	4	3	0.035	0.065	180.00	-180.00
0	4	4	0.180	0.141	0.00	0.00
0	6	5	0.137	0.084	0.00	0.00

1	2	1	0.238	0.371	0.00	0.00
1	4	2	0.071	0.109	0.00	0.00
1	6	3	0.140	0.090	0.00	0.00
2	0	1	0.078	0.085	0.00	0.00
2	1	1	0.080	0.082	0.00	0.00
2	6	3	0.110	0.163	0.00	0.00
4	0	2	0.309	0.278	0.00	0.00

The mean absolute phase error of these newly predicted reflections is 0°. That is a remarkable result, and the magnitudes of the reflections are on average well within reasonable limits, many being close to or on the exact value of the actual measured value. In the worst cases observed and calculated magnitudes differ by a factor of 2.

Figure 4-2 Comparison of ±30° tilt series maps before and after the calculation



4.2.3 ±45° tilt series Copper Chloride results

The following calculation uses reflections and phases from a tilt series up to and including 45° tilt angle. The calculation attempts to predict the magnitude and phase for the missing reflections that are missing from the tilt series and also those from higher tilt angles that have not been observed.

The following table lists the results from the ±45° tilt set of reflections. Each reflection listed has been predicted by maximum entropy, and was not part of the

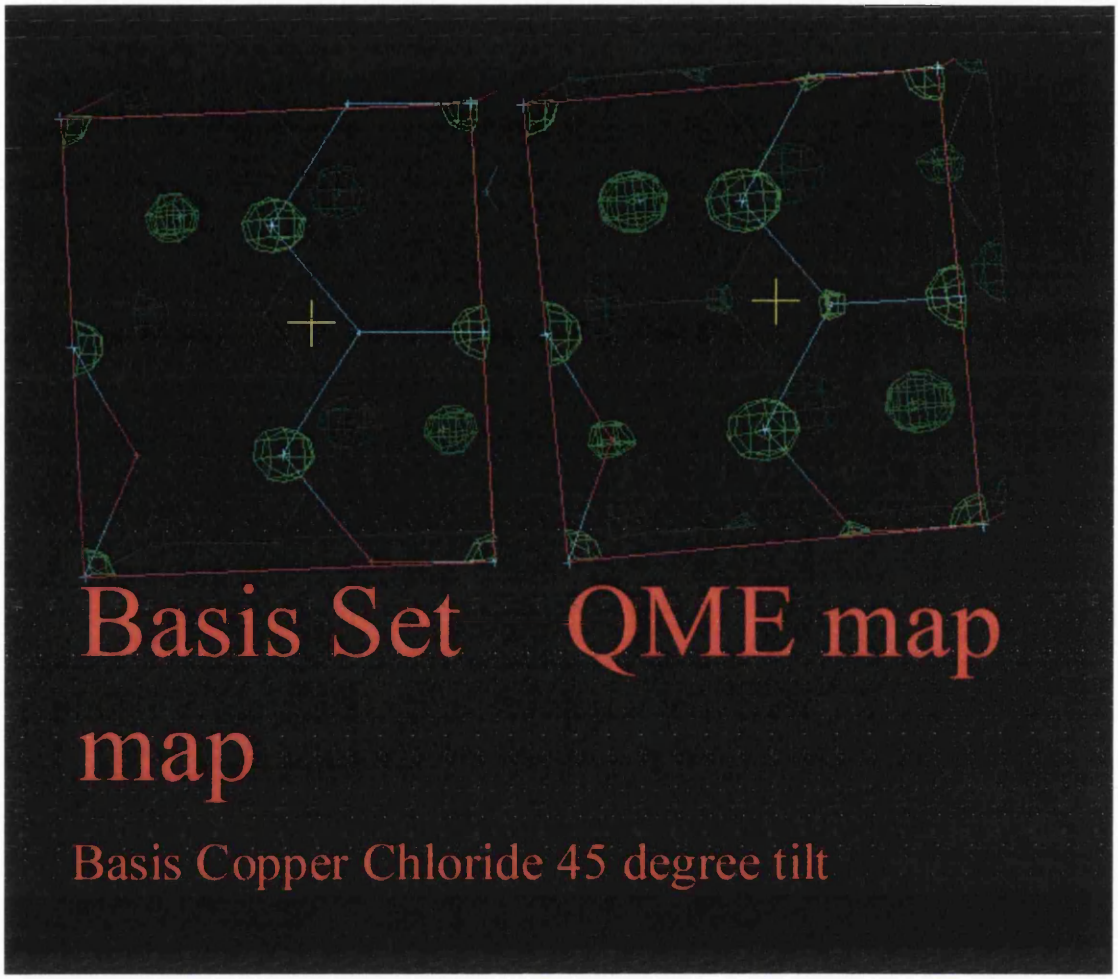
basis set, although the results are compared to the observed values from higher tilt angles.

Table 4-4 Extrapolated magnitudes and phases and their observed values, for ±45° tilt series

H	K	L	U(Obs)	U(ME)	PHASE(Uobs)	PHASE (U(ME))
-1	-2	5	0.156	0.186	0.00	0.00
-1	-2	6	0.026	0.056	0.00	0.00
-1	-1	5	0.152	0.032	180.00	180.00
-1	0	6	0.134	0.164	0.00	0.00
0	1	1	0.144	0.044	180.00	-180.00
0	2	2	0.061	0.045	0.00	0.00
0	3	3	0.185	0.079	180.00	180.00
0	4	4	0.180	0.132	0.00	0.00
0	6	5	0.137	0.072	0.00	0.00

The results show a mean absolute phase error of 0° when compared to the values obtained from their structure solution by direct methods. The precision of the predicted magnitudes has not improved or deteriorated with the increased information available.

Figure 4-3 Comparison of ±45° tilt series maps before and after the calculation



4.2.4 ±60° tilt series Copper Chloride results

The following table lists the results from the ±60° tilt set of reflections. Each reflection listed has been predicted by maximum entropy, and was not part of the basis set.

Table 4-5 Extrapolated magnitudes and phases and their observed values, for ±60° tilt series

H	K	L	U(Obs)	U(ME)	PHASE(Uobs)	PHASE (U(ME))
-1	0	6	0.134	0.159	0.00	0.00
0	1	1	0.144	0.042	180.00	-180.00
0	2	2	0.061	0.032	0.00	0.00
0	3	3	0.185	0.091	180.00	180.00
0	4	4	0.180	0.139	0.00	0.00

The results show a mean absolute phase error of 0° when compared to the values obtained from their structure solution by direct methods. The accuracy of the predicted magnitudes has not improved or deteriorated with the increased information available.

Figure 4-4 Comparison of ±60° tilt series maps before and after the calculation

4.2.5 Isotactic Poly(1-butene) form III

The data set for Isotactic Poly(1-butene) Form III was obtained by electron diffraction by Dorset and co-workers in 1994[8]. They obtained the data set from solution and epitaxially grown crystals. A tilt series was used collect as much three-dimensional data as possible. The orthorhombic space group is $P2_12_12_1$ with unit cell dimensions a = 12.38, b = 8.88, c = 7.56Å. The data set has 125 unique reflections. Z= 4

No $\sigma\left(\left|F_{\underline{h}}\right|\right)$ values were published with the structure so 10% of the $\left|F_{\underline{h}}\right|^{obs}$ was used. In RALF all the following reflections were normalised with their phases to allow the

comparison of predicted and known phases later in the calculation. The $\sigma(|F_{\underline{h}}|)$ values were also scaled so as to produce sensible U-magnitudes and associated $\sigma(|U|)$ values.

Table 4-6 Isotactic Poly(1-butene) form III reflections and phases

H	H	L	F	_F	_	H	K	L	F	_F	_
2	0	0	14.22	1.42	180	1	1	1	19.53	1.95	239
4	0	0	9.28	0.93	180	2	1	1	6.63	0.66	306
6	0	0	4.10	0.41	0	3	1	1	3.74	0.37	186
8	0	0	5.30	0.53	180	4	1	1	7.96	0.80	231
10	0	0	3.98	0.40	180	5	1	1	0.84	0.08	272
1	1	0	17.12	1.71	90	6	1	1	5.91	0.59	47
2	1	0	9.88	0.99	0	7	1	1	3.50	0.35	134
3	1	0	7.84	0.78	270	0	2	1	15.19	1.52	180
4	1	0	9.52	0.95	180	1	2	1	16.27	1.63	1
5	1	0	9.16	0.92	270	2	2	1	4.22	0.42	319
6	1	0	0.84	0.08	0	3	2	1	4.10	0.41	115
7	1	0	2.65	0.27	270	4	2	1	12.78	1.28	122
8	1	0	3.13	0.31	180	5	2	1	3.74	0.37	166
9	1	0	0.84	0.08	270	6	2	1	5.06	0.51	32
10	1	0	4.22	0.42	0	7	2	1	0.84	0.08	128
0	2	0	9.40	0.94	0	8	2	1	3.25	0.33	294
1	2	0	17.00	1.70	90	0	3	1	5.30	0.53	270
2	2	0	6.51	0.65	180	1	3	1	0.84	0.08	285
3	2	0	0.84	0.08	270	2	3	1	4.70	0.47	333
4	2	0	4.34	0.43	180	3	3	1	10.49	1.05	54
5	2	0	6.27	0.63	270	4	3	1	9.16	0.92	351
6	2	0	11.81	1.18	0	5	3	1	2.77	0.28	150
7	2	0	4.46	0.45	90	0	4	1	3.86	0.39	0
8	2	0	2.89	0.29	0	1	4	1	6.39	0.64	334
9	2	0	6.03	0.60	270	2	4	1	6.27	0.63	238
10	2	0	3.86	0.39	180	3	4	1	3.86	0.39	208
1	3	0	7.47	0.75	270	4	4	1	0.84	0.08	236
2	3	0	4.10	0.41	0	5	4	1	10.97	1.10	158
3	3	0	5.79	0.58	90	6	4	1	2.17	0.22	177
4	3	0	3.98	0.40	0	7	4	1	2.29	0.23	310
5	3	0	12.54	1.25	270	1	5	1	3.98	0.40	112
6	3	0	9.64	0.96	0	2	5	1	2.77	0.28	184
7	3	0	2.53	0.25	90	3	5	1	3.25	0.33	323
8	3	0	2.05	0.20	180	4	5	1	3.62	0.36	150
9	3	0	3.13	0.31	90	5	5	1	0.84	0.08	78
0	4	0	6.99	0.70	180	6	5	1	0.84	0.08	236
1	4	0	6.99	0.70	90	7	5	1	3.25	0.33	1
2	4	0	9.64	0.96	0	1	6	1	0.84	0.08	1
3	4	0	7.72	0.77	270	2	6	1	4.46	0.45	176
4	4	0	10.85	1.09	180	3	6	1	0.84	0.08	17
5	4	0	8.56	0.86	90	4	6	1	3.25	0.33	297
6	4	0	0.84	0.08	0	1	7	1	0.84	0.08	25
7	4	0	3.38	0.34	90	2	7	1	3.74	0.37	356

8	4	0	3.50	0.35	0	3	7	1	0.84	0.08	295
9	4	0	3.13	0.31	270	4	7	1	4.94	0.49	131
1	5	0	0.84	0.08	90	1	0	2	0.84	0.08	180
2	5	0	2.77	0.28	180	2	0	2	0.84	0.08	0
3	5	0	6.27	0.63	90	3	0	2	0.84	0.08	180
4	5	0	7.11	0.71	0	4	0	2	3.38	0.34	180
0	6	0	5.42	0.54	0	5	0	2	4.82	0.48	0
1	6	0	5.06	0.51	90	0	1	2	11.09	1.11	270
2	6	0	0.84	0.08	180	1	1	2	13.26	1.33	336
3	6	0	5.06	0.51	270	2	1	2	0.84	0.08	266
4	6	0	3.38	0.34	0	3	1	2	0.84	0.08	161
5	6	0	3.86	0.39	90	4	1	2	2.65	0.27	292
1	7	0	4.70	0.47	90	5	1	2	3.62	0.36	147
2	7	0	7.84	0.78	180	0	2	2	4.82	0.48	0
0	8	0	2.65	0.27	180	1	2	2	3.86	0.39	104
1	0	1	0.84	0.08	90	2	2	2	4.70	0.47	135
2	0	1	28.93	2.89	270	3	2	2	4.94	0.49	248
3	0	1	17.12	1.71	90	4	2	2	3.98	0.40	145
4	0	1	11.57	1.16	90	5	2	2	4.46	0.45	138
5	0	1	7.72	0.77	270	0	3	2	7.35	0.74	90
6	4	2	2.77	0.28	261	1	3	2	0.84	0.08	33
7	4	2	2.65	0.27	28	2	3	2	0.84	0.08	38
1	1	3	10.25	1.02	315	3	3	2	0.84	0.08	276
2	2	3	4.22	0.42	229	4	3	2	4.10	0.41	194
0	3	3	4.22	0.42	90	0	4	2	3.01	0.30	180
1	3	3	3.86	0.39	251	1	4	2	2.77	0.28	144
2	3	3	3.98	0.40	316	2	4	2	4.58	0.46	256
3	3	3	3.38	0.34	118	3	4	2	3.25	0.33	297
0	0	4	19.29	1.93	180	4	4	2	3.50	0.35	282
1	1	4	5.91	0.59	279	5	4	2	0.84	0.08	119

4.2.6 ±30° tilt series Isotactic Poly(1-butene) results

The following table lists the results from the ±30° tilt set of reflections. Each reflection listed has been predicted by maximum entropy, and was not part of the basis set.

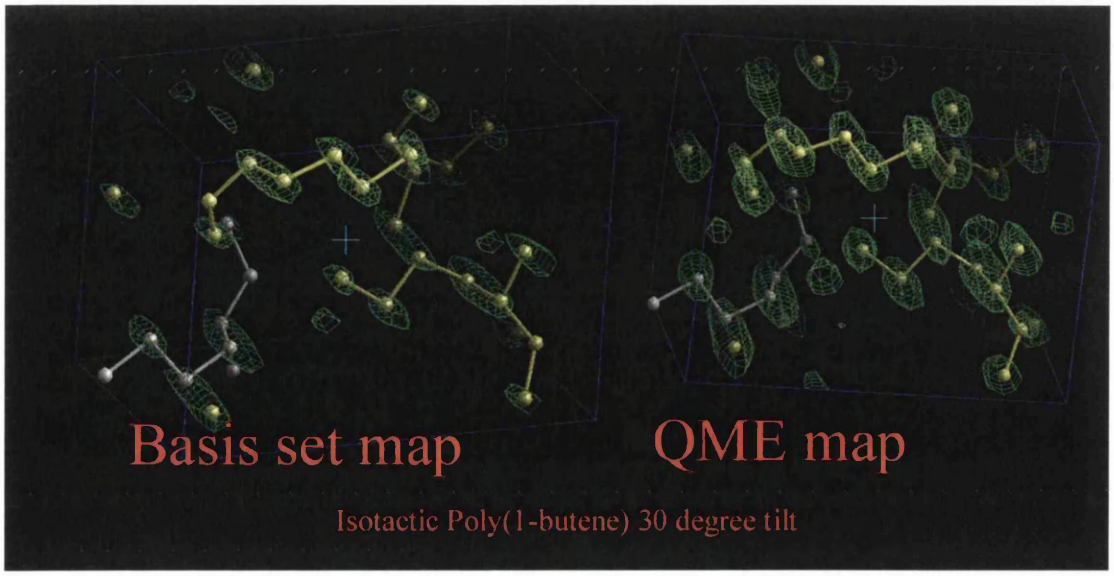
Table 4-7 Extrapolated magnitudes and phases and their observed values, for ±30° tilt series

H	K	L	U(Obs)	U(ME)	PHASE(Uobs)	PHASE (U(ME))
0	0	4	0.319	0.012	180.00	-180.00
0	1	2	0.111	0.054	-90.00	-90.00
0	2	2	0.055	0.057	0.00	0.00
0	3	2	0.096	0.050	90.00	90.00
0	3	3	0.068	0.024	90.00	90.00
1	0	1	0.006	0.023	90.00	90.00
1	0	2	0.008	0.003	180.00	180.00

1	1	1	0.190	0.099	-120.00	-113.47
1	1	2	0.136	0.065	-30.00	-20.72
1	1	3	0.149	0.024	-45.00	-21.71
1	1	4	0.097	0.009	-75.00	-120.19
1	2	2	0.047	0.034	105.00	128.53
1	3	2	0.011	0.024	30.00	66.00
1	3	3	0.069	0.009	-105.00	-129.38
2	0	1	0.220	0.158	-90.00	-90.00
2	0	2	0.009	0.020	0.00	-180.00
2	1	1	0.062	0.040	-60.00	-45.00
2	1	2	0.009	0.016	-90.00	-47.12
2	2	2	0.058	0.014	135.00	157.11
2	2	3	0.053	0.038	-135.00	-114.85
2	3	2	0.011	0.005	45.00	-37.99
2	3	3	0.067	0.014	-45.00	-63.96
3	0	2	0.010	0.021	180.00	-180.00
3	1	2	0.010	0.017	165.00	55.61
3	2	2	0.068	0.047	-105.00	-125.84
3	3	2	0.012	0.024	-90.00	-8.05
3	3	3	0.066	0.016	120.00	84.88
4	0	2	0.045	0.024	180.00	180.00
4	1	2	0.034	0.034	-75.00	-72.45
4	2	2	0.059	0.041	150.00	122.39
5	0	2	0.075	0.040	0.00	0.00
5	1	2	0.054	0.032	150.00	174.00

The prediction of magnitudes and phases is excellent, with the majority of magnitudes again being roughly within a factor of 2 to the observed value. The mean absolute phase value for the predicted phases is 5.6°

Figure 4-5 Comparison of ±30° tilt series maps before and after the calculation



4.2.7 ±45° tilt series Isotactic Poly(1-butene) results

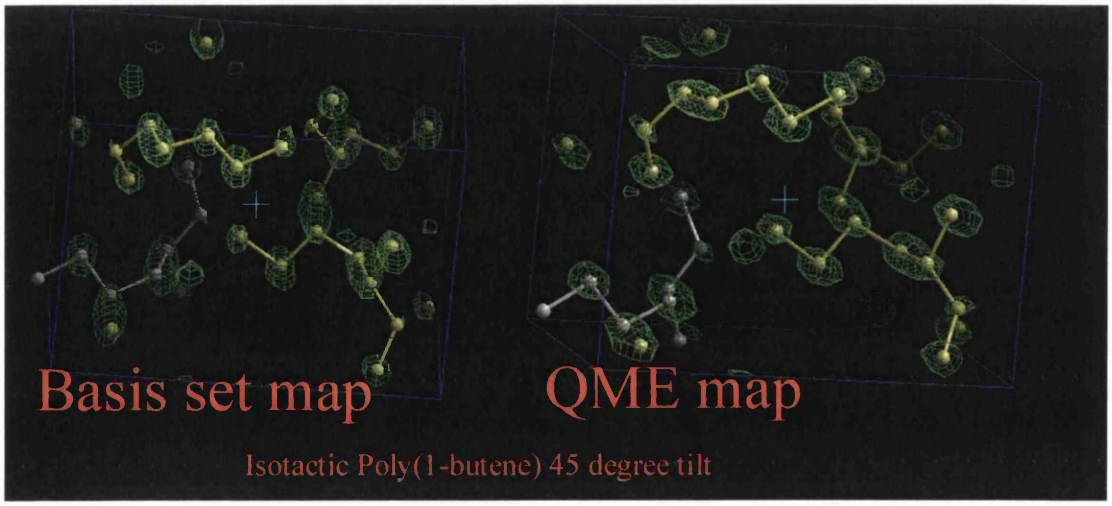
The following table lists the results from the ±45° tilt set of reflections. Each reflection listed has been predicted by maximum entropy, and was not part of the basis set.

Table 4-8 Extrapolated magnitudes and phases and their observed values, for ±45° tilt series

H	K	L	U(Obs)	U(ME)	PHASE(Uobs)	PHASE (U(ME))
0	0	4	0.319	0.034	180.00	-180.00
0	1	2	0.111	0.049	-90.00	-90.00
0	2	2	0.055	0.056	0.00	0.00
0	3	3	0.068	0.028	90.00	90.00
1	0	1	0.006	0.008	90.00	90.01
1	0	2	0.008	0.002	180.00	-179.99
1	1	2	0.136	0.074	-30.00	-15.60
1	1	3	0.149	0.042	-45.00	-36.46
1	1	4	0.097	0.020	-75.00	-87.12
1	2	2	0.047	0.041	105.00	132.40
1	3	3	0.069	0.018	-105.00	-133.48
2	0	2	0.009	0.009	0.00	179.99
2	1	2	0.009	0.018	-90.00	-29.74
2	2	3	0.053	0.040	-135.00	-105.28
2	3	3	0.067	0.022	-45.00	-47.31
3	0	2	0.010	0.019	180.00	-180.00

The prediction of magnitudes and phases is excellent, with the majority of magnitudes again being roughly within a factor of 2 to the observed value. The mean absolute phase value for the predicted phases is 1.6°, an even better phase error with the increased number of known phases.

Figure 4-6 Comparison of ±45° tilt series maps before and after the calculation



The results from the ±60° tilt series yield no new reflections within the given hkl range.

4.2.8 Data from Poly(1,4,trans-cyclohexanediyl dimethylene succinate) (T-cds).

The three-dimensional electron diffraction data set was published by Brisse and co-workers in 1984[9]. The monoclinic space group is $P2_1/n$ with unit cell dimensions a = 6.48, b = 9.48, c = 13.51Å, β = 45.9° The data set has 87 observed reflections. Z=104

No $\sigma(|F_{\underline{h}}|)$ values were published with the structure so 10% of the $|F_{\underline{h}}|^{obs}$ was used. In RALF all the following reflections were normalised with their phases to allow the comparison of predicted and known phases later in the calculation. The $\sigma(|F_{\underline{h}}|)$ values were also scaled so as to produce sensible U-magnitudes and associated $\sigma(|U|)$ values.

Table 4-9 Poly(1,4,trans-cyclohexanediyl dimethylene succinate) (T-cds) magnitudes and phases

H	K	L	F	_F	_	H	K	L	F	_F	_
2	0	0	20.08	2.01	0	1	8	2	2.16	0.22	0
1	1	0	43.25	4.33	0	2	9	2	2.78	0.28	0
2	1	0	5.87	0.59	180	3	0	3	3.71	0.37	180
3	1	0	4.33	0.43	0	3	1	3	3.40	0.34	0
0	2	0	24.72	2.47	0	3	2	3	1.24	0.12	180
1	2	0	9.89	0.99	180	3	3	3	1.24	0.12	180
2	2	0	6.80	0.68	0	3	4	3	1.54	0.15	0
3	2	0	3.40	0.34	180	3	5	3	2.16	0.22	0
1	3	0	6.80	0.68	0	3	9	3	2.78	0.28	180
2	3	0	6.49	0.65	180	2	0	4	7.72	0.77	180
3	3	0	1.85	0.19	0	2	1	4	4.02	0.40	180
0	4	0	6.18	0.62	0	4	1	4	2.16	0.22	180
1	4	0	4.94	0.49	180	2	2	4	4.33	0.43	180
2	4	0	1.24	0.12	0	2	3	4	3.09	0.31	0
3	4	0	1.85	0.19	180	4	3	4	2.47	0.25	0
1	5	0	2.16	0.22	0	2	4	4	1.54	0.15	180
2	5	0	1.54	0.15	180	4	4	4	2.78	0.28	0
0	6	0	3.09	0.31	0	2	5	4	5.25	0.53	0
1	6	0	1.24	0.12	180	4	5	4	4.94	0.49	0
0	8	0	1.54	0.15	180	2	6	4	3.09	0.31	0
1	0	1	6.18	0.62	0	4	6	4	1.54	0.15	0
1	1	1	7.72	0.77	0	2	7	4	2.78	0.28	0
1	2	1	0.93	0.09	0	4	7	4	2.78	0.28	0
1	3	1	4.33	0.43	0	2	8	4	3.09	0.31	0
1	4	1	4.33	0.43	180	3	0	5	12.05	1.21	0
1	5	1	2.16	0.22	0	5	0	5	3.71	0.37	0
1	6	1	2.16	0.22	0	3	1	5	3.09	0.31	180
1	8	1	1.54	0.15	0	3	2	5	5.87	0.59	180

2	0	2	3.71	0.37	0	5	2	5	1.45	0.15	0
1	1	2	7.72	0.77	0	3	4	5	3.71	0.37	180
2	1	2	5.87	0.59	180	3	5	5	2.16	0.22	0
1	2	2	5.25	0.53	180	5	5	5	1.54	0.15	0
2	2	2	4.94	0.49	180	3	1	6	4.33	0.43	180
1	3	2	2.78	0.28	180	3	2	6	5.87	0.59	0
2	3	2	2.16	0.22	0	3	3	6	8.34	0.83	180
1	4	2	1.54	0.15	0	3	4	6	1.54	0.15	0
2	4	2	2.78	0.28	0	3	5	6	3.71	0.37	180
1	5	2	2.16	0.22	0	3	7	6	2.16	0.22	0
2	5	2	0.93	0.09	0	4	1	7	1.24	0.12	0
1	6	2	3.09	0.31	180	4	2	7	2.16	0.22	180
2	6	2	2.78	0.28	0	4	5	7	2.16	0.22	0
2	7	2	0.93	0.09	180	4	6	7	1.85	0.19	180
4	0	8	3.09	0.31	0						

4.2.9 Poly (1,4-trans –cyclohexanediyl dimethylene succinate) ±30° tilt results

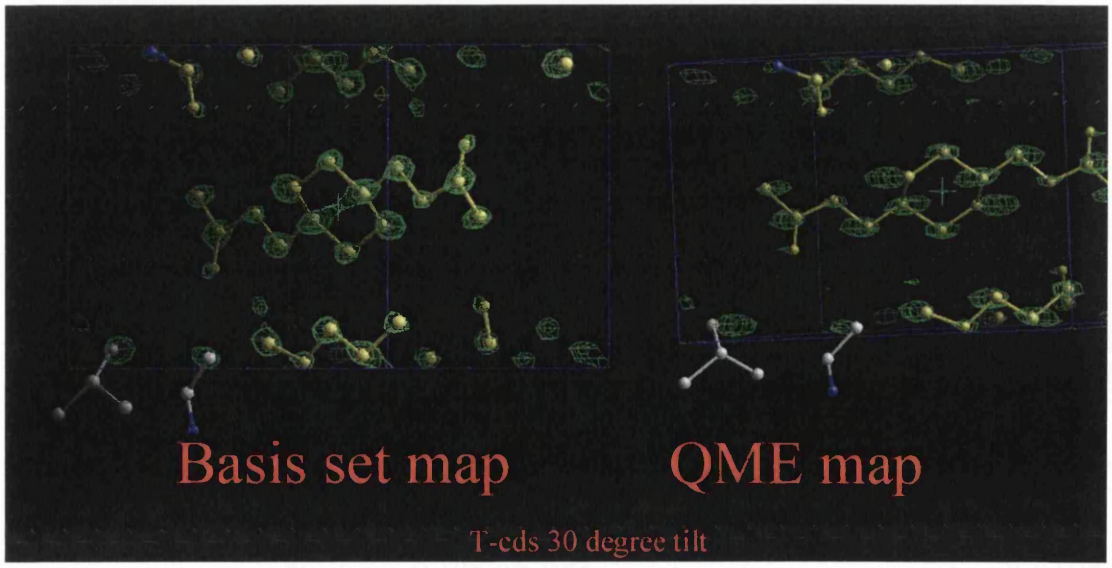
The following table lists the results from the ±30° tilt set of reflections. Each reflection listed has been predicted by maximum entropy, and was not part of the basis set.

Table 4-10 Extrapolated magnitudes and phases and their observed values, for ±30° tilt series

H	K	L	U(Obs)	U(ME)	PHASE(Uobs)	PHASE (U(ME))
0	2	0	0.520	0.156	0.00	0.00
0	4	0	0.189	0.088	0.00	180.00
0	6	0	0.145	0.035	0.00	0.00
0	8	0	0.110	0.088	180.00	180.00
1	3	0	0.104	0.053	0.00	180.00
1	4	0	0.194	0.047	180.00	180.00
1	4	1	0.182	0.121	180.00	180.00
1	5	0	0.049	0.025	0.00	0.00
1	5	1	0.105	0.035	0.00	180.00
1	5	2	0.046	0.130	0.00	0.00
1	6	0	0.073	0.028	180.00	0.00
1	6	1	0.137	0.022	0.00	0.00
1	6	2	0.175	0.062	180.00	180.00
1	8	1	0.147	0.052	0.00	0.00
1	8	2	0.184	0.124	0.00	0.00
2	5	0	0.076	0.006	180.00	180.00
2	7	2	0.059	0.062	180.00	180.00
2	9	2	0.256	0.122	0.00	0.00

The prediction of magnitudes and phases is excellent, with the majority of magnitudes again being roughly within a factor of 2 to the observed value. The mean absolute phase value for the predicted phases is 40°.

Figure 4-7 Comparison of ±30° tilt series maps before and after the calculation



4.2.10 Poly (1,4-trans –cyclohexanediyl dimethylene succinate) ±45° tilt results

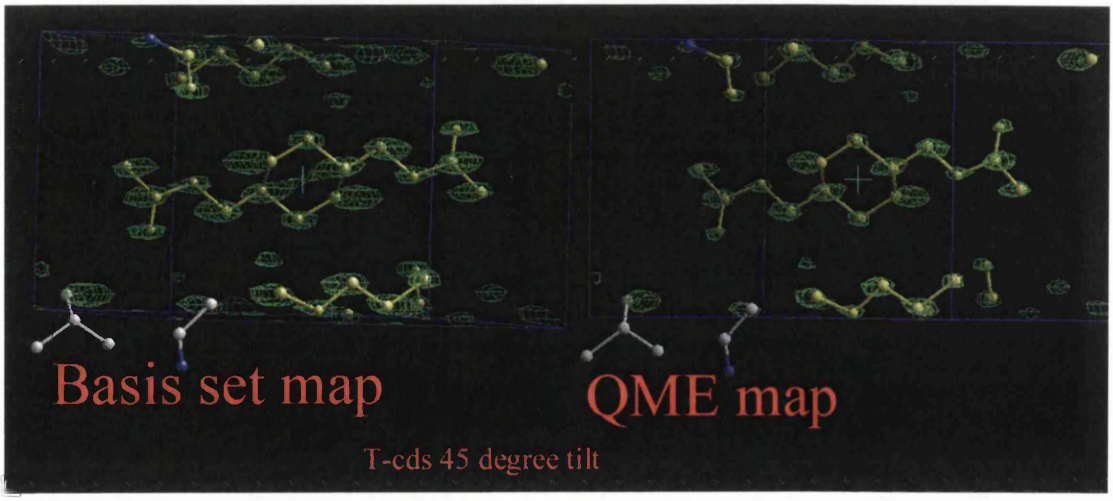
The following table lists the results from the ±45° tilt set of reflections. Each reflection listed has been predicted by maximum entropy, and was not part of the basis set.

Table 4-11 Extrapolated magnitudes and phases and their observed values, for ±45° tilt series

H	K	L	U(Obs)	U(ME)	PHASE(Uobs)	PHASE (U(ME))
0	2	0	0.520	0.218	0.00	0.00
0	4	0	0.189	0.039	0.00	180.00
0	6	0	0.145	0.003	0.00	0.00
0	8	0	0.110	0.128	180.00	180.00
1	5	0	0.049	0.073	0.00	0.00
1	6	0	0.073	0.017	180.00	0.00
1	6	1	0.137	0.002	0.00	0.00
1	8	1	0.147	0.082	0.00	0.00
1	8	2	0.184	0.142	0.00	0.00

The prediction of magnitudes and phases is excellent, with the majority of magnitudes again being roughly within a factor of 2 to the observed value. The mean absolute phase value for the predicted phases is 40°.

Figure 4-8 Comparison of ±45° tilt series maps before and after the calculation



At the ±60°tilt angle for T-cds, there were no new reflections predicted within the given hkl range.

4.2.11 Beyond the observed hkl range

For the molecules looked at so far, we have only considered the reflections that have been predicted where we have observed magnitudes for comparison purpose. The real strength of this technique is that it can predict reflections, not only reflections that are in the same hkl ranges as the observed reflections, but beyond these ranges as well.

In the case where the data is complete to a given tilt angle for example: ±60° tilt angles for the poly(1-butene) and T-cds data, no further extrapolation could occur, within the given hkl range. We can go beyond this using the SUPER command in MICE, which allows the program extrapolated reflections beyond the measured hkl range and present them as usable data within the maps. This can also be used for phase extension. The critical question here is how reliable are the magnitudes and phases within the extended range. Judging from the extrapolation observed above, it should be possible to include a slightly higher resolution shell of reflections, with a high degree of accuracy. The magnitudes should be within a factor of 2 of their real values, and the phase accuracy within the acceptable ±40° mean absolute phase error, for both acentric and centric reflections.

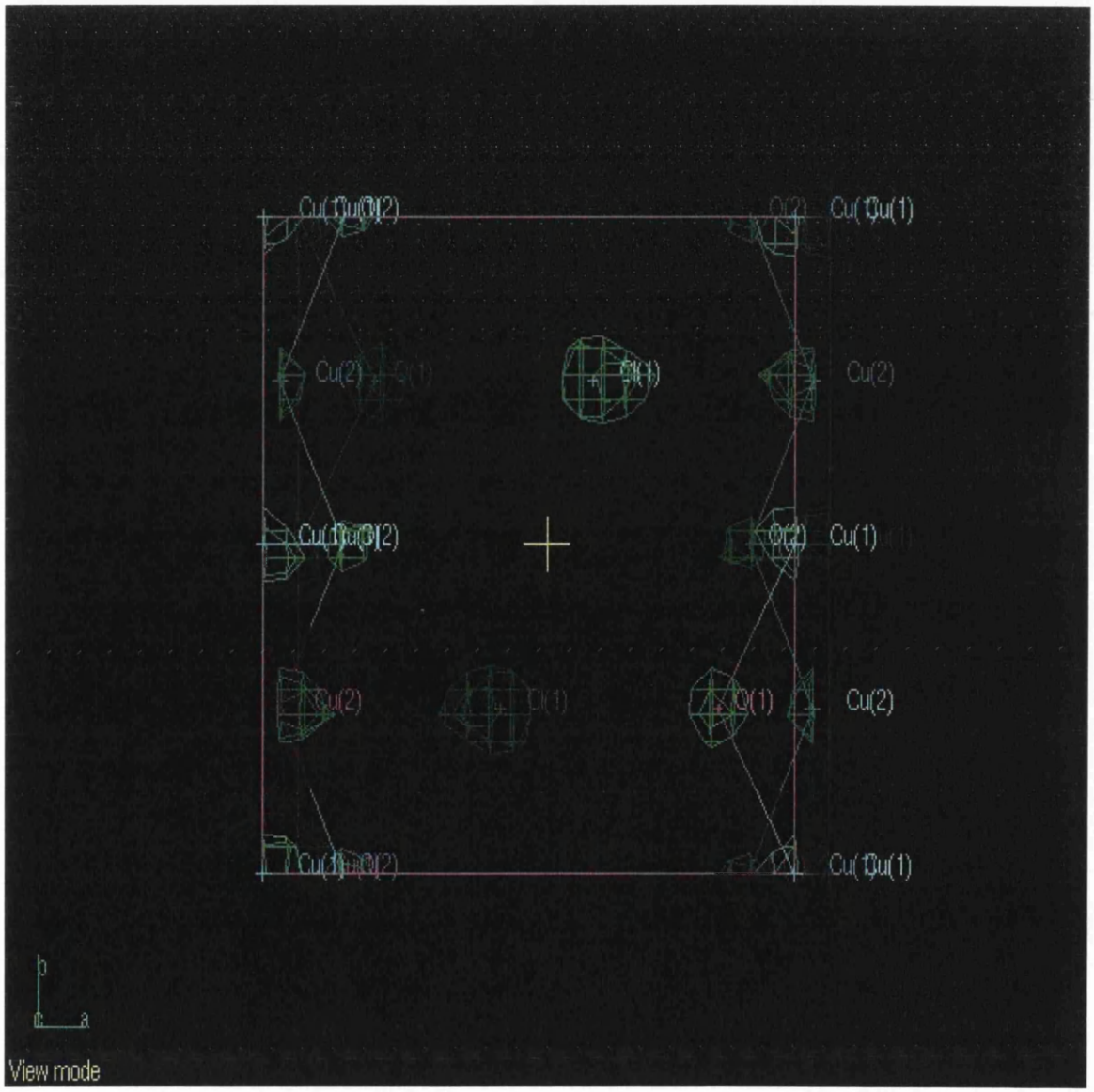
For Basic CuCl , the results were as follows:

Table 4-12 Basic Copper Chloride super resolution predicted reflections

h	k	l	U(ME)	PHASE (U(ME)	h	k	l	U(ME)	PHASE (U(ME)
1	0	7	0.062	0.00	6	0	0	0.141	0.00
1	1	7	0.054	0.00	6	0	1	0.010	180.00
1	2	7	0.061	0.00	6	0	2	0.129	0.00
1	7	0	0.013	0.00	6	0	3	0.030	0.00
1	7	1	0.042	180.00	6	0	4	0.064	0.00
1	7	2	0.054	180.00	6	1	0	0.062	180.00
1	7	3	0.012	180.00	6	1	1	0.035	0.00
1	8	0	0.078	0.00	6	1	2	0.040	0.00
1	8	1	0.002	180.00	6	1	3	0.083	0.00
2	0	5	0.074	0.00	6	1	4	0.034	180.00
2	0	6	0.192	0.00	6	2	0	0.013	180.00
2	0	7	0.064	0.00	6	2	1	0.057	0.00
2	1	5	0.087	0.00	6	2	2	0.035	0.00
2	1	6	0.002	180.00	6	2	3	0.092	0.00
2	1	7	0.028	0.00	6	2	4	0.019	180.00
2	2	5	0.082	0.00	6	3	0	0.061	0.00
2	2	6	0.106	180.00	6	3	1	0.035	180.00
2	2	7	0.060	0.00	6	3	2	0.032	180.00
2	3	5	0.034	180.00	6	3	3	0.070	180.00
2	3	6	0.027	180.00	6	4	0	0.112	0.00
2	4	5	0.014	0.00	6	4	1	0.005	0.00
2	4	6	0.144	0.00	6	4	2	0.077	0.00
2	5	5	0.038	0.00	6	5	0	0.041	180.00
2	7	0	0.067	180.00	6	5	1	0.017	0.00

2	7	1	0.057	180.00	7	0	0	0.119	0.00
2	7	2	0.044	0.00	7	0	1	0.043	0.00
2	7	3	0.068	0.00	7	0	2	0.098	0.00
2	8	0	0.079	0.00	7	1	0	0.003	180.00
3	0	6	0.105	0.00	7	1	1	0.076	0.00
3	1	6	0.068	180.00	7	1	2	0.012	180.00
3	2	6	0.015	180.00	7	2	0	0.006	0.00
3	3	6	0.030	0.00	7	2	1	0.046	0.00
3	7	0	0.006	180.00	7	2	2	0.013	0.00
3	7	1	0.020	0.00	7	3	0	0.023	180.00
3	7	2	0.047	0.00	7	3	1	0.044	180.00
4	0	6	0.034	0.00	4	2	6	0.018	0.00
4	1	6	0.030	0.00					

Figure 4-9 Basic copper chloride super resolution q^{ME} map



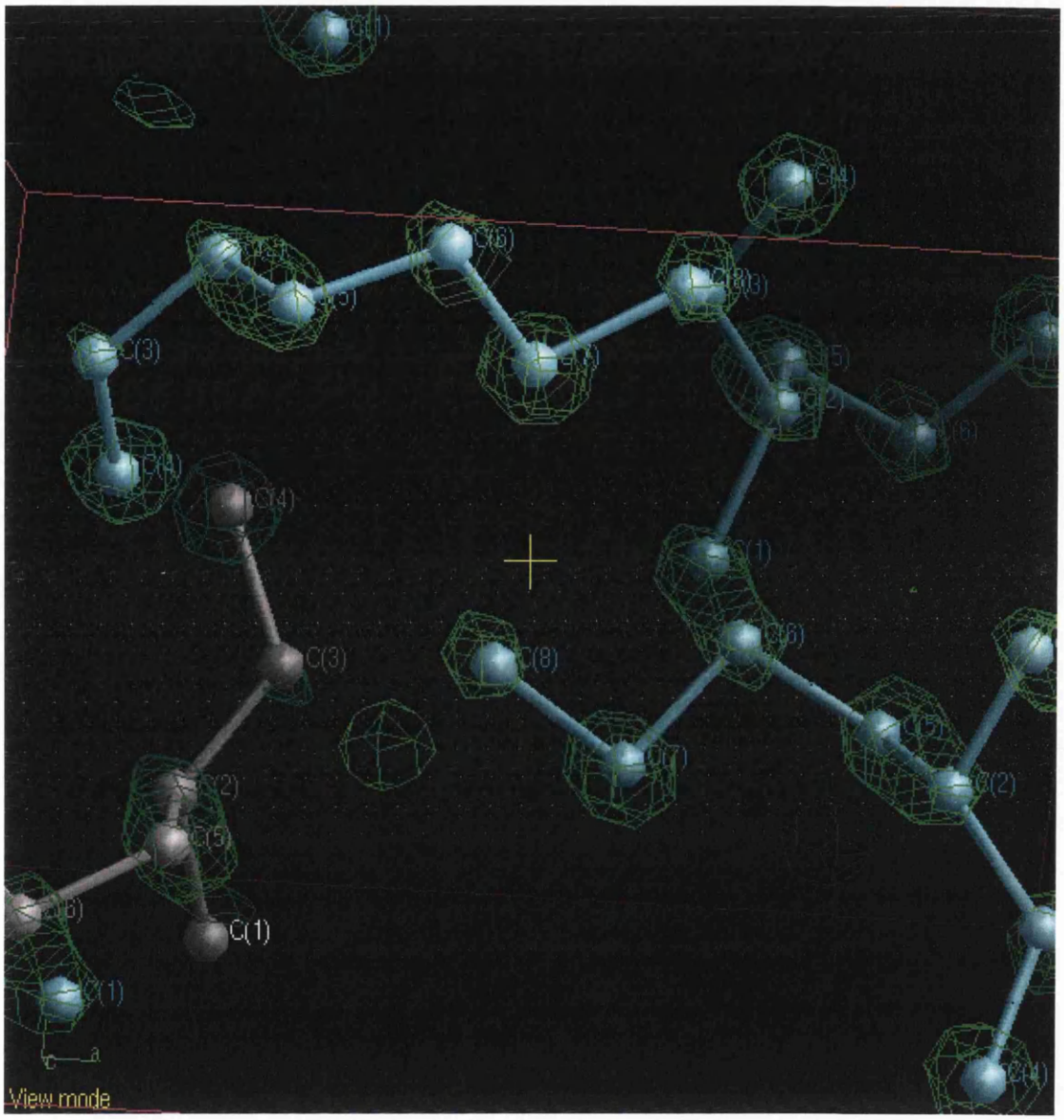
For the Basic Copper Chloride no appreciable difference can be seen in the map quality.

Table 4-13 The Isotactic Poly(1-butene) Form III results for super resolution

H	K	L	U(ME)	PHASE (U(ME))	H	K	L	U(ME)	PHASE (U(ME))
0	0	6	0.009	-180.00	5	2	3	0.023	-177.34
0	1	5	0.024	90.00	5	2	4	0.018	83.62
0	1	6	0.002	-90.00	5	2	5	0.008	-147.59
0	2	5	0.015	0.00	5	3	3	0.037	57.95
0	2	6	0.000	-179.96	5	3	4	0.025	47.32
0	3	5	0.009	90.00	5	3	5	0.013	58.95
0	3	6	0.008	-90.00	5	4	3	0.014	95.59
0	4	5	0.006	0.00	5	4	4	0.012	-135.69
0	4	6	0.004	0.00	5	4	5	0.006	7.03
0	5	5	0.001	90.02	5	5	3	0.020	121.60
1	0	5	0.011	-90.00	5	5	4	0.016	-63.69
1	0	6	0.016	180.00	5	6	3	0.005	-6.86
1	1	5	0.014	-0.70	5	7	0	0.018	90.00
1	1	6	0.006	-172.77	5	7	1	0.028	-111.87
1	2	5	0.003	140.66	6	0	3	0.019	-90.00
1	2	6	0.003	-111.83	6	0	4	0.033	0.00
1	3	5	0.015	-123.98	6	0	5	0.002	-90.00
1	3	6	0.006	7.93	6	1	3	0.026	-105.80
1	4	5	0.008	29.84	6	1	4	0.017	-69.38
1	4	6	0.007	4.88	6	1	5	0.009	-141.91
1	5	5	0.006	-48.46	6	2	3	0.024	8.30
1	8	0	0.024	-90.00	6	2	4	0.010	-156.65
1	8	1	0.035	-77.78	6	2	5	0.004	-86.39
2	0	4	0.020	180.00	6	3	3	0.020	66.13
2	0	5	0.005	90.00	6	3	4	0.020	174.28
2	0	6	0.001	0.00	6	3	5	0.005	94.34
2	1	4	0.015	22.76	6	4	3	0.009	161.89
2	1	5	0.011	152.35	6	4	4	0.012	-133.34
2	1	6	0.007	63.23	6	5	3	0.030	142.62
2	2	4	0.014	-19.10	6	5	4	0.011	97.38
2	2	5	0.008	-128.98	6	6	0	0.055	-180.00
2	2	6	0.009	-83.69	6	6	1	0.025	-65.75
2	3	4	0.007	35.01	6	6	2	0.044	135.07
2	3	5	0.006	-144.66	7	0	3	0.029	-90.00
2	3	6	0.002	-89.57	7	0	4	0.011	-180.00
2	4	4	0.011	-105.47	7	0	5	0.010	90.00
2	4	5	0.008	162.36	7	1	3	0.004	49.57
2	5	4	0.008	-120.45	7	1	4	0.005	178.40
2	5	5	0.017	-10.48	7	1	5	0.010	-83.11
2	6	4	0.004	133.82	7	2	3	0.031	28.90
2	8	0	0.055	-180.00	7	2	4	0.016	128.15
3	0	4	0.009	0.00	7	2	5	0.008	-42.26
3	0	5	0.008	90.00	7	3	3	0.013	-30.47
3	0	6	0.001	0.01	7	3	4	0.009	-65.78
3	1	4	0.004	-170.93	7	4	3	0.009	-19.67
3	1	5	0.010	118.36	7	4	4	0.016	-54.43
3	1	6	0.006	-160.79	7	5	3	0.024	-166.50
3	2	4	0.018	-104.22	7	6	0	0.034	-90.00
3	2	5	0.006	124.99	7	6	1	0.025	-10.95
3	2	6	0.007	-13.03	8	0	2	0.004	179.99
3	3	4	0.025	150.30	8	0	3	0.019	-90.00
3	3	5	0.010	-123.98	8	0	4	0.009	0.00
3	3	6	0.009	170.25	8	1	2	0.024	66.47
3	4	4	0.009	-112.77	8	1	3	0.021	-88.64
3	4	5	0.015	-36.18	8	1	4	0.012	102.45
3	5	4	0.027	-97.40	8	2	2	0.008	-100.70

3	5	5	0.005	11.38	8	2	3	0.006	14.15
3	6	4	0.005	73.79	8	2	4	0.008	-154.68
4	0	3	0.035	90.00	8	3	2	0.020	-24.22
4	0	4	0.000	179.92	8	3	3	0.005	171.35
4	0	5	0.013	90.00	8	3	4	0.013	-5.49
4	0	6	0.005	180.00	8	4	2	0.011	103.86
4	1	3	0.020	12.35	8	4	3	0.016	-169.16
4	1	4	0.011	8.27	8	5	0	0.022	0.00
4	1	5	0.014	-6.51	8	5	1	0.027	13.36
4	1	6	0.003	56.88	8	5	2	0.016	134.83
4	2	3	0.012	-156.01	9	0	1	0.058	-90.00
4	2	4	0.018	145.30	9	0	2	0.037	180.00
4	2	5	0.010	-114.09	9	0	3	0.018	-90.00
4	2	6	0.002	74.98	9	0	4	0.004	0.00
4	3	3	0.021	-23.99	9	1	1	0.058	38.93
4	3	4	0.019	-174.43	9	1	2	0.030	165.31
4	3	5	0.006	-175.64	9	1	3	0.031	153.24
4	4	3	0.018	-138.74	9	1	4	0.009	142.04
4	4	4	0.020	-3.53	9	2	1	0.025	116.27
4	4	5	0.001	-64.11	9	2	2	0.022	-35.97
4	5	3	0.039	178.57	9	2	3	0.012	31.86
4	5	4	0.017	-154.89	9	3	1	0.015	170.15
4	6	3	0.015	-19.19	9	3	2	0.038	175.20
4	6	4	0.014	-15.68	9	3	3	0.011	18.67
5	0	3	0.017	-90.00	9	4	1	0.031	55.75
5	0	4	0.003	-0.01	9	4	2	0.024	-117.59
5	0	5	0.005	-90.00	10	0	1	0.071	90.00
5	0	6	0.002	-179.99	10	0	2	0.030	-180.00
5	1	3	0.008	-91.24	10	0	3	0.011	-89.99
5	1	4	0.015	50.28	10	1	1	0.037	34.67
5	1	5	0.009	88.83	10	1	2	0.015	137.50
5	1	6	0.009	17.68	10	1	3	0.026	-8.19

Figure 4-10 Isotactic Poly(1-butene) Form III super resolution q^{ME}



In this map there is a small improvement in the quality of the data down the c-axis showing an improvement in the electron density of the C5-C2 and C1-C6 bonds. The improved definition of the electron density down the c-axis helps define the model more accurately.

Table 4-14 The Poly(1,4,trans-cyclohexanediyl dimethylene succinate) (T-cds) results for super resolution

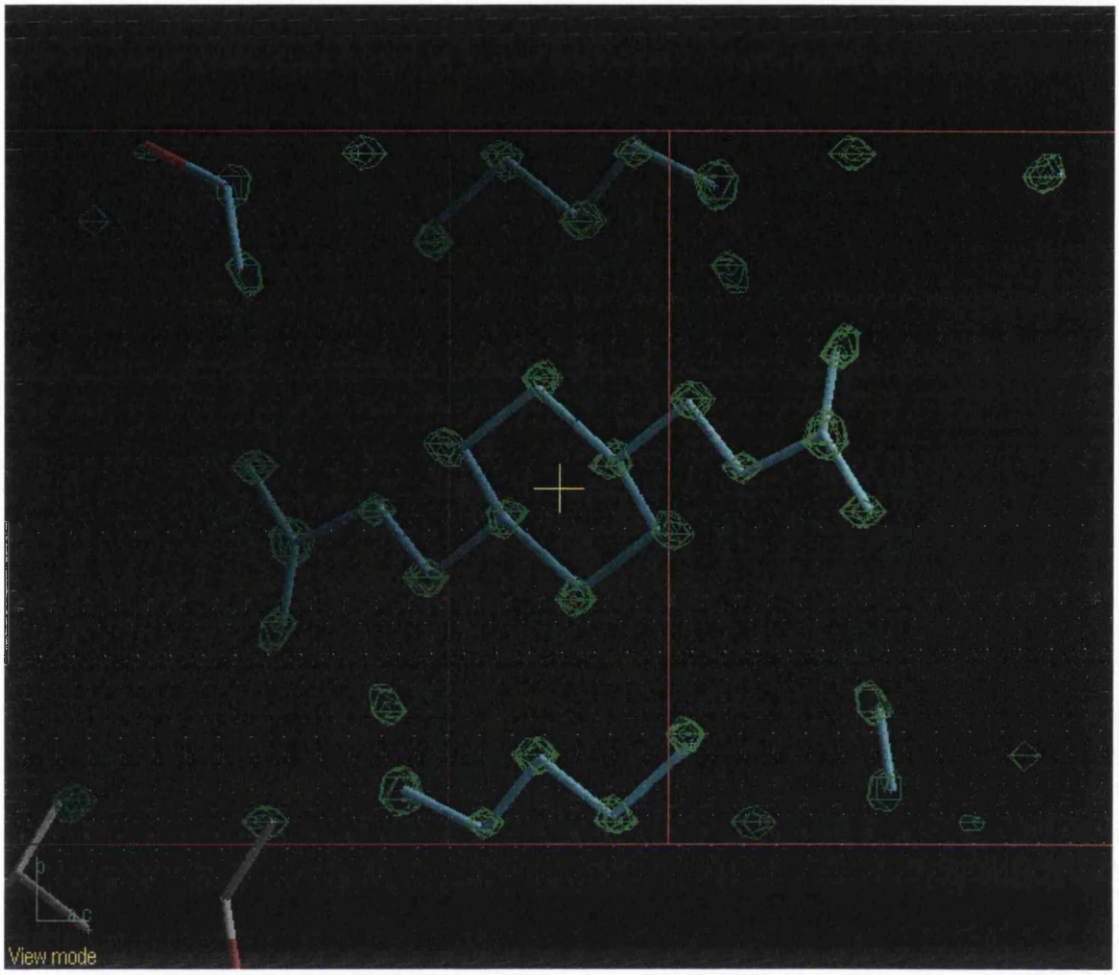
H	K	L	U(ME)	PHASE (U(ME))	H	K	L	U(ME)	PHASE (U(ME))
0	0	1	0.000	-179.34	2	5	9	0.005	180.00
0	0	2	0.066	0.00	2	5	10	0.031	180.00
0	0	3	0.000	-179.32	2	5	11	0.010	0.00
0	0	4	0.103	180.00	2	6	5	0.016	180.00
0	0	6	0.124	180.00	2	6	6	0.039	0.00
0	0	8	0.003	0.00	2	6	7	0.072	180.00
0	0	10	0.012	0.00	2	6	8	0.022	180.00
0	1	1	0.001	0.00	2	6	9	0.007	-180.00
0	1	2	0.060	180.00	2	6	10	0.013	180.00
0	1	3	0.001	0.00	2	7	5	0.016	-180.00
0	1	4	0.024	-180.00	2	7	6	0.024	-180.00
0	1	5	0.024	0.00	2	7	7	0.005	-180.00
0	1	6	0.046	0.00	2	7	8	0.030	0.00
0	1	7	0.015	0.00	2	7	9	0.032	0.00
0	1	8	0.010	0.00	2	7	10	0.014	180.00
0	1	9	0.003	0.00	2	8	5	0.029	0.00
0	1	10	0.006	0.00	2	8	6	0.065	0.00
0	2	1	0.204	0.00	2	8	7	0.052	0.00
0	2	2	0.022	0.00	2	8	8	0.045	180.00
0	2	3	0.024	0.00	2	9	5	0.025	0.00
0	2	4	0.072	-180.00	2	9	6	0.030	180.00
0	2	5	0.073	0.00	2	9	7	0.050	-180.00
0	2	6	0.054	-180.00	3	0	7	0.085	0.00
0	2	7	0.015	180.00	3	0	8	0.000	179.89
0	2	8	0.001	-180.00	3	0	9	0.004	0.00
0	2	9	0.001	180.00	3	0	11	0.027	180.00
0	2	10	0.010	0.00	3	0	13	0.004	180.00
0	3	1	0.016	-180.00	3	1	7	0.065	180.00
0	3	2	0.045	0.00	3	1	8	0.050	0.00
0	3	3	0.028	180.00	3	1	9	0.022	180.00
0	3	4	0.028	-180.00	3	1	10	0.070	0.00
0	3	5	0.064	180.00	3	1	11	0.041	0.00
0	3	6	0.015	0.00	3	1	12	0.011	180.00
0	3	7	0.037	0.00	3	1	13	0.007	0.00
0	3	8	0.008	0.00	3	2	7	0.034	0.00
0	3	9	0.003	0.00	3	2	8	0.015	0.00
0	3	10	0.006	0.00	3	2	9	0.013	0.00
0	4	1	0.012	-180.00	3	2	10	0.002	0.00
0	4	2	0.060	0.00	3	2	11	0.003	-180.00
0	4	3	0.063	180.00	3	2	12	0.011	180.00
0	4	4	0.078	-180.00	3	2	13	0.006	180.00
0	4	5	0.019	180.00	3	3	7	0.012	180.00
0	4	6	0.011	0.00	3	3	8	0.020	180.00
0	4	7	0.010	-180.00	3	3	9	0.026	180.00
0	4	8	0.005	180.00	3	3	10	0.021	180.00
0	4	9	0.003	-180.00	3	3	11	0.032	0.00
0	5	1	0.026	180.00	3	3	12	0.007	0.00
0	5	2	0.086	-180.00	3	3	13	0.004	0.00
0	5	3	0.052	0.00	3	4	7	0.054	0.00
0	5	4	0.013	-180.00	3	4	8	0.037	180.00
0	5	5	0.029	180.00	3	4	9	0.016	-180.00
0	5	6	0.047	-180.00	3	4	10	0.045	180.00

0	5	7	0.023	0.00	3	4	11	0.013	0.00
0	5	8	0.015	0.00	3	4	12	0.015	-180.00
0	5	9	0.003	180.00	3	5	7	0.105	-180.00
0	6	1	0.002	-180.00	3	5	8	0.017	0.00
0	6	2	0.028	0.00	3	5	9	0.020	-180.00
0	6	3	0.055	0.00	3	5	10	0.008	180.00
0	6	4	0.045	0.00	3	5	11	0.009	180.00
0	6	5	0.065	180.00	3	5	12	0.010	0.00
0	6	6	0.023	0.00	3	6	7	0.062	0.00
0	6	7	0.000	180.00	3	6	8	0.040	0.00
0	6	8	0.011	-180.00	3	6	9	0.032	0.00
0	7	1	0.139	0.00	3	6	10	0.050	180.00
0	7	2	0.003	0.00	3	6	11	0.015	0.00
0	7	3	0.003	180.00	3	7	7	0.059	-180.00
0	7	4	0.060	0.00	3	7	8	0.056	180.00
0	7	5	0.017	-180.00	3	7	9	0.009	0.00
0	7	6	0.027	-180.00	3	7	10	0.011	-180.00
0	7	7	0.022	180.00	3	8	7	0.031	-180.00
0	8	1	0.004	0.00	3	8	8	0.006	-180.00
0	8	2	0.100	-180.00	3	8	9	0.045	0.00
0	8	3	0.020	180.00	4	0	10	0.137	0.00
0	8	4	0.113	0.00	4	0	12	0.023	180.00
0	8	5	0.003	-180.00	4	0	14	0.008	180.00
0	8	6	0.026	0.00	4	1	9	0.009	180.00
0	9	0	0.000	-0.60	4	1	10	0.019	180.00
0	9	1	0.038	180.00	4	1	11	0.002	0.00
0	9	2	0.183	0.00	4	1	12	0.007	180.00
0	9	3	0.002	0.00	4	1	13	0.002	-180.00
0	9	4	0.080	0.00	4	1	14	0.003	0.00
0	9	5	0.026	0.00	4	2	9	0.062	180.00
0	10	0	0.066	180.00	4	2	10	0.007	180.00
0	10	1	0.068	0.00	4	2	11	0.084	0.00
0	10	2	0.057	-180.00	4	2	12	0.012	-180.00
1	0	3	0.041	0.00	4	2	13	0.016	0.00
1	0	4	0.000	-179.33	4	2	14	0.001	-180.00
1	0	5	0.128	0.00	4	3	9	0.005	-180.00
1	0	6	0.000	179.89	4	3	10	0.037	180.00
1	0	7	0.026	0.00	4	3	11	0.005	-180.00
1	0	9	0.006	0.00	4	3	12	0.008	180.00
1	0	11	0.018	180.00	4	3	13	0.013	180.00
1	1	3	0.090	0.00	4	4	9	0.024	180.00
1	1	4	0.112	180.00	4	4	10	0.039	180.00
1	1	5	0.039	0.00	4	4	11	0.008	0.00
1	1	6	0.123	-180.00	4	4	12	0.018	0.00
1	1	7	0.018	180.00	4	4	13	0.002	0.00
1	1	8	0.010	0.00	4	5	9	0.003	-180.00
1	1	9	0.003	180.00	4	5	10	0.067	-180.00
1	1	10	0.024	0.00	4	5	11	0.006	0.00
1	1	11	0.007	0.00	4	5	12	0.027	-180.00
1	2	3	0.094	180.00	4	6	9	0.002	0.00
1	2	4	0.033	-180.00	4	6	10	0.010	0.00
1	2	5	0.087	180.00	4	6	11	0.006	180.00
1	2	6	0.130	0.00	4	7	9	0.110	0.00
1	2	7	0.036	0.00	4	7	10	0.042	-180.00
1	2	8	0.021	0.00	4	8	3	0.087	180.00
1	2	9	0.010	0.00	4	8	4	0.048	0.00
1	2	10	0.012	0.00	4	8	5	0.008	0.00
1	2	11	0.001	-180.00	4	8	6	0.059	0.00
1	3	3	0.024	-180.00	4	8	7	0.006	0.00
1	3	4	0.088	-180.00	4	8	8	0.102	180.00

1	3	5	0.047	0.00	5	0	7	0.055	0.00
1	3	6	0.050	-180.00	5	0	8	0.000	179.05
1	3	7	0.016	180.00	5	0	9	0.079	180.00
1	3	8	0.001	180.00	5	0	11	0.024	0.00
1	3	9	0.006	-180.00	5	0	12	0.000	-179.35
1	3	10	0.009	0.00	5	0	13	0.017	0.00
1	3	11	0.006	0.00	5	1	6	0.005	-180.00
1	4	3	0.028	0.00	5	1	7	0.121	180.00
1	4	4	0.008	0.00	5	1	8	0.026	0.00
1	4	5	0.109	180.00	5	1	9	0.096	-180.00
1	4	6	0.009	180.00	5	1	10	0.090	0.00
1	4	7	0.052	0.00	5	1	11	0.084	0.00
1	4	8	0.004	0.00	5	1	12	0.022	180.00
1	4	9	0.003	-180.00	5	1	13	0.018	0.00
1	4	10	0.005	180.00	5	1	14	0.004	180.00
1	4	11	0.011	0.00	5	2	6	0.048	0.00
1	5	3	0.040	0.00	5	2	7	0.038	-180.00
1	5	4	0.057	-180.00	5	2	8	0.057	180.00
1	5	5	0.041	180.00	5	2	9	0.001	-180.00
1	5	6	0.004	-180.00	5	2	10	0.085	180.00
1	5	7	0.035	-180.00	5	2	11	0.001	0.00
1	5	8	0.007	180.00	5	2	12	0.001	180.00
1	5	9	0.010	-180.00	5	2	13	0.018	-180.00
1	5	10	0.002	180.00	5	2	14	0.009	0.00
1	6	3	0.074	0.00	5	3	6	0.049	-180.00
1	6	4	0.062	0.00	5	3	7	0.011	180.00
1	6	5	0.010	-180.00	5	3	8	0.034	180.00
1	6	6	0.043	-180.00	5	3	9	0.025	180.00
1	6	7	0.008	0.00	5	3	10	0.123	180.00
1	6	8	0.030	0.00	5	3	11	0.065	0.00
1	6	9	0.000	0.00	5	3	12	0.010	-180.00
1	7	3	0.043	0.00	5	3	13	0.006	0.00
1	7	4	0.182	0.00	5	3	14	0.003	-180.00
1	7	5	0.014	180.00	5	4	6	0.089	0.00
1	7	6	0.046	0.00	5	4	7	0.059	0.00
1	7	7	0.012	0.00	5	4	8	0.090	180.00
1	7	8	0.022	180.00	5	4	9	0.016	-180.00
1	7	9	0.005	0.00	5	4	10	0.070	-180.00
1	8	3	0.059	180.00	5	4	11	0.031	180.00
1	8	4	0.131	0.00	5	4	12	0.000	-0.11
1	8	5	0.004	0.00	5	4	13	0.020	180.00
1	8	6	0.021	-180.00	6	0	4	0.022	-180.00
1	8	7	0.037	-180.00	6	0	5	0.000	179.79
1	9	0	0.039	180.00	6	0	6	0.010	0.00
1	9	1	0.012	180.00	6	0	8	0.062	-180.00
1	9	2	0.079	180.00	6	0	10	0.132	0.00
1	9	3	0.124	180.00	6	0	12	0.010	180.00
1	9	4	0.069	0.00	6	0	13	0.000	-179.88
1	9	5	0.018	0.00	6	0	14	0.001	0.00
1	9	6	0.018	0.00	6	1	4	0.005	-180.00
1	10	0	0.015	-180.00	6	1	5	0.062	0.00
1	10	1	0.104	180.00	6	1	6	0.002	-180.00
1	10	2	0.147	0.00	6	1	7	0.021	-180.00
1	10	3	0.018	-180.00	6	1	8	0.032	0.00
2	0	6	0.106	-180.00	6	1	9	0.080	180.00
2	0	8	0.054	0.00	6	1	10	0.114	180.00
2	0	10	0.056	0.00	6	1	11	0.077	0.00
2	0	12	0.007	180.00	6	1	12	0.028	180.00
2	1	5	0.128	0.00	6	1	13	0.018	0.00
2	1	6	0.141	0.00	6	1	14	0.002	-180.00

2	1	7	0.045	0.00	6	2	4	0.079	0.00
2	1	8	0.029	0.00	6	2	5	0.111	180.00
2	1	9	0.008	0.00	6	2	6	0.026	0.00
2	1	10	0.013	0.00	6	2	7	0.078	180.00
2	1	11	0.017	-180.00	6	2	8	0.030	180.00
2	1	12	0.007	180.00	6	2	9	0.118	180.00
2	2	5	0.054	0.00	6	2	10	0.103	180.00
2	2	6	0.097	-180.00	6	2	11	0.101	0.00
2	2	7	0.023	180.00	6	2	12	0.027	180.00
2	2	8	0.005	0.00	6	2	13	0.020	0.00
2	2	9	0.012	180.00	6	3	5	0.052	0.00
2	2	10	0.021	0.00	6	3	6	0.036	0.00
2	2	11	0.024	0.00	6	3	7	0.001	0.00
2	2	12	0.001	0.00	6	3	8	0.085	-180.00
2	3	5	0.205	180.00	6	3	9	0.036	0.00
2	3	6	0.087	0.00	6	3	10	0.100	180.00
2	3	7	0.057	0.00	6	3	11	0.033	180.00
2	3	8	0.000	179.99	6	3	12	0.027	0.00
2	3	9	0.003	0.00	6	3	13	0.036	180.00
2	3	10	0.002	0.00	6	4	5	0.027	-180.00
2	3	11	0.011	0.00	6	4	6	0.104	-180.00
2	3	12	0.011	-180.00	6	4	7	0.011	-180.00
2	4	5	0.025	0.00	6	4	8	0.042	0.00
2	4	6	0.091	-180.00	6	4	9	0.068	0.00
2	4	7	0.032	-180.00	6	4	10	0.055	180.00
2	4	8	0.004	180.00	6	4	11	0.043	0.00
2	4	9	0.020	-180.00	6	4	12	0.013	0.00
2	4	10	0.001	180.00	6	5	7	0.073	0.00
2	4	11	0.001	0.00	6	5	8	0.004	180.00
2	4	12	0.009	0.00	6	5	9	0.008	0.00
2	5	5	0.052	-180.00	6	5	10	0.010	0.00
2	5	6	0.028	180.00	7	0	9	0.190	180.00
2	5	7	0.072	0.00	7	0	11	0.066	0.00
2	5	8	0.011	0.00	7	1	10	0.015	-180.00
7	1	11	0.027	0.00					

Figure 4-11 The Poly(1,4,trans-cyclohexanediyl dimethylene succinate) q^{ME} map



The electron density has been improved and shows better defined atom positions in the map.

4.2.12 Discussion and conclusions

Unlike the previous studies that have attempted to resolve the missing cone problem, this study has used real data, resulting in the super resolution reflections having no observed values to be compared with. The only way in which the data can be evaluated is by visual inspection of the map produced, and comparison with the published models.

All the super resolution solutions produced by the SUPER command make the model slightly easier to place in the electron density, although this in itself does not add any thing to these maps. These maps have already been solved and published, so it is unsurprising that it adds little to the solution. Where this technique may be a benefit is

when the electron density is ill defined around a potential atom position, which may make an ill-defined map into one which is more interpretable.

The technique has shown that it can predict magnitudes and phases to a high degree of accuracy compared to the observed values.

An empirical evaluation of the SUPER command results would appear to suggest that the magnitudes and phases from this extrapolation are also of high quality and accuracy, within hkl ranges close to the observed values. The phases would appear to be more accurate than the magnitudes, just as in the case of predicting the known values.

References

(1) Amos, L.A.; Henderson, R.; Unwin, P.N.T. *Progress in Biophysical Molecular Biology* **1982**, *39*, 183-231.

(2) Dorset, D.L. *Structural Electron Crystallography*; Plenum Publishing Coorporation: New York, **1995**.

(3) Sayre, D. *Acta Crystallographica* **1952**, *5*, 60-65.

(4) Dorset, D.L.; McCourt, M.P. *Zeitschrift FÜR Kristallographie* **1999**, *214*, 735-738.

(5) Dorset, D.L.; McCourt, M.P. *Zeitschrift FÜR Kristallographie* **1999**, *214*, 803-807.

(6) Ramachandran, G.N.; Srinivasan, R. *Fourier Methods in Crystallography*; Wiley Interscience: **1970**.

(7) Voronova, A.A.; Vainshtein, B.K. *Soviet Physics--Crystallography* **1958**, *3*, 445-451.

(8) Dorset, D.L.; McCourt, M.P.; Kopp, S.; Wittmann, J.C.; Lotz, B. *Acta Crystallographica* **1994**, *B50*, 201-208.

(9) Brisse, F.; Rémillard, B.; Chanzy, H. *Macromolecules* **1984**, *17*, 1980-1987.

5 LEAST SQUARES REFINEMENT OF ATOMIC MODELS

5.1 Structure refinement

Crystal structure analysis in all its forms - trial and error methods, Patterson methods or direct methods - leads to a proposed model which makes chemical sense and a rough consistency with the diffraction pattern intensities. The process of structure refinement aims to find the optimal agreement between calculated and observed intensities, while still retaining a model which makes chemical sense. This usually takes the form of refining atomic coordinates, occupancies and thermal parameters for the atoms, which account for the motion of the atoms in the crystal structure, normally associated with thermal vibrations.

The success of a structure refinement is usually judged by the R-factor figure of merit.

$$R = \frac{\sum \| F_o | - | F_c \|}{\sum | F_o |}$$

Figure 5-1

(5-1)

or

$$wR2 = \left(\frac{\sum w \left(F_0^2 | - | F_c^2 | \right)^2}{\sum \left(w | F_0^2 | \right)^2} \right) \quad \textbf{(5-2)}$$

The R-factor is a useful indicator for deciding whether a particular model needs adjustment. If the factor is above R = 0.83 for centrosymmetric space group and R = 0.59 of non-centrosymmertic space group, the atoms are essentially random in relation to the diffraction pattern. R-factors of around R ≈ 0.45 indicate the model has some merit. R ≈ 0.35 is taken to mean that the final model can be deduced from the current point. R ≈ 0.25 means that the model is within 0.1Å of the correct atomic positions[1]. The R-factor is generally used as an indicator of whether one proposed model is better or worse than another, but the R-factor is a simple figure of merit that can be easily manipulated to show better agreement between calculated and observed than actually exists. Usually other figures of merit are used for such observations, such as the Goodness of Fit, which takes account of the number of parameters (p)

used in comparison to the number of reflections observed (n). A weight (w) is used on each reflection to model measurement errors. The goodness of fit, S, is defined as[2].

$$S = \frac{\sum_i w_i \Delta F_i^2}{n - m} \quad (5\text{-}3)$$

Linear least squares is a common method of checking the validity of a proposed model against the observed data.

$$a_1 x_1 + a_2 x_2 + \ldots = y_1$$
$$\text{or} \quad (5\text{-}4)$$
$$Ax = y$$

Where in the first equation the terms of the left-hand side are functions of the atomic co-ordinates. The matrix A, is the design matrix. The variance-covariance matrix is the design matrix multiplied by its inverse, the diagonal of which gives the variance of the calculated values, or in other words the errors on the calculated values. The process of least squares refinement is used to adjust parameters in the model, x, to more accurately fit the observed data, y. The calculation is more often than not weighted, helping the calculation converge more efficiently. The weighting matrix W is a positive definite matrix.

If the model is linear ($Ax = y$) then the set of parameters that minimise the function/model are given by,

$$\hat{x} = \left(A^T W A\right)^{-1} A^T W y \quad \textbf{(5-5)}$$

This is known as the least squares estimate, this is unbiased with respect to the choice of weighting matrix because the estimator is itself a random variable from a random population and it has a mean and variance.

In the case of non-linear least squares, which occur in the crystallographic context, the observations do not linearly depend on the coordinates. The left-hand side of the linear least squares equation is replaced by estimated derivatives of the observation from the model parameters. The principle is to find a set of parameters close enough to the least squares estimator so that gradient of the function becomes so small that the calculation is effectively linear over the given parameter.

$$\frac{\partial y_c}{\partial x_1}\partial x_1 + \frac{\partial y_c}{\partial x_2}\partial x_2 + \ldots = y_1 - y_c \quad \textbf{(5-6)}$$

This leads to difficulties in the least squares process, first, it is difficult to choose the terms of the design matrix. Secondly the values of the design matrix are calculated from the model, this can lead to situations where the an essentially incorrect model is likely to have derivatives that will make improvements to the model, or that the model will produce singular or unstable matrices.

Weights are normally calculated from $\left(\left|F_{\underline{h}}\right|\right)$or σ values,

$$w_{\underline{h}} = \frac{1}{\sigma^2\left(\left|F_{\underline{h}}\right|\right)} \quad \textbf{(5-7)}$$

of a particular reflection *i*. Intensity based weighting schemes are based on standard deviations of the intensities, obtained from the inverse of the variance-covariance matrix, and work on a principle of trying to find the minimal variance for a particular reflection intensity. Model based weighting schemes also use the estimated standard deviations found by inverting the covariance matrix V.

Cochran developed a weighting scheme with weights inversely proportional to the reciprocal of the atomic scattering factor, this was further developed by Dunitz and

Seiler[3], making the scheme work when several atom types are present, by using an average atomic scattering factor. This approach allows the coordinates of the maxima of the electron density to be represented as a Fourier series. This works by upweighting the high angle reflections.

The scheme essentially looks at the sigma values of the reflections,

$$\sum_{\underline{h}}\left[F_0(\underline{h})-F_c(\underline{h})\right]^2 = D(0)=\int\delta^2(\underline{x})d\underline{x} \quad \textbf{(5-8)}$$

By minimising the right hand side of this equation, all the positions in the unit cell have equal weight. During a refinement the atomic peaks are being assigned to the maximum electron density, therefore it is preferable to introduce a function that assigns higher weights to density close to atomic centres, and assign low weights to density that for these purposes is of less use to the crystallographer.

If $q_0(\underline{x})$and $q_c(\underline{x})$are the sums of atomic distributions then $q(\underline{x})$becomes,

$$q(\underline{x})=\sum_j q_j(x-x_j) \quad \textbf{(5-9)}$$

Which in turn modifies $D(0)$ to become,

$$D'(0)=\sum_j\int\delta_j^2(x-x_j)\exp\left[-2q(x-x_j)\right]^2 \quad \textbf{(5-10)}$$

The weights can now be taken as a Gaussian function. Allowing for isotropic temperature factors and the weights are modified to account for the higher angle the scheme now looks like the following;

$$Q' = \sum_{\underline{h}} w'(\underline{h})\left[F_o(\underline{h})-F_c(\underline{h})\right]^2 \quad \textbf{(5-11)}$$

Where the weights are calculated as follows,

$$w = \frac{1}{\sigma^2\left(F_o^2\right) + \left(aP\right)^2 + bP} \quad \textbf{(5-12)}$$

Where P is $\left[2F_{calc}^2 + Max\left(F_{obs}^2, 0\right)\right]/3$ and a and b are functions of the Bragg angle and are independent of the measured intensity.

In electron diffraction, the estimated standard deviation values of the intensities are not usually available, because the intensities are often inaccurate due to dynamical and secondary scattering, and the incompleteness of the data sets. This makes calculating accurate esds from the variance-covariance matrix almost impossible. Other factors affecting the problem of refinement for electron diffraction are the poor data to parameter ratios: with only one or two reflections per variable in the structure, where you can expect 100 reflections per atom in an equivalent X-ray refinement experiment. This has lead to many people arguing that the examples of the X-ray community should not be followed. In addition, the inaccuracy and incompleteness of the data makes refinement a tricky process for electron diffraction data.

We have only the R-factor or a similar function to help us find the correct solution to the phase problem. This has lead to many in the electron diffraction community ignoring the problem of refinement. Once a satisfactory R-factor has been approached by the proposed model the structure solution is often deemed satisfactory.

Despite these misgivings, others take the view that refinement is an important part of the structure analysis process, which helps to check for inaccurate or incomplete models. Electron diffraction structure determinations are less accurate than those from X-ray experiments of a non-protein molecule. R-factors of <0.05 are expected for a successful X-ray structure refinement, in order to give accurate bond lengths and geometry. In the case of electron diffraction experiments R-factors of <0.20 are considered exceptional with values being more commonly ≥ 0.25. This means that typical X-ray quality results are unobtainable, and the accuracy of such models should not be expected to be comparable with X-ray experiments. Instead electron diffraction models are used to show conformations and packing is a more pragmatic approach to looking at electron diffraction results.

A notable exception to the view that refinement is not possible with electron diffraction data is a Program called MSLS[4] which refines electron crystallographic

data in two ways: one with respect to the atomic positions, and two, taking into account the crystal thickness *via* multislice calculations which account for dynamical scattering and hence produce more accurate refinement. The molecules being studied here have all been solved using *ab initio* maximum entropy calculations. The assumption is that if the amount of dynamical scattering is small enough to allow an *ab initio* calculation to function then the dynamical scattering should be small enough to allow a least-squares calculation to refine the atomic positions reasonably.

The standard refinement calculations in X-ray diffraction refinement all use a least-squares approach. Although there are other alternatives to this method, small molecule X-ray crystallographers have seen no need to look beyond this to improve their results because the current methodologies are adequate, whereas macromolecular crystallographers have branched out into using log-likelihood refinement as an modification to least-squares for more demanding situations. This approach has not yet been considered for electron diffraction methods. Singular value decomposition or SVD[2,5] is another alternative, but this has only been implemented in a crystallographic context as an undocumented feature in maXus, although the SVD method may be useful with electron diffraction data, due to its sparse nature. SVD will be able to overcome the problems of singular matrices, caused by sparse data sets where traditional least squares methods would fail. SVD allows errors to be estimated *via* an alternative route to least squares methods, and can therefore overcome difficulties that the least squares method has with sparse or ill defined matrices.

Before starting on these newer alternatives to least squares, a fresh approach was taken to the least squares method. Using SHELX-L[6] we consider the use of least squares refinement of electron diffraction data, to ask whether it can produce sensible results with no adjustment to the default parameters, and if it is possible to improve on these results with some restraints and constraints, possibly leading to usable results, and useful geometries. It must also be considered whether or not temperature factors can be allowed to vary. Using the standard X-ray tools for refinement also has the advantage of allowing the electron crystallographer to use freely available tools, with a proven record of reliability.

5.2 Using refinement packages with electron diffraction data

SHELX-L is one of the most widely used refinement packages available, although there are many others in packages such as *e.g.* maXus and CRYSTALS. SHELX-L was chosen in this case because it was easy to introduce the electron scattering factors, *via* the maXus suite of crystallography programs. A strategy was devised for addressing the refinement problem. Each of the following points was incrementally introduced to the refinement process if the previous set of conditions had not yielded a satisfactory result. Points one and two were used as starting conditions.

1. Use electron scattering factors. (maXus[19])

Electron scattering factors from maXus were used in the refinement, because electron diffraction occurs from the electron potential, whereas X-rays are diffracted from the electron density.

2. Refine coordinates and temperature factors without any restraints or constraints.

Restraints are considered to be "soft" method of ensuring a difficult refinement finds its minimum. Restraints impose known molecular knowledge of bond lengths and angles on to a molecular fragment or entire molecule. The restraints take the form of additional equations that become part of the least squares function that has to be minimised. They do not impose absolute values on the bond lengths and angles but allow them to vary within certain limits, using the difference between the ideal and the calculated values as a penalty function.

$$D = \sum_{\underline{h}} w_{\underline{h}} \left(|F_o| - |kF_c| \right)^2 + \sum_{l} w_l \left(d_{ideal}^2 - d_c^2 \right)^2 + \sum_{\phi} w_\phi \left(\phi_{ideal}^2 - \phi_c^2 \right)^2 \quad \textbf{(5-13)}$$

Where d_{ideal}, ϕ_{ideal} are the ideal values of bond distance and phase these are subtracted from current calculated values. The root mean square difference between the two values is used to adjust the weights w_l, w_ϕ as the calculation progresses.

Constraints on the other hand are rigid values that are assigned to the bond length angle or thermal parameters during the refinement, these are known as "hard" methods of helping difficult refinements to converge. Constraints reduce the data parameter ratio of a refinement by setting a value for a particular parameter and reducing the amount of work the calculation has to do.

Unrestrained least-squares refinement is attempted first of all because there is a chance that the model is accurate enough and the data abundant enough to achieve a full least-squares refinement.

3. Fix the temperature factors.

The temperature factors are fixed, this is a constraint reducing the number of parameters for each atom from 6 to 3. If the unrestrained calculation has not work, then lowering the number of parameters to observations that are being calculated, should improve the chance of showing any inaccuracies in the model coordinates, although not the atom type assignments or disorder. There is a disadvantage to this in X-ray diffraction the behaviour of temperature factors can identify mis-placed atoms and/or wrong atom types

4. Damp the diagonal of the least squares matrix.

Damping of the least squares matrix stops the calculation from moving too far in any given shift. These types of shifts are common when dealing with data sets that have very low data parameter ratios, electron diffraction data sets inherently of a this type.

5. Introduce restraints on bond lengths where required.

Bond restraints are introduced when there is a lack of data down a particular axis and a given bond becomes elongated or there is trouble finding a single co-ordinate along an axis.

6. Introduce bond angle restraints where required

Bond angle restraints are used when bond lengths between two atoms are sensible but the bond has trouble finding a reasonable angle, this is often due to lack of data down a particular axis.

7. Introduce bond length and angle constraints as required.

Constraints are used when restraining the bond angles and lengths is sufficient, but does appear to be improving the solution. Constraints should be used with extreme caution, particularly as a constraint may stop an inappropriate model being refined.

The reasoning behind each of these points is reasonably straightforward. First, use electron scattering factors, because we are looking at electron diffraction data.

Initially the refinement should be attempted with no additional restraints or constraints being added to the calculation. This allows particularly exceptional three dimensional electron diffraction sets to refine freely, if possible.

Holding the temperature factors constant may seem strange to an X-ray crystallographer initially, until closer attention to the detail of electron diffraction experimental set-up is brought to bear. The small volume of the crystal is one of the main reasons for wishing to attempt electron crystallography. This is also one of its many drawbacks, because of such high concentration of energy entering the crystal and the small dimensions of the crystal. Heat caused by the electrons passing through often causes large thermal vibrations, due to the lack of energy dissipation. The small nature of these crystals and the thermal vibrations of the local area are sometimes not averaged out over the crystal adequately and consequently there is a justification to hold the temperature factors constant in the refinement of electron diffraction data. This has the obvious disadvantage of meaning that the temperature factors cannot be used to show the wrong assignment of atom type or chemically labile groups, but there are other techniques for identifying the atom types in the electron microscope.

Damping of the least squares matrix diagonal is a commonly used feature from X-ray crystallography, where the diagonal of the least squares matrix is constrained to small increments between cycles of refinement. This stops the calculation jumping out of the minimum, which is especially easily done with sparse data sets, which are normal with electron crystallographic data.

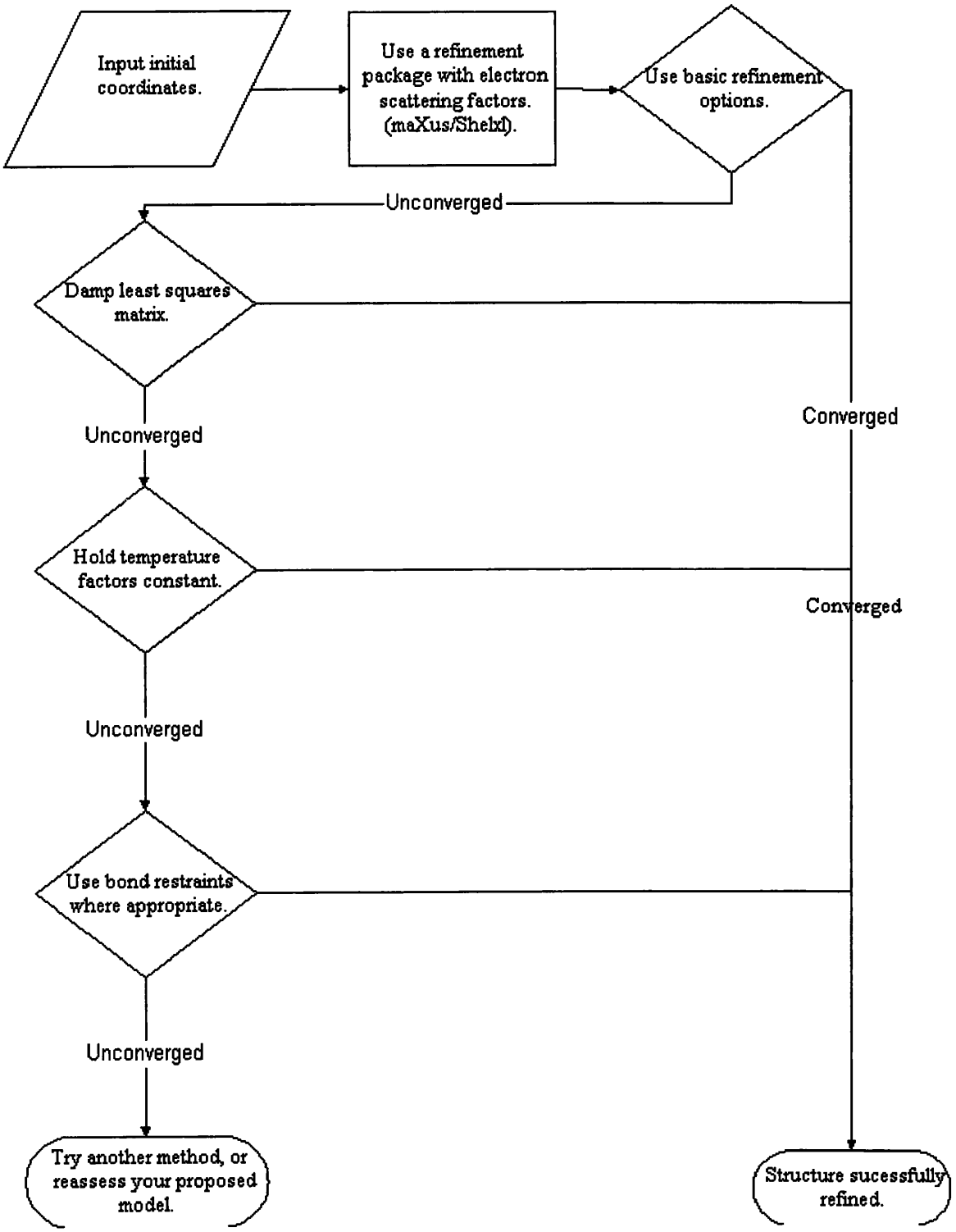
If damping has not succeeded in refining a structure successfully, generally some of the bonds are elongated or the distance between the atoms is too long to be considered a chemical bond. This is usually as a consequence of missing data down one axis, invariably the result of the missing cone or lack of an epitaxially grown crystal. This lack of data down one axis has a profound effect on the refinement process. The calculation will not be able to constrain the atom position accurately; therefore, it is argued that to restrain the atom position within known limits is reasonable. The same argument can be used to apply restraints to the angles of the molecule, and moreover if the restraints fail to hold a reasonable model together during refinement, then constraints that are more forceful can be used, following on from the aforementioned argument.

All obtainable three-dimensional data sets were used to test this methodology. They were the following: aluminium iron alloy, Brucite ($Mg(OH)_2$), CNBA, DMABC, DMACB, copper perchlorophthalocyanine, poly(1-butene) form III, polyethylene, silicon surface, poly (1,4- trans-cyclohexanediyl dimethylene succinate) (T-cds), .

5.3 Refinement experimental

The following procedure was implemented when refining electron crystallographic structures. The choice of refinement program was influenced by the availability of programs with access to electron scattering factors, and are fully featured, stable least squares program. The crystallographic suite of programs, maXus[19] was one such place where all the necessary requirements were meet. Electron scattering factors are available in the program database and the popular refinement program SHELX-L, has an easy to use graphical user interface, provided by the maXus suite.

Figure 5-2 Flow diagram of the refinement process



5.4 Results

5.4.1 Table of all results

Table 5-1 Summary of all refinement calculations

Compound	No. Of Reflections	No. of Atoms in a.u.	R1 for $F_o>4F_o$	No. Of Parameters	No. of restraints
Al_mFe	743	13	0.46	32	0
Brucite, Mg(OH)	59	3	0.11	4	0
10-cyano-9,9'-Bianthryl CNBA	144	30	0.29	123	57
Basic Copper Chloride	119	12	0.27	15	0
DMABC, C_{20} N_2 O_1	133	14	0.26	42	18
DMACB, C_{15} N_2	118	17	0.26	53	24
Mannan I	58	11	0.20	34	9
Copper Perchlorophthalocyanine	197	16	0.20	45	0
Isotactic Poly(1-butene) Form III	146	8	0.29	27	28
Poly(_-caprolactone)	47	10	0.20	33	2
Polyethylene	51	3	0.14	6	0
Poly(1,4,trans-cyclohexanediyl dimethylene succinate)	85	8	0.16	33	0

5.4.2 Al_mFe

The Al_mFe structure was determined by Gønnjes and co-workers[7], in tetragonal space group $I\bar{4}2m$ with a = 8.84 c = 21.6Å. The cell contains 90Al and 20Fe atoms. The reflections were collected using a novel precession technique, instead of selected area diffraction. The precession technique moves the electron beam in a circular movement around the sample, and resulting reflections are more akin to the single-crystal X-ray diffraction because they integrate through the Bragg angle. The data extend further out in the Laue zones, while dynamical effects reduced by the tilt out of the zone axis. Thickness variations will be damped considerably by the integrating motion, which also has the benefit of giving more kinematic data than selected area diffraction spots. Thus the technique can help to overcome some of the more difficult aspects of electron crystallography, even when heavy atoms are present.

A three-dimensional least squares refinement carried out using SHELXL97, but a very poor fit to the intensities was obtained, WR2 = 0.758, with R1 = 0.42, although the atom positions themselves had shifted very little during the refinement process. An attempt to refine the occupancy factors failed, due to the dynamical effects making the reliability of such numbers too uncertain.

The following are the results of a least squares refinement by SHELX-L on the Al_mFe data. The refinement required damping to the least squares matrix, bond restraints were also added to the calculation, but were found to make no improvement to the structure, or R-factor.

Table 5-2 Aluminium Iron atom positions

Atom	x	y	z	U(iso)
Fe(1)	0.3162	0.3162	0.0792	0.05
Fe(2)	0.1740	0.1740	0.2750	0.05
Fe(3)	0.1830	0.1830	-0.1284	0.05
Al(4)	0.0310	0.3370	0.0729	0.05
Al(5)	0.5440	0.2050	0.1343	0.05
Al(6)	0.0440	0.2950	-0.2271	0.05
Al(7)	0.0000	0.0000	-0.0748	0.05
Al(8)	0.2150	0.2150	-0.0201	0.05
Al(9)	0.2330	0.2330	0.1730	0.05
Al(10)	0.1970	0.5000	0.0000	0.05
Al(11)	0.5000	0.5000	0.1512	0.05
Al(12)	0.0000	0.0000	-0.2110	0.05
Al(13)	0.5000	0.5000	0.0000	0.05

Table 5-3 Aluminium Iron bond lengths (Å)

Atom1	Atom2	Bond length (Å)
Fe(1)	Al(9)	2.27
Fe(2)	Al(6)	2.43
Fe(2)	Al(9)	2.32
Fe(3)	Al(5)	2.42
Fe(3)	Al(8)	2.37

Table 5-4 Aluminium Iron bond angles

Atom1	Atom2	Atom3	Angle(°)
Al(6)	Fe(2)	Al(6)	120.9
Al(6)	Fe(2)	Al(9)	71.1
Al(5)	Fe(3)	Al(5)	80.6
Al(5)	Fe(3)	Al(8)	85.6
Al(5)	Fe(3)	Al(8)	85.6
Fe(1)	Al(9)	Fe(2)	171.3

5.4.3 Brucite ($Mg(OH)_2$)

The Brucite structure was originally solved by Zvyagin and co-workers [8]. The compound crystallised in the tetragonal space group $P\bar{3}m1$ with unit cell dimensions a = 3.149 c = 4.769Å, with one molecule in the asymmetric unit. There were 70 reflections obtained by electron diffraction, no attempt was made to refine the original structure by least squares methods in the original paper.

The Brucite structure was refined successfully by damping the diagonal of the least squares matrix, The instruction DAMP 1000 was the only addition requirement needed in the SHELX-L input file. The following are the results of the least squares refinement.

Table 5-5 Brucite atom positions

Atom	x	y	z	U(iso)
O(2)	0.3333	0.6666	0.2228(12)	0.0134(7)
H(2)	0.3333	0.6666	0.4307	0.0160
Mg(1)	0.0000	0.0000	0.0000	0.0080(7)

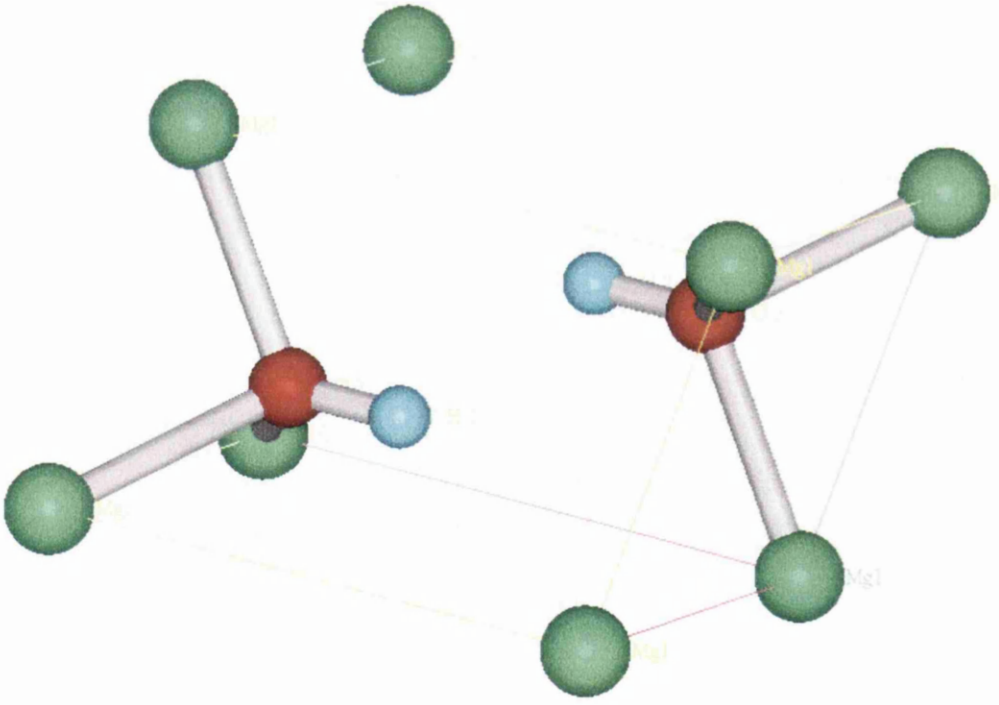
Table 5-6 Brucite bond lengths (Å)

Atom1	Atom2	Bond length (Å)
O(2)	H(2)	0.992(6)
O(2)	Mg(1)	2.106(4)

Table 5-7 Brucite bond angles

Atom1	Atom2	Atom3	Angle(°)
Mg(1)	O(2)	Mg(1)	96.8(2)

Figure 5-3 Refined Brucite model



5.4.4 10-Cyano-9,9'-Bianthryl (CNBA)

The CNBA structure was originally solved by Voigt-Martin and co-workers [9] and has the cell dimensions a = 14.70 b = 9.47 c = 15.42Å, β = 112°. The structure crystallised in space group $P2_1/c$, and 150 reflections were recorded by electron diffraction. No attempt was made to refine the structure by least squares.

The following are the results of the least squares refinement from SHELX-L. The structure was successfully refined by damping the matrix diagonal with the instruction DAMP 500, and fixing the temperature factors.

Table 5-8 CNBA atom positions

Atom	x	y	Z	U(iso)
C(1)	0.2932	0.0973	-0.0193	0.050
C(2)	0.2216	0.0003	-0.0723	0.050
C(3)	0.1606	-0.0618	-0.0367	0.050
C(4)	0.1661	-0.0297	0.0563	0.050
C(5)	0.1030	-0.0889	0.0952	0.050
C(6)	0.1079	-0.0537	0.1858	0.050
C(7)	0.0435	-0.1128	0.2259	0.050
C(8)	0.0507	-0.0757	0.3132	0.050
C(9)	0.1224	0.0225	0.3661	0.050
C(10)	0.1847	0.0812	0.3310	0.050
C(11)	0.1812	0.0454	0.2389	0.050
C(12)	0.2460	0.1039	0.2016	0.050
C(13)	0.2388	0.0688	0.1103	0.050
C(14)	0.3020	0.1301	0.0689	0.050
C(15)	0.2833	0.6446	0.2217	0.050
C(16)	0.3699	0.5932	0.2830	0.050
C(17)	0.3860	0.4469	0.2968	0.050
C(18)	0.4745	0.3923	0.3591	0.050
C(19)	0.4905	0.2493	0.3717	0.050
C(20)	0.5825	0.1924	0.4340	0.050
C(21)	0.5972	0.0533	0.4448	0.050
C(22)	0.5230	-0.0437	0.3956	0.050
C(23)	0.4341	0.0040	0.3347	0.050
C(24)	0.4150	0.1507	0.3206	0.050
C(25)	0.3248	0.2047	0.2576	0.050
C(26)	0.3099	0.3502	0.2452	0.050
C(27)	0.2206	0.4093	0.1813	0.050
C(28)	0.2087	0.5499	0.1700	0.050
C(29)	0.0306	-0.1919	0.0400	0.050
N(30)	-0.0229	-0.2737	-0.0021	0.050

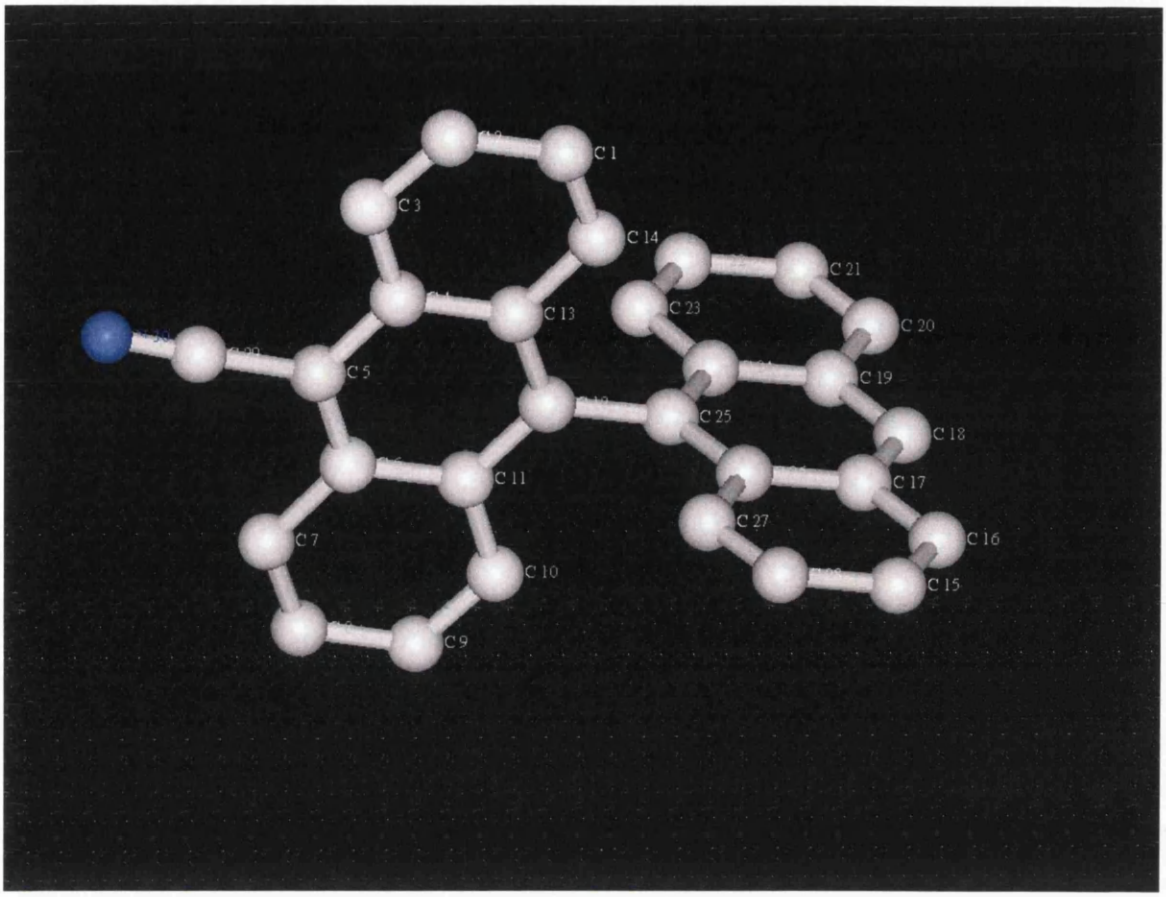
Table 5-9 CNBA bond lengths (Å)

Atom1	Atom2	Bond length (Å)	Atom1	Atom2	Bond length (Å)
C(1)	C(2)	1.40	C(1)	C(14)	1.35
C(2)	C(3)	1.34	C(3)	C(4)	1.43
C(4)	C(5)	1.39	C(4)	C(13)	1.43
C(5)	C(6)	1.41	C(5)	C(29)	1.46
C(6)	C(7)	1.42	C(6)	C(11)	1.43
C(7)	C(8)	1.35	C(8)	C(9)	1.41
C(9)	C(10)	1.34	C(10)	C(11)	1.44
C(11)	C(12)	1.39	C(12)	C(13)	1.41
C(12)	C(25)	1.50	C(13)	C(14)	1.43
C(15)	C(16)	1.36	C(15)	C(28)	1.41
C(16)	C(17)	1.40	C(17)	C(18)	1.39
C(17)	C(26)	1.43	C(18)	C(19)	1.37
C(19)	C(20)	1.44	C(19)	C(24)	1.44
C(20)	C(21)	1.33	C(21)	C(22)	1.41
C(22)	C(23)	1.37	C(23)	C(24)	1.41
C(24)	C(25)	1.41	C(25)	C(26)	1.39
C(26)	C(27)	1.43	C(27)	C(28)	1.34
C(29)	N(30)	1.12			

Table 5-10 CNBA bond angles

Atom 1	Atom 2	Atom 2	Angle (°)	Atom 1	Atom 2	Atom 3	Angle (°)
C(2)	C(1)	C(14)	120.3	C(1)	C(2)	C(3)	121.0
C(2)	C(3)	C(4)	120.9	C(3)	C(4)	C(5)	122.8
C(3)	C(4)	C(13)	118.1	C(5)	C(4)	C(13)	118.9
C(4)	C(5)	C(6)	122.1	C(4)	C(5)	C(29)	118.3
C(6)	C(5)	C(29)	119.5	C(5)	C(6)	C(7)	122.6
C(5)	C(6)	C(11)	118.1	C(7)	C(6)	C(11)	119.2
C(6)	C(7)	C(8)	120.6	C(7)	C(8)	C(9)	120.6
C(8)	C(9)	C(10)	120.8	C(9)	C(10)	C(11)	121.2
C(6)	C(11)	C(10)	117.4	C(6)	C(11)	C(12)	120.4
C(10)	C(11)	C(12)	122.0	C(11)	C(12)	C(13)	120.4
C(11)	C(12)	C(25)	120.9	C(13)	C(12)	C(25)	118.6
C(4)	C(13)	C(12)	119.8	C(4)	C(13)	C(14)	118.2
C(12)	C(13)	C(14)	121.8	C(1)	C(14)	C(13)	121.2
C(16)	C(15)	C(28)	119.6	C(15)	C(16)	C(17)	121.1
C(16)	C(17)	C(18)	121.9	C(16)	C(17)	C(26)	119.4
C(18)	C(17)	C(26)	118.6	C(17)	C(18)	C(19)	121.8
C(18)	C(19)	C(20)	122.0	C(18)	C(19)	C(24)	120.2
C(20)	C(19)	C(24)	117.6	C(19)	C(20)	C(21)	121.3
C(20)	C(21)	C(22)	121.1	C(21)	C(22)	C(23)	120.1
C(22)	C(23)	C(24)	120.7	C(19)	C(24)	C(23)	118.8
C(19)	C(24)	C(25)	118.4	C(23)	C(24)	C(25)	122.7
C(12)	C(25)	C(24)	119.3	C(12)	C(25)	C(26)	119.9
C(24)	C(25)	C(26)	120.6	C(17)	C(26)	C(25)	120.1
C(17)	C(26)	C(27)	117.3	C(25)	C(26)	C(27)	122.4
C(26)	C(27)	C(28)	121.2	C(15)	C(28)	C(27)	121.2
C(5)	C(29)	N(30)	178.0				

Figure 5-4 Refined 10-Cyano-9,9'-Bianthryl model



5.4.5 Basic Copper Chloride

Basic copper chloride structure was solved by electron diffraction in 1958 by Vainstein and co-workers[10]. The structure crystallised in space group $P2_1/m$ with cell dimensions a = 5.73 b = 6.1, c = 5.63Å, β = 93.75°. The data set contains 120 unique reflections. Again no least squares refinement was attempted.

The following are the results from a SHELX-L refinement. The model was successfully refined to R = 0.27 without any additional restraints or constraints. Further improvement was achieved by damping the least squares matrix using the DAMP 500 command, then R = 0.25, although there appeared to be no change in the geometry or atom positions.

Table 5-11 Basic copper chloride atom positions

Atom	x	y	Z	U(iso)
Cu(1)	0.0000	0.0000	1.0000	0.008(4)
Cu(2)	0.015(5)	0.2500	0.488(3)	0.013(4)
Cl(3)	0.405(7)	0.2500	0.442(6)	0.052(11)
O(4)	-0.223(17)	0.2500	1.178(15)	0.07(2)
O(5)	-0.147(6)	-0.008(6)	0.348(4)	0.011(7)

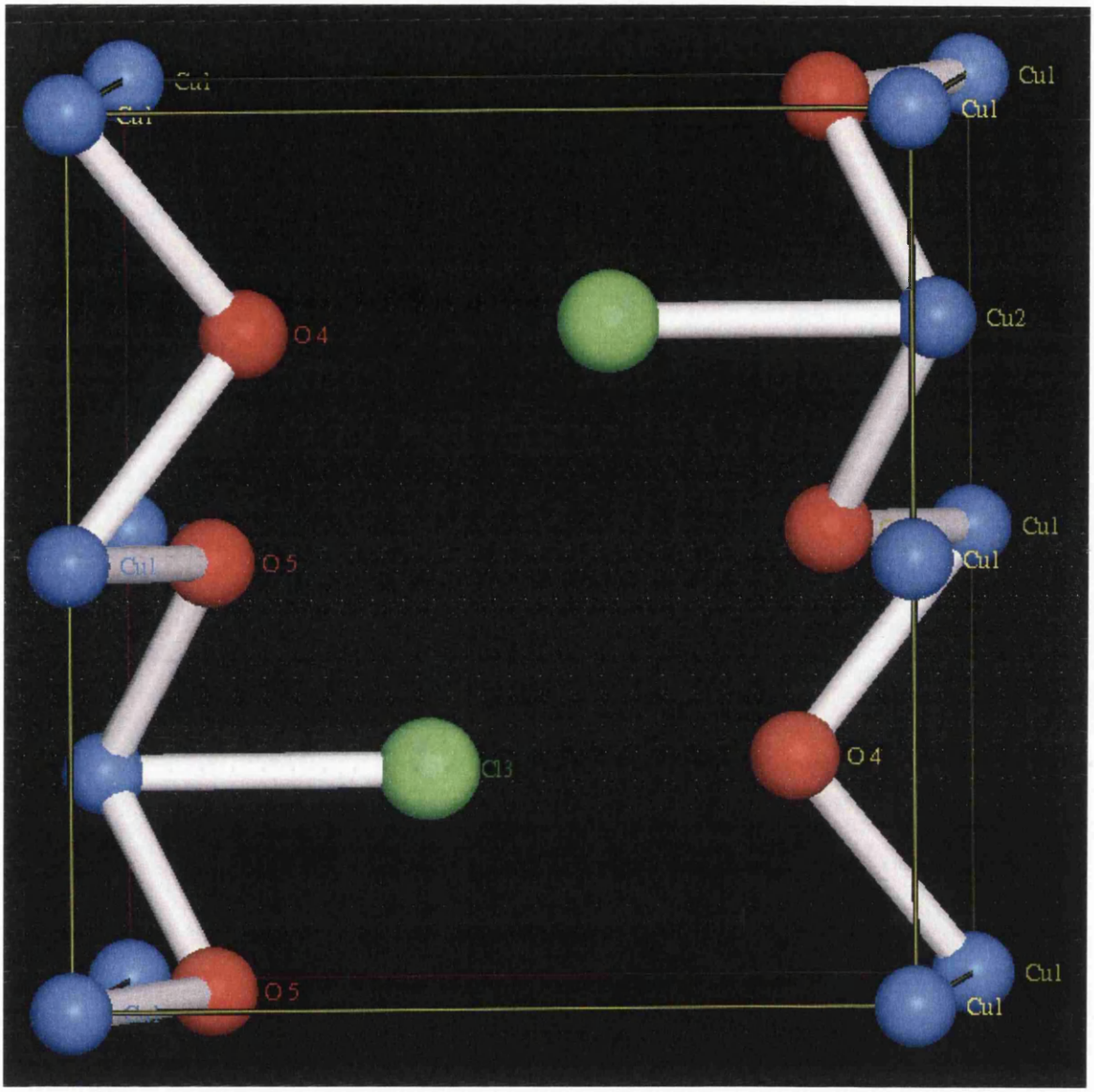
Table 5-12 Basic copper chloride bond lengths(Å)

Atom1	Atom2	Bond length (Å)	Atom1	Atom2	Bond length (Å)
Cu(1)	O(4)	2.27(7)	Cu(1)	O(5)	2.18(2)
Cu(2)	Cl(3)	2.26(5)	Cu(2)	O(4)	2.14(9)
Cu(2)	O(5)	1.97(4)	Cu(2)	O(5)	1.88(4)
Cu(2)	O(5)	1.97(4)	Cl(3)	O(4)	2.67(10)
Cl(3)	O(5)	2.45(4)	O(4)	O(5)	1.88(6)
O(5)	O(5)	2.32(4)			

Table 5-13 Basic copper chloride bond angles

Atom1	Atom2	Atom3	Angle(°)	Atom1	Atom2	Atom3	Angle(°)
O(4)	Cu(1)	O(4)	180	Cu(1)	O(4)	Cl(3)	138.(2)
O(4)	Cu(1)	O(5)	130.(2)	Cu(1)	O(4)	Cu(1)	85.(3)
O(4)	Cu(1)	O(5)	50.(2)	O(4)	Cl(3)	O(5)	142.7(10)
O(5)	Cu(1)	O(5)	179	Cu(2)	Cl(3)	O(5)	46.7(11)
Cl(3)	Cu(2)	O(5)	113.3(13)	Cu(2)	O(5)	O(5)	54.8(13)
O(4)	Cu(2)	O(5)	54.3(11)	Cu(2)	O(5)	O(5)	51.1(13)
O(4)	Cu(2)	O(5)	127.9(11)	Cu(2)	O(5)	Cl(3)	125.1(14)
O(5)	Cu(2)	O(5)	74.1(13)	Cu(2)	Cl(3)	O(4)	153.(2)
O(5)	Cu(2)	O(5)	107.(2)	Cu(1)	O(5)	O(5)	111.(2)
O(5)	Cu(2)	O(5)	173.(2)	O(5)	O(4)	O(5)	114.(5)
Cu(1)	O(4)	O(5)	132.(5)	O(5)	Cl(3)	O(5)	74.(2)
Cu(1)	O(4)	Cu(2)	91.(3)	O(5)	Cu(2)	O(5)	104.(2)
Cu(1)	O(4)	O(5)	63.(2)	Cl(3)	Cu(2)	O(5)	71.9(14)
Cu(2)	O(4)	Cl(3)	92.(3)	Cl(3)	Cu(2)	O(4)	119.(3)
Cl(3)	O(4)	O(5)	84.(3)	O(4)	O(5)	O(5)	118.(3)
Cu(1)	O(5)	Cu(2)	98.(2)	Cl(3)	O(5)	O(5)	95.8(15)
Cu(1)	O(5)	Cl(3)	137.(2)	Cl(3)	O(5)	O(4)	129.(3)
Cu(1)	O(5)	O(4)	67.(3)	Cu(2)	O(5)	O(4)	170.(4)
Cu(2)	O(5)	Cu(2)	106.(2)	Cu(2)	O(5)	Cl(3)	61.4(14)
Cu(2)	O(5)	O(4)	68.(3)				

Figure 5-5 Refined Basic copper chloride model



5.4.6 2,6-bis[4-(dimethylamino)benzylidene]cyclohexanone DMABC

The DMABC structure was solved by Voigt-Martin and co-workers in 1997[11] using a combination of approaches to ensure that the electron diffraction intensities were correct and that they would lead to a correct structure solution, and not lead to predicting the wrong properties for the molecule in question.

The compound crystallised in the orthorhombic space group $Cmc2_1$ with cell parameters a = 21.8575, b = 9.315, c = 9.6437Å, with 133 reflections recorded by electron diffraction. The intensities from the diffraction were compared to simulations and found to have very good agreement between the simulated and observed diffraction patterns, although a small amount of secondary scattering was detected for

the 100 and 002 reflections. although there is a possibility this is an artifact of the simulation procedure. No least squares refinement was attempted for this solution in the original paper.

The following are the results of a least squares refinement using SHELX-L. The structure refined successfully when the least squares matrix was damped, the temperature factors were held constant and bond restraints were added to all the bonds.

Table 5-14 DMABC atomic positions

Atom	x	y	z	U(iso)
O(1)	0.5000	0.2819	0.1339	0.050
N(13)	0.1380	0.3781	0.4050	0.050
C(2)	0.5000	0.3279	0.2552	0.050
C(3)	0.4412	0.3581	0.3264	0.050
C(4)	0.4434	0.3984	0.4782	0.050
C(5)	0.5000	0.3397	0.5486	0.0500
C(6)	0.3892	0.3408	0.2545	0.0500
C(7)	0.3253	0.3614	0.2952	0.0500
C(8)	0.2804	0.2877	0.2186	0.0500
C(9)	0.2195	0.2943	0.2529	0.0500
C(10)	0.1990	0.3752	0.3638	0.0500
C(11)	0.2429	0.4588	0.4345	0.0500
C(12)	0.3044	0.4479	0.4023	0.0500
C(14)	0.0931	0.2977	0.3279	0.0500
C(15)	0.1157	0.4880	0.4964	0.0500

Table 5-15 DMABC bond lengths (Å)

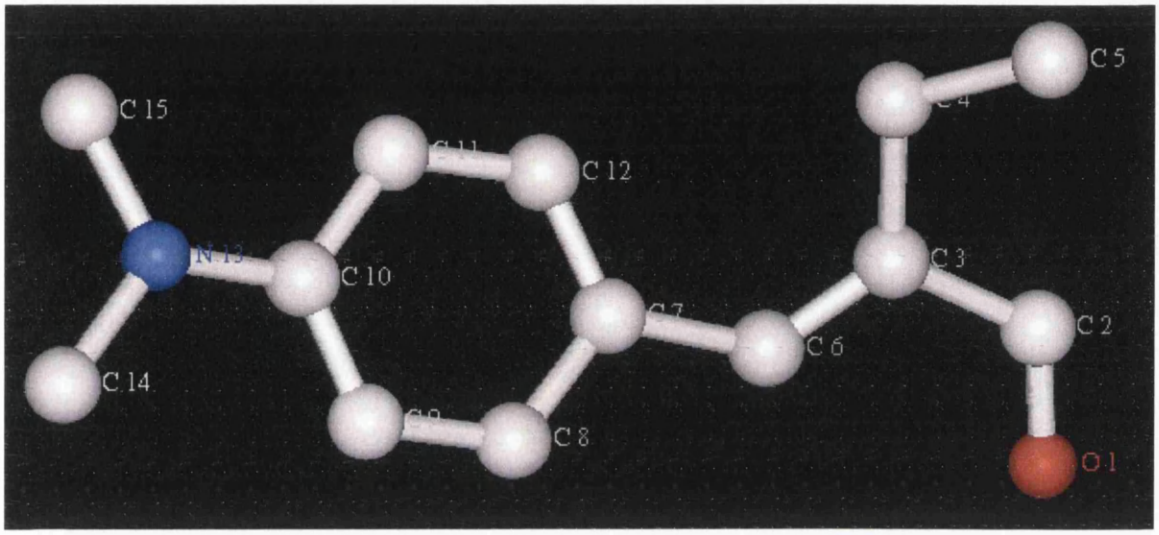
Atom1	Atom2	Bond length (Å)	Atom1	Atom2	Bond length (Å)
O(1)	C(2)	1.24	O(1)	C(2)	1.24
N(13)	C(10)	1.39	N(13)	C(14)	1.44
N(13)	C(15)	1.43	C(2)	C(3)	1.48
C(2)	C(3)	1.48	C(3)	C(4)	1.51
C(3)	C(6)	1.34	C(4)	C(5)	1.51
C(4)	C(5)	1.51	C(6)	C(7)	1.46
C(7)	C(8)	1.40	C(7)	C(12)	1.38
C(8)	C(9)	1.37	C(9)	C(10)	1.38
C(10)	C(11)	1.41	C(11)	C(12)	1.38

Table 5-16 DMABC Bond angles

Atom1	Atom2	Atom3	Angle (°)	Atom1	Atom2	Atom3	Angle (°)
C(10)	N(13)	C(14)	119.6	C(10)	N(13)	C(15)	120.9
C(14)	N(13)	C(15)	117.1	C(7)	C(12)	C(11)	121.9
O(1)	C(2)	C(3)	119.9	O(1)	C(2)	C(3)	119.9
O(1)	C(2)	C(3)	119.9	C(3)	C(2)	C(3)	119.9
C(2)	C(3)	C(4)	117.8	C(2)	C(3)	C(6)	118.1

C(2)	C(3)	C(4)	117.8	C(4)	C(5)	C(4)	109.6
C(4)	C(3)	C(6)	123.8	C(3)	C(4)	C(5)	111.7
C(3)	C(6)	C(7)	130.8	C(6)	C(7)	C(8)	117.4
C(6)	C(7)	C(12)	126.1	C(8)	C(7)	C(12)	116.3
C(7)	C(8)	C(9)	121.8	C(8)	C(9)	C(10)	121.6
N(13)	C(10)	C(9)	122.8	N(13)	C(10)	C(11)	120.1
C(9)	C(10)	C(11)	116.9	C(10)	C(11)	C(12)	120.7

Figure 5-6 Refined 2,6-bis[4-(dimethylamino)benzylidene]cyclohexanone model



5.4.7 4-dimethylamino-3-cyanobiphenyl (DMACB)

Voigt-Martin and co-workers solved the DMACB structure using Maximum entropy methods (MICE) in 1997[12]; there was no attempt to refine the proposed final structure. The crystal formed in the orthorhombic space group $Pna2_1$ a = 10.28, b = 22.64, c = 5.77Å, with 117 measured reflections. The data set has dynamical effects taking place, which was demonstrated by comparing the calculated and observed intensities. The 400 reflection was omitted on the grounds that it was too strong to be reasonably included in a direct methods program.

The following are results of a least squares refinement in SHELX-L, where the least squares matrix was damped and the temperature factors were fixed U_{ij}= 0.05, and bond restraints were added to all the bonds *via* the DFIX command.

Table 5-17 DMACB atomic positions

Atom	x	y	z	U(iso)
N(7)	0.4828	0.1834	0.5004	0.050
N(11)	0.2905	0.1839	-0.1277	0.050
C(1)	0.2656	0.0652	0.2456	0.050

C(2)	0.3355	0.1177	0.2572	0.050
C(3)	0.4146	0.1294	0.4652	0.050
C(4)	0.4238	0.0886	0.6616	0.050
C(5)	0.3538	0.0361	0.6500	0.050
C(6)	0.2747	0.0244	0.4420	0.050
C(8)	0.6103	0.1797	0.6268	0.050
C(9)	0.4359	0.2423	0.4522	0.050
C(10)	0.3146	0.1552	0.0545	0.050
C(12)	0.1932	-0.0310	0.4447	0.050
C(13)	0.0802	-0.0357	0.2898	0.050
C(14)	0.0062	-0.0866	0.2994	0.050
C(15)	0.0345	-0.1325	0.4557	0.050
C(16)	0.1479	-0.1306	0.6071	0.050
C(17)	0.2224	-0.0795	0.5997	0.050

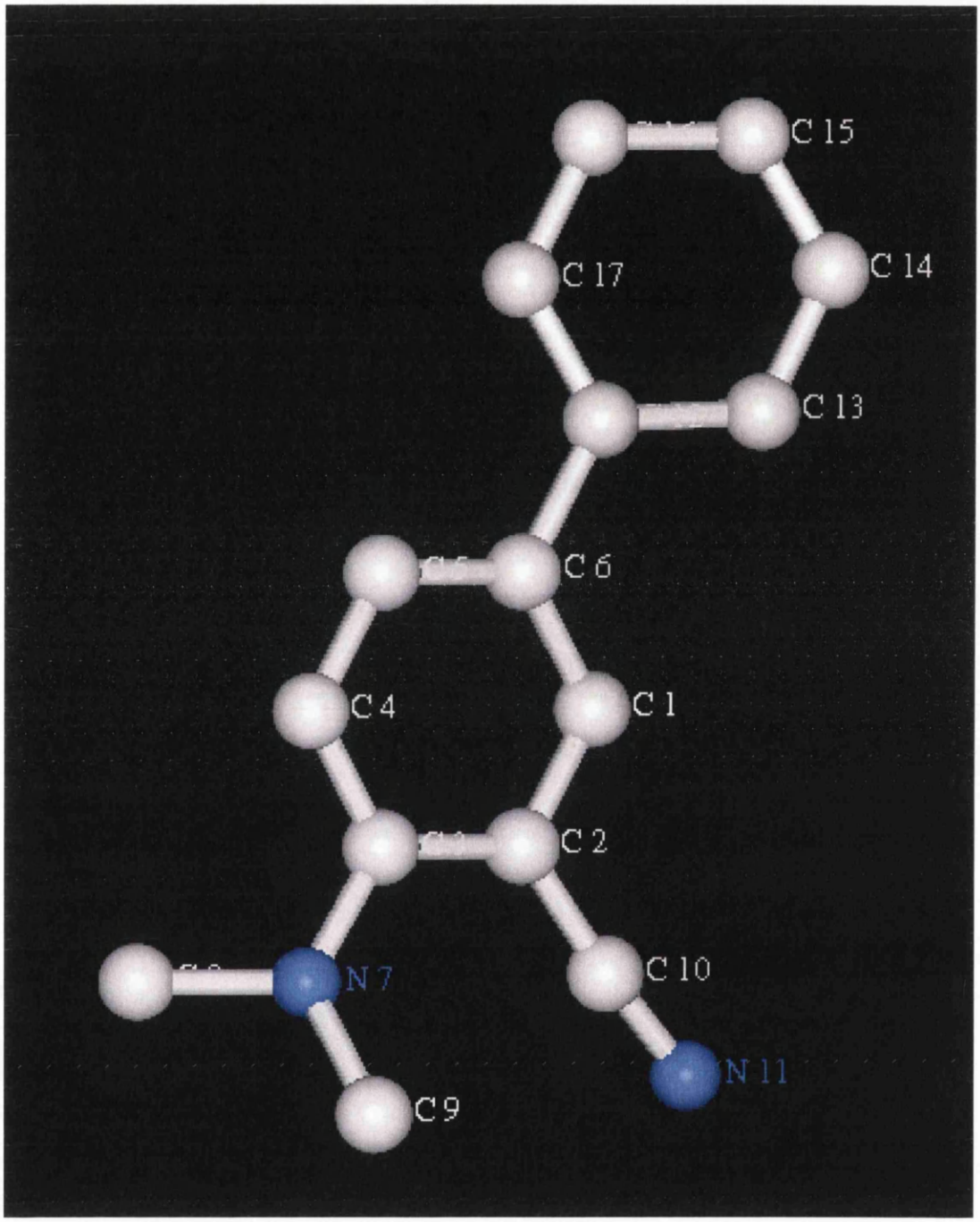
Table 5-18 DMACB bond lengths (Å)

Atom1	Atom2	Bond length (Å)	Atom1	Atom2	Bond length (Å)
N(7)	C(3)	1.42	N(7)	C(9)	1.44
N(11)	C(10)	1.18	C(1)	C(2)	1.39
C(1)	C(6)	1.39	C(2)	C(3)	1.39
C(2)	C(10)	1.38	C(3)	C(4)	1.39
C(4)	C(5)	1.39	C(5)	C(6)	1.39
C(6)	C(12)	1.51	C(12)	C(13)	1.42
C(12)	C(17)	1.40	C(13)	C(14)	1.38
C(14)	C(15)	1.35	C(15)	C(16)	1.41
C(16)	C(17)	1.38			

Table 5-19 DMACB bond angles

Atom1	Atom2	Atom3	Angle (°)	Atom1	Atom2	Atom3	Angle (°)
C(3)	N(7)	C(9)	127.4	C(2)	C(1)	C(6)	119.9
C(1)	C(2)	C(3)	120.0	C(1)	C(2)	C(10)	114.2
C(3)	C(2)	C(10)	125.7	N(7)	C(3)	C(2)	123.7
N(7)	C(3)	C(4)	116.1	C(2)	C(3)	C(4)	119.9
C(3)	C(4)	C(5)	119.9	C(4)	C(5)	C(6)	120.0
C(1)	C(6)	C(5)	119.9	C(1)	C(6)	C(12)	121.4
C(5)	C(6)	C(12)	118.4	N(11)	C(10)	C(2)	174.7
C(6)	C(12)	C(13)	120.6	C(6)	C(12)	C(17)	122.5
C(13)	C(12)	C(17)	116.7	C(12)	C(13)	C(14)	119.4
C(13)	C(14)	C(15)	122.8	C(14)	C(15)	C(16)	119.7
C(15)	C(16)	C(17)	117.6	C(12)	C(17)	C(16)	123.4

Figure 5-7 Refined 4-dimethylamino-3-cyanobiphenyl model



5.4.8 Mannan I

The structure of Mannan I was solved using electron diffraction by Chanzy and co-workers in 1987[13]. The crystal formed in the orthorhombic space group is $P2_12_12_1$ with unit cell dimensions a = 8.92, b = 7.21, c = 10.27Å. The data set has 58 unique reflections and 12 atoms in the asymmetric unit. The original paper made no attempt to refine the structure using least squares methods.

The following tables are results from a least squares calculation in SHELX-L, where the least squares matrix has been damped by the DAMP 1000 command and the temperature factors are held constant at $U_{ij} = 0.05$. The bond lengths of the carbon-carbon and carbon-oxygen bonds are restrained using the DFIX command.

Table 5-20 Mannan I atom positions

Atom	x	y	z	U(iso)
O(7)	-0.0120	-0.0960	0.6300	0.050
O(8)	0.1230	-0.2020	0.4000	0.050
O(9)	0.1740	-0.1110	0.8710	0.050
O(10)	0.2330	0.1220	0.7180	0.050
O(11)	0.3280	0.4550	0.7180	0.050
C(1)	0.2230	-0.0720	0.7440	0.050
C(2)	0.1340	-0.1660	0.6340	0.050
C(3)	0.2130	-0.1290	0.5040	0.050
C(4)	0.2640	-0.0780	1.0160	0.050
C(5)	0.3160	0.1610	0.6020	0.050
C(6)	0.3310	0.3700	0.5930	0.050

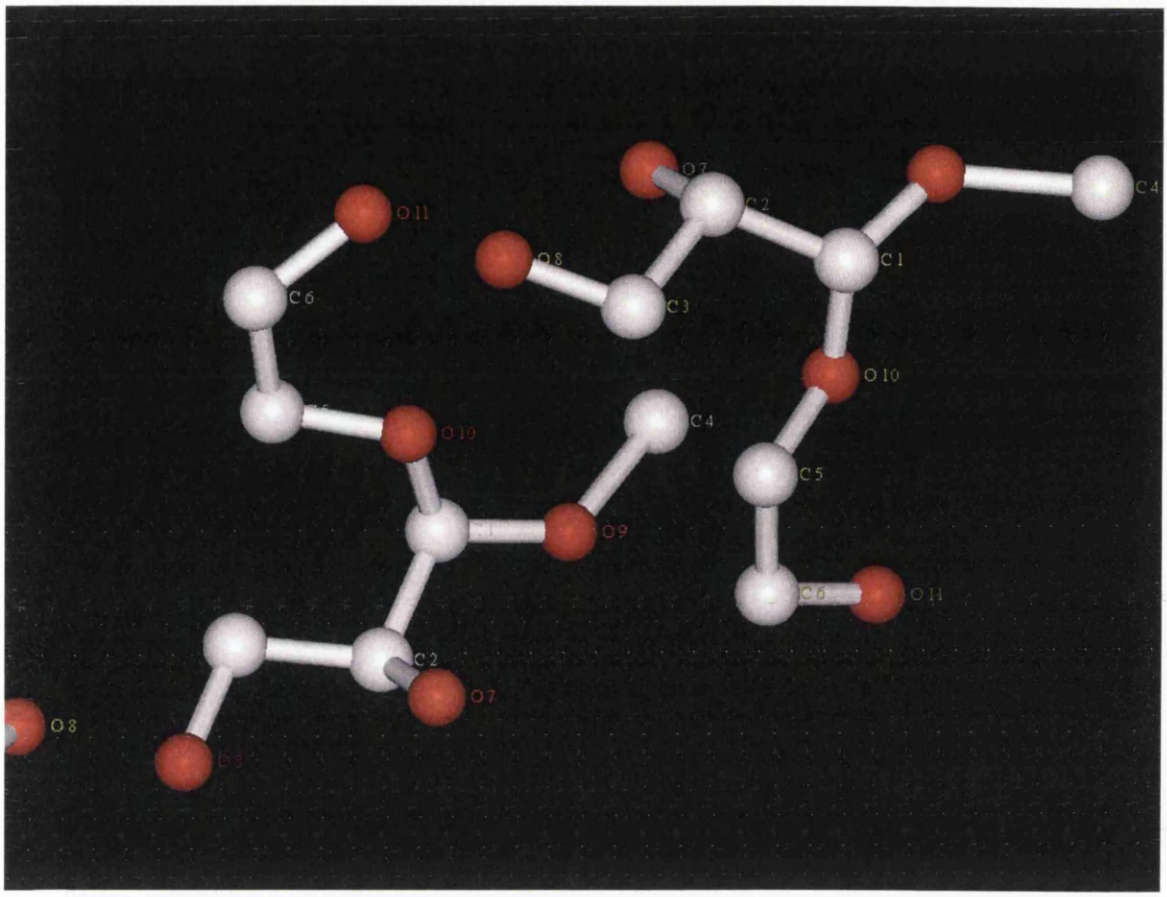
Table 5-21Mannan I bond lengths (Å)

Atom1	Atom2	Bond length (Å)	Atom1	Atom2	Bond length (Å)
O(7)	C(2)	1.39	O(9)	C(1)	1.40
O(10)	C(1)	1.42	O(11)	C(6)	1.42
C(1)	C(2)	1.53	C(2)	C(3)	1.53
C(3)	C(4)	1.51	C(4)	C(5)	1.28
C(5)	C(6)	1.51			

Table 5-22 Mannan I bond angles

Atom1	Atom2	Atom3	Angle (°)	Atom1	Atom2	Atom3	Angle (°)
O(9)	C(1)	O(10)	112.9	O(9)	C(1)	C(2)	115.6
O(10)	C(1)	C(2)	109.0	O(7)	C(2)	C(1)	110.0
O(7)	C(2)	C(3)	109.8	C(1)	C(2)	C(3)	109.0
C(2)	C(3)	C(4)	99.3	C(3)	C(4)	C(5)	126.2
C(4)	C(5)	C(6)	118.0	O(11)	C(6)	C(5)	111.8

Figure 5-8 Refined Mannan I model



5.4.9 Copper Perchlorophthalocyanine

This structure was solved by Dorset and co-workers[14] using electron diffraction and the tangent formula. The crystal formed in the space group C2/c, with unit cell dimensions a = 19.62Å, b = 26.08Å, c = 3.79Å, there are 198 unique reflections in the data set. The final structure was Fourier refined to a final R-factor of 28% for all data with an isotropic B = 3.0Å^2, although there was no least squares refinement attempted.

The following tables are the results from a least squares refinement in the program SHELX-L, The solution required the least squares matrix to be damped with the instruction DAMP 1000.

Table 5-23 Copper Perchlorophthalocyanine atom positions

Atom	x	y	z	U(iso)
Cu(1)	0.000(2)	0.0000(18)	0.0000	0.098(15)
Cl(2)	0.077(2)	0.3067(18)	0.0000	0.071(15)
Cl(3)	0.164(2)	0.2078(18)	0.0000	0.055(15)
Cl(4)	0.272(2)	0.1159(18)	0.0000	0.048(15)
Cl(5)	0.411(2)	0.0584(18)	0.0000	0.028(15)
N(6)	0.000(2)	0.0691(18)	0.0000	0.023(15)
N(7)	0.094(2)	0.0000(18)	0.0000	0.017(15)
N(8)	0.120(2)	0.0924(18)	0.0000	0.021(15)
C(9)	0.136(2)	0.0423(18)	0.0000	0.022(15)
C(10)	0.213(2)	0.0283(18)	0.0000	0.052(15)
C(11)	0.279(2)	0.0586(18)	0.0000	0.056(15)
C(12)	0.341(2)	0.0220(18)	0.0000	0.008(15)
C(13)	0.058(2)	0.1064(18)	0.0000	0.029(15)
C(14)	0.035(2)	0.1594(18)	0.0000	0.015(15)
C(15)	0.071(2)	0.2030(18)	0.0000	0.028(15)
C(16)	0.035(2)	0.2514(18)	0.0000	0.036(15)

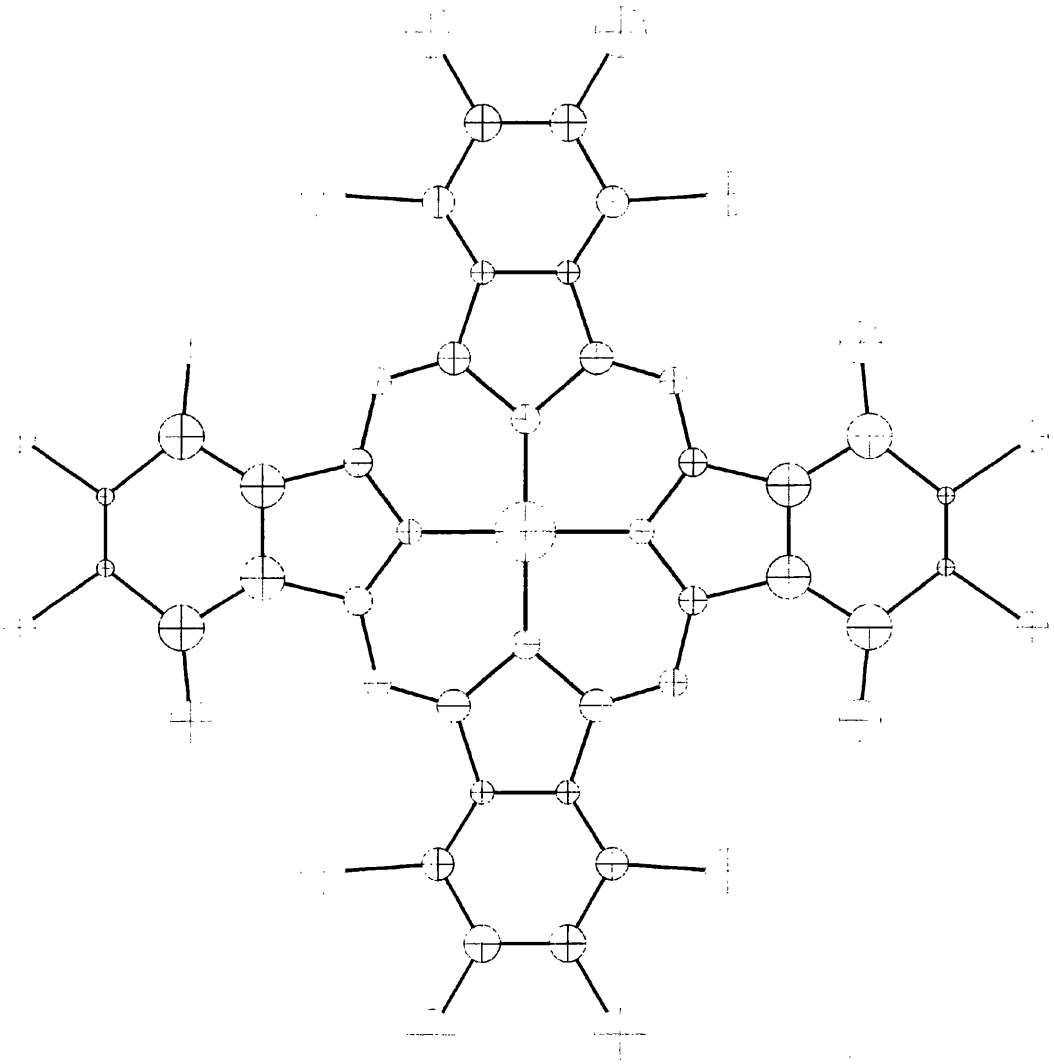
Table 5-24 Copper Perchlorophthalocyanine bond lengths (Å)

Atom1	Atom2	Bondlength(Å)	Atom1	Atom2	Bondlength(Å)
Cu(1)	N(6)	1.80(7)	Cu(1)	N(6)	1.80(7)
Cu(1)	N(6)	1.80(7)	Cu(1)	N(6)	1.80(7)
Cu(1)	N(7)	1.85(6)	Cu(1)	N(7)	1.85(6)
Cu(1)	N(7)	1.85(6)	Cu(1)	N(7)	1.85(6)
Cl(2)	C(16)	1.67(6)	Cl(4)	C(11)	1.50(7)
Cl(5)	C(12)	1.67(6)	N(7)	C(9)	1.37(6)
N(7)	C(9)	1.37(6)	N(8)	C(9)	1.34(7)
N(8)	C(13)	1.27(6)	C(10)	C(10)	1.47(7)
C(10)	C(11)	1.52(6)	C(11)	C(12)	1.54(6)
C(12)	C(12)	1.15(7)	C(13)	C(14)	1.45(7)
C(14)	C(14)	1.36(6)	C(14)	C(15)	1.34(6)
C(15)	C(16)	1.45(6)	C(16)	C(16)	1.36(6)

Table 5-25 Copper Perchlorophthalocyanine bond angles

Atom1	Atom2	Atom3	Angle(°)	Atom1	Atom2	Atom3	Angle(°)
N(6)	Cu(1)	N(6)	180.0	N(6)	Cu(1)	N(7)	90.(3)
N(7)	Cu(1)	N(7)	180.0	Cu(1)	N(7)	C(9)	126.(4)
C(9)	N(7)	C(9)	107.(4)	C(9)	N(8)	C(13)	120.(4)
N(7)	C(9)	N(8)	130.(4)	C(15)	C(16)	C(16)	119.(4)
C(10)	C(10)	C(11)	121.(4)	Cl(4)	C(11)	C(10)	116.(3)
Cl(4)	C(11)	C(12)	134.(4)	C(10)	C(11)	C(12)	110.(4)
Cl(5)	C(12)	C(11)	107.(4)	Cl(5)	C(12)	C(12)	125.(4)
C(11)	C(12)	C(12)	128.(4)	N(8)	C(13)	C(14)	125.(4)
C(13)	C(14)	C(14)	108.(4)	C(13)	C(14)	C(15)	130.(4)
C(14)	C(14)	C(15)	122.(4)	C(14)	C(15)	C(16)	119.(4)
Cl(2)	C(16)	C(15)	121.(3)	Cl(2)	C(16)	C(16)	120.(4)

Figure 5-9 Refined Copper Perchlorophthalocyanine model



5.4.10 Isotactic Poly(1-butene) Form III

The data set for isotactic poly(1-butene) Form III was obtained by electron diffraction by Dorset and co-workers in 1994[15]. They obtained the data set from solution and epitaxially grown crystals. A tilt series was used to collect as much three-dimensional data as possible. The orthorhombic space group is $P2_12_12_1$ with unit cell dimensions a = 12.38, b = 8.88, c = 7.56Å. The data set has 125 unique reflections. A full-matrix least squares calculation was attempted, with a fixed B= 4.0Å^2 and a damping factor of the shifts, but no geometrically reasonable structure could be found, even though the R-factors were reduced during the calculations. This was attributed to three large intensities, which dynamical diffraction was believed to have caused the intensity values to be too high. Upon remeasuring some reflections, a further three reflections were found to be too weak, also attributed to dynamical diffraction problems. These six reflections were removed from the data set[20]. Fourier refinement was found to produce a stable refinement upon removal of the problematic reflections.

The following tables are the results from a least squares refinement in the program SHELX-L. The solution required the least squares matrix to be damped and the temperature factors to be held constant, with U_{ij} = 0.05. The carbon-carbon bond lengths were all restrained with the DFIX command. After the first cycle of refinement two reflections were removed from the data set because they were making the refinement unstable and could easily be attributed to dynamical effects.

Table 5-26 Poly(1-butene) Form III atom positions

Atom	x	y	z	U(iso)
C(1)	0.3460	-0.0380	0.1080	0.050
C(2)	0.2980	0.1250	0.1630	0.050
C(3)	0.3810	0.2480	0.1030	0.050
C(4)	0.3250	0.4020	0.1200	0.050
C(5)	0.2790	0.1340	0.3580	0.050
C(6)	0.3400	-0.0670	-0.0870	0.050
C(7)	0.4280	-0.1830	-0.1470	0.050
C(8)	0.5380	-0.1050	-0.1300	0.050

Table 5-27 Poly(1-butene) Form III bond lengths (Å)

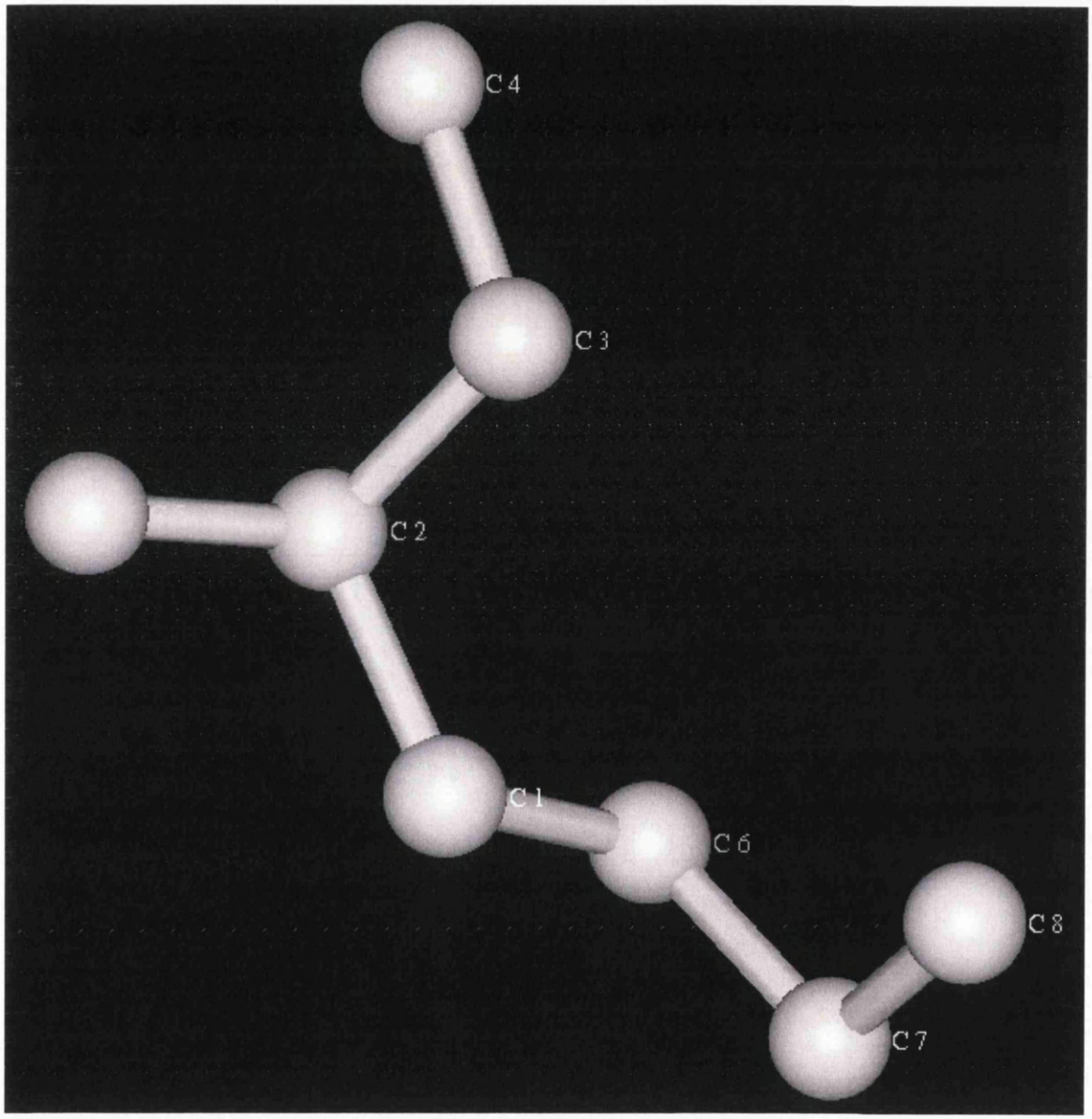
Atom1	Atom2	Bond length (Å)	Atom1	Atom2	Bond length (Å)
C(1)	C(6)	1.49	C(2)	C(5)	1.49
C(3)	C(4)	1.53	C(7)	C(8)	1.53

Table 5-28 Poly(1-butene) Form III bond angles

Atom1	Atom2	Atom3	Angle (°)	Atom1	Atom2	Atom3	Angle (°)

C(4)	C(3)	C(5)	116.2	C(3)	C(4)	C(6)	116.2
O(2)	C(6)	C(4)	170.4	O(1)	C(7)	C(8)	128.2
C(7)	C(8)	C(8)	110.8				

Figure 5-10 Refined Isotactic Poly(1-butene) Form III model



5.4.11 Poly(_-caprolactone)

The structure was solved by electron diffraction in 1990 by Dorset and co-workers[16]. The structure crystallised in the orthorhombic space group $P2_12_12_1$, a = 7.48Å, b = 4.98Å, c = 17.26Å, a set of 47 reflections was recorded. No attempt to refine the structure was made by Fourier or least squares refinement.

The following tables are the results from a least squares refinement in the program SHELX-L. The refinement solution required damping of the least squares matrix with the DAMP 1000 command, and fixed temperature factors of $U_{ij} = 0.05$. The bond lengths of the carbon-carbon bonds and the carbon-oxygen bonds are all restrained.

Table 5-29 Poly(_-caprolactone)atom positions

Atom	x	y	z	U(iso)
O(7)	0.3960	0.2050	0.1340	0.050
O(8)	0.1970	-0.0320	0.2020	0.050
C(1)	0.2910	0.0320	0.1340	0.050
C(2)	0.2660	0.0970	0.2720	0.050
C(3)	0.1970	-0.0410	0.3440	0.050
C(4)	0.2660	0.0840	0.4140	0.050
C(5)	0.1970	-0.0350	0.4890	0.050
C(6)	0.2220	-0.0810	0.0630	0.050

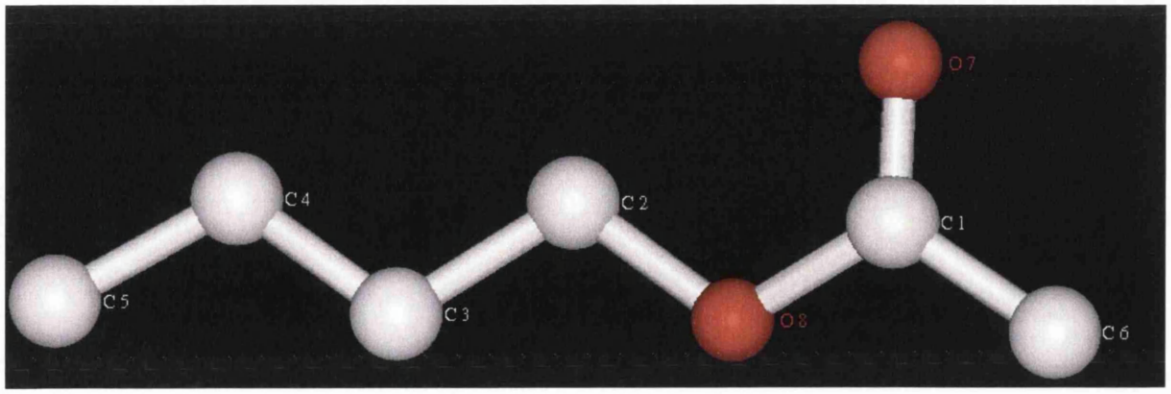
Table 5-30 Poly(e-caprolactone) bond lengths (Å)

Atom1	Atom2	Bond length (Å)	Atom1	Atom2	Bond length (Å)
O(7)	C(1)	1.16	O(8)	C(1)	1.40
C(1)	C(6)	1.44	C(2)	C(3)	1.51
C(3)	C(4)	1.45	C(4)	C(5)	1.51
C(5)	C(6)	1.52			

Table 5-31 Poly(e-caprolactone) bond angles

Atom1	Atom2	Atom3	Angle (°)	Atom1	Atom2	Atom3	Angle (°)
O(7)	C(1)	O(8)	120.3	O(7)	C(1)	C(6)	121.9
O(8)	C(1)	C(6)	116.2	C(2)	C(3)	C(4)	111.5
C(3)	C(4)	C(5)	114.9	C(4)	C(5)	C(6)	115.5
C(1)	C(6)	C(5)	114.8				

Figure 5-11 Refined Poly(e-caprolactone) model



5.4.12 Polyethylene

The structure was determined by electron diffraction by Dorset and co-workers in 1991[17]. The compound crystallised in the orthorhombic space group Pnma, with unit cell dimensions a = 7.48Å, b = 2.55Å, c = 4.97Å, 50 reflections were recorded. No attempt was made to the refine this structure by any method.

The following tables are the results from a least squares refinement in the program SHELX-L. the refinement solution required damping of the least squares matrix, using the DAMP 1000 command, holding the temperature factors constant at U_{ij} = 0.05, and the carbon-carbon bond lengths were restrained.

Table 5-32 Polyethylene atom positions

Atom	x	y	z	U(iso)
C(1)	0.0477	0.2500	0.0617	0.050
H(1A)	0.1930	0.2500	0.0506	0.050
H(1B)	0.0371	0.2500	0.2593	0.050

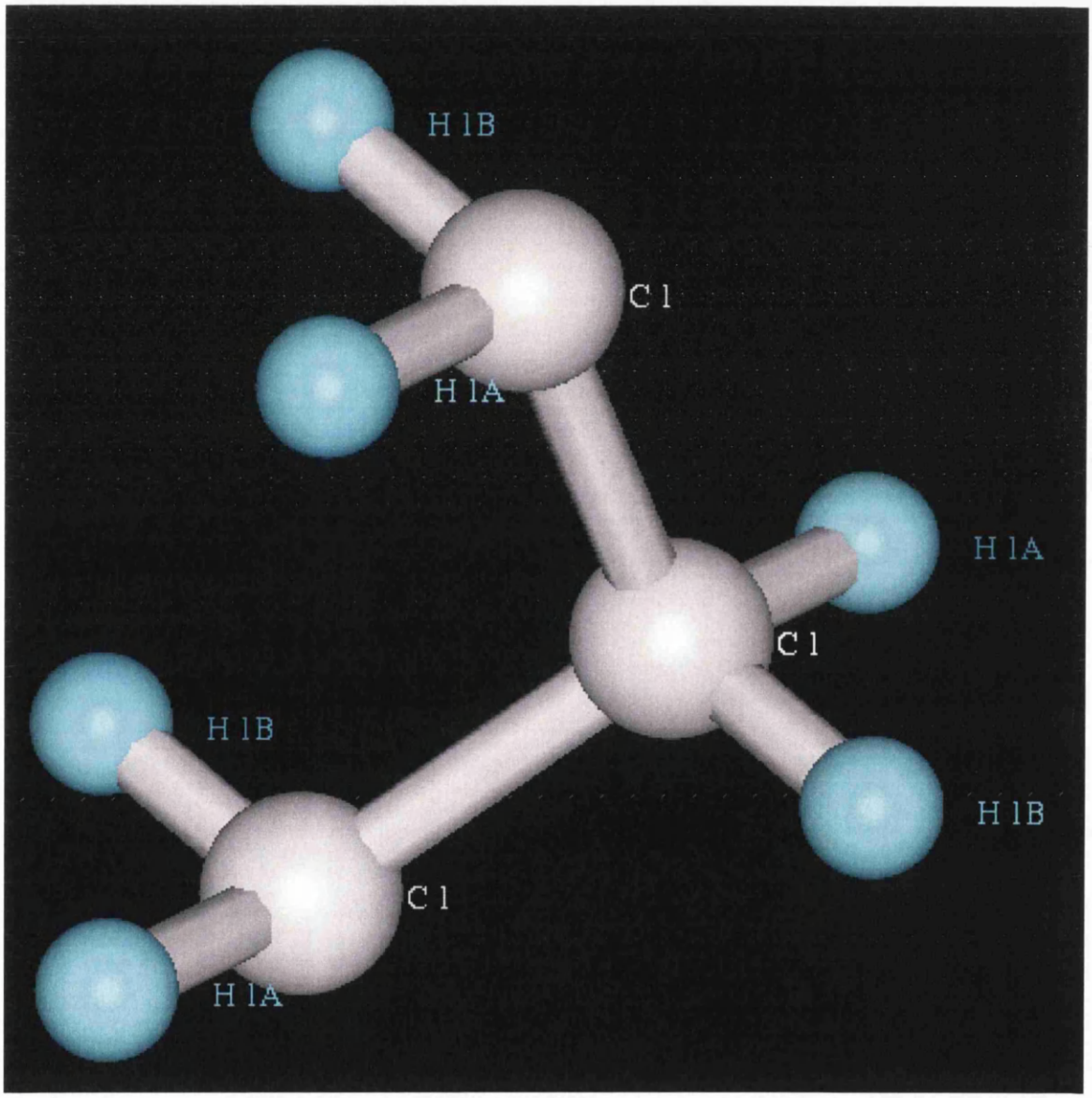
Table 5-33 Polyethylene bond lengths (Å)

Atom1	Atom2	Bond length (Å)	Atom1	Atom2	Bond length (Å)
C(1)	C(1)	1.58	C(1)	H(1A)	1.08
C(1)	H(1B)	0.99			

Table 5-34 Polyethylene bond angles

Atom1	Atom2	Atom3	Angle (°)
C(1)	C(1)	C(1)	107.5
C(1)	C(1)	H(1B)	110.0
C(1)	C(1)	H(1A)	115.7
H(1A)	C(1)	H(1B)	97.1

Figure 5-12 Refined Polyethylene model



5.4.13 Poly(1,4,trans-cyclohexanediyl dimethylene succinate) (T-cds)

The three-dimensional electron diffraction data set was published by Brisse and co-workers in 1984[18]. The monoclinic space group is $P2_1/n$ with unit cell dimensions a =

6.48, b = 9.48, c = 13.51Å, β = 45.9° The data set has 87 observed reflections. No attempt was made to refine this structure.

The following tables are the results from a least squares refinement in the program SHELX-L. The refinement required damping of the least squares matrix, holding the temperature factors constant at B = 0.05 $Å^2$, and bond restraints on the carbon-oxygen bonds.

Table 5-35 T-cds atom positions

Atom	x	y	z	U(iso)
O(1)	0.4750	0.3080	0.1260	0.050
O(2)	0.5500	0.3080	0.2050	0.050
C(3)	0.3380	0.5580	0.4740	0.050
C(4)	0.5750	0.4600	0.3690	0.050
C(5)	0.6250	0.3590	0.4460	0.050
C(6)	0.5380	0.3790	0.2930	0.050
C(7)	0.5000	0.4230	0.1270	0.050
C(8)	0.5250	0.5310	0.0420	0.050

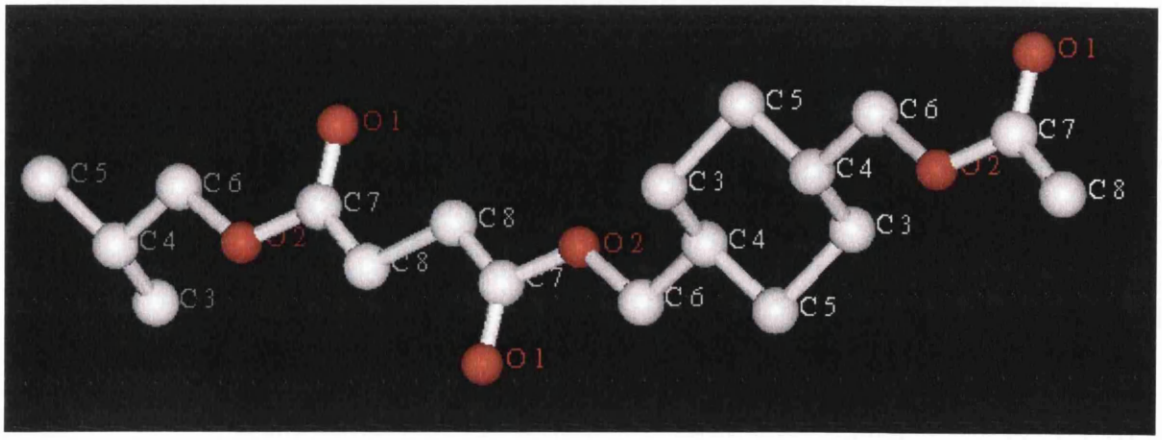
Table 5-36 T-cds bond lengths (Å)

Atom1	Atom2	Bond length (Å)	Atom1	Atom2	Bond length (Å)
O(1)	C(7)	1.10	O(2)	C(6)	1.32
C(3)	C(4)	1.48	C(3)	C(5)	1.48
C(4)	C(6)	1.42	C(7)	C(8)	1.46

Table 5-37 T-cds bond angles

Atom1	Atom2	Atom3	Angle (°)	Atom1	Atom2	Atom3	Angle (°)
C(4)	C(3)	C(5)	116.2	C(3)	C(4)	C(6)	116.2
O(2)	C(6)	C(4)	170.4	O(1)	C(7)	C(8)	128.2
C(7)	C(8)	C(8)	110.8				

Figure 5-13 Refined Poly(1,4,trans-cyclohexanediyl dimethylene succinate) model



5.5 Refinement discussion

The twelve examples of refinement of electron diffraction data were all refineable without exception. There is a definite trend in the difficulty of the refinement, namely that the lower the data parameter ratio the more difficult the refinement becomes. In cases where the data parameter ratios were relatively high, as in brucite and copper perchlorophthalocyanine, the refinement was stable and the concept of refinement is akin to X-ray refinement, but with slightly less well determined data.

Structures such as CNBA, DMABC, and DMACB all suffered from lack of data down one axis and a poor data parameter ratio, hence the large number of restraints that had to be used while refining these structures. Another structure that suffers from lack of data down a particular axis is polybut-1-ene form III. The problem of under-sampling down a particular axis causes problems with the refinement especially if the data parameter ratios are already low, there is then a danger that an incorrect solution may be made to fit a small set of reflections by adding more and more parameters. Hence a stepwise approach to the refinement process can only be used if, it has been determined that the structure is essentially correct, and the limiting factor to the solution is under sampling down one axis. It would be prudent to point out that because temperature factors are one of the first things to be held fixed in the refinement process, an atom could misassigned. While this is possible within the refinement process, microscopists have other tools available to them that can be used to determine atom types in a sample and where they occur, such as electron loss spectroscopy, or energy dispersive X-ray experiments.

Some of the structures have estimated standard deviations (esds) and others do not, this arises from whether or not the temperature factors have been held constant or not. If the temperature factors have been fixed then no esds are calculated by the program, this is also the case when we have atoms lieing on special positions. If esd's were to be calculated they would have no real meaning because the fixed temperature factors are distorting the calculated values, and would potentially mask or any errors that would be highlighted by the errors on the atom positions, bond lengths and angles.

The R-factors for all the structures except one are below 30% indicating that the final proposed models are correct and are close to 0.1Å of the final structure. This is as close as can be reasonably expected for electron diffraction experiments.

The one structure that is over the 30% threshold is the aluminium-iron alloy. It is known from the original paper that some of the data is dynamical, and, will therefore affect the R-factor making it seem larger than it actually is. This arises because some reflections will have much larger measured values than the calculated ones. The data set is also large for an electron diffraction data set so there will be a large an accumulation of small errors as well.

The Si(111) surface structure was also refined for the purposes of this study to an R-factor of 7%, but all the atoms sit on special positions, so it was felt that to include the structure would be a misleading example of how well electron diffraction data can be refined.

5.6 Refinement conclusions

The refinement of electron diffraction data is a legitimate method of structure validation, although more care needs to be taken than when using its X-ray counterpart. The refinement of electron diffraction structures should become a routine part of structure solution. A further improvement to the methodology could be to use maximum likelihood refinement and singular value decomposition to refine the structures when $\sigma(|F|)$ values are available. Currently there is a lack of data available with σ values to test this method.

References

(1) Dorset, D.L. *Structural Electron Crystallography*; Plenum Publishing Coorporation: New York, **1995**.

(2) Watkin, D. *Acta Crystallographica* **1994**, *A50*, 411-437.

(3) Dunitz, J.D.; Seiler, P. *Acta Crystallographica* **1973**, *B29*, 589-595.

(4) Jansen, J.; Tang, D.; Zandbergen, H.; Schenk, H. *Acta Crystallographica* **1998**, *A54*, 91-101.

(5) Press, W.H.; Teukolsky, S.A.; Vetterling, W.T.; Flannery, B.P. *Numerical Recipices in C*; Cambridge University Press: **1996**.

(6) Shelx97. Program for the refinement of crystal structures, University of Göttingen, Germany. Sheldrick, G.M. **1997**.

(7) Gjønnes, J.; Hansen, V.; Berg, B.S.; Runde, P.; Cheng, Y.F.; Dorset, D.L.; Gilmore, C.J. *Acta Crystallographica* **1998**, *A54*, 306-319.

(8) Zhukhlistov, A.P.; Zvyagin, B.B.; Avilov, A.S.; Ferraris, D.; Poltnikov, V.P. *Crystallography Reports* **1997**, *42*, 774-777.

(9) Voigt-Martin, I.G.; Yan, D.H.; Yakimansky, A.; Schollmeyer, D.; Gilmore, C.J.; Bricogne, G. *Acta Crystallographica* **1995**, *A51*, 849-868.

(10) Voronova, A.A.; Vainshtein, B.K. *Soviet Physics--Crystallography* **1958**, *3*, 445-451.

(11) Yakimansky, A.; Kolb, U.; Matveeva, G.N.; Voigt-Martin, I.G.; Tenkovtsev, A.V. *Acta Crystallographica* **1997**, *A53*, 603-614.

(12) Voigt-Martin, I.G.; Zhang, Z.X.; Kolb, U.; Gilmore, C.J. *Ultramicroscopy* **1997**, *68*, 43-59.

(13) Chanzy, H.; Pérez, S.; Miller, D.P.; Paradossi, G.; Winter, W.T. *Macromolecules* **1987**, *20*, 2407-2413.

(14) Dorset, D.L.; Tivol, W.F.; Turner, J.N. *Ultramicroscopy* **1991**, *38*, 41-45.

(15) Dorset, D.L.; McCourt, M.P.; Kopp, S.; Wittmann, J.C.; Lotz, B. *Acta Crystallographica* **1994**, *B50*, 201-208.

(16) Hu, H.; Dorset, D.L. *Macromolecules* **1990**, *23*, 4606-4607.

(17) Dorset, D.L.; Hu, H.; Jäger, J. *Acta Crystallographica* **1991**, *A47*, 543-549.

(18) Brisse, F.; Rémillard, B.; Chanzy, H. *Macromolecules* **1984**, *17*, 1980-1987.

(19) 'maXus: Integrated Crystallography Software' S.Mackay, C.Edwards, C.J. Gilmore, N.Stewart, A.Donald & K.Shankland *Holland: Nonius, Japan: MacScience*, **1997**)

(20) Dorset, D. L.; Gilmore, C. *Acta Crystallographica* **2000,** A56, 62-67.

6 THE $P(\delta q)$ FUNCTION

6.1 The $P(\delta q)$ Function

The $P(\delta q)$ function (pdq) has been apart of the Mice program from its inception[1,2]. Although potentially a highly effective function, it has been mainly ignored because of its lack of discrimination between nodes. The figures of merit used by the function are crude and quick to calculate, and although crude its strength lies in the ability to quickly identify promising phase combinations having established a current maximum entropy map $q^{ME}(\underline{x})$.

The $P(\delta q)$ function works *via* the EXTEND command in MICE. The basis set of reflections is used to generate a maximum entropy map $q^{ME}(\underline{x})$. The NEXT command is used to find a set of reflections {E} to add to the basis set {H}. Phases of the Reflections in the set {E} are permuted a phase. Each set of permuted reflections {E} are added to the set {H}, to give $E\bigcup H$. We now compute

$$\delta q(\underline{x}) \text{as the FT of} \left|U_{\underline{h}\in H}\right|^{obs} - \left|U_{\underline{h}\in H}\right|^{ME} \quad \textbf{(6-1)}$$

where the $\left|U_{\underline{h}\in H}\right|^{ME}$ is found from a Fourier transform of the current version of $q^{ME}(\underline{x})$. The calculation

$$P(\delta q) = -\int_V \frac{\delta q^2(\underline{x})}{q^{ME}(\underline{x})} d^3x \quad \textbf{(6-2)}$$

is the performed.

This process should highlight new features in the maximum entropy maps where $q^{ME}(\underline{x})$ is large and not where it is small. The $P(\delta q)$value is then used as a rough guide to which nodes should be kept and which ones should be thrown away.Solutions with a minimum value of $P(\delta q)$ are preffered. The lower the differences the more likely a correct solution will be calculated in a full maximum entropy calculation. A FILTER command allows the user to set a value at which the selection of the nodes that are to be kept is set.

The $P(\delta q)$ has certain limitations in its use: $q^{ME}(\underline{x})$ must have developed sufficient detail to allow sufficient contrast between the U-magnitudes, otherwise the $P(\delta q)$ values will appear almost constant. In addition the reflections used in $q^{ME}(\underline{x})$ should be chosen on the same basis as the NEXT command, in that they should optimally increase the second-neighbourhood.

There are four methods currently for phase permutation with this method: Magic integers (MI)[3] permutations which do not evenly sample phase space. Quadrant permutation has the obvious disadvantages of a quick build up of phase error, due to the crude sampling. As an alternative to the magic integers, error-correcting codes[4] (E.C.C.s) are excellent. They span phase space efficiently and are evenly spaced, therefore avoiding the problems of MI permutation. The principle of quadrant fixing is now being applied in a more economic manner. With the Golay code there will be one solution with a maximum of 4 wrong phases in one of the 4096 sets generated, or 3 if 23 degrees are freedom are used.

The use of error-correcting codes with the $P(\delta q)$function could help improve the discrimination between nodes in the calculation, through the more efficient methods of spanning phase space. Another aid to the process of making the nodes more distinguishable is to add an analysis process to the results, for looking at phase relationships and asking whether they can be analysed in the same way as LLGs.

6.2 Experimental

The $P(\delta q)$ function was linked to the error correcting codes as a source of phase permutation. Then an analysis package was added to assess the quality of the results in terms of triplet phases and phase relationships.

The codes assign phases in a straightforward manner, for centric phases one code bit is used, 0 or 1. 0 = 0° and 1= π, for reflections restricted to 0 or π.

Acentric reflections require two bits to assign a quadrant, one bit assign the real and the other bit assigns the imaginary part, *i.e.* 0,0 = π/4; 1,0 = 3π/4; 1,1 = 5 π/4; 0,1 = 7 π/4.

The analysis process consists of calculating a Student t-test to measure the significance of a particular phase choice, The level of significance is then used to

calculate a weight that is then used to assign a score to a given basis set. The score is identical to the score that is calculated in the maximum likelihood analysis except that we now use the $P(\delta q)$ value instead of LLG.

. The score essentially counts the number of significant reflections in a given set and uses that number to calculate a relative weight *via* a ratio of Bessel functions

$\text{Relative weight} = 1 - \frac{I_1(\text{number of significant reflections})}{I_0(\text{number of significant reflections})}$ This relative weight is then added to the figure of merit to produce a score.

The $P(\delta q)$ function has been shown to work when there is sufficient detail in the $q^{ME}(\underline{x})$ maps[1], but is believed to be much less effective when there is insufficient detail in the maps for the calculation to distinguish significantly between the nodes. So the addition of ECCs and analysis to the calculation, should improve these situations, making a more significant contrast between nodes when there are few phased reflections in the basis map.

The following calculations were use to investigate the feasibility of the P(δq) function in conjunction with ECCs and analysis, being more effective in situations that are presently considered unfeasible. The P(δq) function is potentially of most use at low resolution where, due to is speed and ease of use it could be used to do large searches of phase space to find a molecular envelope, or develop detail within a molecular envelope.

All the following examples only used the origin defining reflections as a basis set. A set of reflections is then chosen and permuted via an error correcting code, for each new node created the $q^{ME}(\underline{x})$ is calculated. This is then used to distinguish between each of the nodes using the P(δq) function. If there is not enough variation between the nodes in terms of the $q^{ME}(\underline{x})$ map then the P(δq) function will give almost constant figures of merit. If only a little variation occurs between $q^{ME}(\underline{x})$ maps then the P(δq) function is ineffectual, the idea being that the error correcting codes should assign the phases to the reflections in such a manner that the maps should have sufficiently different details to allow the function to work efficiently.

6.3 Results

6.3.1 Al_mFe Aluminium alloy

The Al_mFe structure was determined by Gønnjes and co-workers[5], in tetragonal space group $I\bar{4}2m$ with a = 8.84Å c = 21.6Å. The cell contains 90Al and 20Fe atoms. The reflections were collected using the precession technique, instead of selected area diffraction. The precession technique moves the electron beam in a circular movement around the sample, the resulting reflections are more akin to the single-crystal X-ray diffraction because they integrate through the Bragg angle. The data will extend further out in the Laue zones, while dynamical effects will be reduced by the tilt out of the zone axis. Thickness variations will be damped considerably by the integrating

motion, this also has the benefit of being more kinematic like than selected area diffraction spots. So the precession technique should give more kiematic data than selected area diffraction and be more tolerant of thickness variation, hopefully helping overcome some of the more difficult aspects of electron crystallography, even when heavy atoms are present.

The NEXT command was used to select the reflections that were permuted in the EXTEND command, and the analysis scores used to rank the results. The FILTER command selects a cut off from the $P(\delta q)$ figure of merit that is nearest the number of nodes requested by the user. The top 512 nodes were selected for full entropy maximisation. The nodes were then analysed by the LLG analysis and the top nodes inspected visually for potential solutions.

Table 6-1 Al_mFe origin reflection used to generate $q^{ME}(\underline{x})$

h	k	l	Phase(°)
5	0	7	360

Table 6-2 Al_mFe reflections selected by next command for permutation by P(δq) function

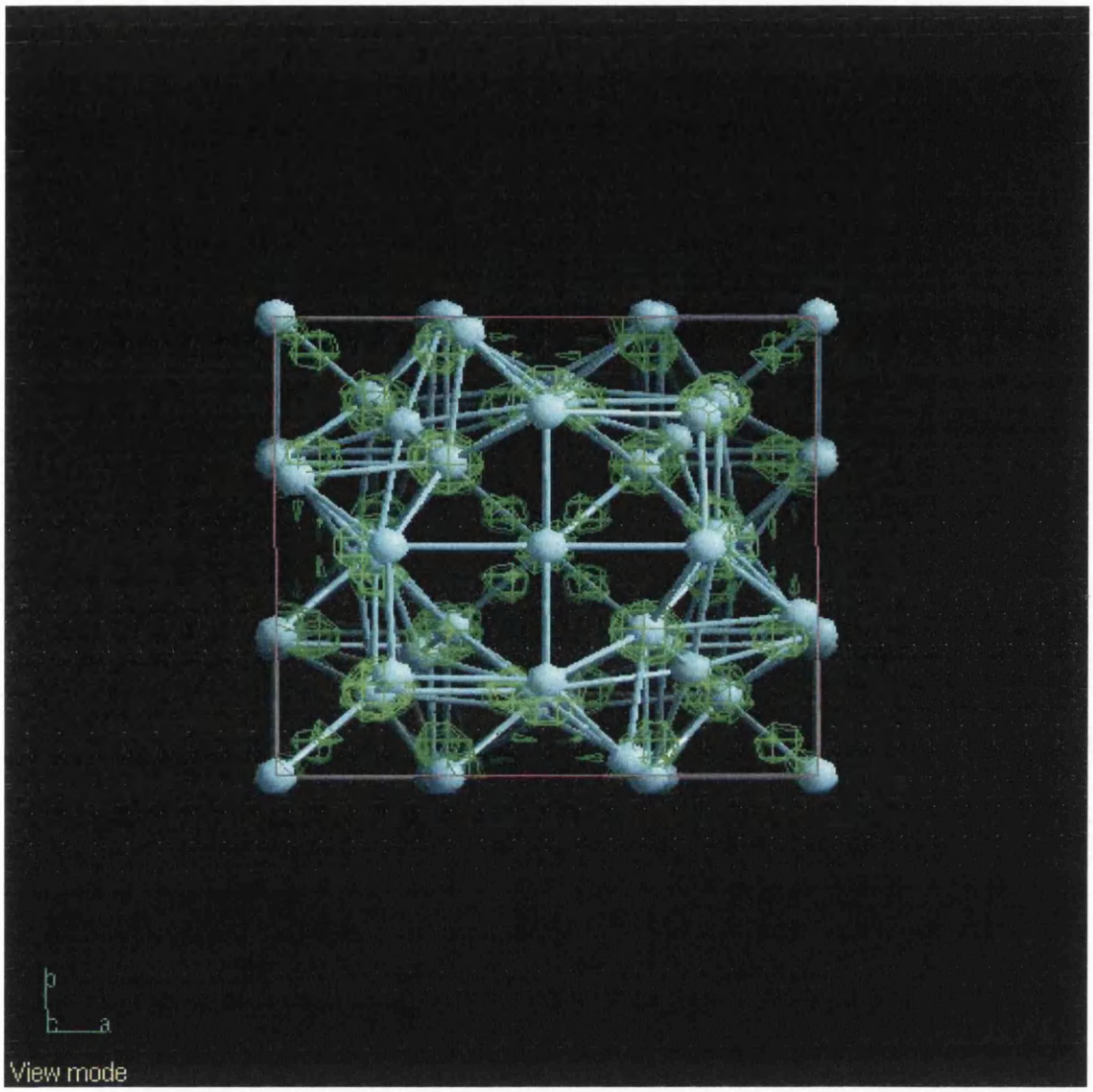
h	k	l	Uobs
7	1	10	0.4021
8	0	0	0.3897
0	0	20	0.3788
3	3	0	0.3759
6	0	0	0.3691
1	1	10	0.3628
4	1	17	0.3468

Table 6-3 Al_mFe results from analysis of entropy maximised nodes

Set No.	LLG	Entropy	Score	No of violations
235	4.081	-2.656	0.690	1
266	4.081	-2.656	0.690	1
270	3.084	-2.903	0.521	1
239	3.084	-2.903	0.521	1
336	4.345	-2.893	0.510	2
177	4.345	-2.893	0.510	2
497	3.565	-2.849	0.419	2
16	3.565	-2.849	0.419	2

The top nodes found from this procedure were of just as high quality and had reasonable figures of merit and more importantly the procedure yielded a number of nodes that are potential solutions in the top 8 of the LLG analysis. The time taken to go from 1024 starting sets to a final top 8 from LLG analysis was less than half the time of a normal entropy maximisation of all 1024 nodes then analysis. The rate-limiting step of such calculations is definitely the entropy maximisation of each node.

Figure 6-1 Al_mFe centroid map of node 235



6.3.2 10-Cyano-9,9'-Bianthryl (CNBA)

The CNBA structure was originally solved by Voigt-Martin and co-workers [6] and has the cell dimensions a = 14.70, b = 9.47, c = 15.42Å, β= 112°. The structure crystallised in space group $P2_1/c$, and 150 reflections were recorded by electron diffraction.

The NEXT command was used to select the reflections that were permuted in the EXTEND command, and the analysis scores used to rank the results. The FILTER command then selected the top 512 nodes for full entropy maximisation. The nodes were then analysed by the LLG analysis and the top nodes inspected visually for potential solutions.

Table 6-4 10-Cyano-9,9'-Bianthryl origin reflections used to generate $q^{ME}(\underline{x})$

h	k	l	phase(°)
0	3	7	360
3	4	-6	360
1	1	-2	360

Table 6-5 10-Cyano-9,9'-Bianthryl reflections selected by next command for permutation by P(_q) function

h	k	l	Uobs
-5	0	6	0.3947
-4	0	6	0.3123
0	2	0	0.3034
-2	-5	6	0.1743
-2	0	6	0.1419
-2	0	8	0.1402
-1	-1	1	0.1388
0	4	0	0.1354
-4	-1	8	0.1227
0	1	1	0.1191

Table 6-6 10-Cyano-9,9'-Bianthryl results from analysis of entropy maximised nodes

Set No.	LLG	Entropy	Score	No of violations
21	-0.235	-0.358	0.025	8
469	-0.238	-0.357	0.025	8
477	-0.257	-0.356	0.027	8
29	-0.258	-0.358	0.027	8
221	-0.263	-0.358	0.028	8
661	-0.271	-0.356	0.028	8
669	-0.273	-0.357	0.029	8
213	-0.273	-0.358	0.029	8

No particular node was an obvious candidate for further permutation and entropy maximisation. All the LLGs being negative and low scores being indicator that there was inadequate detail in the $q^{ME}(\underline{x})$map for the P(δq) function to be effective here. So the advantage of a faster over all calculation has been negated by the inability to continue from this point.

6.3.3 2,6-bis[4-(dimethylamino)benzylidene]cyclohexanone (DMABC)

The DMABC structure was solved by Voigt-Martin and co-workers in 1997 [7] using a combination of approaches to ensure that the electron diffraction intensities were correct and that they would lead to a correct structure solution, and not lead to predicting the wrong properties for the molecule in question.

The compound crystallised in the orthorhombic space group $Cmc2_1$ with cell parameters a = 21.8575, b = 9.315, c = 9.6437Å, with 133 reflections recorded by electron diffraction. The intensities from the diffraction were compared to simulations and found to have very good agreement between the simulated and observed diffraction patterns, although a small amount of secondary scattering was detected for the 100 and 002 reflections. Although there is a possibility this is an artifact of the simulation procedure.

The NEXT command was used to select the reflections that were permuted in the EXTEND command, and the analysis scores used to rank the results. The FILTER command then selected the top 512 nodes for full entropy maximisation. The nodes were then analysed by the LLG analysis and the top nodes inspected visually for potential solutions.

Table 6-7 2,6-bis[4-(dimethylamino)benzylidene]cyclohexanone origin reflections used to generate $q^{ME}(\underline{x})$

h	k	l	phase(°)
9	1	3	360
9	1	4	360

Table 6-8 2,6-bis[4-(dimethylamino)benzylidene]cyclohexanone reflections selected by next command for permutation by P(_q) function

h	k	l	Uobs
0	2	2	0.3347
18	2	3	0.2643
4	0	10	0.2356
18	0	2	0.2342
18	0	4	0.2240
16	0	0	0.2211
9	1	5	0.2177
16	0	2	0.2139
10	0	2	0.2134
2	2	4	0.2019

Table 6-9 2,6-bis[4-(dimethylamino)benzylidene]cyclohexanone results from analysis of entropy maximised nodes

Set No.	LLG	Entropy	Score	No of violations
385	3.420	-1.188	0.873	1
513	2.868	-1.038	0.732	1
383	2.672	-1.044	0.682	1
511	2.643	-1.044	0.675	1
449	2.633	-1.131	0.672	1
257	2.536	-1.043	0.647	1
129	2.494	-1.040	0.637	1
127	2.389	-1.115	0.610	1

While the output from this calculation looks promising, with high LLG values and scores, with very few violations indicated from the student t-test, the final maps are all very poor, showing no sensible atomic density that could be used to further the calculation. When compared to the known model the electron density that was present did not locate any part of the structure. Again this is due to inadequate detail in the basis set for the P(_q) function to be effective, as a figure of merit even with the ECC code and analysis.

6.3.4 4-dimethylamino-3-cyanobiphenyl (DMACB)

Voigt-Martin and co-workers solved the DMACB structure using Maximum entropy methods (MICE) in 1997[8], there was no attempt to refine the proposed final structure. The crystal formed in the orthorhombic space group $Pna2_1$ a = 10.28, b = 22.64, c = 5.77Å, with 117 measured reflections. The data set has dynamical effects taking place, this was demonstrated by comparing the calculated and observed intensities. The 400 reflection was omitted on the grounds that it was too strong to be reasonably

included in direct methods program as a kinematic reflection, and is probably due to dynamical diffraction.

The NEXT command was used to select the reflections that were permuted in the EXTEND command, and the analysis scores used to rank the results. The FILTER command then selected the top 500 nodes for full entropy maximisation. The nodes were then analysed by the LLG analysis and the top nodes inspected visually for potential solutions.

Table 6-10 4-dimethylamino-3-cyanobiphenyl origin reflection used to generate $q^{ME}(\underline{x})$

H	k	l	phase(°)
2	1	0	360

Table 6-11 4-dimethylamino-3-cyanobiphenyl reflections selected by next command for permutation by P(_q) function

h	k	l	Uobs
1	3	0	0.3243
6	2	0	0.3046
4	0	0	0.2666
1	2	0	0.2481
2	3	0	0.2477
6	0	0	0.1758
3	2	0	0.1593
6	1	0	0.1505
3	3	0	0.1461
2	2	0	0.1446

Table 6-12 4-dimethylamino-3-cyanobiphenyl results from analysis of entropy maximised nodes

Set No.	LLG	Entropy	Score	No of violations
2	4.968	-1.216	0.444	5
429	3.516	-0.927	0.381	4
49	3.315	-1.080	0.296	5
204	2.568	-1.010	0.278	4
179	3.758	-1.187	0.272	6
212	3.251	-1.149	0.236	6
148	2.385	-0.947	0.213	5
320	1.740	-0.865	0.188	4

Unsurprisingly the maps show no sensible electron density or fragments of the molecule. The low scores and relatively high number of violations, all indicate that the calculation has yielded no promising results.

6.3.5 Copper Perchlorophthalocyanine

This structure was solved by Dorset and co-workers[9] using electron diffraction and the tangent formula. The crystal formed in the space group C2/c, with unit cell dimensions a = 19.62, b = 26.08, c = 3.79Å, there are 198 unique reflections in the data set.

The NEXT command was used to select the reflections that were permuted in the EXTEND command, and the analysis scores used to rank the results. The FILTER command then selected the top 500 nodes for full entropy maximisation. The nodes were then analysed by the LLG analysis and the top nodes inspected visually for potential solutions.

Table 6-13 Copper Perchlorophthalocyanine origin reflection used to generate $q^{ME}(\underline{x})$

h	k	l	phase(°)
9	3	0	360

Table 6-14 Copper Perchlorophthalocyanine reflections selected by next command for permutation by P(δq) function

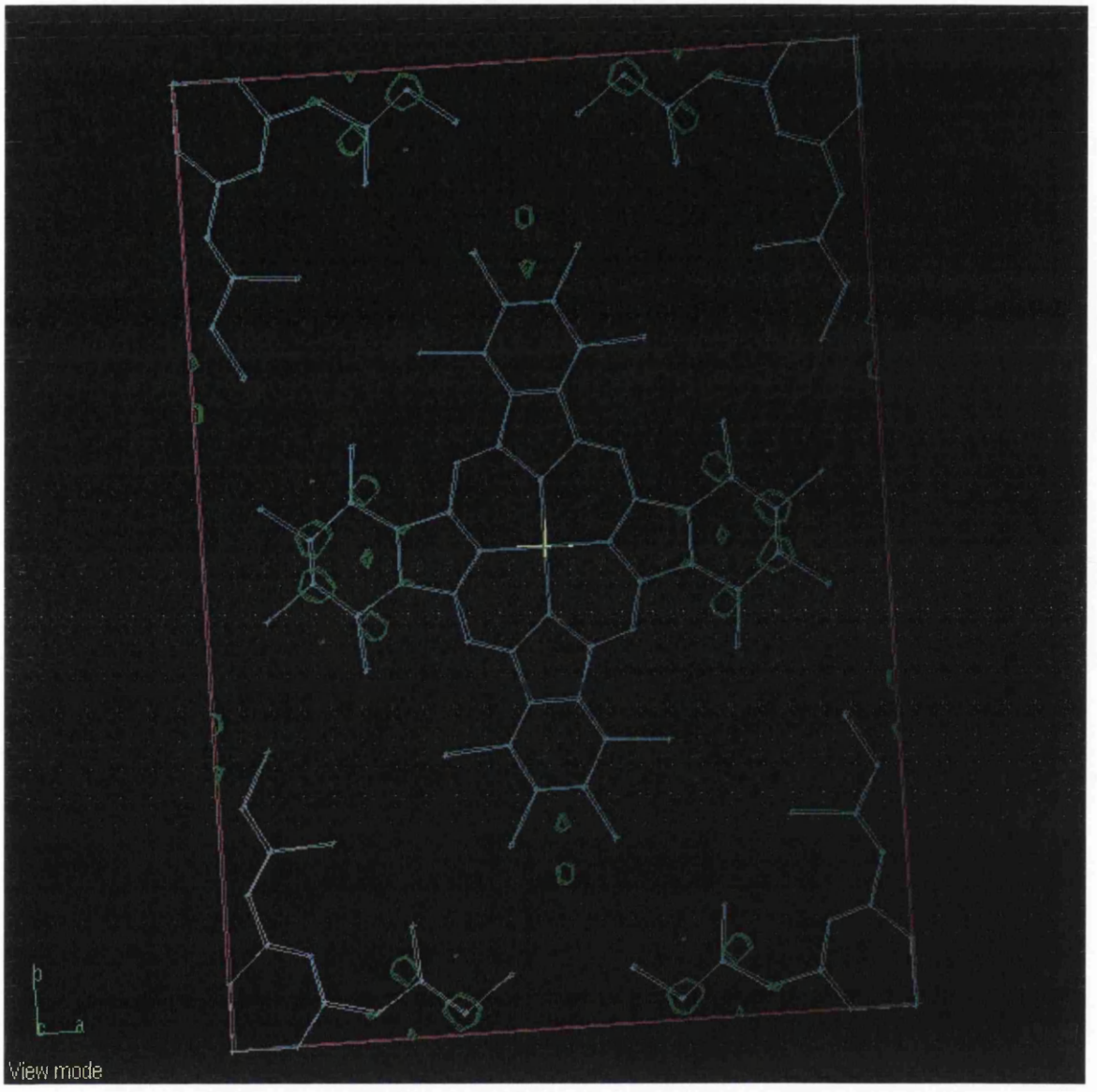
h	k	l	Uobs
9	19	0	0.6185
0	20	0	0.5596
11	9	0	0.5264
15	1	0	0.5246
2	12	0	0.4928
6	16	0	0.4917
12	14	0	0.4868
4	8	0	0.4866
13	1	0	0.4822
7	15	0	0.4672

Table 6-15 Copper Perchlorophthalocyanine results from analysis of entropy maximised nodes

Set No.	LLG	Entropy	Score	No of violations
272	4.043	-2.974	0.518	0
144	3.835	-3.624	0.491	0
429	3.571	-3.306	0.457	0
108	6.176	-3.413	0.445	1
50	3.188	-3.741	0.408	0
69	5.397	-3.320	0.389	1
162	5.213	-3.603	0.375	1
437	2.738	-3.699	0.351	0

The figures of merit indicate a mixed bag of results from this particular calculation, with good LLG but very mediocre scores from the analysis, but with the plus side of there have been very few violations detected by the student t-tests. The resulting maps have all got excellent detail.

Figure 6-2 Copper Perchlorophthalocyanine centroid map of node 272



6.3.6 Isotactic Poly(1-butene) Form III

The data set for Isotactic Poly(1-butene) Form III was obtained by electron diffraction by Dorset and co-workers in 1994[10]. They obtained the data set from solution and epitaxially grown crystals. A tilt series was used collect as much three-dimensional data as possible. The orthorhombic space group is $P2_12_12_1$ with unit cell dimensions a = 12.38, b = 8.88, c = 7.56Å. The data set has 125 unique reflections.

The NEXT command was used to select the reflections that were permuted in the EXTEND command, and the analysis scores used to rank the results. The FILTER command then selected the top 512 nodes for full entropy maximisation. The nodes

were then analysed by the LLG analysis and the top nodes inspected visually for potential solutions.

Table 6-16 Isotactic Poly(1-butene) Form III origin reflections

h	k	l	phase(°)
4	5	3	90
9	6	3	360
2	2	0	90

Table 6-17 Isotactic Poly(1-butene) Form III reflections selected by next command for permutation by P(δq) function

h	k	l	Uobs
1	1	1	0.3800
4	4	0	0.3763
3	3	1	0.3154
1	1	3	0.2985
4	7	1	0.2919
4	2	1	0.2827

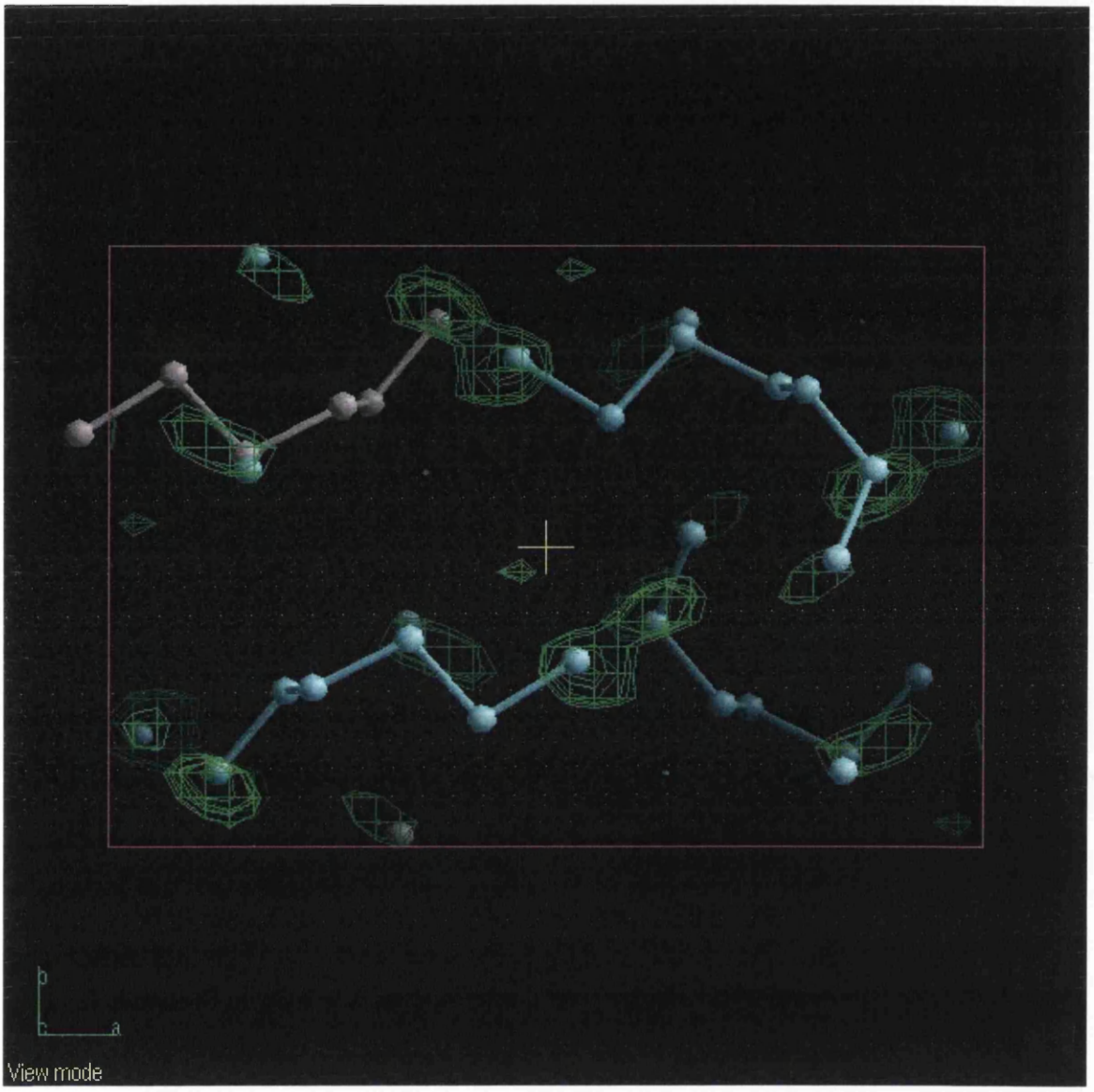
Table 6-18 Isotactic Poly(1-butene) Form III results from analysis of entropy maximised nodes

Set No.	LLG	Entropy	Score	No of violations
207	10.376	-2.435	0.640	1
117	10.376	-2.435	0.640	1
251	10.083	-2.139	0.622	1
73	10.083	-2.138	0.622	1
446	9.109	-2.010	0.562	1
260	9.107	-2.009	0.562	1
219	8.745	-2.159	0.539	1
105	8.744	-2.160	0.539	1

The LLGs, scores and number of violations are all showing that the outcome of this calculation is very good, and the maps bear this fact out. All the maps show excellent electron density, with some maps having almost complete solutions.

Figure 6-3 Isotactic Poly(1-butene) Form III centroid map of node 73 All the maps show excellent electron density, with some maps having almost complete solutions.

Figure 6-4Isotactic Poly(1-butene) Form III centroid map of node 73



6.4 Conclusions of $P(\delta q)$ results

The P(δq) function has a definite usefulness in speeding up the time taken to do the large tree calculations that are attempted in the permutation step of an *ab initio* calculation. This has been known for some time, although it was assumed that the function was only useful when there was considerable detail in the maps, position when maximum entropy generally does not need too much more help in finding the solution. Here we have demonstrated that with some care the P(δq) function can be of use in the early stages of an *ab initio* calculation, when coupled with error correcting codes and an analysis process.

The P(δq) function could be further improved by using a larger error correcting code where there would be more detail in the initial maps and therefore a greater difference in the native P(δq) figure of merit, which could lead to even fewer nodes being taken to the entropy maximisation and time consuming step.

Other possible methods for improving the effectiveness of this technique would be to use larger ECCs and permuting more reflections at a time thus increasing the basis set and therefore the effectiveness of the P(δq) functions figure of merit. There is also another figure of merit within the program the $P(\delta q) = \int_V \frac{\delta q^3}{q^{ME}(\underline{x})^2} d^3\underline{x}$ that has not been investigated here, also has the potential to give better discrimination between the nodes. Although the function has been shown to be moderately successful here it is unlikely to have success with larger more complex models, because the maps would be too flat to allow the function to distinguish sufficiently between them.

References

(1) Gilmore, C.J.; Bricogne, G. *Acta Crystallographica* **1990**, *A46*, 297-308.

(2) Gilmore, C.J.; Bricogne, G. *Methods in Enzmology* **1997**, *277*, 65-78.

(3) Main, P. *Acta Crystallographica* **1977**, *A33*, 750-757.

(4) Gilmore, C.J.; Dong, W.; Bricogne, G. *Acta Crystallographica* **1999**, *A55*, 70-83.

(5) Gjønnes, J.; Hansen, V.; Berg, B.S.; Runde, P.; Cheng, Y.F.; Dorset, D.L.; Gilmore, C.J. *Acta Crystallographica* **1998**, *A54*, 306-319.

(6) Voigt-Martin, I.G.; Yan, D.H.; Yakimansky, A.; Schollmeyer, D.; Gilmore, C.J.; Bricogne, G. *Acta Crystallographica* **1995**, *A51*, 849-868.

(7) Yakimansky, A.; Kolb, U.; Matveeva, G.N.; Voigt-Martin, I.G.; Tenkovtsev, A.V. *Acta Crystallographica* **1997**, *A53*, 603-614.

(8) Voigt-Martin, I.G.; Zhang, Z.X.; Kolb, U.; Gilmore, C.J. *Ultramicroscopy* **1997**, *68*, 43-59.

(9) Dorset, D.L.; Tivol, W.F.; Turner, J.N. *Ultramicroscopy* **1991**, *38*, 41-45.

(10) Dorset, D.L.; McCourt, M.P.; Kopp, S.; Wittmann, J.C.; Lotz, B. *Acta Crystallographica* **1994**, *B50*, 201-208.

7 CONCLUSIONS

The use of maximum entropy as a method of solving electron crystallographic structure *ab initio* has been demonstrated to be an invaluable approach, especially when, no or very few phases are available from electron microscopy images. Even with very incomplete data the maximum entropy approach in MICE was able to solve the structure *e.g.* poly (1,4,-trans-cyclohexanediyl dimethylene succinate).

In addition to the ability to solve structure *ab initio*, we have established a viable method for overcoming the missing cone problem. Using maximum entropy techniques to help predict the missing reflections and phases, with more accuracy than other techniques proposed. The usefulness of this technique has only been established with small molecules at atomic resolution. Further investigation is needed to establish if the technique will be of use at lower resolution and/or with larger molecules, with potentially larger pieces of missing information.

The $P(\delta q)$ function, with the addition of the analysis and error correcting codes, has become a more useful form of phase permutation. The speed of the $P(\delta q)$ function with respect to the more traditional form of phase permutation, allows for considerably more efficient method of calculating basis sets. With no need calculate the maximum entropy of all the nodes from the permutation to achieve a figure of merit, the technique gives considerable time savings, with very little or no loss of discrimination between the nodes, due to the use of the analysis process. The analysis added to the function will also allow bigger error correcting codes or incomplete factorial designs to be used in the future. The bigger ECCs and IFDs will also have the beneficial effect of improving the accuracy of the $P(\delta q)$ figure of merit.

The refinement of electron diffraction data, has also been shown to work with data close to the kinematical approximation. This is a useful tool for checking model validity. The requirement for the data to be approximately kinematical is important and if there is any significant dynamical scattering or the need for an accurate refinement then the programme MSLS should be used to account for crystal thickness and dynamical scattering. The framework proposed is a guide for structure refinement of the kinematical and sparse data sets.

Although some progress has been made here towards the goal of making electron crystallography a more routine method in terms of data analysis, there are still many avenues to be investigated. These include maximum likelihood refinement, density modification, especially in the case of proteins. The maximum likelihood refinement formulas that are currently available, and new modified versions can potentially be a significant addition to the electron crystallographers tool box. The use of molecular envelopes for solving small molecules and proteins has already been used, but this could be exploited further by making use of the high quality phase information that can be obtained from lower resolution TEM images, at say 15Å. This information could be readily incorporated into the structure solution process, akin to the way that molecular envelopes are utilised in protein crystallography.

Currently electron crystallography is a small field with relatively few participants compared to its X-ray and neutron cousins, while most people within the field know each other and collaborate or exchange ideas freely this environment can make it difficult for new groups to enter the field. The area may well benefit greatly from a central resource similar to protein crystallography's CCP4, where data formats are standardised and computer programs are gathered for redistribution in a larger suite. This allows groups with little or no experience in the area to gain a foothold within the field rapidly and productively. Electron crystallography may well benefit from a similar resource as more and more people try to enter the field from a large diversity of backgrounds such as, protein crystallography, structural chemistry, microscopy and polymer science.

Another area of potential growth for electron crystallography maybe in the area of density fitting or Molecular replacement, these methods have been automated by protein crystallographers, in programs like Solve, Resolve, Arp/warp, Amore, and AutoSharp. This type of approach could be utilised in a brut force effort to solve difficult and problematic electron crystallographic structures, especially in combination with a low-resolution envelope *via* an image.

These are only a few suggestions for the advancement of electron crystallography from the perspective of a structure solution and direct methods stand point, there are also many other practical and theoretical aspects to the microscope setup and use, that can be improved to make the process of achieving an electron diffraction data set

more practical. These issues will not be addressed here, but instead left to others for further investigation.

All the methodology and adapted algorithms should help make the elucidation of electron diffraction data sets easier and help move electron crystallography towards becoming a more routine method of solving nanometer and micrometer sized crystals, that cannot be approached by X-ray or neutron diffraction by powders.

8 APPENDIX A: CIF FILES FROM DATABASE

During the passage of my PhD I have acquired several sets of electron diffraction data. I found these to be in a large and often disordered format, which required a great deal of time and effort being spent formatting and often inputting the data into programs.

To save other researchers having to repeat this mindless task I have created a Crystallographic Information File each of the data sets I have acquired through out the course my PhD and made them available *via* a crude ftp database site.

ftp.chem.gla.ac.uk

dir: eddata

password: boarest

Each of the following structure has the CIF file available from this site in plain text format. The database is also on a CD accompanying this thesis.

STRUCTURE LIST:

AL_MFE

BRUCITE

10-CYANO-9,9'-BIANTHRYL

BASIC COPPER CHLORIDE

2,6-BIS[4-(DIMETHYLAMINO)BENZYLIDENE]CYCLOHEXANONE

4-DIMETHYLAMINO-3-CYANOBIPHENYL

MANNAN I

COPPER PERCHLOROPHTHALOCYANINE

ISOTACTIC POLY(1-BUTENE) FORM III

POLY(ε-CAPROLACTONE)

POLYETHYLENE

POLY(1,4,TRANS-CYCLOHEXANEDIYL DIMETHYLENE SUCCINATE)

8.1 Al_mFe CIF

```
data_shelxl

_audit_creation_method            SHELXL-97
_chemical_name_systematic
;
 ?
;
_chemical_name_common             AlmFe
_chemical_formula_moiety          ?
_chemical_formula_sum
 'Al10 Fe3'
_chemical_formula_weight          437.35

_citation_journal_full Acta Cryst
_citation_journal_volume A54
_citation_langauge English
_citation_page_first 306
_citation_page_last 319
_citation_year 1998
_citation_author_name 'Gjonnes,J., Hansen, V., Berg, B.S., Runde, P.,
Cheng,Y.F., Gjonnes, K., Dorset, D.L., Gilmore, C.J.'
loop_
 _atom_type_symbol
 _atom_type_description
 _atom_type_scat_dispersion_real
 _atom_type_scat_dispersion_imag
 _atom_type_scat_source
 'Fe'  'Fe'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
 'Al'  'Al'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'

_symmetry_cell_setting            ?
_symmetry_space_group_name_H-M    'I -4 2 m'

loop_
 _symmetry_equiv_pos_as_xyz
 'x, y, z'
 '-x, -y, z'
 'x, -y, -z'
 '-x, y, -z'
 'y, x, z'
 '-y, -x, z'
 '-y, x, -z'
 'y, -x, -z'
 'x+1/2, y+1/2, z+1/2'
 '-x+1/2, -y+1/2, z+1/2'
 'x+1/2, -y+1/2, -z+1/2'
 '-x+1/2, y+1/2, -z+1/2'
 'y+1/2, x+1/2, z+1/2'
 '-y+1/2, -x+1/2, z+1/2'
 '-y+1/2, x+1/2, -z+1/2'
 'y+1/2, -x+1/2, -z+1/2'

_cell_length_a                    8.840
```

```
_cell_length_b                    8.840
_cell_length_c                    21.600
_cell_angle_alpha                 90.00
_cell_angle_beta                  90.00
_cell_angle_gamma                 90.00
_cell_volume                      1687.9
_cell_formula_units_Z             8
_cell_measurement_temperature     293(2)

_exptl_crystal_density_diffrn     3.442
_exptl_crystal_density_method     'not measured'
_exptl_crystal_F_000              642
_exptl_absorpt_coefficient_mu     0.006


_diffrn_ambient_temperature       293(2)
_diffrn_radiation_wavelength      ?
_diffrn_radiation_type            Electron
_diffrn_radiation_source          Microscope
_diffrn_reflns_number             743
_diffrn_reflns_av_R_equivalents   0.0000
_diffrn_reflns_av_sigmaI/netI     0.0155
_diffrn_reflns_limit_h_min        0
_diffrn_reflns_limit_h_max        9
_diffrn_reflns_limit_k_min        0
_diffrn_reflns_limit_k_max        13
_diffrn_reflns_limit_l_min        0
_diffrn_reflns_limit_l_max        32
_diffrn_reflns_theta_min          3.26
_diffrn_reflns_theta_max          68.24
_reflns_number_total              743
_reflns_number_gt                 743
_reflns_threshold_expression      >2sigma(I)

_computing_structure_solution     'SHELXS-97 (Sheldrick, 1990)'
_computing_structure_refinement   'SHELXL-97 (Sheldrick, 1997)'
_computing_molecular_graphics     ?
_computing_publication_material   maXus

_refine_special_details
;
 Refinement of F^2^ against ALL reflections.  The weighted R-factor wR and
 goodness of fit S are based on F^2^, conventional R-factors R are based
 on F, with F set to zero for negative F^2^. The threshold expression of
 F^2^ > 2sigma(F^2^) is used only for calculating R-factors(gt) etc. and is
 not relevant to the choice of reflections for refinement.  R-factors based
 on F^2^ are statistically about twice as large as those based on F, and R-
 factors based on ALL data will be even larger.
;

_refine_ls_structure_factor_coef  Fsqd
_refine_ls_matrix_type            full
_refine_ls_weighting_scheme
 'calc w=1/[\s^2^(Fo^2^)+(0.1000P)^2^+0.0000P] where P=(Fo^2^+2Fc^2^)/3'
_atom_sites_solution_primary      direct
_atom_sites_solution_secondary    difmap
_atom_sites_solution_hydrogens    geom
```

```
_refine_ls_hydrogen_treatment     mixed
_refine_ls_extinction_method      none
_refine_ls_extinction_coef        ?
_refine_ls_abs_structure_details
 'Flack H D (1983), Acta Cryst. A39, 876-881'
_refine_ls_abs_structure_Flack    0(10)
_refine_ls_number_reflns          743
_refine_ls_number_parameters      34
_refine_ls_number_restraints      0
_refine_ls_R_factor_all           0.4475
_refine_ls_R_factor_gt            0.4475
_refine_ls_wR_factor_ref          0.8113
_refine_ls_wR_factor_gt           0.8113
_refine_ls_goodness_of_fit_ref    9.433
_refine_ls_restrained_S_all       9.433
_refine_ls_shift/su_max           2.596
_refine_ls_shift/su_mean          0.113

loop_
 _atom_site_label
 _atom_site_type_symbol
 _atom_site_fract_x
 _atom_site_fract_y
 _atom_site_fract_z
 _atom_site_U_iso_or_equiv
 _atom_site_adp_type
 _atom_site_occupancy
 _atom_site_calc_flag
 _atom_site_refinement_flags
 _atom_site_disorder_assembly
 _atom_site_disorder_group
Fe1 Fe 0.3162 0.3162 0.0783(6) 0.038(3) Uiso 1 d S . .
Fe2 Fe 0.1740 0.1740 0.2731(9) 0.049(4) Uiso 1 d S . .
Fe3 Fe 0.1830 0.1830 -0.1280(9) 0.062(6) Uiso 1 d S . .
Al4 Al 0.0335(18) 0.3493(17) 0.0759(6) 0.021(3) Uiso 1 d . . .
Al5 Al 0.521(4) 0.185(3) 0.1369(11) 0.057(6) Uiso 1 d . . .
Al6 Al 0.042(3) 0.312(3) -0.2272(12) 0.055(6) Uiso 1 d . . .
Al7 Al 0.0000 0.0000 -0.069(2) 0.061(14) Uiso 1 d S . .
Al8 Al 0.2150 0.2150 -0.0217(11) 0.031(5) Uiso 1 d S . .
Al9 Al 0.2330 0.2330 0.1732(11) 0.044(7) Uiso 1 d S . .
Al10 Al 0.207(4) 0.5000 0.0000 0.040(6) Uiso 1 d S . .
Al11 Al 0.5000 0.5000 0.1648(19) 0.051(10) Uiso 1 d S . .
Al12 Al 0.0000 0.0000 -0.214(3) 0.074(17) Uiso 1 d S . .
Al13 Al 0.5000 0.5000 0.0000 0.11(4) Uiso 1 d S . .

_geom_special_details
;
 All esds (except the esd in the dihedral angle between two l.s. planes)
 are estimated using the full covariance matrix.  The cell esds are taken
 into account individually in the estimation of esds in distances, angles
 and torsion angles; correlations between esds in cell parameters are only
 used when they are defined by crystal symmetry.  An approximate (isotropic)
 treatment of cell esds is used for estimating esds involving l.s. planes.
;

loop_
 _geom_bond_atom_site_label_1
```

```
 _geom_bond_atom_site_label_2
 _geom_bond_distance
 _geom_bond_site_symmetry_2
 _geom_bond_publ_flag
Fe1 Al9 2.30(2) . ?
Fe1 Al5 2.49(3) . ?
Fe1 Al5 2.49(3) 5 ?
Fe1 Al8 2.50(2) . ?
Fe1 Al4 2.516(15) . ?
Fe1 Al4 2.516(15) 5 ?
Fe1 Al10 2.537(16) 5 ?
Fe1 Al10 2.537(16) . ?
Fe1 Al13 2.853(8) . ?
Fe1 Al11 2.96(3) . ?
Fe2 Al9 2.28(3) . ?
Fe2 Al6 2.47(3) 8 ?
Fe2 Al6 2.47(3) 4 ?
Fe2 Al6 2.51(3) 14 ?
Fe2 Al6 2.51(3) 10 ?
Fe2 Al12 2.53(3) 3 ?
Fe2 Al11 2.56(2) 11_455 ?
Fe2 Al5 2.68(3) 11_455 ?
Fe2 Al5 2.68(3) 15_545 ?
Fe2 Fe3 2.79(2) 10 ?
Fe3 Al8 2.33(3) . ?
Fe3 Al5 2.62(3) 8_565 ?
Fe3 Al5 2.62(3) 4_655 ?
Fe3 Al7 2.62(2) . ?
Fe3 Al4 2.663(18) 8 ?
Fe3 Al4 2.663(18) 4 ?
Fe3 Al6 2.73(3) 5 ?
Fe3 Al6 2.73(3) . ?
Fe3 Fe2 2.79(2) 10_554 ?
Fe3 Al12 2.94(4) . ?
Al4 Al5 2.42(3) 5 ?
Al4 Al10 2.61(3) . ?
Al4 Al5 2.61(3) 6_565 ?
Al4 Fe3 2.663(18) 4 ?
Al4 Al4 2.73(3) 2_565 ?
Al4 Al8 2.758(18) 4 ?
Al4 Al8 2.90(2) . ?
Al4 Al9 2.93(2) . ?
Al4 Al10 3.00(3) 2_565 ?
Al5 Al4 2.42(3) 5 ?
Al5 Fe3 2.62(3) 4_655 ?
Al5 Al4 2.60(3) 6_655 ?
Al5 Fe2 2.68(3) 11 ?
Al5 Al9 2.70(3) . ?
Al5 Al6 2.75(4) 7_655 ?
Al5 Al11 2.85(3) . ?
Al5 Al10 2.97(2) 5 ?
Al5 Al6 2.99(4) 10 ?
Al6 Fe2 2.47(3) 4 ?
Al6 Fe2 2.51(3) 10_554 ?
Al6 Al6 2.60(4) 15_554 ?
Al6 Al6 2.60(4) 16_454 ?
Al6 Al5 2.75(4) 8_565 ?
```

```
Al6 Al9 2.79(3) 4 ?
Al6 Al12 2.80(3) . ?
Al6 Al9 2.96(3) 10_554 ?
Al6 Al5 2.99(4) 10_554 ?
Al7 Fe3 2.62(2) 2 ?
Al7 Al8 2.872(18) 2 ?
Al7 Al8 2.872(18) . ?
Al8 Al10 2.564(4) . ?
Al8 Al10 2.564(4) 5 ?
Al8 Al4 2.758(18) 8 ?
Al8 Al4 2.758(18) 4 ?
Al8 Al4 2.90(2) 5 ?
Al9 Al5 2.70(3) 5 ?
Al9 Al6 2.79(3) 8 ?
Al9 Al6 2.79(3) 4 ?
Al9 Al4 2.93(2) 5 ?
Al9 Al6 2.96(3) 10 ?
Al9 Al6 2.96(3) 14 ?
Al10 Fe1 2.537(16) 3_565 ?
Al10 Al8 2.564(4) 3_565 ?
Al10 Al4 2.61(3) 3_565 ?
Al10 Al13 2.59(3) . ?
Al10 Al5 2.97(2) 8_565 ?
Al10 Al5 2.97(2) 5 ?
Al10 Al4 3.00(3) 4 ?
Al10 Al4 3.00(3) 2_565 ?
Al11 Fe2 2.56(2) 12 ?
Al11 Fe2 2.56(2) 11 ?
Al11 Al12 2.63(7) 9 ?
Al11 Al5 2.85(3) 5 ?
Al11 Al5 2.85(3) 6_665 ?
Al11 Al5 2.85(3) 2_665 ?
Al11 Fe1 2.96(3) 2_665 ?
Al12 Fe2 2.53(3) 4 ?
Al12 Fe2 2.53(3) 3 ?
Al12 Al11 2.63(7) 9_444 ?
Al12 Al6 2.80(3) 6 ?
Al12 Al6 2.80(3) 2 ?
Al12 Al6 2.80(3) 5 ?
Al12 Fe3 2.94(4) 2 ?
Al13 Al10 2.59(3) 6_665 ?
Al13 Al10 2.59(3) 5 ?
Al13 Al10 2.59(3) 2_665 ?
Al13 Fe1 2.853(8) 3_565 ?
Al13 Fe1 2.853(8) 4_655 ?
Al13 Fe1 2.853(8) 2_665 ?

loop_
 _geom_angle_atom_site_label_1
 _geom_angle_atom_site_label_2
 _geom_angle_atom_site_label_3
 _geom_angle
 _geom_angle_site_symmetry_1
 _geom_angle_site_symmetry_3
 _geom_angle_publ_flag
Al9 Fe1 Al5 68.4(7) . . ?
Al9 Fe1 Al5 68.4(7) . 5 ?
```

```
Al5 Fe1 Al5 114.6(13) . 5 ?
Al9 Fe1 Al8 122.7(4) . . ?
Al5 Fe1 Al8 122.1(6) . . ?
Al5 Fe1 Al8 122.1(6) 5 . ?
Al9 Fe1 Al4 74.8(4) . . ?
Al5 Fe1 Al4 141.8(8) . . ?
Al5 Fe1 Al4 57.7(7) 5 . ?
Al8 Fe1 Al4 70.7(4) . . ?
Al9 Fe1 Al4 74.8(4) . 5 ?
Al5 Fe1 Al4 57.7(7) . 5 ?
Al5 Fe1 Al4 141.8(8) 5 5 ?
Al8 Fe1 Al4 70.7(4) . 5 ?
Al4 Fe1 Al4 103.3(7) . 5 ?
Al9 Fe1 Al10 132.6(6) . 5 ?
Al5 Fe1 Al10 72.3(7) . 5 ?
Al5 Fe1 Al10 155.8(10) 5 5 ?
Al8 Fe1 Al10 61.1(3) . 5 ?
Al4 Fe1 Al10 131.8(6) . 5 ?
Al4 Fe1 Al10 62.2(8) 5 5 ?
Al9 Fe1 Al10 132.6(6) . . ?
Al5 Fe1 Al10 155.8(10) . . ?
Al5 Fe1 Al10 72.3(7) 5 . ?
Al8 Fe1 Al10 61.1(3) . . ?
Al4 Fe1 Al10 62.2(8) . . ?
Al4 Fe1 Al10 131.8(6) 5 . ?
Al10 Fe1 Al10 92.5(10) 5 . ?
Al9 Fe1 Al13 153.2(5) . . ?
Al5 Fe1 Al13 98.7(7) . . ?
Al5 Fe1 Al13 98.7(7) 5 . ?
Al8 Fe1 Al13 84.0(4) . . ?
Al4 Fe1 Al13 119.2(5) . . ?
Al4 Fe1 Al13 119.2(5) 5 . ?
Al10 Fe1 Al13 57.1(7) 5 . ?
Al10 Fe1 Al13 57.1(7) . . ?
Al9 Fe1 Al11 77.8(7) . . ?
Al5 Fe1 Al11 62.4(7) . . ?
Al5 Fe1 Al11 62.4(7) 5 . ?
Al8 Fe1 Al11 159.5(7) . . ?
Al4 Fe1 Al11 119.6(5) . . ?
Al4 Fe1 Al11 119.6(5) 5 . ?
Al10 Fe1 Al11 106.2(5) 5 . ?
Al10 Fe1 Al11 106.2(5) . . ?
Al13 Fe1 Al11 75.5(6) . . ?
Al9 Fe2 Al6 71.6(7) . 8 ?
Al9 Fe2 Al6 71.6(7) . 4 ?
Al6 Fe2 Al6 127.1(14) 8 4 ?
Al9 Fe2 Al6 76.0(7) . 14 ?
Al6 Fe2 Al6 138.3(6) 8 14 ?
Al6 Fe2 Al6 62.7(12) 4 14 ?
Al9 Fe2 Al6 76.0(7) . 10 ?
Al6 Fe2 Al6 62.7(12) 8 10 ?
Al6 Fe2 Al6 138.3(6) 4 10 ?
Al6 Fe2 Al6 84.5(12) 14 10 ?
Al9 Fe2 Al12 78.3(13) . 3 ?
Al6 Fe2 Al12 68.2(7) 8 3 ?
Al6 Fe2 Al12 68.2(7) 4 3 ?
Al6 Fe2 Al12 129.5(10) 14 3 ?
```

```
Al6 Fe2 Al12 129.5(10) 10 3 ?
Al9 Fe2 Al11 140.6(9) . 11_455 ?
Al6 Fe2 Al11 92.5(8) 8 11_455 ?
Al6 Fe2 Al11 92.5(8) 4 11_455 ?
Al6 Fe2 Al11 129.1(9) 14 11_455 ?
Al6 Fe2 Al11 129.1(9) 10 11_455 ?
Al12 Fe2 Al11 62.2(14) 3 11_455 ?
Al9 Fe2 Al5 134.2(6) . 11_455 ?
Al6 Fe2 Al5 154.1(11) 8 11_455 ?
Al6 Fe2 Al5 70.8(9) 4 11_455 ?
Al6 Fe2 Al5 64.0(9) 14 11_455 ?
Al6 Fe2 Al5 118.9(11) 10 11_455 ?
Al12 Fe2 Al5 110.2(10) 3 11_455 ?
Al11 Fe2 Al5 66.0(9) 11_455 11_455 ?
Al9 Fe2 Al5 134.2(6) . 15_545 ?
Al6 Fe2 Al5 70.8(9) 8 15_545 ?
Al6 Fe2 Al5 154.1(11) 4 15_545 ?
Al6 Fe2 Al5 118.9(11) 14 15_545 ?
Al6 Fe2 Al5 64.0(9) 10 15_545 ?
Al12 Fe2 Al5 110.2(10) 3 15_545 ?
Al11 Fe2 Al5 66.0(9) 11_455 15_545 ?
Al5 Fe2 Al5 86.6(14) 11_455 15_545 ?
Al9 Fe2 Fe3 121.2(4) . 10 ?
Al6 Fe2 Fe3 115.6(6) 8 10 ?
Al6 Fe2 Fe3 115.6(6) 4 10 ?
Al6 Fe2 Fe3 61.8(7) 14 10 ?
Al6 Fe2 Fe3 61.8(7) 10 10 ?
Al12 Fe2 Fe3 160.5(14) 3 10 ?
Al11 Fe2 Fe3 98.2(10) 11_455 10 ?
Al5 Fe2 Fe3 57.3(8) 11_455 10 ?
Al5 Fe2 Fe3 57.3(8) 15_545 10 ?
Al8 Fe3 Al5 87.1(7) . 8_565 ?
Al8 Fe3 Al5 87.1(7) . 4_655 ?
Al5 Fe3 Al5 88.8(12) 8_565 4_655 ?
Al8 Fe3 Al7 70.6(10) . . ?
Al5 Fe3 Al7 131.0(8) 8_565 . ?
Al5 Fe3 Al7 131.0(8) 4_655 . ?
Al8 Fe3 Al4 66.7(5) . 8 ?
Al5 Fe3 Al4 138.0(8) 8_565 8 ?
Al5 Fe3 Al4 59.1(7) 4_655 8 ?
Al7 Fe3 Al4 72.0(5) . 8 ?
Al8 Fe3 Al4 66.7(5) . 4 ?
Al5 Fe3 Al4 59.1(7) 8_565 4 ?
Al5 Fe3 Al4 138.0(8) 4_655 4 ?
Al7 Fe3 Al4 72.0(5) . 4 ?
Al4 Fe3 Al4 128.0(10) 8 4 ?
Al8 Fe3 Al6 141.1(6) . 5 ?
Al5 Fe3 Al6 113.2(10) 8_565 5 ?
Al5 Fe3 Al6 61.9(8) 4_655 5 ?
Al7 Fe3 Al6 111.2(9) . 5 ?
Al4 Fe3 Al6 76.8(6) 8 5 ?
Al4 Fe3 Al6 152.2(10) 4 5 ?
Al8 Fe3 Al6 141.1(6) . . ?
Al5 Fe3 Al6 61.9(8) 8_565 . ?
Al5 Fe3 Al6 113.2(10) 4_655 . ?
Al7 Fe3 Al6 111.2(9) . . ?
Al4 Fe3 Al6 152.2(10) 8 . ?
```

```
Al4 Fe3 Al6 76.8(6) 4 . ?
Al6 Fe3 Al6 76.5(13) 5 . ?
Al8 Fe3 Fe2 130.2(3) . 10_554 ?
Al5 Fe3 Fe2 59.2(6) 8_565 10_554 ?
Al5 Fe3 Fe2 59.2(6) 4_655 10_554 ?
Al7 Fe3 Fe2 159.2(11) . 10_554 ?
Al4 Fe3 Fe2 113.5(4) 8 10_554 ?
Al4 Fe3 Fe2 113.5(4) 4 10_554 ?
Al6 Fe3 Fe2 54.2(7) 5 10_554 ?
Al6 Fe3 Fe2 54.2(7) . 10_554 ?
Al8 Fe3 Al12 138.8(9) . . ?
Al5 Fe3 Al12 120.4(8) 8_565 . ?
Al5 Fe3 Al12 120.4(8) 4_655 . ?
Al7 Fe3 Al12 68.2(12) . . ?
Al4 Fe3 Al12 100.0(5) 8 . ?
Al4 Fe3 Al12 100.0(5) 4 . ?
Al6 Fe3 Al12 59.1(8) 5 . ?
Al6 Fe3 Al12 59.1(8) . . ?
Fe2 Fe3 Al12 91.0(11) 10_554 . ?
Al5 Al4 Fe1 60.7(8) 5 . ?
Al5 Al4 Al10 72.3(9) 5 . ?
Fe1 Al4 Al10 59.3(7) . . ?
Al5 Al4 Al5 82.0(12) 5 6_565 ?
Fe1 Al4 Al5 141.1(10) . 6_565 ?
Al10 Al4 Al5 121.9(9) . 6_565 ?
Al5 Al4 Fe3 121.1(10) 5 4 ?
Fe1 Al4 Fe3 129.9(7) . 4 ?
Al10 Al4 Fe3 165.7(8) . 4 ?
Al5 Al4 Fe3 59.7(8) 6_565 4 ?
Al5 Al4 Al4 60.5(9) 5 2_565 ?
Fe1 Al4 Al4 109.2(9) . 2_565 ?
Al10 Al4 Al4 68.2(6) . 2_565 ?
Al5 Al4 Al4 53.8(8) 6_565 2_565 ?
Fe3 Al4 Al4 112.5(8) 4 2_565 ?
Al5 Al4 Al8 160.7(10) 5 4 ?
Fe1 Al4 Al8 138.5(7) . 4 ?
Al10 Al4 Al8 114.9(8) . 4 ?
Al5 Al4 Al8 79.2(8) 6_565 4 ?
Fe3 Al4 Al8 50.9(7) 4 4 ?
Al4 Al4 Al8 104.3(7) 2_565 4 ?
Al5 Al4 Al8 110.2(9) 5 . ?
Fe1 Al4 Al8 54.4(5) . . ?
Al10 Al4 Al8 55.1(4) . . ?
Al5 Al4 Al8 163.1(9) 6_565 . ?
Fe3 Al4 Al8 118.6(6) 4 . ?
Al4 Al4 Al8 121.3(6) 2_565 . ?
Al8 Al4 Al8 87.5(4) 4 . ?
Al5 Al4 Al9 59.7(8) 5 . ?
Fe1 Al4 Al9 49.2(5) . . ?
Al10 Al4 Al9 106.1(8) . . ?
Al5 Al4 Al9 103.8(8) 6_565 . ?
Fe3 Al4 Al9 86.3(6) 4 . ?
Al4 Al4 Al9 118.3(6) 2_565 . ?
Al8 Al4 Al9 129.3(6) 4 . ?
Al8 Al4 Al9 92.6(6) . . ?
Al5 Al4 Al10 114.4(10) 5 2_565 ?
Fe1 Al4 Al10 140.1(8) . 2_565 ?
```

```
Al10 Al4 Al10 81.1(11) . 2_565 ?
Al5 Al4 Al10 63.6(7) 6_565 2_565 ?
Fe3 Al4 Al10 88.1(7) 4 2_565 ?
Al4 Al4 Al10 54.0(4) 2_565 2_565 ?
Al8 Al4 Al10 52.7(4) 4 2_565 ?
Al8 Al4 Al10 100.1(6) . 2_565 ?
Al9 Al4 Al10 167.3(7) . 2_565 ?
Al4 Al5 Fe1 61.6(8) 5 . ?
Al4 Al5 Fe3 125.6(13) 5 4_655 ?
Fe1 Al5 Fe3 133.7(12) . 4_655 ?
Al4 Al5 Al4 65.7(9) 5 6_655 ?
Fe1 Al5 Al4 114.0(10) . 6_655 ?
Fe3 Al5 Al4 61.2(8) 4_655 6_655 ?
Al4 Al5 Fe2 166.3(15) 5 11 ?
Fe1 Al5 Fe2 121.3(11) . 11 ?
Fe3 Al5 Fe2 63.5(8) 4_655 11 ?
Al4 Al5 Fe2 119.4(12) 6_655 11 ?
Al4 Al5 Al9 69.6(10) 5 . ?
Fe1 Al5 Al9 52.4(8) . . ?
Fe3 Al5 Al9 164.5(12) 4_655 . ?
Al4 Al5 Al9 132.7(12) 6_655 . ?
Fe2 Al5 Al9 101.1(10) 11 . ?
Al4 Al5 Al6 117.8(13) 5 7_655 ?
Fe1 Al5 Al6 164.0(14) . 7_655 ?
Fe3 Al5 Al6 61.0(9) 4_655 7_655 ?
Al4 Al5 Al6 77.4(10) 6_655 7_655 ?
Fe2 Al5 Al6 55.1(7) 11 7_655 ?
Al9 Al5 Al6 111.7(12) . 7_655 ?
Al4 Al5 Al11 128.0(13) 5 . ?
Fe1 Al5 Al11 66.9(9) . . ?
Fe3 Al5 Al11 95.1(9) 4_655 . ?
Al4 Al5 Al11 149.3(13) 6_655 . ?
Fe2 Al5 Al11 55.0(8) 11 . ?
Al9 Al5 Al11 74.0(8) . . ?
Al6 Al5 Al11 109.5(12) 7_655 . ?
Al4 Al5 Al10 56.9(8) 5 5 ?
Fe1 Al5 Al10 54.5(6) . 5 ?
Fe3 Al5 Al10 89.5(10) 4_655 5 ?
Al4 Al5 Al10 64.7(9) 6_655 5 ?
Fe2 Al5 Al10 136.6(13) 11 5 ?
Al9 Al5 Al10 102.7(10) . 5 ?
Al6 Al5 Al10 140.3(13) 7_655 5 ?
Al11 Al5 Al10 98.3(11) . 5 ?
Al4 Al5 Al6 114.9(13) 5 10 ?
Fe1 Al5 Al6 111.1(13) . 10 ?
Fe3 Al5 Al6 105.0(11) 4_655 10 ?
Al4 Al5 Al6 125.7(11) 6_655 10 ?
Fe2 Al5 Al6 51.5(7) 11 10 ?
Al9 Al5 Al6 62.4(10) . 10 ?
Al6 Al5 Al6 53.6(8) 7_655 10 ?
Al11 Al5 Al6 76.9(11) . 10 ?
Al10 Al5 Al6 165.0(14) 5 10 ?
Fe2 Al6 Fe2 142.7(11) 4 10_554 ?
Fe2 Al6 Al6 128.1(15) 4 15_554 ?
Fe2 Al6 Al6 57.9(11) 10_554 15_554 ?
Fe2 Al6 Al6 59.3(12) 4 16_454 ?
Fe2 Al6 Al6 139.3(12) 10_554 16_454 ?
```

```
Al6 Al6 Al6 81.8(8) 15_554 16_454 ?
Fe2 Al6 Fe3 117.3(10) 4 . ?
Fe2 Al6 Fe3 64.0(8) 10_554 . ?
Al6 Al6 Fe3 113.5(14) 15_554 . ?
Al6 Al6 Fe3 149.5(16) 16_454 . ?
Fe2 Al6 Al5 154.5(13) 4 8_565 ?
Fe2 Al6 Al5 60.9(9) 10_554 8_565 ?
Al6 Al6 Al5 67.8(11) 15_554 8_565 ?
Al6 Al6 Al5 111.3(11) 16_454 8_565 ?
Fe3 Al6 Al5 57.1(8) . 8_565 ?
Fe2 Al6 Al9 50.9(8) 4 4 ?
Fe2 Al6 Al9 151.9(12) 10_554 4 ?
Al6 Al6 Al9 142.0(11) 15_554 4 ?
Al6 Al6 Al9 66.5(12) 16_454 4 ?
Fe3 Al6 Al9 87.9(9) . 4 ?
Al5 Al6 Al9 103.8(11) 8_565 4 ?
Fe2 Al6 Al12 56.8(9) 4 . ?
Fe2 Al6 Al12 100.3(9) 10_554 . ?
Al6 Al6 Al12 151.2(16) 15_554 . ?
Al6 Al6 Al12 115.1(16) 16_454 . ?
Fe3 Al6 Al12 64.2(12) . . ?
Al5 Al6 Al12 120.9(15) 8_565 . ?
Al9 Al6 Al12 65.9(8) 4 . ?
Fe2 Al6 Al9 99.3(9) 4 10_554 ?
Fe2 Al6 Al9 48.4(7) 10_554 10_554 ?
Al6 Al6 Al9 59.8(11) 15_554 10_554 ?
Al6 Al6 Al9 108.5(14) 16_454 10_554 ?
Fe3 Al6 Al9 102.0(10) . 10_554 ?
Al5 Al6 Al9 106.2(11) 8_565 10_554 ?
Al9 Al6 Al9 149.0(10) 4 10_554 ?
Al12 Al6 Al9 91.8(11) . 10_554 ?
Fe2 Al6 Al5 57.7(9) 4 10_554 ?
Fe2 Al6 Al5 100.9(11) 10_554 10_554 ?
Al6 Al6 Al5 73.5(12) 15_554 10_554 ?
Al6 Al6 Al5 58.6(13) 16_454 10_554 ?
Fe3 Al6 Al5 149.2(12) . 10_554 ?
Al5 Al6 Al5 141.2(13) 8_565 10_554 ?
Al9 Al6 Al5 104.5(12) 4 10_554 ?
Al12 Al6 Al5 94.9(15) . 10_554 ?
Al9 Al6 Al5 54.0(8) 10_554 10_554 ?
Fe3 Al7 Fe3 121.5(19) 2 . ?
Fe3 Al7 Al8 49.9(6) 2 2 ?
Fe3 Al7 Al8 171.4(18) . 2 ?
Fe3 Al7 Al8 171.4(18) 2 . ?
Fe3 Al7 Al8 49.9(6) . . ?
Al8 Al7 Al8 138.7(19) 2 . ?
Fe3 Al8 Fe1 159.5(4) . . ?
Fe3 Al8 Al10 107.2(6) . . ?
Fe1 Al8 Al10 60.1(5) . . ?
Fe3 Al8 Al10 107.2(6) . 5 ?
Fe1 Al8 Al10 60.1(5) . 5 ?
Al10 Al8 Al10 91.3(15) . 5 ?
Fe3 Al8 Al4 62.5(5) . 8 ?
Fe1 Al8 Al4 119.8(5) . 8 ?
Al10 Al8 Al4 150.8(8) . 8 ?
Al10 Al8 Al4 68.4(7) 5 8 ?
Fe3 Al8 Al4 62.5(5) . 4 ?
```

```
Fe1 Al8 Al4 119.8(5) . 4 ?
Al10 Al8 Al4 68.4(7) . 4 ?
Al10 Al8 Al4 150.8(8) 5 4 ?
Al4 Al8 Al4 120.4(9) 8 4 ?
Fe3 Al8 Al7 59.5(9) . . ?
Fe1 Al8 Al7 141.0(11) . . ?
Al10 Al8 Al7 134.1(8) . . ?
Al10 Al8 Al7 134.1(8) 5 . ?
Al4 Al8 Al7 66.9(5) 8 . ?
Al4 Al8 Al7 66.9(5) 4 . ?
Fe3 Al8 Al4 134.3(4) . . ?
Fe1 Al8 Al4 54.9(5) . . ?
Al10 Al8 Al4 56.6(6) . . ?
Al10 Al8 Al4 114.9(9) 5 . ?
Al4 Al8 Al4 150.7(7) 8 . ?
Al4 Al8 Al4 72.1(5) 4 . ?
Al7 Al8 Al4 99.3(7) . . ?
Fe3 Al8 Al4 134.3(4) . 5 ?
Fe1 Al8 Al4 54.9(5) . 5 ?
Al10 Al8 Al4 114.9(9) . 5 ?
Al10 Al8 Al4 56.6(6) 5 5 ?
Al4 Al8 Al4 72.1(5) 8 5 ?
Al4 Al8 Al4 150.7(7) 4 5 ?
Al7 Al8 Al4 99.3(7) . 5 ?
Al4 Al8 Al4 85.7(9) . 5 ?
Fe2 Al9 Fe1 172.0(5) . . ?
Fe2 Al9 Al5 117.1(8) . 5 ?
Fe1 Al9 Al5 59.2(6) . 5 ?
Fe2 Al9 Al5 117.1(8) . . ?
Fe1 Al9 Al5 59.2(6) . . ?
Al5 Al9 Al5 102.1(13) 5 . ?
Fe2 Al9 Al6 57.4(6) . 8 ?
Fe1 Al9 Al6 124.9(7) . 8 ?
Al5 Al9 Al6 170.0(9) 5 8 ?
Al5 Al9 Al6 75.4(8) . 8 ?
Fe2 Al9 Al6 57.4(6) . 4 ?
Fe1 Al9 Al6 124.9(7) . 4 ?
Al5 Al9 Al6 75.4(8) 5 4 ?
Al5 Al9 Al6 170.0(9) . 4 ?
Al6 Al9 Al6 105.3(12) 8 4 ?
Fe2 Al9 Al4 128.4(4) . 5 ?
Fe1 Al9 Al4 56.0(5) . 5 ?
Al5 Al9 Al4 114.5(10) 5 5 ?
Al5 Al9 Al4 50.7(6) . 5 ?
Al6 Al9 Al4 71.7(6) 8 5 ?
Al6 Al9 Al4 139.3(7) 4 5 ?
Fe2 Al9 Al4 128.4(4) . . ?
Fe1 Al9 Al4 56.0(5) . . ?
Al5 Al9 Al4 50.7(6) 5 . ?
Al5 Al9 Al4 114.5(10) . . ?
Al6 Al9 Al4 139.3(7) 8 . ?
Al6 Al9 Al4 71.7(6) 4 . ?
Al4 Al9 Al4 84.7(8) 5 . ?
Fe2 Al9 Al6 55.6(7) . 10 ?
Fe1 Al9 Al6 118.5(6) . 10 ?
Al5 Al9 Al6 116.4(8) 5 10 ?
Al5 Al9 Al6 63.6(7) . 10 ?
```

```
Al6 Al9 Al6 53.7(9) 8 10 ?
Al6 Al9 Al6 108.4(8) 4 10 ?
Al4 Al9 Al6 101.8(6) 5 10 ?
Al4 Al9 Al6 167.0(5) . 10 ?
Fe2 Al9 Al6 55.6(7) . 14 ?
Fe1 Al9 Al6 118.5(6) . 14 ?
Al5 Al9 Al6 63.6(7) 5 14 ?
Al5 Al9 Al6 116.4(8) . 14 ?
Al6 Al9 Al6 108.4(8) 8 14 ?
Al6 Al9 Al6 53.7(9) 4 14 ?
Al4 Al9 Al6 167.0(5) 5 14 ?
Al4 Al9 Al6 101.8(6) . 14 ?
Al6 Al9 Al6 69.7(12) 10 14 ?
Fe1 Al10 Fe1 135.2(14) 3_565 . ?
Fe1 Al10 Al8 119.8(7) 3_565 . ?
Fe1 Al10 Al8 58.8(5) . . ?
Fe1 Al10 Al8 58.8(5) 3_565 3_565 ?
Fe1 Al10 Al8 119.8(7) . 3_565 ?
Al8 Al10 Al8 176.8(15) . 3_565 ?
Fe1 Al10 Al4 58.5(4) 3_565 3_565 ?
Fe1 Al10 Al4 165.9(13) . 3_565 ?
Al8 Al10 Al4 113.8(8) . 3_565 ?
Al8 Al10 Al4 68.3(5) 3_565 3_565 ?
Fe1 Al10 Al4 165.9(13) 3_565 . ?
Fe1 Al10 Al4 58.5(4) . . ?
Al8 Al10 Al4 68.3(5) . . ?
Al8 Al10 Al4 113.8(8) 3_565 . ?
Al4 Al10 Al4 108.1(13) 3_565 . ?
Fe1 Al10 Al13 67.6(7) 3_565 . ?
Fe1 Al10 Al13 67.6(7) . . ?
Al8 Al10 Al13 88.4(7) . . ?
Al8 Al10 Al13 88.4(7) 3_565 . ?
Al4 Al10 Al13 125.9(7) 3_565 . ?
Al4 Al10 Al13 125.9(7) . . ?
Fe1 Al10 Al5 53.2(7) 3_565 8_565 ?
Fe1 Al10 Al5 130.4(8) . 8_565 ?
Al8 Al10 Al5 76.0(8) . 8_565 ?
Al8 Al10 Al5 104.2(8) 3_565 8_565 ?
Al4 Al10 Al5 50.9(7) 3_565 8_565 ?
Al4 Al10 Al5 123.8(10) . 8_565 ?
Al13 Al10 Al5 93.7(8) . 8_565 ?
Fe1 Al10 Al5 130.4(8) 3_565 5 ?
Fe1 Al10 Al5 53.2(7) . 5 ?
Al8 Al10 Al5 104.2(8) . 5 ?
Al8 Al10 Al5 76.0(8) 3_565 5 ?
Al4 Al10 Al5 123.8(10) 3_565 5 ?
Al4 Al10 Al5 50.9(7) . 5 ?
Al13 Al10 Al5 93.7(8) . 5 ?
Al5 Al10 Al5 172.7(16) 8_565 5 ?
Fe1 Al10 Al4 101.0(4) 3_565 4 ?
Fe1 Al10 Al4 110.5(5) . 4 ?
Al8 Al10 Al4 58.9(6) . 4 ?
Al8 Al10 Al4 123.9(9) 3_565 4 ?
Al4 Al10 Al4 57.8(7) 3_565 4 ?
Al4 Al10 Al4 72.6(8) . 4 ?
Al13 Al10 Al4 135.2(5) . 4 ?
Al5 Al10 Al4 51.8(6) 8_565 4 ?
```

```
Al5 Al10 Al4 121.9(10) 5 4 ?
Fe1 Al10 Al4 110.5(5) 3_565 2_565 ?
Fe1 Al10 Al4 101.0(4) . 2_565 ?
Al8 Al10 Al4 123.9(9) . 2_565 ?
Al8 Al10 Al4 58.9(6) 3_565 2_565 ?
Al4 Al10 Al4 72.6(8) 3_565 2_565 ?
Al4 Al10 Al4 57.8(7) . 2_565 ?
Al13 Al10 Al4 135.2(5) . 2_565 ?
Al5 Al10 Al4 121.9(10) 8_565 2_565 ?
Al5 Al10 Al4 51.8(6) 5 2_565 ?
Al4 Al10 Al4 89.7(11) 4 2_565 ?
Fe2 Al11 Fe2 116.6(17) 12 11 ?
Fe2 Al11 Al12 58.3(8) 12 9 ?
Fe2 Al11 Al12 58.3(8) 11 9 ?
Fe2 Al11 Al5 59.0(6) 12 5 ?
Fe2 Al11 Al5 137.5(9) 11 5 ?
Al12 Al11 Al5 102.2(9) 9 5 ?
Fe2 Al11 Al5 137.5(9) 12 6_665 ?
Fe2 Al11 Al5 59.0(6) 11 6_665 ?
Al12 Al11 Al5 102.2(9) 9 6_665 ?
Al5 Al11 Al5 155.6(17) 5 6_665 ?
Fe2 Al11 Al5 59.0(6) 12 2_665 ?
Fe2 Al11 Al5 137.5(9) 11 2_665 ?
Al12 Al11 Al5 102.2(9) 9 2_665 ?
Al5 Al11 Al5 80.1(13) 5 2_665 ?
Al5 Al11 Al5 94.8(13) 6_665 2_665 ?
Fe2 Al11 Al5 137.5(9) 12 . ?
Fe2 Al11 Al5 59.0(6) 11 . ?
Al12 Al11 Al5 102.2(9) 9 . ?
Al5 Al11 Al5 94.8(13) 5 . ?
Al5 Al11 Al5 80.1(13) 6_665 . ?
Al5 Al11 Al5 155.6(17) 2_665 . ?
Fe2 Al11 Fe1 109.4(3) 12 2_665 ?
Fe2 Al11 Fe1 109.4(3) 11 2_665 ?
Al12 Al11 Fe1 129.1(6) 9 2_665 ?
Al5 Al11 Fe1 111.5(11) 5 2_665 ?
Al5 Al11 Fe1 50.8(6) 6_665 2_665 ?
Al5 Al11 Fe1 50.8(6) 2_665 2_665 ?
Al5 Al11 Fe1 111.4(11) . 2_665 ?
Fe2 Al11 Fe1 109.4(3) 12 . ?
Fe2 Al11 Fe1 109.4(3) 11 . ?
Al12 Al11 Fe1 129.1(6) 9 . ?
Al5 Al11 Fe1 50.8(6) 5 . ?
Al5 Al11 Fe1 111.5(11) 6_665 . ?
Al5 Al11 Fe1 111.5(11) 2_665 . ?
Al5 Al11 Fe1 50.8(6) . . ?
Fe1 Al11 Fe1 101.8(13) 2_665 . ?
Fe2 Al12 Fe2 119(2) 4 3 ?
Fe2 Al12 Al11 59.4(12) 4 9_444 ?
Fe2 Al12 Al11 59.4(12) 3 9_444 ?
Fe2 Al12 Al6 55.1(7) 4 6 ?
Fe2 Al12 Al6 117.8(14) 3 6 ?
Al11 Al12 Al6 84.0(14) 9_444 6 ?
Fe2 Al12 Al6 117.8(14) 4 2 ?
Fe2 Al12 Al6 55.1(7) 3 2 ?
Al11 Al12 Al6 84.0(14) 9_444 2 ?
Al6 Al12 Al6 74.2(12) 6 2 ?
```

```
Fe2 Al12 Al6 117.8(14) 4 5 ?
Fe2 Al12 Al6 55.1(7) 3 5 ?
Al11 Al12 Al6 84.0(14) 9_444 5 ?
Al6 Al12 Al6 168(3) 6 5 ?
Al6 Al12 Al6 104.5(13) 2 5 ?
Fe2 Al12 Al6 55.1(7) 4 . ?
Fe2 Al12 Al6 117.8(14) 3 . ?
Al11 Al12 Al6 84.0(14) 9_444 . ?
Al6 Al12 Al6 104.5(13) 6 . ?
Al6 Al12 Al6 168(3) 2 . ?
Al6 Al12 Al6 74.2(11) 5 . ?
Fe2 Al12 Fe3 108.6(4) 4 . ?
Fe2 Al12 Fe3 108.6(4) 3 . ?
Al11 Al12 Fe3 129.0(9) 9_444 . ?
Al6 Al12 Fe3 132.9(15) 6 . ?
Al6 Al12 Fe3 132.9(15) 2 . ?
Al6 Al12 Fe3 56.7(8) 5 . ?
Al6 Al12 Fe3 56.7(8) . . ?
Fe2 Al12 Fe3 108.6(4) 4 2 ?
Fe2 Al12 Fe3 108.6(4) 3 2 ?
Al11 Al12 Fe3 129.0(9) 9_444 2 ?
Al6 Al12 Fe3 56.7(8) 6 2 ?
Al6 Al12 Fe3 56.7(8) 2 2 ?
Al6 Al12 Fe3 132.9(15) 5 2 ?
Al6 Al12 Fe3 132.9(15) . 2 ?
Fe3 Al12 Fe3 102.1(18) . 2 ?
Al10 Al13 Al10 180.0 6_665 5 ?
Al10 Al13 Al10 90.000(4) 6_665 2_665 ?
Al10 Al13 Al10 90.0 5 2_665 ?
Al10 Al13 Al10 90.0 6_665 . ?
Al10 Al13 Al10 90.000(2) 5 . ?
Al10 Al13 Al10 180.0 2_665 . ?
Al10 Al13 Fe1 55.28(11) 6_665 3_565 ?
Al10 Al13 Fe1 124.72(11) 5 3_565 ?
Al10 Al13 Fe1 124.72(11) 2_665 3_565 ?
Al10 Al13 Fe1 55.28(11) . 3_565 ?
Al10 Al13 Fe1 124.72(11) 6_665 4_655 ?
Al10 Al13 Fe1 55.28(11) 5 4_655 ?
Al10 Al13 Fe1 55.28(11) 2_665 4_655 ?
Al10 Al13 Fe1 124.72(11) . 4_655 ?
Fe1 Al13 Fe1 107.3(4) 3_565 4_655 ?
Al10 Al13 Fe1 124.72(11) 6_665 . ?
Al10 Al13 Fe1 55.28(11) 5 . ?
Al10 Al13 Fe1 124.72(11) 2_665 . ?
Al10 Al13 Fe1 55.28(11) . . ?
Fe1 Al13 Fe1 110.6(2) 3_565 . ?
Fe1 Al13 Fe1 110.6(2) 4_655 . ?
Al10 Al13 Fe1 55.28(11) 6_665 2_665 ?
Al10 Al13 Fe1 124.72(11) 5 2_665 ?
Al10 Al13 Fe1 55.28(11) 2_665 2_665 ?
Al10 Al13 Fe1 124.72(11) . 2_665 ?
Fe1 Al13 Fe1 110.6(2) 3_565 2_665 ?
Fe1 Al13 Fe1 110.6(2) 4_655 2_665 ?
Fe1 Al13 Fe1 107.3(4) . 2_665 ?

_diffrn_measured_fraction_theta_max    0.395
_diffrn_reflns_theta_full              68.24
```

```
_diffrn_measured_fraction_theta_full    0.395
_refine_diff_density_max     1.152
_refine_diff_density_min    -1.574
_refine_diff_density_rms     0.289
```

8.2 Brucite CIF

```
data_shelxl

_audit_creation_method            SHELXL-97
_chemical_name_systematic
;
 ?
;
_chemical_name_common             Brucite
_chemical_formula_moiety          ?
_chemical_formula_sum
 'H Mg O'
_chemical_formula_weight          41.32

_citation_journal_full Crystallography Reports
_citation_journal_volume 42
_citation_langauge English
_citation_page_first  774
_citation_page_last 777
_citation_year 1997
_citation_author_name 'Zhukhlistov AP, Avilov AS, Ferraris D'


loop_
 _atom_type_symbol
 _atom_type_description
 _atom_type_scat_dispersion_real
 _atom_type_scat_dispersion_imag
 _atom_type_scat_source
 'Mg'  'Mg'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
 'O'  'O'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
 'H'  'H'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'

_symmetry_cell_setting            ?
_symmetry_space_group_name_H-M    'P -3 m 1'

loop_
 _symmetry_equiv_pos_as_xyz
 'x, y, z'
 '-y, x-y, z'
 '-x+y, -x, z'
 'x-y, -y, -z'
 '-x, -x+y, -z'
 'y, x, -z'
 '-x, -y, -z'
 'y, -x+y, -z'
 'x-y, x, -z'
 '-x+y, y, z'
 'x, x-y, z'
 '-y, -x, z'

_cell_length_a                    3.149
```

```
_cell_length_b                    3.149
_cell_length_c                    4.769
_cell_angle_alpha                 90.00
_cell_angle_beta                  90.00
_cell_angle_gamma                 120.00
_cell_volume                      41.0
_cell_formula_units_Z             12
_cell_measurement_temperature     293(2)

_exptl_crystal_density_diffrn     20.104
_exptl_crystal_density_method     'not measured'
_exptl_crystal_F_000              93
_exptl_absorpt_coefficient_mu     0.088

_diffrn_ambient_temperature       293(2)
_diffrn_radiation_wavelength      ?
_diffrn_radiation_type            Electron
_diffrn_radiation_source          Microscope
_diffrn_reflns_number             59
_diffrn_reflns_av_R_equivalents   0.0000
_diffrn_reflns_av_sigmaI/netI     0.0121
_diffrn_reflns_limit_h_min        0
_diffrn_reflns_limit_h_max        4
_diffrn_reflns_limit_k_min        0
_diffrn_reflns_limit_k_max        4
_diffrn_reflns_limit_l_min        0
_diffrn_reflns_limit_l_max        6
_diffrn_reflns_theta_min          13.00
_diffrn_reflns_theta_max          81.52
_reflns_number_total              59
_reflns_number_gt                 57
_reflns_threshold_expression      >2sigma(I)

_computing_structure_solution     'SHELXS-97 (Sheldrick, 1990)'
_computing_structure_refinement   'SHELXL-97 (Sheldrick, 1997)'
_computing_molecular_graphics     ?
_computing_publication_material   maXus

_refine_special_details
;
 Refinement of F^2^ against ALL reflections.  The weighted R-factor wR and
 goodness of fit S are based on F^2^, conventional R-factors R are based
 on F, with F set to zero for negative F^2^. The threshold expression of
 F^2^ > 2sigma(F^2^) is used only for calculating R-factors(gt) etc. and is
 not relevant to the choice of reflections for refinement.  R-factors based
 on F^2^ are statistically about twice as large as those based on F, and R-
 factors based on ALL data will be even larger.
;

_refine_ls_structure_factor_coef  Fsqd
_refine_ls_matrix_type            full
_refine_ls_weighting_scheme
 'calc w=1/[\s^2^(Fo^2^)+(0.1000P)^2^+0.0000P] where P=(Fo^2^+2Fc^2^)/3'
_atom_sites_solution_primary      direct
_atom_sites_solution_secondary    difmap
_atom_sites_solution_hydrogens    geom
_refine_ls_hydrogen_treatment     mixed
```

```
_refine_ls_extinction_method      none
_refine_ls_extinction_coef        ?
_refine_ls_number_reflns          59
_refine_ls_number_parameters      4
_refine_ls_number_restraints      0
_refine_ls_R_factor_all           0.1108
_refine_ls_R_factor_gt            0.1060
_refine_ls_wR_factor_ref          0.4823
_refine_ls_wR_factor_gt           0.4322
_refine_ls_goodness_of_fit_ref    4.892
_refine_ls_restrained_S_all       4.892
_refine_ls_shift/su_max           0.020
_refine_ls_shift/su_mean          0.013

loop_
 _atom_site_label
 _atom_site_type_symbol
 _atom_site_fract_x
 _atom_site_fract_y
 _atom_site_fract_z
 _atom_site_U_iso_or_equiv
 _atom_site_adp_type
 _atom_site_occupancy
 _atom_site_calc_flag
 _atom_site_refinement_flags
 _atom_site_disorder_assembly
 _atom_site_disorder_group
Mg1 Mg 0.0000 0.0000 0.0000 0.0080(6) Uiso 1 d S . .
O2 O 0.3330 0.6670 0.2228(12) 0.0134(7) Uiso 1 d S . .
H2 H 0.3330 0.6670 0.4308 0.016 Uiso 1 d SR . .

_geom_special_details
;
 All esds (except the esd in the dihedral angle between two l.s. planes)
 are estimated using the full covariance matrix.  The cell esds are taken
 into account individually in the estimation of esds in distances, angles
 and torsion angles; correlations between esds in cell parameters are only
 used when they are defined by crystal symmetry.  An approximate (isotropic)
 treatment of cell esds is used for estimating esds involving l.s. planes.
;

loop_
 _geom_bond_atom_site_label_1
 _geom_bond_atom_site_label_2
 _geom_bond_distance
 _geom_bond_site_symmetry_2
 _geom_bond_publ_flag
Mg1 O2 2.104(3) 7_565 ?
Mg1 O2 2.104(3) 1_545 ?
Mg1 O2 2.107(3) 7 ?
Mg1 O2 2.107(3) 7_665 ?
Mg1 O2 2.107(3) 1_445 ?
Mg1 O2 2.107(3) . ?
Mg1 Mg1 3.1490 1_445 ?
Mg1 Mg1 3.1490 1_665 ?
Mg1 Mg1 3.1490 1_455 ?
Mg1 Mg1 3.1490 1_655 ?
```

```
Mg1 Mg1 3.1490 1_545 ?
Mg1 Mg1 3.1490 1_565 ?
O2 Mg1 2.104(3) 1_565 ?
O2 Mg1 2.107(3) 1_665 ?
O2 H2 0.9920 . ?

loop_
 _geom_angle_atom_site_label_1
 _geom_angle_atom_site_label_2
 _geom_angle_atom_site_label_3
 _geom_angle
 _geom_angle_site_symmetry_1
 _geom_angle_site_symmetry_3
 _geom_angle_publ_flag
O2 Mg1 O2 180.0 7_565 1_545 ?
O2 Mg1 O2 96.81(18) 7_565 7 ?
O2 Mg1 O2 83.19(18) 1_545 7 ?
O2 Mg1 O2 96.81(18) 7_565 7_665 ?
O2 Mg1 O2 83.19(18) 1_545 7_665 ?
O2 Mg1 O2 96.74(18) 7 7_665 ?
O2 Mg1 O2 83.19(18) 7_565 1_445 ?
O2 Mg1 O2 96.81(18) 1_545 1_445 ?
O2 Mg1 O2 83.26(18) 7 1_445 ?
O2 Mg1 O2 180.0 7_665 1_445 ?
O2 Mg1 O2 83.19(18) 7_565 . ?
O2 Mg1 O2 96.81(18) 1_545 . ?
O2 Mg1 O2 180.0 7 . ?
O2 Mg1 O2 83.26(18) 7_665 . ?
O2 Mg1 O2 96.74(18) 1_445 . ?
O2 Mg1 Mg1 90.0 7_565 1_445 ?
O2 Mg1 Mg1 90.0 1_545 1_445 ?
O2 Mg1 Mg1 41.63(9) 7 1_445 ?
O2 Mg1 Mg1 138.37(9) 7_665 1_445 ?
O2 Mg1 Mg1 41.63(9) 1_445 1_445 ?
O2 Mg1 Mg1 138.37(9) . 1_445 ?
O2 Mg1 Mg1 90.0 7_565 1_665 ?
O2 Mg1 Mg1 90.0 1_545 1_665 ?
O2 Mg1 Mg1 138.37(9) 7 1_665 ?
O2 Mg1 Mg1 41.63(9) 7_665 1_665 ?
O2 Mg1 Mg1 138.37(9) 1_445 1_665 ?
O2 Mg1 Mg1 41.63(9) . 1_665 ?
Mg1 Mg1 Mg1 180.0 1_445 1_665 ?
O2 Mg1 Mg1 41.62(9) 7_565 1_455 ?
O2 Mg1 Mg1 138.38(9) 1_545 1_455 ?
O2 Mg1 Mg1 90.0 7 1_455 ?
O2 Mg1 Mg1 138.43(9) 7_665 1_455 ?
O2 Mg1 Mg1 41.57(9) 1_445 1_455 ?
O2 Mg1 Mg1 90.0 . 1_455 ?
Mg1 Mg1 Mg1 60.0 1_445 1_455 ?
Mg1 Mg1 Mg1 120.0 1_665 1_455 ?
O2 Mg1 Mg1 138.38(9) 7_565 1_655 ?
O2 Mg1 Mg1 41.62(9) 1_545 1_655 ?
O2 Mg1 Mg1 90.0 7 1_655 ?
O2 Mg1 Mg1 41.57(9) 7_665 1_655 ?
O2 Mg1 Mg1 138.43(9) 1_445 1_655 ?
O2 Mg1 Mg1 90.0 . 1_655 ?
Mg1 Mg1 Mg1 120.0 1_445 1_655 ?
```

```
Mg1 Mg1 Mg1 60.0 1_665 1_655 ?
Mg1 Mg1 Mg1 180.0 1_455 1_655 ?
O2 Mg1 Mg1 138.38(9) 7_565 1_545 ?
O2 Mg1 Mg1 41.62(9) 1_545 1_545 ?
O2 Mg1 Mg1 41.57(9) 7 1_545 ?
O2 Mg1 Mg1 90.0 7_665 1_545 ?
O2 Mg1 Mg1 90.0 1_445 1_545 ?
O2 Mg1 Mg1 138.43(9) . 1_545 ?
Mg1 Mg1 Mg1 60.0 1_445 1_545 ?
Mg1 Mg1 Mg1 120.0 1_665 1_545 ?
Mg1 Mg1 Mg1 120.0 1_455 1_545 ?
Mg1 Mg1 Mg1 60.0 1_655 1_545 ?
O2 Mg1 Mg1 41.62(9) 7_565 1_565 ?
O2 Mg1 Mg1 138.38(9) 1_545 1_565 ?
O2 Mg1 Mg1 138.43(9) 7 1_565 ?
O2 Mg1 Mg1 90.0 7_665 1_565 ?
O2 Mg1 Mg1 90.0 1_445 1_565 ?
O2 Mg1 Mg1 41.57(9) . 1_565 ?
Mg1 Mg1 Mg1 120.0 1_445 1_565 ?
Mg1 Mg1 Mg1 60.0 1_665 1_565 ?
Mg1 Mg1 Mg1 60.0 1_455 1_565 ?
Mg1 Mg1 Mg1 120.0 1_655 1_565 ?
Mg1 Mg1 Mg1 180.0 1_545 1_565 ?
Mg1 O2 Mg1 96.81(18) 1_565 1_665 ?
Mg1 O2 Mg1 96.81(18) 1_565 . ?
Mg1 O2 Mg1 96.74(18) 1_665 . ?
Mg1 O2 H2 120.3 1_565 . ?
Mg1 O2 H2 120.3 1_665 . ?
Mg1 O2 H2 120.3 . . ?

_diffrn_measured_fraction_theta_max    0.648
_diffrn_reflns_theta_full              81.52
_diffrn_measured_fraction_theta_full   0.648
_refine_diff_density_max    0.445
_refine_diff_density_min   -0.143
_refine_diff_density_rms    0.054
```

8.3 10-Cyano-9,9',-Bianthryl

```
data_shelxl

_audit_creation_method            SHELXL-97
_chemical_name_systematic         '[9,9'-bianthryl}-10-carbonitrile'
;
 ?
;
_chemical_name_common             CNBA
_chemical_formula_moiety          ?
_chemical_formula_sum
 'C29 N'
_chemical_formula_weight          362.30

_citation_journal_full Acta Cryst.
_citation_journal_volume A51
_citation_langauge English
_citation_page_first 849
_citation_page_last 868
_citation_year 1995
_citation_author_name 'Voigt-Martin, I.G. Yan, D.H., Yakimansky, A.,
Schollmeyer, D., Gilmore, C., Bricogne, G.'

loop_
 _atom_type_symbol
 _atom_type_description
 _atom_type_scat_dispersion_real
 _atom_type_scat_dispersion_imag
 _atom_type_scat_source
 'C'  'C'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
 'N'  'N'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'

_symmetry_cell_setting            ?
_symmetry_space_group_name_H-M    'P21/c'

loop_
 _symmetry_equiv_pos_as_xyz
 'x, y, z'
 '-x, y+1/2, -z+1/2'
 '-x, -y, -z'
 'x, -y-1/2, z-1/2'

_cell_length_a                    14.702
_cell_length_b                    9.474
_cell_length_c                    15.422
_cell_angle_alpha                 90.00
_cell_angle_beta                  111.77
_cell_angle_gamma                 90.00
_cell_volume                      1994.9
_cell_formula_units_Z             4
_cell_measurement_temperature     293(2)

_exptl_crystal_density_diffrn     1.206
_exptl_crystal_density_method     'not measured'
```

```
_exptl_crystal_F_000              300
_exptl_absorpt_coefficient_mu     0.006

_diffrn_ambient_temperature       293(2)
_diffrn_radiation_wavelength      ?
_diffrn_radiation_type            Electron
_diffrn_radiation_source          Microscope
_diffrn_reflns_number             144
_diffrn_reflns_av_R_equivalents   0.0000
_diffrn_reflns_av_sigmaI/netI     0.0022
_diffrn_reflns_limit_h_min        -7
_diffrn_reflns_limit_h_max        6
_diffrn_reflns_limit_k_min        0
_diffrn_reflns_limit_k_max        6
_diffrn_reflns_limit_l_min        -7
_diffrn_reflns_limit_l_max        10
_diffrn_reflns_theta_min          4.45
_diffrn_reflns_theta_max          26.00
_reflns_number_total              144
_reflns_number_gt                 144
_reflns_threshold_expression      >2sigma(I)

_computing_structure_solution     'SHELXS-97 (Sheldrick, 1990)'
_computing_structure_refinement   'SHELXL-97 (Sheldrick, 1997)'
_computing_molecular_graphics     ?
_computing_publication_material   maXus

_refine_special_details
;
 Refinement of F^2^ against ALL reflections.  The weighted R-factor wR and
 goodness of fit S are based on F^2^, conventional R-factors R are based
 on F, with F set to zero for negative F^2^. The threshold expression of
 F^2^ > 2sigma(F^2^) is used only for calculating R-factors(gt) etc. and is
 not relevant to the choice of reflections for refinement.  R-factors based
 on F^2^ are statistically about twice as large as those based on F, and R-
 factors based on ALL data will be even larger.
;

_refine_ls_structure_factor_coef  Fsqd
_refine_ls_matrix_type            full
_refine_ls_weighting_scheme
 'calc w=1/[\s^2^(Fo^2^)+(0.1000P)^2^+0.0000P] where P=(Fo^2^+2Fc^2^)/3'
_atom_sites_solution_primary      direct
_atom_sites_solution_secondary    difmap
_atom_sites_solution_hydrogens    geom
_refine_ls_hydrogen_treatment     mixed
_refine_ls_extinction_method      none
_refine_ls_extinction_coef        ?
_refine_ls_number_reflns          144
_refine_ls_number_parameters      93
_refine_ls_number_restraints      57
_refine_ls_R_factor_all           0.3782
_refine_ls_R_factor_gt            0.3782
_refine_ls_wR_factor_ref          0.7524
_refine_ls_wR_factor_gt           0.7524
_refine_ls_goodness_of_fit_ref    17.870
_refine_ls_restrained_S_all       12.548
```

```
_refine_ls_shift/su_max           0.107
_refine_ls_shift/su_mean          0.024

loop_
 _atom_site_label
 _atom_site_type_symbol
 _atom_site_fract_x
 _atom_site_fract_y
 _atom_site_fract_z
 _atom_site_U_iso_or_equiv
 _atom_site_adp_type
 _atom_site_occupancy
 _atom_site_calc_flag
 _atom_site_refinement_flags
 _atom_site_disorder_assembly
 _atom_site_disorder_group
C1 C 0.2930(8) 0.1013(14) -0.014(2) 0.050 Uiso 1 d D . .
C2 C 0.2228(7) 0.0009(12) -0.063(3) 0.050 Uiso 1 d D . .
C3 C 0.1567(15) -0.071(2) -0.035(2) 0.050 Uiso 1 d D . .
C4 C 0.169(3) -0.029(4) 0.055(2) 0.050 Uiso 1 d D . .
C5 C 0.107(3) -0.098(5) 0.094(3) 0.050 Uiso 1 d D . .
C6 C 0.111(3) -0.067(4) 0.183(2) 0.050 Uiso 1 d D . .
C7 C 0.051(3) -0.122(6) 0.227(3) 0.050 Uiso 1 d D . .
C8 C 0.047(4) -0.095(5) 0.316(3) 0.050 Uiso 1 d D . .
C9 C 0.122(3) -0.001(5) 0.361(4) 0.050 Uiso 1 d D . .
C10 C 0.188(4) 0.071(5) 0.331(3) 0.050 Uiso 1 d D . .
C11 C 0.174(3) 0.033(4) 0.241(3) 0.050 Uiso 1 d D . .
C12 C 0.237(3) 0.101(4) 0.204(3) 0.050 Uiso 1 d D . .
C13 C 0.237(3) 0.071(4) 0.114(3) 0.050 Uiso 1 d D . .
C14 C 0.301(4) 0.139(5) 0.079(3) 0.050 Uiso 1 d D . .
C15 C 0.2809(13) 0.658(4) 0.2267(11) 0.050 Uiso 1 d D . .
C16 C 0.3661(10) 0.595(4) 0.2865(9) 0.050 Uiso 1 d D . .
C17 C 0.3834(11) 0.448(4) 0.2974(9) 0.050 Uiso 1 d D . .
C18 C 0.470(3) 0.382(4) 0.360(3) 0.050 Uiso 1 d D . .
C19 C 0.495(3) 0.239(4) 0.374(3) 0.050 Uiso 1 d D . .
C20 C 0.589(4) 0.202(4) 0.441(3) 0.050 Uiso 1 d D . .
C21 C 0.589(4) 0.053(4) 0.439(3) 0.050 Uiso 1 d D . .
C22 C 0.521(3) -0.050(5) 0.391(3) 0.050 Uiso 1 d D . .
C23 C 0.430(4) -0.005(4) 0.327(3) 0.050 Uiso 1 d D . .
C24 C 0.420(3) 0.143(4) 0.318(3) 0.050 Uiso 1 d D . .
C25 C 0.328(3) 0.194(4) 0.256(3) 0.050 Uiso 1 d D . .
C26 C 0.317(3) 0.343(4) 0.249(3) 0.050 Uiso 1 d D . .
C27 C 0.227(4) 0.396(4) 0.185(3) 0.050 Uiso 1 d D . .
C28 C 0.220(4) 0.544(4) 0.183(3) 0.050 Uiso 1 d D . .
C29 C 0.016(4) -0.187(6) 0.056(4) 0.050 Uiso 1 d D . .
N30 N -0.065(5) -0.269(6) -0.008(4) 0.050 Uiso 1 d D . .

_geom_special_details
;
 All esds (except the esd in the dihedral angle between two l.s. planes)
 are estimated using the full covariance matrix.  The cell esds are taken
 into account individually in the estimation of esds in distances, angles
 and torsion angles; correlations between esds in cell parameters are only
 used when they are defined by crystal symmetry.  An approximate (isotropic)
 treatment of cell esds is used for estimating esds involving l.s. planes.
;
```

```
loop_
 _geom_bond_atom_site_label_1
 _geom_bond_atom_site_label_2
 _geom_bond_distance
 _geom_bond_site_symmetry_2
 _geom_bond_publ_flag
C1 C14 1.43(4) . ?
C1 C2 1.40(2) . ?
C2 C3 1.39(4) . ?
C3 C4 1.39(4) . ?
C4 C5 1.41(4) . ?
C4 C13 1.43(4) . ?
C5 C6 1.39(4) . ?
C5 C29 1.51(4) . ?
C6 C7 1.39(5) . ?
C6 C11 1.39(4) . ?
C7 C8 1.42(5) . ?
C8 C9 1.38(5) . ?
C9 C10 1.39(5) . ?
C10 C11 1.38(4) . ?
C11 C12 1.41(5) . ?
C12 C13 1.42(4) . ?
C12 C25 1.56(4) . ?
C13 C14 1.41(5) . ?
C15 C16 1.38(3) . ?
C15 C28 1.40(4) . ?
C16 C17 1.41(4) . ?
C17 C18 1.43(4) . ?
C17 C26 1.40(4) . ?
C18 C19 1.40(4) . ?
C19 C24 1.44(4) . ?
C19 C20 1.42(4) . ?
C20 C21 1.41(4) . ?
C21 C22 1.40(5) . ?
C22 C23 1.40(5) . ?
C23 C24 1.41(4) . ?
C24 C25 1.42(4) . ?
C25 C26 1.42(4) . ?
C26 C27 1.41(4) . ?
C27 C28 1.41(4) . ?
C29 N30 1.45(5) . ?

loop_
 _geom_angle_atom_site_label_1
 _geom_angle_atom_site_label_2
 _geom_angle_atom_site_label_3
 _geom_angle
 _geom_angle_site_symmetry_1
 _geom_angle_site_symmetry_3
 _geom_angle_publ_flag
C14 C1 C2 119(3) . . ?
C3 C2 C1 128(4) . . ?
C2 C3 C4 109(3) . . ?
C5 C4 C13 117(4) . . ?
C5 C4 C3 115(3) . . ?
C13 C4 C3 128(4) . . ?
C4 C5 C6 122(4) . . ?
```

```
C4 C5 C29 135(4) . . ?
C6 C5 C29 102(4) . . ?
C5 C6 C7 127(4) . . ?
C5 C6 C11 123(4) . . ?
C7 C6 C11 110(4) . . ?
C8 C7 C6 131(5) . . ?
C7 C8 C9 107(5) . . ?
C10 C9 C8 132(5) . . ?
C9 C10 C11 110(5) . . ?
C12 C11 C6 116(4) . . ?
C12 C11 C10 115(4) . . ?
C6 C11 C10 130(4) . . ?
C11 C12 C13 123(4) . . ?
C11 C12 C25 128(4) . . ?
C13 C12 C25 108(3) . . ?
C12 C13 C4 120(4) . . ?
C12 C13 C14 122(4) . . ?
C4 C13 C14 118(4) . . ?
C1 C14 C13 117(4) . . ?
C16 C15 C28 105(4) . . ?
C15 C16 C17 125(2) . . ?
C18 C17 C16 126(3) . . ?
C18 C17 C26 109(3) . . ?
C16 C17 C26 125(2) . . ?
C19 C18 C17 131(4) . . ?
C18 C19 C24 114(3) . . ?
C18 C19 C20 119(3) . . ?
C24 C19 C20 127(4) . . ?
C21 C20 C19 104(4) . . ?
C20 C21 C22 135(5) . . ?
C23 C22 C21 118(5) . . ?
C22 C23 C24 114(4) . . ?
C25 C24 C23 116(4) . . ?
C25 C24 C19 121(3) . . ?
C23 C24 C19 123(4) . . ?
C26 C25 C24 116(3) . . ?
C26 C25 C12 118(3) . . ?
C24 C25 C12 126(3) . . ?
C25 C26 C27 118(3) . . ?
C25 C26 C17 128(3) . . ?
C27 C26 C17 114(3) . . ?
C28 C27 C26 114(4) . . ?
C27 C28 C15 136(5) . . ?
N30 C29 C5 161(6) . . ?

_diffrn_measured_fraction_theta_max     0.171
_diffrn_reflns_theta_full               26.00
_diffrn_measured_fraction_theta_full    0.171
_refine_diff_density_max    0.176
_refine_diff_density_min   -0.200
_refine_diff_density_rms    0.052
```

8.4 Basic Copper Chloride

```
data_shelxl

_audit_creation_method            SHELXL-97
_chemical_name_systematic
;
 ?
;
_chemical_name_common             Basic copper chloride
_chemical_formula_moiety          ?
_chemical_formula_sum
 'Cl Cu2 O2'
_chemical_formula_weight          194.55

_citation_journal_full Sov. Phys.-Crystallogr.
_citation_journal_volume 3
_citation_langauge Russian
_citation_page_first 445
_citation_page_last 451
_citation_year 1958
_citation_author_name 'Voronova, A.A., Vainshtein, B.K.'

loop_
 _atom_type_symbol
 _atom_type_description
 _atom_type_scat_dispersion_real
 _atom_type_scat_dispersion_imag
 _atom_type_scat_source
 'Cu'  'Cu'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
 'Cl'  'Cl'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
 'O'  'O'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'

_symmetry_cell_setting            ?
_symmetry_space_group_name_H-M    'P21/m

loop_
 _symmetry_equiv_pos_as_xyz
 'x, y, z'
 '-x, y+1/2, -z'
 '-x, -y, -z'
 'x, -y-1/2, z'

_cell_length_a                    5.730
_cell_length_b                    6.120
_cell_length_c                    5.630
_cell_angle_alpha                 90.00
_cell_angle_beta                  93.75
_cell_angle_gamma                 90.00
_cell_volume                      197.0
_cell_formula_units_Z             4
_cell_measurement_temperature     293(2)

_exptl_crystal_density_diffrn     6.559
```

```
_exptl_crystal_density_method     'not measured'
_exptl_crystal_F_000              80
_exptl_absorpt_coefficient_mu     0.010

_diffrn_ambient_temperature       293(2)
_diffrn_radiation_wavelength      ?
_diffrn_radiation_type            Electron
_diffrn_radiation_source          Microscope
_diffrn_radiation_monochromator   graphite
_diffrn_reflns_number             119
_diffrn_reflns_av_R_equivalents   0.0000
_diffrn_reflns_av_sigmaI/netI     0.0521
_diffrn_reflns_limit_h_min        0
_diffrn_reflns_limit_h_max        4
_diffrn_reflns_limit_k_min        0
_diffrn_reflns_limit_k_max        8
_diffrn_reflns_limit_l_min        -6
_diffrn_reflns_limit_l_max        7
_diffrn_reflns_theta_min          6.27
_diffrn_reflns_theta_max          54.33
_reflns_number_total              119
_reflns_number_gt                 117
_reflns_threshold_expression      >2sigma(I)

_computing_structure_solution     'SHELXS-97 (Sheldrick, 1990)'
_computing_structure_refinement   'SHELXL-97 (Sheldrick, 1997)'
_computing_molecular_graphics     ?
_computing_publication_material   maXus

_refine_special_details
;
 Refinement of F^2^ against ALL reflections.  The weighted R-factor wR and
 goodness of fit S are based on F^2^, conventional R-factors R are based
 on F, with F set to zero for negative F^2^. The threshold expression of
 F^2^ > 2sigma(F^2^) is used only for calculating R-factors(gt) etc. and is
 not relevant to the choice of reflections for refinement.  R-factors based
 on F^2^ are statistically about twice as large as those based on F, and R-
 factors based on ALL data will be even larger.
;

_refine_ls_structure_factor_coef  Fsqd
_refine_ls_matrix_type            full
_refine_ls_weighting_scheme
 'calc w=1/[\s^2^(Fo^2^)+(0.1000P)^2^+0.0000P] where P=(Fo^2^+2Fc^2^)/3'
_atom_sites_solution_primary      direct
_atom_sites_solution_secondary    difmap
_atom_sites_solution_hydrogens    geom
_refine_ls_hydrogen_treatment     mixed
_refine_ls_extinction_method      none
_refine_ls_extinction_coef        ?
_refine_ls_number_reflns          119
_refine_ls_number_parameters      17
_refine_ls_number_restraints      0
_refine_ls_R_factor_all           0.2567
_refine_ls_R_factor_gt            0.2569
_refine_ls_wR_factor_ref          0.5633
_refine_ls_wR_factor_gt           0.5634
```

```
_refine_ls_goodness_of_fit_ref     5.027
_refine_ls_restrained_S_all        5.027
_refine_ls_shift/su_max            0.017
_refine_ls_shift/su_mean           0.002

loop_
 _atom_site_label
 _atom_site_type_symbol
 _atom_site_fract_x
 _atom_site_fract_y
 _atom_site_fract_z
 _atom_site_U_iso_or_equiv
 _atom_site_adp_type
 _atom_site_occupancy
 _atom_site_calc_flag
 _atom_site_refinement_flags
 _atom_site_disorder_assembly
 _atom_site_disorder_group
Cu1 Cu 0.0000 0.0000 1.0000 0.008(4) Uiso 1 d S . .
Cu2 Cu 0.015(5) 0.2500 0.488(3) 0.013(4) Uiso 1 d S . .
Cl3 Cl 0.405(7) 0.2500 0.442(6) 0.052(11) Uiso 1 d S . .
O4 O -0.223(17) 0.2500 1.178(15) 0.07(2) Uiso 1 d S . .
O5 O -0.147(6) -0.008(6) 0.348(4) 0.011(7) Uiso 1 d . . .

_geom_special_details
;
 All esds (except the esd in the dihedral angle between two l.s. planes)
 are estimated using the full covariance matrix.  The cell esds are taken
 into account individually in the estimation of esds in distances, angles
 and torsion angles; correlations between esds in cell parameters are only
 used when they are defined by crystal symmetry.  An approximate (isotropic)
 treatment of cell esds is used for estimating esds involving l.s. planes.
;

loop_
 _geom_bond_atom_site_label_1
 _geom_bond_atom_site_label_2
 _geom_bond_distance
 _geom_bond_site_symmetry_2
 _geom_bond_publ_flag
Cu1 O4 2.27(7) 3_557 ?
Cu1 O4 2.27(7) . ?
Cu1 O5 2.18(3) 3_556 ?
Cu1 O5 2.18(3) 1_556 ?
Cu1 Cu1 3.0600 2_557 ?
Cu1 Cu1 3.0600 2_547 ?
Cu2 O5 1.97(4) 4_565 ?
Cu2 O5 1.97(4) . ?
Cu2 O5 1.88(4) 2_556 ?
Cu2 O5 1.88(4) 3_556 ?
Cu2 O4 4.20(9) . ?
Cu2 Cl3 2.26(5) . ?
Cu2 Cu2 3.068(3) 3_556 ?
Cu2 Cu2 3.068(3) 3_566 ?
O4 Cu1 2.27(7) 2_557 ?
O5 Cu2 1.88(4) 3_556 ?
O5 Cu1 2.18(3) 1_554 ?
```

```
loop_
 _geom_angle_atom_site_label_1
 _geom_angle_atom_site_label_2
 _geom_angle_atom_site_label_3
 _geom_angle
 _geom_angle_site_symmetry_1
 _geom_angle_site_symmetry_3
 _geom_angle_publ_flag
O4 Cu1 O4 180.000(17) 3_557 . ?
O4 Cu1 O5 50.0(19) 3_557 3_556 ?
O4 Cu1 O5 130.0(19) . 3_556 ?
O4 Cu1 O5 130.0(19) 3_557 1_556 ?
O4 Cu1 O5 50.0(19) . 1_556 ?
O5 Cu1 O5 180.000(7) 3_556 1_556 ?
O4 Cu1 Cu1 132.3(16) 3_557 2_557 ?
O4 Cu1 Cu1 47.7(16) . 2_557 ?
O5 Cu1 Cu1 88.7(9) 3_556 2_557 ?
O5 Cu1 Cu1 91.3(9) 1_556 2_557 ?
O4 Cu1 Cu1 47.7(16) 3_557 2_547 ?
O4 Cu1 Cu1 132.3(16) . 2_547 ?
O5 Cu1 Cu1 91.3(9) 3_556 2_547 ?
O5 Cu1 Cu1 88.7(9) 1_556 2_547 ?
Cu1 Cu1 Cu1 180.0 2_557 2_547 ?
O5 Cu2 O5 107(2) 4_565 . ?
O5 Cu2 O5 74.1(14) 4_565 2_556 ?
O5 Cu2 O5 173.2(13) . 2_556 ?
O5 Cu2 O5 173.2(13) 4_565 3_556 ?
O5 Cu2 O5 74.1(14) . 3_556 ?
O5 Cu2 O5 104(2) 2_556 3_556 ?
O5 Cu2 O4 101.3(13) 4_565 . ?
O5 Cu2 O4 101.3(13) . . ?
O5 Cu2 O4 71.9(12) 2_556 . ?
O5 Cu2 O4 71.9(12) 3_556 . ?
O5 Cu2 Cl3 113.3(13) 4_565 . ?
O5 Cu2 Cl3 113.3(13) . . ?
O5 Cu2 Cl3 71.9(14) 2_556 . ?
O5 Cu2 Cl3 71.9(14) 3_556 . ?
O4 Cu2 Cl3 119.3(17) . . ?
O5 Cu2 Cu2 142.2(18) 4_565 3_556 ?
O5 Cu2 Cu2 36.0(10) . 3_556 ?
O5 Cu2 Cu2 142.0(17) 2_556 3_556 ?
O5 Cu2 Cu2 38.1(10) 3_556 3_556 ?
O4 Cu2 Cu2 86.4(7) . 3_556 ?
Cl3 Cu2 Cu2 93.7(10) . 3_556 ?
O5 Cu2 Cu2 36.0(10) 4_565 3_566 ?
O5 Cu2 Cu2 142.2(18) . 3_566 ?
O5 Cu2 Cu2 38.1(10) 2_556 3_566 ?
O5 Cu2 Cu2 142.0(17) 3_556 3_566 ?
O4 Cu2 Cu2 86.4(7) . 3_566 ?
Cl3 Cu2 Cu2 93.7(10) . 3_566 ?
Cu2 Cu2 Cu2 171.5(17) 3_556 3_566 ?
Cu1 O4 Cu1 85(3) 2_557 . ?
Cu1 O4 Cu2 50.5(17) 2_557 . ?
Cu1 O4 Cu2 50.5(17) . . ?
Cu2 O5 Cu2 105.9(14) . 3_556 ?
Cu2 O5 Cu1 98.1(15) . 1_554 ?
```

```
Cu2 O5 Cu1 107.1(17) 3_556 1_554 ?

_diffrn_measured_fraction_theta_max    0.228
_diffrn_reflns_theta_full              54.33
_diffrn_measured_fraction_theta_full   0.228
_refine_diff_density_max    0.943
_refine_diff_density_min   -0.764
_refine_diff_density_rms    0.176
```

8.5 2,6-bis[4(dimethylamino)benzylidene]cyclohexanone

```
data_shelxl

_audit_creation_method            SHELXL-97
_chemical_name_systematic
2,6-bis(4-dimethylamino-benzlidene)-cyclohexanone
;
 ?
;
_chemical_name_common             DMABC
_chemical_formula_moiety          ?
_chemical_formula_sum
 'C13 N O'
_chemical_formula_weight          186.14

_citation_journal_full Ultramicroscopy
_citation_journal_volume
_citation_langauge English
_citation_page_first
_citation_page_last
_citation_year
_citation_author_name 'Voigt-Martin, I.G., Kothe, H., Yakimansky, A.V.,
Tenkovtsev, A.V., Zandbergen, H., Jansen, J., Gilmore, C.'

loop_
 _atom_type_symbol
 _atom_type_description
 _atom_type_scat_dispersion_real
 _atom_type_scat_dispersion_imag
 _atom_type_scat_source
 'O'  'O'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
 'C'  'C'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
 'N'  'N'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'

_symmetry_cell_setting            ?
_symmetry_space_group_name_H-M    'Cmc21

loop_
 _symmetry_equiv_pos_as_xyz
 'x, y, z'
 '-x, -y, z+1/2'
 '-x, y, z'
 'x, -y, z+1/2'
 'x+1/2, y+1/2, z'
 '-x+1/2, -y+1/2, z+1/2'
 '-x+1/2, y+1/2, z'
 'x+1/2, -y+1/2, z+1/2'

_cell_length_a                    21.857
_cell_length_b                    9.315
_cell_length_c                    9.644
_cell_angle_alpha                 90.00
_cell_angle_beta                  90.00
```

```
_cell_angle_gamma                 90.00
_cell_volume                      1963.5
_cell_formula_units_Z             8
_cell_measurement_temperature     293(2)

_exptl_crystal_density_diffrn     1.259
_exptl_crystal_density_method     'not measured'
_exptl_crystal_F_000              294
_exptl_absorpt_coefficient_mu     0.006

_diffrn_ambient_temperature       293(2)
_diffrn_radiation_wavelength      ?
_diffrn_radiation_type            Electron
_diffrn_radiation_source          Microscope
_diffrn_reflns_number             133
_diffrn_reflns_av_R_equivalents   0.0000
_diffrn_reflns_av_sigmaI/netI     0.0141
_diffrn_reflns_limit_h_min        0
_diffrn_reflns_limit_h_max        20
_diffrn_reflns_limit_k_min        0
_diffrn_reflns_limit_k_max        4
_diffrn_reflns_limit_l_min        0
_diffrn_reflns_limit_l_max        10
_diffrn_reflns_theta_min          3.22
_diffrn_reflns_theta_max          45.82
_reflns_number_total              133
_reflns_number_gt                 133
_reflns_threshold_expression      >2sigma(I)

_computing_structure_solution     'SHELXS-97 (Sheldrick, 1990)'
_computing_structure_refinement   'SHELXL-97 (Sheldrick, 1997)'
_computing_molecular_graphics     ?
_computing_publication_material   maXus

_refine_special_details
;
 Refinement of F^2^ against ALL reflections.  The weighted R-factor wR and
 goodness of fit S are based on F^2^, conventional R-factors R are based
 on F, with F set to zero for negative F^2^. The threshold expression of
 F^2^ > 2sigma(F^2^) is used only for calculating R-factors(gt) etc. and is
 not relevant to the choice of reflections for refinement.  R-factors based
 on F^2^ are statistically about twice as large as those based on F, and R-
 factors based on ALL data will be even larger.
;

_refine_ls_structure_factor_coef  Fsqd
_refine_ls_matrix_type            full
_refine_ls_weighting_scheme
 'calc w=1/[\s^2^(Fo^2^)+(0.1000P)^2^+0.0000P] where P=(Fo^2^+2Fc^2^)/3'
_atom_sites_solution_primary      direct
_atom_sites_solution_secondary    difmap
_atom_sites_solution_hydrogens    geom
_refine_ls_hydrogen_treatment     mixed
_refine_ls_extinction_method      none
_refine_ls_extinction_coef        ?
_refine_ls_abs_structure_details
 'Flack H D (1983), Acta Cryst. A39, 876-881'
```

```
_refine_ls_abs_structure_Flack     0(10)
_refine_ls_number_reflns           133
_refine_ls_number_parameters       42
_refine_ls_number_restraints       18
_refine_ls_R_factor_all            0.2618
_refine_ls_R_factor_gt             0.2618
_refine_ls_wR_factor_ref           0.5319
_refine_ls_wR_factor_gt            0.5319
_refine_ls_goodness_of_fit_ref     7.095
_refine_ls_restrained_S_all        6.562
_refine_ls_shift/su_max            0.007
_refine_ls_shift/su_mean           0.002

loop_
 _atom_site_label
 _atom_site_type_symbol
 _atom_site_fract_x
 _atom_site_fract_y
 _atom_site_fract_z
 _atom_site_U_iso_or_equiv
 _atom_site_adp_type
 _atom_site_occupancy
 _atom_site_calc_flag
 _atom_site_refinement_flags
 _atom_site_disorder_assembly
 _atom_site_disorder_group
O1 O 0.5000 0.279(7) 0.1339 0.050 Uiso 1 d SD . .
C2 C 0.5000 0.370(7) 0.251(3) 0.050 Uiso 1 d SD . .
C3 C 0.4400(7) 0.388(3) 0.3336(16) 0.050 Uiso 1 d D . .
C4 C 0.4499(10) 0.424(7) 0.4889(18) 0.050 Uiso 1 d D . .
C5 C 0.5000 0.426(13) 0.602(3) 0.050 Uiso 1 d SD . .
C6 C 0.3947(7) 0.294(3) 0.252(3) 0.050 Uiso 1 d D . .
C7 C 0.3362(5) 0.3717(3) 0.2995(5) 0.050 Uiso 1 d D . .
C8 C 0.2872(6) 0.2873(4) 0.2202(7) 0.050 Uiso 1 d D . .
C9 C 0.2223(7) 0.3057(7) 0.2748(10) 0.050 Uiso 1 d D . .
C10 C 0.2003(9) 0.3662(9) 0.3597(13) 0.050 Uiso 1 d D . .
C11 C 0.2382(7) 0.4618(10) 0.4581(14) 0.050 Uiso 1 d D . .
C12 C 0.3055(7) 0.4556(11) 0.4174(12) 0.050 Uiso 1 d D . .
N13 N 0.1360(12) 0.344(4) 0.404(3) 0.050 Uiso 1 d D . .
C14 C 0.0982(16) 0.233(6) 0.329(4) 0.050 Uiso 1 d D . .
C15 C 0.1218(16) 0.463(6) 0.506(4) 0.050 Uiso 1 d D . .

_geom_special_details
;
 All esds (except the esd in the dihedral angle between two l.s. planes)
 are estimated using the full covariance matrix.  The cell esds are taken
 into account individually in the estimation of esds in distances, angles
 and torsion angles; correlations between esds in cell parameters are only
 used when they are defined by crystal symmetry.  An approximate (isotropic)
 treatment of cell esds is used for estimating esds involving l.s. planes.
;

loop_
 _geom_bond_atom_site_label_1
 _geom_bond_atom_site_label_2
 _geom_bond_distance
 _geom_bond_site_symmetry_2
```

```
 _geom_bond_publ_flag
O1 C2 1.41(4) . ?
C2 C3 1.541(13) 3_655 ?
C2 C3 1.541(13) . ?
C3 C6 1.537(19) . ?
C3 C4 1.549(18) . ?
C4 C5 1.546(18) . ?
C5 C4 1.546(18) 3_655 ?
C6 C7 1.537(17) . ?
C7 C8 1.533(11) . ?
C7 C12 1.535(13) . ?
C8 C9 1.523(16) . ?
C9 C10 1.104(19) . ?
C10 N13 1.48(3) . ?
C10 C11 1.542(15) . ?
C11 C12 1.523(17) . ?
N13 C15 1.51(4) . ?
N13 C14 1.51(4) . ?

loop_
 _geom_angle_atom_site_label_1
 _geom_angle_atom_site_label_2
 _geom_angle_atom_site_label_3
 _geom_angle
 _geom_angle_site_symmetry_1
 _geom_angle_site_symmetry_3
 _geom_angle_publ_flag
O1 C2 C3 118.5(19) . 3_655 ?
O1 C2 C3 118.5(19) . . ?
C3 C2 C3 117(2) 3_655 . ?
C6 C3 C2 102.9(17) . . ?
C6 C3 C4 135(2) . . ?
C2 C3 C4 113.7(16) . . ?
C5 C4 C3 142(3) . . ?
C4 C5 C4 90(2) 3_655 . ?
C3 C6 C7 96.7(14) . . ?
C8 C7 C12 108.9(9) . . ?
C8 C7 C6 101.1(9) . . ?
C12 C7 C6 145.4(13) . . ?
C9 C8 C7 114.9(7) . . ?
C10 C9 C8 136.0(14) . . ?
C9 C10 N13 123.8(19) . . ?
C9 C10 C11 121.1(18) . . ?
N13 C10 C11 114.3(17) . . ?
C12 C11 C10 109.8(13) . . ?
C11 C12 C7 129.2(12) . . ?
C10 N13 C15 106(2) . . ?
C10 N13 C14 118(3) . . ?
C15 N13 C14 135(3) . . ?

_diffrn_measured_fraction_theta_max    0.073
_diffrn_reflns_theta_full              45.82
_diffrn_measured_fraction_theta_full   0.073
_refine_diff_density_max    0.218
_refine_diff_density_min   -0.178
_refine_diff_density_rms    0.062
```

8.6 4-dimethylamino-3-cyanobiphenyl

```
data_shelxl

_audit_creation_method            SHELXL-97
_chemical_name_systematic         4-dimethylamino-3-cyanobiphenyl
;
_chemical_name_common             DMACB
_chemical_formula_moiety          ?
_chemical_formula_sum
 'C15 N2'
_chemical_formula_weight          208.17

_citation_journal_full Ultramicroscopy
_citation_journal_volume 68
_citation_langauge English
_citation_page_first 43
_citation_page_last 59
_citation_year 1997
_citation_author_name ' Voigt-Martin, I.G., Zhang, Z.X., Kolb, U., Gilmore,
C.'

loop_
 _atom_type_symbol
 _atom_type_description
 _atom_type_scat_dispersion_real
 _atom_type_scat_dispersion_imag
 _atom_type_scat_source
 'C'  'C'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
 'N'  'N'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'

_symmetry_cell_setting            ?
_symmetry_space_group_name_H-M    'Pna 21'

loop_
 _symmetry_equiv_pos_as_xyz
 'x, y, z'
 '-x, -y, z+1/2'
 '-x+1/2, y+1/2, z+1/2'
 'x+1/2, -y+1/2, z'

_cell_length_a                    10.280
_cell_length_b                    22.640
_cell_length_c                    5.270
_cell_angle_alpha                 90.00
_cell_angle_beta                  90.00
_cell_angle_gamma                 90.00
_cell_volume                      1226.5
_cell_formula_units_Z             4
_cell_measurement_temperature     293(2)
_
_exptl_crystal_density_diffrn     1.127
_exptl_crystal_density_method     'not measured'
_exptl_crystal_F_000              168
_exptl_absorpt_coefficient_mu     0.006
```

```
_diffrn_ambient_temperature       293(2)
_diffrn_radiation_wavelength      ?
_diffrn_radiation_type            Electron
_diffrn_radiation_source          Microscope
_diffrn_reflns_number             118
_diffrn_reflns_av_R_equivalents   0.0000
_diffrn_reflns_av_sigmaI/netI     0.1000
_diffrn_reflns_limit_h_min        0
_diffrn_reflns_limit_h_max        7
_diffrn_reflns_limit_k_min        0
_diffrn_reflns_limit_k_max        8
_diffrn_reflns_limit_l_min        0
_diffrn_reflns_limit_l_max        3
_diffrn_reflns_theta_min          3.76
_diffrn_reflns_theta_max          29.37
_reflns_number_total              118
_reflns_number_gt                 118
_reflns_threshold_expression      >2sigma(I)

_computing_structure_solution     'SHELXS-97 (Sheldrick, 1990)'
_computing_structure_refinement   'SHELXL-97 (Sheldrick, 1997)'
_computing_molecular_graphics     ?
_computing_publication_material   maXus

_refine_special_details
;
 Refinement of F^2^ against ALL reflections.  The weighted R-factor wR and
 goodness of fit S are based on F^2^, conventional R-factors R are based
 on F, with F set to zero for negative F^2^. The threshold expression of
 F^2^ > 2sigma(F^2^) is used only for calculating R-factors(gt) etc. and is
 not relevant to the choice of reflections for refinement.  R-factors based
 on F^2^ are statistically about twice as large as those based on F, and R-
 factors based on ALL data will be even larger.
;

_refine_ls_structure_factor_coef  Fsqd
_refine_ls_matrix_type            full
_refine_ls_weighting_scheme
 'calc w=1/[\s^2^(Fo^2^)+(0.1000P)^2^+0.0000P] where P=(Fo^2^+2Fc^2^)/3'
_atom_sites_solution_primary      direct
_atom_sites_solution_secondary    difmap
_atom_sites_solution_hydrogens    geom
_refine_ls_hydrogen_treatment     mixed
_refine_ls_extinction_method      none
_refine_ls_extinction_coef        ?
_refine_ls_abs_structure_details
 'Flack H D (1983), Acta Cryst. A39, 876-881'
_refine_ls_abs_structure_Flack    0(10)
_refine_ls_number_reflns          118
_refine_ls_number_parameters      53
_refine_ls_number_restraints      24
_refine_ls_R_factor_all           0.2624
_refine_ls_R_factor_gt            0.2624
_refine_ls_wR_factor_ref          0.5491
_refine_ls_wR_factor_gt           0.5491
```

```
_refine_ls_goodness_of_fit_ref    5.624
_refine_ls_restrained_S_all       4.882
_refine_ls_shift/su_max           0.039
_refine_ls_shift/su_mean          0.010

loop_
 _atom_site_label
 _atom_site_type_symbol
 _atom_site_fract_x
 _atom_site_fract_y
 _atom_site_fract_z
 _atom_site_U_iso_or_equiv
 _atom_site_adp_type
 _atom_site_occupancy
 _atom_site_calc_flag
 _atom_site_refinement_flags
 _atom_site_disorder_assembly
 _atom_site_disorder_group
C1 C 0.7522(3) 0.9385(4) 0.7698(16) 0.050 Uiso 1 d D . .
C2 C 0.6798(5) 0.8861(5) 0.810(2) 0.050 Uiso 1 d D . .
C3 C 0.5906(9) 0.8688(9) 1.004(3) 0.050 Uiso 1 d D . .
C4 C 0.5625(14) 0.9082(11) 1.202(4) 0.050 Uiso 1 d D . .
C5 C 0.6300(14) 0.9610(14) 1.177(5) 0.050 Uiso 1 d D . .
C6 C 0.7171(13) 0.9752(11) 0.979(4) 0.050 Uiso 1 d D . .
N7 N 0.524(2) 0.8106(15) 1.0005 0.050 Uiso 1 d D . .
C8 C 0.392(2) 0.826(4) 1.091(9) 0.050 Uiso 1 d D . .
C9 C 0.607(4) 0.757(2) 0.959(6) 0.050 Uiso 1 d D . .
C10 C 0.690(3) 0.848(3) 0.572(7) 0.050 Uiso 1 d D . .
N11 N 0.693(3) 0.823(3) 0.310(6) 0.050 Uiso 1 d D . .
C12 C 0.7892(5) 1.0348(4) 0.9446(14) 0.050 Uiso 1 d D . .
C13 C 0.9040(7) 1.0565(5) 0.828(2) 0.050 Uiso 1 d D . .
C14 C 1.0010(17) 1.0978(10) 0.795(4) 0.050 Uiso 1 d D . .
C15 C 0.981(2) 1.1427(14) 0.955(4) 0.050 Uiso 1 d D . .
C16 C 0.870(2) 1.1251(15) 1.080(6) 0.050 Uiso 1 d D . .
C17 C 0.773(2) 1.0829(13) 1.110(5) 0.050 Uiso 1 d D . .

_geom_special_details
;
 All esds (except the esd in the dihedral angle between two l.s. planes)
 are estimated using the full covariance matrix.  The cell esds are taken
 into account individually in the estimation of esds in distances, angles
 and torsion angles; correlations between esds in cell parameters are only
 used when they are defined by crystal symmetry.  An approximate (isotropic)
 treatment of cell esds is used for estimating esds involving l.s. planes.
;

loop_
 _geom_bond_atom_site_label_1
 _geom_bond_atom_site_label_2
 _geom_bond_distance
 _geom_bond_site_symmetry_2
 _geom_bond_publ_flag
C1 C2 1.416(12) . ?
C1 C6 1.43(3) . ?
C2 C3 1.428(16) . ?
C2 C10 1.53(4) . ?
C3 C4 1.40(3) . ?
```

```
C3 N7 1.49(3) . ?
C4 C5 1.39(3) . ?
C5 C6 1.41(3) . ?
C6 C12 1.55(3) . ?
N7 C8 1.48(3) . ?
N7 C9 1.49(4) . ?
C10 N11 1.49(4) . ?
C12 C17 1.41(3) . ?
C12 C13 1.419(10) . ?
C13 C14 1.38(2) . ?
C14 C15 1.34(3) . ?
C15 C16 1.38(3) . ?
C16 C17 1.39(4) . ?

loop_
 _geom_angle_atom_site_label_1
 _geom_angle_atom_site_label_2
 _geom_angle_atom_site_label_3
 _geom_angle
 _geom_angle_site_symmetry_1
 _geom_angle_site_symmetry_3
 _geom_angle_publ_flag
C2 C1 C6 103.8(11) . . ?
C3 C2 C1 132.5(13) . . ?
C3 C2 C10 118(2) . . ?
C1 C2 C10 108(2) . . ?
C2 C3 C4 119.4(16) . . ?
C2 C3 N7 122.0(16) . . ?
C4 C3 N7 118.6(14) . . ?
C5 C4 C3 111.8(19) . . ?
C4 C5 C6 126(3) . . ?
C5 C6 C1 127(2) . . ?
C5 C6 C12 126(2) . . ?
C1 C6 C12 107.2(13) . . ?
C3 N7 C8 102(4) . . ?
C3 N7 C9 117(2) . . ?
C8 N7 C9 140(4) . . ?
N11 C10 C2 167(5) . . ?
C17 C12 C13 95.7(13) . . ?
C17 C12 C6 123.1(15) . . ?
C13 C12 C6 138.6(10) . . ?
C14 C13 C12 153.0(14) . . ?
C15 C14 C13 109.3(19) . . ?
C14 C15 C16 102(3) . . ?
C17 C16 C15 148(3) . . ?
C16 C17 C12 112(2) . . ?

_diffrn_measured_fraction_theta_max     0.306
_diffrn_reflns_theta_full               29.37
_diffrn_measured_fraction_theta_full    0.306
_refine_diff_density_max    0.151
_refine_diff_density_min   -0.171
_refine_diff_density_rms    0.044
```

8.7 Mannan I

```
data_shelxl

_audit_creation_method            SHELXL-97

_chemical_name_common             'Mannan 1'
_chemical_formula_moiety          ?
_chemical_formula_sum             'C6 O5'
_chemical_formula_weight          152.06

_citation_journal_full Macromolecules
_citation_journal_volume 20
_citation_langauge English
_citation_page_first 2407
_citation_page_last 2413
_citation_year 1987
_citation_author_name 'Chanzy, H., Perez, S., Miller, D.P., Paradossi, G.,
Winter, W.T.'

loop_
 _atom_type_symbol
 _atom_type_description
 _atom_type_scat_dispersion_real
 _atom_type_scat_dispersion_imag
 _atom_type_scat_source
 'C'  'C'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
 'O'  'O'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'

_symmetry_cell_setting            ?
_symmetry_space_group_name_H-M    'P 21 21 21'

loop_
 _symmetry_equiv_pos_as_xyz
 'x, y, z'
 '-x+1/2, -y, z+1/2'
 'x+1/2, -y+1/2, -z'
 '-x, y+1/2, -z+1/2'

_cell_length_a                    8.920
_cell_length_b                    7.210
_cell_length_c                    10.270
_cell_angle_alpha                 90.00
_cell_angle_beta                  90.00
_cell_angle_gamma                 90.00
_cell_volume                      660.5
_cell_formula_units_Z             4
_cell_measurement_temperature     293(2)

_exptl_crystal_density_diffrn     1.529
_exptl_crystal_density_method     'not measured'
_exptl_crystal_F_000              100
_exptl_absorpt_coefficient_mu     0.007
```

```
_diffrn_ambient_temperature       293(2)
_diffrn_radiation_wavelength      ?
_diffrn_radiation_type            electron
_diffrn_radiation_source          microscope
_diffrn_reflns_number             58
_diffrn_reflns_av_R_equivalents   0.0000
_diffrn_reflns_av_sigmaI/netI     0.0146
_diffrn_reflns_limit_h_min        0
_diffrn_reflns_limit_h_max        6
_diffrn_reflns_limit_k_min        0
_diffrn_reflns_limit_k_max        5
_diffrn_reflns_limit_l_min        0
_diffrn_reflns_limit_l_max        4
_diffrn_reflns_theta_min          5.23
_diffrn_reflns_theta_max          33.17
_reflns_number_total              58
_reflns_number_gt                 58
_reflns_threshold_expression      >2sigma(I)

_computing_structure_solution     'SHELXS-97 (Sheldrick, 1990)'
_computing_structure_refinement   'SHELXL-97 (Sheldrick, 1997)'

_refine_special_details
;
 Refinement of F^2^ against ALL reflections.  The weighted R-factor wR and
 goodness of fit S are based on F^2^, conventional R-factors R are based
 on F, with F set to zero for negative F^2^. The threshold expression of
 F^2^ > 2sigma(F^2^) is used only for calculating R-factors(gt) etc. and is
 not relevant to the choice of reflections for refinement.  R-factors based
 on F^2^ are statistically about twice as large as those based on F, and R-
 factors based on ALL data will be even larger.
;

_refine_ls_structure_factor_coef  Fsqd
_refine_ls_matrix_type            full
_refine_ls_weighting_scheme
 'calc w=1/[\s^2^(Fo^2^)+(0.1000P)^2^+0.0000P] where P=(Fo^2^+2Fc^2^)/3'
_atom_sites_solution_primary      direct
_atom_sites_solution_secondary    difmap
_atom_sites_solution_hydrogens    geom
_refine_ls_hydrogen_treatment     mixed
_refine_ls_extinction_method      none
_refine_ls_extinction_coef        ?
_refine_ls_abs_structure_details
 'Flack H D (1983), Acta Cryst. A39, 876-881'
_refine_ls_abs_structure_Flack    0(10)
_refine_ls_number_reflns          58
_refine_ls_number_parameters      34
_refine_ls_number_restraints      6
_refine_ls_R_factor_all           0.1774
_refine_ls_R_factor_gt            0.1774
_refine_ls_wR_factor_ref          0.5489
_refine_ls_wR_factor_gt           0.5489
_refine_ls_goodness_of_fit_ref    8.463
_refine_ls_restrained_S_all       7.613
_refine_ls_shift/su_max           0.257
_refine_ls_shift/su_mean          0.083
```

```
loop_
 _atom_site_label
 _atom_site_type_symbol
 _atom_site_fract_x
 _atom_site_fract_y
 _atom_site_fract_z
 _atom_site_U_iso_or_equiv
 _atom_site_adp_type
 _atom_site_occupancy
 _atom_site_calc_flag
 _atom_site_refinement_flags
 _atom_site_disorder_assembly
 _atom_site_disorder_group
O7 O -0.029(4) -0.119(6) 0.686(9) 0.050 Uiso 1 d D . .
O8 O 0.103(5) -0.225(6) 0.432(7) 0.050 Uiso 1 d D . .
O9 O 0.204(5) -0.098(8) 0.898(6) 0.050 Uiso 1 d D . .
O10 O 0.259(5) 0.135(6) 0.717(5) 0.050 Uiso 1 d D . .
O11 O 0.324(5) 0.439(8) 0.694(5) 0.050 Uiso 1 d D . .
C1 C 0.235(6) -0.064(6) 0.705(14) 0.050 Uiso 1 d D . .
C2 C 0.126(5) -0.160(8) 0.661(10) 0.050 Uiso 1 d D . .
C3 C 0.204(6) -0.141(8) 0.524(8) 0.050 Uiso 1 d D . .
C4 C 0.279(6) -0.069(8) 1.018(7) 0.050 Uiso 1 d D . .
C5 C 0.323(7) 0.128(10) 0.588(6) 0.050 Uiso 1 d D . .
C6 C 0.354(7) 0.395(8) 0.559(6) 0.050 Uiso 1 d D . .

_geom_special_details
;
 All esds (except the esd in the dihedral angle between two l.s. planes)
 are estimated using the full covariance matrix.  The cell esds are taken
 into account individually in the estimation of esds in distances, angles
 and torsion angles; correlations between esds in cell parameters are only
 used when they are defined by crystal symmetry.  An approximate (isotropic)
 treatment of cell esds is used for estimating esds involving l.s. planes.
;

loop_
 _geom_bond_atom_site_label_1
 _geom_bond_atom_site_label_2
 _geom_bond_distance
 _geom_bond_site_symmetry_2
 _geom_bond_publ_flag
O7 C2 1.44(5) . ?
O8 C3 1.44(6) . ?
O9 C1 2.03(16) . ?
O9 C4 1.41(6) . ?
O10 C1 1.45(5) . ?
O10 C5 1.45(6) . ?
O11 C6 1.45(6) . ?
C1 C2 1.27(9) . ?
C2 C3 1.58(11) . ?
C3 C4 1.52(8) 2_554 ?
C4 C5 1.24(8) 2 ?
C4 C3 1.52(8) 2 ?
C5 C4 1.24(8) 2_554 ?
C5 C6 1.96(9) . ?
```

```
loop_
 _geom_angle_atom_site_label_1
 _geom_angle_atom_site_label_2
 _geom_angle_atom_site_label_3
 _geom_angle
 _geom_angle_site_symmetry_1
 _geom_angle_site_symmetry_3
 _geom_angle_publ_flag
C1 O9 C4 141(4) . . ?
C1 O10 C5 87(6) . . ?
O9 C1 O10 93(7) . . ?
O9 C1 C2 100(8) . . ?
O10 C1 C2 132(6) . . ?
O7 C2 C3 124(6) . . ?
O7 C2 C1 124(7) . . ?
C3 C2 C1 86(7) . . ?
O8 C3 C4 117(5) . 2_554 ?
O8 C3 C2 106(5) . . ?
C4 C3 C2 99(5) 2_554 . ?
C5 C4 C3 113(6) 2 2 ?
C5 C4 O9 96(5) 2 . ?
C3 C4 O9 103(6) 2 . ?
C4 C5 O10 104(5) 2_554 . ?
C4 C5 C6 111(5) 2_554 . ?
O10 C5 C6 99(5) . . ?
O11 C6 C5 93(5) . . ?

_diffrn_measured_fraction_theta_max    0.186
_diffrn_reflns_theta_full              33.17
_diffrn_measured_fraction_theta_full   0.186
_refine_diff_density_max    0.125
_refine_diff_density_min   -0.129
_refine_diff_density_rms    0.039
```

8.8 Copper Perchlorophthalocyanine

```
data_shelxl

_audit_creation_method            SHELXL-97
_chemical_name_common             'Copper Perchlorophthalocyanine'
_chemical_formula_moiety          ?
_chemical_formula_sum             'C8 Cl4 Cu N3'
_chemical_formula_weight          343.46

_citation_journal_full Ultramicroscopy
_citation_journal_volume 38
_citation_langauge English
_citation_page_first 41
_citation_page_last 45
_citation_year 1991
_citation_author_name 'Dorset, D.L., Tivol, W.F., Turner, J.N.'

loop_
 _atom_type_symbol
 _atom_type_description
 _atom_type_scat_dispersion_real
 _atom_type_scat_dispersion_imag
 _atom_type_scat_source
 'Cu'  'Cu'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
 'Cl'  'Cl'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
 'N'  'N'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
 'C'  'C'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'

_symmetry_cell_setting            ?
_symmetry_space_group_name_H-M    C2/c

loop_
 _symmetry_equiv_pos_as_xyz
 'x, y, z'
 '-x, y, -z'
 'x+1/2, y+1/2, z'
 '-x+1/2, y+1/2, -z'
 '-x, -y, -z'
 'x, -y, z'
 '-x+1/2, -y+1/2, -z'
 'x+1/2, -y+1/2, z'

_cell_length_a                    19.620
_cell_length_b                    26.080
_cell_length_c                    3.760
_cell_angle_alpha                 90.00
_cell_angle_beta                  116.50
_cell_angle_gamma                 90.00
_cell_volume                      1721.8
_cell_formula_units_Z             2
_cell_measurement_temperature     293(2)
```

```
_exptl_crystal_density_diffrn     0.662
_exptl_crystal_density_method     'not measured'
_exptl_crystal_F_000              103
_exptl_absorpt_coefficient_mu     0.002


_diffrn_ambient_temperature       293(2)
_diffrn_radiation_wavelength      ?
_diffrn_radiation_type            Electron
_diffrn_radiation_source          Microscope
_diffrn_reflns_number             197
_diffrn_reflns_av_R_equivalents   0.0000
_diffrn_reflns_av_sigmaI/netI     0.1830
_diffrn_reflns_limit_h_min        0
_diffrn_reflns_limit_h_max        18
_diffrn_reflns_limit_k_min        0
_diffrn_reflns_limit_k_max        25
_diffrn_reflns_limit_l_min        0
_diffrn_reflns_limit_l_max        0
_diffrn_reflns_theta_min          2.41
_diffrn_reflns_theta_max          44.64
_reflns_number_total              197
_reflns_number_gt                 146
_reflns_threshold_expression      >2sigma(I)


_computing_structure_solution     'SHELXS-97 (Sheldrick, 1990)'
_computing_structure_refinement   'SHELXL-97 (Sheldrick, 1997)'
_computing_molecular_graphics     ?
_computing_publication_material   maXus

_refine_special_details
;
 Refinement of F^2^ against ALL reflections.  The weighted R-factor wR and
 goodness of fit S are based on F^2^, conventional R-factors R are based
 on F, with F set to zero for negative F^2^. The threshold expression of
 F^2^ > 2sigma(F^2^) is used only for calculating R-factors(gt) etc. and is
 not relevant to the choice of reflections for refinement.  R-factors based
 on F^2^ are statistically about twice as large as those based on F, and R-
 factors based on ALL data will be even larger.
;

_refine_ls_structure_factor_coef  Fsqd
_refine_ls_matrix_type            full
_refine_ls_weighting_scheme
 'calc w=1/[\s^2^(Fo^2^)+(0.1000P)^2^+0.0000P] where P=(Fo^2^+2Fc^2^)/3'
_atom_sites_solution_primary      direct
_atom_sites_solution_secondary    difmap
_atom_sites_solution_hydrogens    geom
_refine_ls_hydrogen_treatment     mixed
_refine_ls_extinction_method      none
_refine_ls_extinction_coef        ?
_refine_ls_number_reflns          197
_refine_ls_number_parameters      45
_refine_ls_number_restraints      0
_refine_ls_R_factor_all           0.2220
_refine_ls_R_factor_gt            0.2016
```

```
_refine_ls_wR_factor_ref          0.4167
_refine_ls_wR_factor_gt           0.4028
_refine_ls_goodness_of_fit_ref    1.761
_refine_ls_restrained_S_all       1.761
_refine_ls_shift/su_max           0.030
_refine_ls_shift/su_mean          0.009

loop_
 _atom_site_label
 _atom_site_type_symbol
 _atom_site_fract_x
 _atom_site_fract_y
 _atom_site_fract_z
 _atom_site_U_iso_or_equiv
 _atom_site_adp_type
 _atom_site_occupancy
 _atom_site_calc_flag
 _atom_site_refinement_flags
 _atom_site_disorder_assembly
 _atom_site_disorder_group
Cu1 Cu 0.0000 0.0000 0.0000 0.10(2) Uiso 1 d S . .
Cl2 Cl 0.0770(16) 0.3067(9) 0.0000 0.071(12) Uiso 1 d . . .
Cl3 Cl 0.1639(15) 0.2078(9) 0.0000 0.055(10) Uiso 1 d . . .
Cl4 Cl 0.2717(13) 0.1159(9) 0.0000 0.048(10) Uiso 1 d . . .
Cl5 Cl 0.4106(12) 0.0584(7) 0.0000 0.028(8) Uiso 1 d . . .
N6 N 0.0000 0.069(2) 0.0000 0.023(18) Uiso 1 d S . .
N7 N 0.094(3) 0.0000 0.0000 0.017(16) Uiso 1 d S . .
N8 N 0.120(2) 0.0924(15) 0.0000 0.021(12) Uiso 1 d . . .
C9 C 0.136(2) 0.0423(15) 0.0000 0.022(14) Uiso 1 d . . .
C10 C 0.213(3) 0.0283(17) 0.0000 0.05(2) Uiso 1 d . . .
C11 C 0.279(3) 0.0586(19) 0.0000 0.06(2) Uiso 1 d . . .
C12 C 0.341(2) 0.0220(12) 0.0000 0.008(12) Uiso 1 d . . .
C13 C 0.058(2) 0.1064(15) 0.0000 0.029(16) Uiso 1 d . . .
C14 C 0.035(2) 0.1594(15) 0.0000 0.015(12) Uiso 1 d . . .
C15 C 0.071(3) 0.2030(16) 0.0000 0.028(14) Uiso 1 d . . .
C16 C 0.035(2) 0.2514(18) 0.0000 0.036(15) Uiso 1 d . . .

_geom_special_details
;
 All esds (except the esd in the dihedral angle between two l.s. planes)
 are estimated using the full covariance matrix.  The cell esds are taken
 into account individually in the estimation of esds in distances, angles
 and torsion angles; correlations between esds in cell parameters are only
 used when they are defined by crystal symmetry.  An approximate (isotropic)
 treatment of cell esds is used for estimating esds involving l.s. planes.
;

loop_
 _geom_bond_atom_site_label_1
 _geom_bond_atom_site_label_2
 _geom_bond_distance
 _geom_bond_site_symmetry_2
 _geom_bond_publ_flag
Cu1 N7 1.85(6) . ?
Cu1 N7 1.85(6) 5 ?
Cu1 N6 1.80(5) 5 ?
Cu1 N6 1.80(5) . ?
```

```
Cl2 C16 1.67(5) . ?
Cl3 C15 1.84(6) . ?
Cl4 C11 1.50(5) . ?
Cl5 C12 1.67(4) . ?
N6 C13 1.50(5) . ?
N6 C13 1.50(5) 2 ?
N7 C9 1.37(5) . ?
N7 C9 1.37(5) 6 ?
N8 C13 1.27(6) . ?
N8 C9 1.34(5) . ?
C9 C10 1.57(8) . ?
C10 C10 1.47(9) 6 ?
C10 C11 1.52(7) . ?
C11 C12 1.54(6) . ?
C12 C12 1.15(6) 6 ?
C13 C14 1.45(5) . ?
C14 C15 1.34(5) . ?
C14 C14 1.36(8) 2 ?
C15 C16 1.45(6) . ?
C16 C16 1.36(9) 2 ?

loop_
 _geom_angle_atom_site_label_1
 _geom_angle_atom_site_label_2
 _geom_angle_atom_site_label_3
 _geom_angle
 _geom_angle_site_symmetry_1
 _geom_angle_site_symmetry_3
 _geom_angle_publ_flag
N7 Cu1 N7 180.0 . 5 ?
N7 Cu1 N6 90.0 . 5 ?
N7 Cu1 N6 90.0 5 5 ?
N7 Cu1 N6 90.0 . . ?
N7 Cu1 N6 90.0 5 . ?
N6 Cu1 N6 180.0 5 . ?
C13 N6 C13 99(5) . 2 ?
C13 N6 Cu1 131(2) . . ?
C13 N6 Cu1 131(2) 2 . ?
C9 N7 C9 107(5) . 6 ?
C9 N7 Cu1 126(3) . . ?
C9 N7 Cu1 126(3) 6 . ?
C13 N8 C9 120(4) . . ?
N8 C9 N7 130(5) . . ?
N8 C9 C10 117(4) . . ?
N7 C9 C10 113(4) . . ?
C10 C10 C9 104(2) 6 . ?
C10 C10 C11 121(3) 6 . ?
C9 C10 C11 135(4) . . ?
C10 C11 C12 110(4) . . ?
C10 C11 Cl4 116(4) . . ?
C12 C11 Cl4 134(4) . . ?
C12 C12 C11 128(2) 6 . ?
C12 C12 Cl5 124.7(14) 6 . ?
C11 C12 Cl5 107(3) . . ?
N8 C13 C14 125(4) . . ?
N8 C13 N6 123(4) . . ?
C14 C13 N6 112(4) . . ?
```

```
C15 C14 C13 130(4) . . ?
C15 C14 C14 122(3) . 2 ?
C13 C14 C14 108(3) . 2 ?
C16 C15 C14 119(4) . . ?
C16 C15 Cl3 115(3) . . ?
C14 C15 Cl3 126(4) . . ?
C15 C16 C16 119(3) . 2 ?
C15 C16 Cl2 121(4) . . ?
C16 C16 Cl2 119.9(19) 2 . ?

_diffrn_measured_fraction_theta_max    0.070
_diffrn_reflns_theta_full              44.64
_diffrn_measured_fraction_theta_full   0.070
_refine_diff_density_rms    0.031
```

8.9 Isotactic Poly(1-butene) Form III

```
data_shelxl

_audit_creation_method            SHELXL-97
_chemical_name_common              'Poly 1-butene FormIII'
_chemical_formula_moiety          ?
_chemical_formula_sum             'C8'
_chemical_formula_weight          96.08

_citation_journal_full Polymer
_citation_journal_volume 35
_citation_langauge English
_citation_page_first 908
_citation_page_last 915
_citation_year 1993
_citation_author_name 'Kopp, S., Wittmann, J. C., Lotz, B.'

loop_
 _atom_type_symbol
 _atom_type_description
 _atom_type_scat_dispersion_real
 _atom_type_scat_dispersion_imag
 _atom_type_scat_source
 'C'  'C'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'

_symmetry_cell_setting            ?
_symmetry_space_group_name_H-M    'P 21 21 21'

loop_
 _symmetry_equiv_pos_as_xyz
 'x, y, z'
 '-x+1/2, -y, z+1/2'
 'x+1/2, -y+1/2, -z'
 '-x, y+1/2, -z+1/2'

_cell_length_a                    12.380
_cell_length_b                    8.880
_cell_length_c                    7.560
_cell_angle_alpha                 90.00
_cell_angle_beta                  90.00
_cell_angle_gamma                 90.00
_cell_volume                      831.1
_cell_formula_units_Z             4
_cell_measurement_temperature     293(2)

_exptl_crystal_density_diffrn     0.768
_exptl_crystal_density_method     'not measured'
_exptl_crystal_F_000              80
_exptl_absorpt_coefficient_mu     0.004



_diffrn_ambient_temperature       293(2)
_diffrn_radiation_wavelength      ?
_diffrn_radiation_type            Electron
```

```
_diffrn_radiation_source          Microscope
_diffrn_reflns_number             146
_diffrn_reflns_av_R_equivalents   0.0000
_diffrn_reflns_av_sigmaI/netI     0.0998
_diffrn_reflns_limit_h_min        0
_diffrn_reflns_limit_h_max        10
_diffrn_reflns_limit_k_min        0
_diffrn_reflns_limit_k_max        8
_diffrn_reflns_limit_l_min        0
_diffrn_reflns_limit_l_max        4
_diffrn_reflns_theta_min          4.88
_diffrn_reflns_theta_max          33.56
_reflns_number_total              146
_reflns_number_gt                 146
_reflns_threshold_expression      >2sigma(I)


_computing_structure_solution     'SHELXS-97 (Sheldrick, 1990)'
_computing_structure_refinement   'SHELXL-97 (Sheldrick, 1997)'
_computing_molecular_graphics     ?
_computing_publication_material   maXus

_refine_special_details
;
 Refinement of F^2^ against ALL reflections.  The weighted R-factor wR and
 goodness of fit S are based on F^2^, conventional R-factors R are based
 on F, with F set to zero for negative F^2^. The threshold expression of
 F^2^ > 2sigma(F^2^) is used only for calculating R-factors(gt) etc. and is
 not relevant to the choice of reflections for refinement.  R-factors based
 on F^2^ are statistically about twice as large as those based on F, and R-
 factors based on ALL data will be even larger.
;

_refine_ls_structure_factor_coef  Fsqd
_refine_ls_matrix_type            full
_refine_ls_weighting_scheme
 'calc w=1/[\s^2^(Fo^2^)+(0.1000P)^2^+0.0000P] where P=(Fo^2^+2Fc^2^)/3'
_atom_sites_solution_primary      direct
_atom_sites_solution_secondary    difmap
_atom_sites_solution_hydrogens    geom
_refine_ls_hydrogen_treatment     mixed
_refine_ls_extinction_method      none
_refine_ls_extinction_coef        ?
_refine_ls_abs_structure_details
 'Flack H D (1983), Acta Cryst. A39, 876-881'
_refine_ls_abs_structure_Flack    0(10)
_refine_ls_number_reflns          146
_refine_ls_number_parameters      35
_refine_ls_number_restraints      0
_refine_ls_R_factor_all           0.2415
_refine_ls_R_factor_gt            0.2415
_refine_ls_wR_factor_ref          0.5284
_refine_ls_wR_factor_gt           0.5284
_refine_ls_goodness_of_fit_ref    4.620
_refine_ls_restrained_S_all       4.620
_refine_ls_shift/su_max           0.369
_refine_ls_shift/su_mean          0.026
```

```
loop_
 _atom_site_label
 _atom_site_type_symbol
 _atom_site_fract_x
 _atom_site_fract_y
 _atom_site_fract_z
 _atom_site_U_iso_or_equiv
 _atom_site_adp_type
 _atom_site_occupancy
 _atom_site_calc_flag
 _atom_site_refinement_flags
 _atom_site_disorder_assembly
 _atom_site_disorder_group
C1 C 0.333(3) -0.053(4) 0.125(9) 0.056(15) Uiso 1 d . . .
C2 C 0.318(2) 0.141(3) 0.210(6) 0.000(9) Uiso 1 d . . .
C3 C 0.369(2) 0.242(2) 0.048(5) 0.000(8) Uiso 1 d . . .
C4 C 0.318(5) 0.411(7) 0.066(9) 0.14(2) Uiso 1 d . . .
C5 C 0.276(3) 0.145(4) 0.376(7) 0.044(13) Uiso 1 d . . .
C6 C 0.167(3) 0.055(4) 0.340(9) 0.068(17) Uiso 1 d . . .
C7 C 0.083(2) 0.212(3) 0.388(7) 0.031(11) Uiso 1 d . . .
C8 C -0.042(3) 0.103(5) 0.373(8) 0.073(16) Uiso 1 d . . .

_geom_special_details
;
 All esds (except the esd in the dihedral angle between two l.s. planes)
 are estimated using the full covariance matrix.  The cell esds are taken
 into account individually in the estimation of esds in distances, angles
 and torsion angles; correlations between esds in cell parameters are only
 used when they are defined by crystal symmetry.  An approximate (isotropic)
 treatment of cell esds is used for estimating esds involving l.s. planes.
;

loop_
 _geom_bond_atom_site_label_1
 _geom_bond_atom_site_label_2
 _geom_bond_distance
 _geom_bond_site_symmetry_2
 _geom_bond_publ_flag
C1 C6 2.16(8) 2_554 ?
C1 C2 1.85(5) . ?
C2 C5 1.36(6) . ?
C2 C3 1.65(5) . ?
C3 C4 1.63(6) . ?
C5 C6 1.60(5) . ?
C6 C1 2.16(8) 2 ?
C6 C7 1.77(5) . ?
C7 C8 1.83(5) . ?

loop_
 _geom_angle_atom_site_label_1
 _geom_angle_atom_site_label_2
 _geom_angle_atom_site_label_3
 _geom_angle
 _geom_angle_site_symmetry_1
 _geom_angle_site_symmetry_3
 _geom_angle_publ_flag
```

```
C6 C1 C2 111(3) 2_554 . ?
C5 C2 C3 145(3) . . ?
C5 C2 C1 113(3) . . ?
C3 C2 C1 102(3) . . ?
C4 C3 C2 107(3) . . ?
C2 C5 C6 99(4) . . ?
C1 C6 C7 78(3) 2 . ?
C1 C6 C5 80(3) 2 . ?
C7 C6 C5 94(2) . . ?
C8 C7 C6 94(2) . . ?

_diffrn_measured_fraction_theta_max    0.365
_diffrn_reflns_theta_full              33.56
_diffrn_measured_fraction_theta_full   0.365
_refine_diff_density_max    0.049
_refine_diff_density_min   -0.063
_refine_diff_density_rms    0.017
```

8.10 Poly (ε-caprolactone)

```
data_shelxl

_audit_creation_method            SHELXL-97

_chemical_name_common             Polycaprolactone
_chemical_formula_moiety          ?
_chemical_formula_sum             'C6 O2'
_chemical_formula_weight          104.06

_citation_journal_full Macromolecules
_citation_journal_volume 23
_citation_langauge English
_citation_page_first 4604
_citation_page_last 4607
_citation_year 1990
_citation_author_name 'Hu, H., Dorset, D.'

loop_
 _atom_type_symbol
 _atom_type_description
 _atom_type_scat_dispersion_real
 _atom_type_scat_dispersion_imag
 _atom_type_scat_source
 'C'  'C'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
 'O'  'O'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'

_symmetry_cell_setting            ?
_symmetry_space_group_name_H-M    'P 21 21 21'

loop_
 _symmetry_equiv_pos_as_xyz
 'x, y, z'
 '-x+1/2, -y, z+1/2'
 'x+1/2, -y+1/2, -z'
 '-x, y+1/2, -z+1/2'

_cell_length_a                    7.480
_cell_length_b                    4.980
_cell_length_c                    17.260
_cell_angle_alpha                 90.00
_cell_angle_beta                  90.00
_cell_angle_gamma                 90.00
_cell_volume                      642.9
_cell_formula_units_Z             4
_cell_measurement_temperature     293(2)


_exptl_crystal_density_diffrn     1.075
_exptl_crystal_density_method     'not measured'
_exptl_crystal_F_000              76
_exptl_absorpt_coefficient_mu     0.005

_diffrn_ambient_temperature       293(2)
```

```
_diffrn_radiation_wavelength      ?
_diffrn_radiation_type            electron
_diffrn_radiation_source          microscope
_diffrn_reflns_number             47
_diffrn_reflns_av_R_equivalents   0.0000
_diffrn_reflns_av_sigmaI/netI     0.0026
_diffrn_reflns_limit_h_min        0
_diffrn_reflns_limit_h_max        5
_diffrn_reflns_limit_k_min        0
_diffrn_reflns_limit_k_max        3
_diffrn_reflns_limit_l_min        0
_diffrn_reflns_limit_l_max        14
_diffrn_reflns_theta_min          7.37
_diffrn_reflns_theta_max          33.73
_reflns_number_total              47
_reflns_number_gt                 45
_reflns_threshold_expression      >2sigma(I)

_computing_structure_solution     'SHELXS-97 (Sheldrick, 1990)'
_computing_structure_refinement   'SHELXL-97 (Sheldrick, 1997)'
_computing_molecular_graphics     ?
_computing_publication_material   maXus

_refine_special_details
;
 Refinement of F^2^ against ALL reflections.  The weighted R-factor wR and
 goodness of fit S are based on F^2^, conventional R-factors R are based
 on F, with F set to zero for negative F^2^. The threshold expression of
 F^2^ > 2sigma(F^2^) is used only for calculating R-factors(gt) etc. and is
 not relevant to the choice of reflections for refinement.  R-factors based
 on F^2^ are statistically about twice as large as those based on F, and R-
 factors based on ALL data will be even larger.
;

_refine_ls_structure_factor_coef  Fsqd
_refine_ls_matrix_type            full
_refine_ls_weighting_scheme
 'calc w=1/[\s^2^(Fo^2^)+(0.1000P)^2^+0.0000P] where P=(Fo^2^+2Fc^2^)/3'
_atom_sites_solution_primary      direct
_atom_sites_solution_secondary    difmap
_atom_sites_solution_hydrogens    geom
_refine_ls_hydrogen_treatment     mixed
_refine_ls_extinction_method      none
_refine_ls_extinction_coef        ?
_refine_ls_abs_structure_details
 'Flack H D (1983), Acta Cryst. A39, 876-881'
_refine_ls_abs_structure_Flack    0(10)
_refine_ls_number_reflns          47
_refine_ls_number_parameters      33
_refine_ls_number_restraints      2
_refine_ls_R_factor_all           0.1989
_refine_ls_R_factor_gt            0.1988
_refine_ls_wR_factor_ref          0.5053
_refine_ls_wR_factor_gt           0.5052
_refine_ls_goodness_of_fit_ref    9.937
_refine_ls_restrained_S_all       9.299
_refine_ls_shift/su_max           0.121
```

```
_refine_ls_shift/su_mean          0.047

loop_
 _atom_site_label
 _atom_site_type_symbol
 _atom_site_fract_x
 _atom_site_fract_y
 _atom_site_fract_z
 _atom_site_U_iso_or_equiv
 _atom_site_adp_type
 _atom_site_occupancy
 _atom_site_calc_flag
 _atom_site_refinement_flags
 _atom_site_disorder_assembly
 _atom_site_disorder_group
C1 C 0.683(16) 0.543(6) 0.3637(18) 0.076(14) Uiso 1 d D . .
C2 C 0.726(19) 0.654(12) 0.2287(12) 0.080(19) Uiso 1 d . . .
C3 C 0.78(4) 0.474(8) 0.158(3) 0.084(17) Uiso 1 d . . .
C4 C 0.729(16) 0.600(8) 0.0829(16) 0.072(10) Uiso 1 d D . .
C5 C 0.818(18) 0.462(6) 0.012(2) 0.066(15) Uiso 1 d D . .
C6 C 0.78(3) 0.436(7) 0.4393(15) 0.054(12) Uiso 1 d . . .
O7 O 0.58(2) 0.722(10) 0.366(2) 0.094(19) Uiso 1 d . . .
O8 O 0.77(2) 0.497(6) 0.293(2) 0.054(13) Uiso 1 d D . .

_geom_special_details
;
 All esds (except the esd in the dihedral angle between two l.s. planes)
 are estimated using the full covariance matrix.  The cell esds are taken
 into account individually in the estimation of esds in distances, angles
 and torsion angles; correlations between esds in cell parameters are only
 used when they are defined by crystal symmetry.  An approximate (isotropic)
 treatment of cell esds is used for estimating esds involving l.s. planes.
;

loop_
 _geom_bond_atom_site_label_1
 _geom_bond_atom_site_label_2
 _geom_bond_distance
 _geom_bond_site_symmetry_2
 _geom_bond_publ_flag
C1 O7 1.16(14) . ?
C1 O8 1.39(7) . ?
C1 C6 1.59(12) . ?
C2 O8 1.40(7) . ?
C2 C3 1.55(9) . ?
C3 C4 1.49(9) . ?
C4 C5 1.55(7) . ?
C5 C6 1.55(13) 2_664 ?
C6 C5 1.55(13) 2_665 ?

loop_
 _geom_angle_atom_site_label_1
 _geom_angle_atom_site_label_2
 _geom_angle_atom_site_label_3
 _geom_angle
 _geom_angle_site_symmetry_1
 _geom_angle_site_symmetry_3
```

```
 _geom_angle_publ_flag
O7 C1 O8 117(5) . . ?
O7 C1 C6 122(5) . . ?
O8 C1 C6 116(8) . . ?
O8 C2 C3 104(6) . . ?
C4 C3 C2 113(8) . . ?
C3 C4 C5 114(8) . . ?
C4 C5 C6 107(7) . 2_664 ?
C1 C6 C5 109(10) . 2_665 ?
C1 O8 C2 120(8) . . ?

_diffrn_measured_fraction_theta_max    0.144
_diffrn_reflns_theta_full              33.73
_diffrn_measured_fraction_theta_full   0.144
_refine_diff_density_max    0.194
_refine_diff_density_min   -0.206
_refine_diff_density_rms    0.081
```

8.11 Polyethylene

```
data_shelxl

_audit_creation_method            SHELXL-97
_chemical_name_systematic
;
 ?
;
_chemical_name_common             Polyethylene
_chemical_formula_moiety          (CH2)n
_chemical_formula_sum
 'C H2'
_chemical_formula_weight          14.03

_citation_journal_full Acta Cryst
_citation_journal_volume B45
_citation_langauge English
_citation_page_first 283
_citation_page_last 290
_citation_year 1989
_citation_author_name 'Hu,H., Dorset, D.'

loop_
 _atom_type_symbol
 _atom_type_description
 _atom_type_scat_dispersion_real
 _atom_type_scat_dispersion_imag
 _atom_type_scat_source
 'C'  'C'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
 'H'  'H'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'

_symmetry_cell_setting            ?
_symmetry_space_group_name_H-M    Pnma

loop_
 _symmetry_equiv_pos_as_xyz
 'x, y, z'
 '-x+1/2, -y, z+1/2'
 'x+1/2, -y+1/2, -z+1/2'
 '-x, y+1/2, -z'
 '-x, -y, -z'
 'x-1/2, y, -z-1/2'
 '-x-1/2, y-1/2, z-1/2'
 'x, -y-1/2, z'

_cell_length_a                    7.480
_cell_length_b                    2.550
_cell_length_c                    4.970
_cell_angle_alpha                 90.00
_cell_angle_beta                  90.00
_cell_angle_gamma                 90.00
_cell_volume                      94.8
_cell_formula_units_Z             8
_cell_measurement_temperature     293(2)
```

```
_exptl_crystal_F_000               29
_exptl_absorpt_coefficient_mu      0.025



_diffrn_ambient_temperature        293(2)
_diffrn_radiation_wavelength       ?
_diffrn_radiation_type             electron
_diffrn_radiation_source           microscope
_diffrn_reflns_number              51
_diffrn_reflns_av_R_equivalents    0.0000
_diffrn_reflns_av_sigmaI/netI      90.7853
_diffrn_reflns_limit_h_min         0
_diffrn_reflns_limit_h_max         7
_diffrn_reflns_limit_k_min         0
_diffrn_reflns_limit_k_max         2
_diffrn_reflns_limit_l_min         0
_diffrn_reflns_limit_l_max         6
_diffrn_reflns_theta_min           8.52
_diffrn_reflns_theta_max           47.79
_reflns_number_total               51
_reflns_number_gt                  22
_reflns_threshold_expression       >2sigma(I)

_computing_structure_solution      'SHELXS-97 (Sheldrick, 1990)'
_computing_structure_refinement    'SHELXL-97 (Sheldrick, 1997)'
_computing_molecular_graphics      ?
_computing_publication_material    maXus

_refine_special_details
;
 Refinement of F^2^ against ALL reflections.  The weighted R-factor wR and
 goodness of fit S are based on F^2^, conventional R-factors R are based
 on F, with F set to zero for negative F^2^. The threshold expression of
 F^2^ > 2sigma(F^2^) is used only for calculating R-factors(gt) etc. and is
 not relevant to the choice of reflections for refinement.  R-factors based
 on F^2^ are statistically about twice as large as those based on F, and R-
 factors based on ALL data will be even larger.
;

_refine_ls_structure_factor_coef   Fsqd
_refine_ls_matrix_type             full
_refine_ls_weighting_scheme
 'calc w=1/[\s^2^(Fo^2^)+(0.1000P)^2^+0.0000P] where P=(Fo^2^+2Fc^2^)/3'
_atom_sites_solution_primary       direct
_atom_sites_solution_secondary     difmap
_atom_sites_solution_hydrogens     geom
_refine_ls_hydrogen_treatment      mixed
_refine_ls_extinction_method       none
_refine_ls_extinction_coef         ?
_refine_ls_number_reflns           51
_refine_ls_number_parameters       6
_refine_ls_number_restraints       0
_refine_ls_R_factor_all            0.2095
```

```
_refine_ls_R_factor_gt            0.1358
_refine_ls_wR_factor_ref          0.3916
_refine_ls_wR_factor_gt           0.2725
_refine_ls_goodness_of_fit_ref    1.230
_refine_ls_restrained_S_all       1.230
_refine_ls_shift/su_max           0.000
_refine_ls_shift/su_mean          0.000

loop_
 _atom_site_label
 _atom_site_type_symbol
 _atom_site_fract_x
 _atom_site_fract_y
 _atom_site_fract_z
 _atom_site_U_iso_or_equiv
 _atom_site_adp_type
 _atom_site_occupancy
 _atom_site_calc_flag
 _atom_site_refinement_flags
 _atom_site_disorder_assembly
 _atom_site_disorder_group
C1 C 0.047(3) 0.2500 0.060(4) 0.063(9) Uiso 1 d S . .
H1A H 0.1924 0.2500 0.0503 0.12(5) Uiso 1 d SR . .
H1B H 0.0365 0.2500 0.2590 0.14(5) Uiso 1 d SR . .

_geom_special_details
;
 All esds (except the esd in the dihedral angle between two l.s. planes)
 are estimated using the full covariance matrix.  The cell esds are taken
 into account individually in the estimation of esds in distances, angles
 and torsion angles; correlations between esds in cell parameters are only
 used when they are defined by crystal symmetry.  An approximate (isotropic)
 treatment of cell esds is used for estimating esds involving l.s. planes.
;

loop_
 _geom_bond_atom_site_label_1
 _geom_bond_atom_site_label_2
 _geom_bond_distance
 _geom_bond_site_symmetry_2
 _geom_bond_publ_flag
C1 C1 1.57(2) 5 ?
C1 C1 1.57(2) 5_565 ?
C1 H1A 1.0879 . ?
C1 H1B 0.9922 . ?

loop_
 _geom_angle_atom_site_label_1
 _geom_angle_atom_site_label_2
 _geom_angle_atom_site_label_3
 _geom_angle
 _geom_angle_site_symmetry_1
 _geom_angle_site_symmetry_3
 _geom_angle_publ_flag
C1 C1 C1 108(2) 5 5_565 ?
C1 C1 H1A 115.5 5 . ?
C1 C1 H1A 115.5 5_565 . ?
```

```
C1 C1 H1B 110.0 5 . ?
C1 C1 H1B 110.0 5_565 . ?
H1A C1 H1B 97.1 . . ?

_diffrn_measured_fraction_theta_max    0.425
_diffrn_reflns_theta_full              47.79
_diffrn_measured_fraction_theta_full   0.425
_refine_diff_density_max    0.114
_refine_diff_density_min   -0.081
_refine_diff_density_rms    0.030
```

8.12 Poly(1,4,trans-cyclohexanediyl dimethylene succinate)

```
data_shelxl

_audit_creation_method            SHELXL-97
_chemical_name_common             t-cds
_chemical_formula_moiety          'trans(1,4-trans-
clclohexanediyldimethylene
succinate'
_chemical_formula_sum            'C6 02'
_chemical_formula_weight          104.06

_citation_journal_full Macromolecules
_citation_journal_volume 17
_citation_langauge English
_citation_page_first 1980
_citation_page_last 1987
_citation_year 1984
_citation_author_name 'Fancois Brisse, Bruno Remillard, Henri Chanzy'

loop_
 _atom_type_symbol
 _atom_type_description
 _atom_type_scat_dispersion_real
 _atom_type_scat_dispersion_imag
 _atom_type_scat_source
 'O'  'O'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'
 'C'  'C'   0.0000   0.0000
 'International Tables Vol C Tables 4.2.6.8 and 6.1.1.4'

_symmetry_cell_setting            ?
_symmetry_space_group_name_H-M    ? P21/n

loop_
 _symmetry_equiv_pos_as_xyz
 'x, y, z'
 '-x+1/2, y+1/2, -z+1/2'
 '-x, -y, -z'
 'x-1/2, -y-1/2, z-1/2'

_cell_length_a                    6.490
_cell_length_b                    9.480
_cell_length_c                    13.510
_cell_angle_alpha                 90.00
_cell_angle_beta                  45.90
_cell_angle_gamma                 90.00
_cell_volume                      596.9
_cell_formula_units_Z             4
_cell_measurement_temperature     293(2)




_diffrn_ambient_temperature       293(2)
_diffrn_radiation_wavelength      ?
_diffrn_radiation_type            electron
```

```
_diffrn_radiation_source          microscope
_diffrn_reflns_number             85
_diffrn_reflns_av_R_equivalents   0.0000
_diffrn_reflns_av_sigmaI/netI     0.0175
_diffrn_reflns_limit_h_min        0
_diffrn_reflns_limit_h_max        5
_diffrn_reflns_limit_k_min        0
_diffrn_reflns_limit_k_max        9
_diffrn_reflns_limit_l_min        0
_diffrn_reflns_limit_l_max        8
_diffrn_reflns_theta_min          5.66
_diffrn_reflns_theta_max          40.80
_reflns_number_total              85
_reflns_number_gt                 85
_reflns_threshold_expression      >2sigma(I)

_computing_structure_solution     'SHELXS-97 (Sheldrick, 1990)'
_computing_structure_refinement   'SHELXL-97 (Sheldrick, 1997)'
_computing_molecular_graphics     ?
_computing_publication_material   maXus

_refine_special_details
;
 Refinement of F^2^ against ALL reflections.  The weighted R-factor wR and
 goodness of fit S are based on F^2^, conventional R-factors R are based
 on F, with F set to zero for negative F^2^. The threshold expression of
 F^2^ > 2sigma(F^2^) is used only for calculating R-factors(gt) etc. and is
 not relevant to the choice of reflections for refinement.  R-factors based
 on F^2^ are statistically about twice as large as those based on F, and R-
 factors based on ALL data will be even larger.
;

_refine_ls_structure_factor_coef  Fsqd
_refine_ls_matrix_type            full
_refine_ls_weighting_scheme
 'calc w=1/[\s^2^(Fo^2^)+(0.1000P)^2^+0.0000P] where P=(Fo^2^+2Fc^2^)/3'
_atom_sites_solution_primary      direct
_atom_sites_solution_secondary    difmap
_atom_sites_solution_hydrogens    geom
_refine_ls_hydrogen_treatment     mixed
_refine_ls_extinction_method      none
_refine_ls_extinction_coef        ?
_refine_ls_number_reflns          85
_refine_ls_number_parameters      33
_refine_ls_number_restraints      0
_refine_ls_R_factor_all           0.1614
_refine_ls_R_factor_gt            0.1614
_refine_ls_wR_factor_ref          0.4696
_refine_ls_wR_factor_gt           0.4696
_refine_ls_goodness_of_fit_ref    5.311
_refine_ls_restrained_S_all       5.311
_refine_ls_shift/su_max           0.109
_refine_ls_shift/su_mean          0.030

loop_
 _atom_site_label
 _atom_site_type_symbol
```

```
 _atom_site_fract_x
 _atom_site_fract_y
 _atom_site_fract_z
 _atom_site_U_iso_or_equiv
 _atom_site_adp_type
 _atom_site_occupancy
 _atom_site_calc_flag
 _atom_site_refinement_flags
 _atom_site_disorder_assembly
 _atom_site_disorder_group
O1 O 0.049(4) -0.195(2) 0.370(4) 0.076(10) Uiso 1 d . . .
O2 O -0.077(4) -0.026(2) 0.308(3) 0.058(9) Uiso 1 d . . .
C3 C 0.184(5) 0.038(3) 0.015(4) 0.074(10) Uiso 1 d . . .
C4 C -0.054(4) -0.043(2) 0.113(3) 0.035(8) Uiso 1 d . . .
C5 C -0.120(5) -0.152(2) 0.058(4) 0.053(10) Uiso 1 d . . .
C6 C -0.069(5) -0.121(3) 0.216(4) 0.074(11) Uiso 1 d . . .
C7 C 0.005(4) -0.073(2) 0.368(3) 0.024(7) Uiso 1 d . . .
C8 C -0.011(5) 0.027(3) 0.444(4) 0.057(10) Uiso 1 d . . .

_geom_special_details
;
 All esds (except the esd in the dihedral angle between two l.s. planes)
 are estimated using the full covariance matrix.  The cell esds are taken
 into account individually in the estimation of esds in distances, angles
 and torsion angles; correlations between esds in cell parameters are only
 used when they are defined by crystal symmetry.  An approximate (isotropic)
 treatment of cell esds is used for estimating esds involving l.s. planes.
;

loop_
 _geom_bond_atom_site_label_1
 _geom_bond_atom_site_label_2
 _geom_bond_distance
 _geom_bond_site_symmetry_2
 _geom_bond_publ_flag
O1 C7 1.20(3) . ?
O2 C7 1.31(4) . ?
O2 C6 1.50(5) . ?
C3 C4 1.37(3) . ?
C3 C5 1.70(5) 3 ?
C4 C6 1.52(5) . ?
C4 C5 1.50(4) . ?
C5 C3 1.70(5) 3 ?
C7 C8 1.35(5) . ?
C8 C8 1.68(8) 3_556 ?

loop_
 _geom_angle_atom_site_label_1
 _geom_angle_atom_site_label_2
 _geom_angle_atom_site_label_3
 _geom_angle
 _geom_angle_site_symmetry_1
 _geom_angle_site_symmetry_3
 _geom_angle_publ_flag
C7 O2 C6 121(2) . . ?
C4 C3 C5 111(3) . 3 ?
C6 C4 C3 118(3) . . ?
```

```
C6 C4 C5 105(2) . . ?
C3 C4 C5 115(3) . . ?
C3 C5 C4 97(2) 3 . ?
C4 C6 O2 114(2) . . ?
O1 C7 O2 123(3) . . ?
O1 C7 C8 123(4) . . ?
O2 C7 C8 113(2) . . ?
C7 C8 C8 118(3) . 3_556 ?

_diffrn_measured_fraction_theta_max    0.106
_diffrn_reflns_theta_full              40.80
_diffrn_measured_fraction_theta_full   0.106
_refine_diff_density_max    0.094
_refine_diff_density_min   -0.107
_refine_diff_density_rms    0.028
```