

Prediction of molecular crystal structures

Theresa Beyer

A thesis submitted to the University of London in partial fulfilment
of the requirements for the degree of Doctor of Philosophy, April 2001.

Centre for Theoretical and Computational Chemistry
Department of Chemistry
University College London
20 Gordon Street
London WC1H 0AJ
United Kingdom

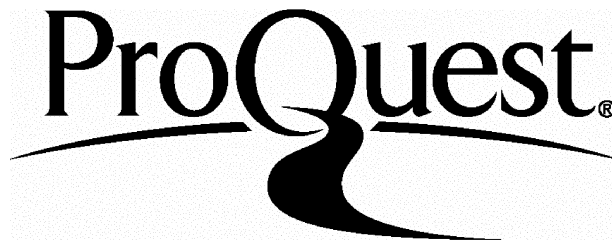
ProQuest Number: 10014432

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10014432

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Academic acknowledgements

I would like to thank my supervisor Professor Sarah L. Price for her enthusiasm and continued support over the course of my postgraduate work. It has been a great pleasure to work with her and I am grateful for her commitment and advice.

Dr W. D. Sam Motherwell and Dr Jos P. M. Lommerse are thanked for collaboration, many helpful and enjoyable discussions and for organising the CCDC crystal structure prediction workshop for small organic molecules in May 1999 and the annual CCDC student days.

Furthermore I would like to thank Dr Herman L. Ammon for help with the program MOLPAK and Professor Christopher S. Frampton (Roche Discovery) and Dr C. C. Wilson (CLRC Rutherford Appelton Laboratories) for providing the unpublished atomic coordinates of many of the paracetamol crystal structures (chapter 6).

Mr. Amer Ramzan, a UCL project student, is thanked for preliminary work on the crystal structure prediction of paracetamol (chapter 6).

I would like to thank all past and present members of our research group for their friendship and their support, especially Dr Tanja van Mourik for her advice and help with *ab initio* techniques. Graeme Day and Michael Brunsteiner are thanked for an introduction into elastic properties and morphology calculations, respectively (chapter 6 and 7).

I am grateful to NATO Science (Euroconference, Erice/Italy, 1998), the British Crystallographic Association (IUCr XVIIIth Congress and General Assembly, Glasgow/United Kingdom, 1999) and the Bursary Committee of the 19th European Crystallographic Meeting (ECM19, Nancy/France, 2000) for conference grants.

Finally I would like to thank the EPSRC and the Cambridge Crystallographic Data Centre for a generous Ph. D. case studentship.

Personal acknowledgements

I would like to thank my family – especially my parents, and Michael’s family for their continued support.

Michael for his talent to make me smile and much more ...

Leaving a familiar place always means leaving friends behind. I would like to thank my friends for the strength and dedication to keep in touch. There are many more years to come ...

All Sisters of St Philomena’s for a warm place in the heart of the city.

London – for attracting people from all over the world and therefore offering a unique learning experience. It has taught me a lesson or two ...

Finally, I would like to thank the UCL Graduate School, the Research Council's Graduate School and the Higher Education and Research and Development Unit (HERDU) / UCL for the non-scientific part of my postgraduate education.

Abstract

The *ab initio* prediction of molecular crystal structures is a scientific challenge. Reliability of first-principle prediction calculations would mirror a fundamental understanding of crystallisation. Crystal structure prediction is also of considerable practical importance as different crystalline arrangements of the same molecule in the solid state (polymorphs) are likely to have different physical properties. A method of crystal structure prediction based on lattice energy minimisation has been developed in this work. The choice of the intermolecular potential and of the molecular model is crucial for the results of such studies and both of these criteria have been investigated. An empirical atom-atom repulsion-dispersion potential for carboxylic acids has been derived and applied in a crystal structure prediction study of formic, benzoic and the polymorphic system of tetrolic acid. As many experimental crystal structure determinations at different temperatures are available for the polymorphic system of paracetamol (acetaminophen), the influence of the variations of the molecular model on the crystal structure lattice energy minima, has also been studied.

The general problem of prediction methods based on the assumption that the experimental thermodynamically stable polymorph corresponds to the global lattice energy minimum, is that more hypothetical low lattice energy structures are found within a few kJ mol^{-1} of the global minimum than are likely to be experimentally observed polymorphs. This is illustrated by the results for molecule I, 3-oxabicyclo(3.2.0)hepta-1,4-diene, studied for the first international blindtest for small organic crystal structures organised by the Cambridge Crystallographic Data Centre (CCDC) in May 1999.

To reduce the number of predicted polymorphs, additional factors to thermodynamic criteria have to be considered. Therefore the elastic constants and vapour growth morphologies have been calculated for the lowest lattice energy structures of paracetamol and the carboxylic acids. These provide approximate mechanical and kinetic models to refine polymorph prediction.

Table of contents

Academic acknowledgements	2
Personal acknowledgements	3
Abstract	4
Table of contents	5
List of tables and figures	8

Chapter 0.

Overview	12
References for chapter 0	16

Chapter 1.

Molecular crystals and polymorphism	17
1.1 Molecular crystals	18
1.2 The phenomenon of polymorphism	19
1.2.1 Discovery and the definition of polymorphism	19
1.2.2 Conformational polymorphs	20
1.2.3 The importance of polymorphism	21
1.2.4 Thermodynamic stability of polymorphs	23
1.2.5 Disappearing polymorphs	24
1.3 Crystal engineering	25
1.4 Databases of molecular crystals and their properties	26
1.4.1 The Cambridge Crystallographic Database (CSD)	26
1.4.2 The WebBook by the National Institute of Standards and Technology (NIST)	27
1.5 Tools for a theoretical comparison of molecular crystal structures and polymorphs	28
1.5.1 Visual comparison	28
1.5.2 Graph set theory	28
1.5.3 Reduced cells	29
1.5.4 Similarity searches	30
1.5.5 X-ray powder diagrams	30
1.5.6 Hirshfeld surfaces	30
1.6 Conclusions	31
References for chapter 1	32

Chapter 2.

The theory of intermolecular forces	41
2.1 Introduction	42
2.2 Classification of the forces between molecules	43
2.2.1 Long range contributions	44
2.2.2 Short range contributions	48
2.3 Models for calculating intermolecular interactions	49
2.3.1 Modelling the electrostatic contribution	49
2.3.1.1 Point charge models	49
2.3.1.2 Central multipole models	50
2.3.1.3 Distributed multipole models	51
2.3.1.3.1 Distributed Multipole Analysis	52
2.3.2 Modelling the induction energy	54

2.3.3	Modelling the dispersion energy.....	55
2.3.4	Modelling the exchange-repulsion energy.....	55
2.3.5	Damping of long range potential terms.....	56
2.4	Development of model intermolecular potentials.....	57
2.4.1	Model potentials for organic crystal structures.....	57
2.5	Conclusions.....	62
	References for chapter 2.....	63
 Chapter 3.		
	Prediction of molecular crystal structures.....	70
3.1	Introduction.....	71
3.2	Crystal structure prediction based on lattice energy methods.....	72
3.2.1	Implementations of crystal structure prediction methods.....	74
3.3	Prediction methods based on statistical distributions of CSD data.....	77
3.4	Studies based on lattice energy minimisation.....	78
3.5	Additional approaches towards solving the prediction problem.....	79
3.6	Structure determination from powder data.....	80
3.7	Conclusions.....	81
	References for chapter 3.....	102
 Chapter 4.		
	Calculation of crystal structures and their properties.....	112
4.1	The lattice energy minimisation program DMAREL.....	113
4.2	Mechanical stability of crystal structures.....	114
4.3	Morphology calculations.....	116
4.3.1	Bravais-Friedel-Donnay-Harker (BFDH) model.....	116
4.3.2	Attachment energy (AE) model.....	117
	References for chapter 4.....	119
 Chapter 5.		
	Cambridge Crystallographic Data Centre (CCDC) crystal structure prediction workshop.....	123
5.1	The blindtest challenge.....	124
5.2	Geometry optimisation of the molecular structures for all compounds.....	125
5.3	Testing the intermolecular potential for compound I.....	126
5.4	The attempt to predict the crystal structure of compound I.....	127
5.5	Comparison with the experimental structures and discussion.....	127
5.6	Comparison with the results of other groups and conclusions.....	129
	References for chapter 5.....	139
 Chapter 6.		
	Polymorphs of paracetamol – sensitivity of the lattice energy minima to experimental variations in the molecular structure crystal structure prediction, morphology and mechanical properties.....	142
6.1	Introduction.....	143
6.2	Method.....	146

6.3	Results	147
6.3.1	Analysis of the molecular structures	147
6.3.2	Analysis of the experimental temperature effects	150
6.3.3	Analysis of the lattice energy minima	150
6.3.4	Crystal structure prediction search	152
6.3.5	Elastic constants and morphology calculations	152
6.4	Discussion	155
6.5	Conclusions	159
	References for chapter 6	182

Chapter 7.

Dimer or catemer? Calculations of low energy packings, morphology and mechanical properties of small carboxylic acids

		188
7.1	Introduction	189
7.2	Method	192
7.2.1	Choice of small carboxylic acids for investigation	192
7.2.2	Development of the intermolecular potential	192
7.2.3	Crystal structure prediction searches	193
7.3	Results	194
7.3.1	Optimisation of the empirical repulsion potential parameters	194
7.3.2	Effect of intramolecular structure variations	195
7.3.3	Generation of hypothetical crystal structures	196
7.3.4	Elastic constants and morphology calculations	197
7.4	Discussion	199
7.4.1	Catemer versus dimer?	199
7.4.2	Additional criteria to lattice energy calculations	199
7.5	Conclusions	200
	References for chapter 7	241

Chapter 8.

	Conclusions and suggestions for further work	246
	References for chapter 8	250

List of tables and figures**Chapter 3**

Table 3.1	82
Methods based on lattice energy minimisation, plus methods based on the statistical distribution of CSD data.	
Table 3.2	83
Summary of published crystal structure prediction studies.	

Chapter 5

Table 5.1	131
The compounds of the CCDC crystal structure prediction blindtest.	
Table 5.2	131
Test molecules retrieved from the CSD (Version December 1998) for the determination of starting parameters for the <i>ab initio</i> optimisation of the molecular structures of the blindtest compounds.	
Table 5.3	132
Repulsion-dispersion potential parameters, FIT (FIT/HA).	
Table 5.4	132
Comparison of the performance of the repulsion-dispersion potentials FIT and FIT/HA in the reproduction of the experimental structures of the two test molecules furane (FURANE10) and (2.2)(3,4) furanophane (WEYXON).	
Table 5.5	133
Results of the blindtest prediction and post analysis for compound I in reduced cells.	
Figure 5.1	134
Overlay of the experimental and lattice energy minimised (FIT) (black) crystal structure of furane (FURANE10).	
Figure 5.2	135
The minima in lattice energy (black symbols) for blindtest compound I found by the MOLPAK/DMAREL search.	
Figure 5.3	136
Overlay of the experimental and minimised (FIT) (black) crystal structure of the 2 nd polymorph (P2 ₁ /c) of compound I.	
Appendix 5.1	137
Fractional coordinates of the three submitted crystal structures for the first blindtest molecule.	

Chapter 6

Scheme 6.1	143
Molecular structure of paracetamol.	
Table 6.1	161
Differences between the experimental and the <i>ab initio</i> optimised molecular structure of paracetamol.	
Table 6.2	162
Experimental crystal structures of paracetamol and the changes in parameters on lattice energy minimisation.	
Table 6.3	163
Low energy crystal structures of paracetamol.	
Table 6.4	164
The mechanical properties for the lowest temperature forms of paracetamol and closest structures found in the search with the <i>ab initio</i> molecular structure.	
Table 6.5	165
Calculated elastic properties (in GPa) for the energetically feasible, hypothetical structures of paracetamol.	
Table 6.6	166
Growth rate and morphology predictions for the known and hypothetical forms of paracetamol.	
Table 6.7	167
Summary of properties of the lowest energy crystal structures of paracetamol found in the lattice energy search.	
Figure 6.1	168
The crystal structures of paracetamol in (a) form I (determined at 20K) and (b) form II (123 K).	
Figure 6.2	169
An overlay of experimental and <i>ab initio</i> molecular structures of paracetamol.	
Figure 6.3	169
Variation of the experimental and lattice energy minimised cell lengths of form I with temperature and determination.	
Figure 6.4	170
Overlay of the experimental (123K) and prediction model (blue) crystal structures.	
Figure 6.5	171
The minima in lattice energy of paracetamol found by the MOLPAK/DMAREL search.	

Figure 6.6	172
BFDH and attachment energy (AE) predicted morphologies for the known and hypothetical crystal structures of paracetamol.	
Appendix 6.1	174
Fractional coordinates of the lowest energy structures of paracetamol found by the MOLPAK/DMAREL search.	
 Chapter 7	
Scheme 7.1	189
Hydrogen bonding motifs observed in carboxylic acids.	
Table 7.1	202
Crystal structures used for deriving the repulsion-dispersion potential for carboxylic acids.	
Table 7.2	203
Validation of the optimised repulsion-dispersion potential for carboxylic acids.	
Table 7.3	204
The repulsion-dispersion potential parameters used in conjunction with a DMA electrostatic model for the carboxylic acids.	
Table 7.4	204
Comparison of the experimental and the <i>ab initio</i> optimised molecular structures of formic, benzoic and tetrolic acid.	
Table 7.5	205
Evaluation of the search procedure.	
Table 7.6.1	206
Low energy crystal structures for formic acid.	
Table 7.6.2	209
Low energy crystal structures for benzoic acid.	
Table 7.6.3	210
Low energy crystal structures for tetrolic acid.	
Table 7.7.1	213
Calculated elastic properties (in GPa) for the minimised experimental structures (Min _{expt} and Min _{opt}) and for the energetically feasible, hypothetical structures for formic acid.	
Table 7.7.2	214
Calculated elastic properties (in GPa) for the minimised experimental structures (Min _{expt} and Min _{opt}) and for the energetically feasible, hypothetical structures for benzoic acid.	

Table 7.7.3	215
Calculated elastic properties (in GPa) for the minimised experimental structures (Min _{expt} and Min _{opt}) and for the energetically feasible, hypothetical structures for tetrolic acid.	
Table 7.8.1	217
Growth rate and morphology predictions for the known and hypothetical forms for formic acid.	
Table 7.8.2	218
Growth rate and morphology predictions for the known and hypothetical forms for benzoic acid.	
Table 7.8.3	219
Growth rate and morphology predictions for the known and hypothetical forms for tetrolic acid.	
Figure 7.1	221
R.m.s. percentage error in the (a) lattice parameters and in the (b) O-(H)...O bond lengths for variations in the polar hydrogen repulsion potential.	
Figure 7.2	223
Comparison of the simulated powder diffraction patterns for the experimental structure (bottom, green), the corresponding minimum in the lattice energy using the optimised molecular structure (middle, blue, offset 20), and the closest structure found in the search procedure shown in table 7.5 (top, red, offset 40) for (a) formic acid, (b) benzoic acid, (c) the α -form of tetrolic acid, and (d) the β -form of tetrolic acid.	
Figure 7.3	224
Minima in lattice energy found by the MOLPAK/DMAREL search, using the DMA+6-exp potential and the <i>ab initio</i> molecular structure for (a) formic acid, (b) benzoic acid and (c) tetrolic acid.	
Figure 7.4.1	228
BFDH and attachment energy (AE) predicted morphologies for the known and hypothetical crystal structures for formic acid.	
Figure 7.4.2	233
BFDH and attachment energy (AE) predicted morphologies for the known and hypothetical crystal structures for benzoic acid.	
Figure 7.4.3	235
BFDH and attachment energy (AE) predicted morphologies for the known and hypothetical crystal structures for tetrolic acid.	

Chapter 0. Overview

*‘ [...] So one has here a big theoretical challenge
going hand in hand with big business.’¹*

The phenomenon of polymorphism of organic crystal structures, the occurrence of different packing arrangements of the same molecule in the solid state², was known at the beginning of the 19th century³. (Chapter 1 provides an introduction into polymorphism and molecular crystals). Due to its significant influence on the physical properties of molecular materials, it has experienced a scientific renaissance on an international scale from the late 1960s when the pharmaceutical industry was confronted with regulations on the bioavailability of new drug substances⁴. These days, a tight control of quality in manufacture and product reliability has become important in *any* application industry and the need to identify polymorphic behaviour at an early

stage of development has fostered the scientific desire of understanding the underlying concepts of crystal formation.

This progress almost naturally faced crystallographers and computational chemists with another challenge, probably one of the scientifically most ambitious goals in this research area, the *ab initio* prediction of molecular crystal structures. The reliable determination of the full crystal structure, including possible polymorphs, using just the atomic connectivity for the organic compound, would mirror a fundamental understanding of the factors that govern the relationship between the molecular and the crystal structure.

A method of molecular crystal structure prediction based on lattice energy minimisation has been developed in this work. The underlying assumption that the experimental thermodynamically stable polymorph corresponds to the global lattice energy minimum has been utilised by many structure prediction programmes (reviewed in chapter 3). All of them have four main components in common, a model for the molecular structure, a method for generating and searching for initial crystal structures and an intermolecular potential for evaluating the lattice energy using a minimisation procedure.

The majority of currently available prediction programmes based on this general concept competed in the first international blind test for crystal structure prediction of small organic molecules organised by the Cambridge Crystallographic Data Centre (CCDC) in May 1999.⁵ The crystal structures of three organic compounds had to be proposed by the workshop participants and the details of the study and the results for molecule I, 3-oxabicyclo(3.2.0)hepta-1,4-diene are presented in chapter 5. Overall none of the programmes, including methods⁶ based on statistical data derived from the experimental coordinates stored in the Cambridge Structural Database (CSD)⁷, gave consistently reliable results, though there were some correct predictions.

Two *major* components of the crystal structure prediction procedure, the molecular model and the intermolecular potential are crucial for the results of such studies. Even if the experimental data is available, an *ab initio* optimised molecular model is used for genuine *ab initio* crystal structure prediction in this work, to not bias the search towards one polymorphic structure. The influence of the variations of the molecular model, including an *ab initio* optimised molecular structure, on the static crystal structure lattice energy minima has been studied (chapter 6) in this thesis for the

polymorphic system of paracetamol (acetaminophen). This has been feasible as a variety of experimental crystal structure determinations (X-ray and neutron) at different temperatures are available for both polymorphic forms of this system.

A variety of theoretically based and empirically fitted potentials for modelling the intermolecular interactions in molecular crystal structures can be found in the literature (see chapter 2). Throughout this work, an accurate model for the intermolecular forces based on a Distributed Multipole Analysis (DMA) of an *ab initio* charge distribution combined with an empirical atom-atom repulsion-dispersion potential has been used. An optimised intermolecular potential for carboxylic acids which reproduces the crystal structures and the available heats of sublimation of several carboxylic acid structures has been derived and applied in a crystal structure prediction study of formic, benzoic and the polymorphic system of tetrolic acid (see chapter 7). As carboxylic acids in general show a variety of different packing motifs, based on chains and dimers, the generated low lattice energy minima give an indication for which compounds the chain and dimer motifs are thermodynamically competitive.

Exact figures for the frequency of occurrence of polymorphism, let alone of ‘disappearing polymorphs’⁸, among molecular crystal structures are not known, and there is still scientific dispute if polymorphism is a pervasive phenomenon⁴ or restricted to a few cases. However large the number of polymorphic molecular crystal structures, crystal structure prediction methods based on lattice energy minimisation will generally find *more* hypothetical low lattice energy structures within a few kJ mol^{-1} of the global minimum, with numbers very probably increasing with the thoroughness of the search⁹, than could be considered as potential polymorphs. This is shown by all prediction searches in this work. Being at or near the global lattice energy minimum is therefore a necessary, but not sufficient condition for a crystal structure to be experimentally observed.

In the prediction search for paracetamol (see chapter 6) both experimental structures were found as low energy forms, reproducing the order of thermodynamical stability. For this pharmaceutical drug, also the difference in mechanical behaviour of the experimentally known polymorphs has been investigated theoretically as the compression ability of a crystal form is a crucial physical property in tablet production.

In this work additional properties to thermodynamic criteria have been considered to reduce the number of hypothetical structures being suggested as predicted

polymorphs. The elastic constants and vapour growth morphologies (for an introduction into the calculation of elastic constants and morphologies, see chapter 4) have been calculated for the lowest lattice energy structures for paracetamol (chapter 6) and for the three carboxylic acids (chapter 7). The mechanical stability of a generally stable crystal can be very low which could enable an easy transformation into a related, more stable, crystal structure. The relative growth rates of the slowest growing faces for each of the low lattice energy crystal structures of one compound, can give some indication on the likelihood with which a structure is going to be observed experimentally.

Thus, in this thesis, current concepts in crystal structure prediction are critically investigated and additional approaches to thermodynamical methods are being developed.

References for chapter 0

- (1) Gavezzotti, A., *Acc. Chem. Res.*, **1994**, 27, 309-314. Are crystal structures predictable?
- (2) Bernstein, J. in *Organic Crystal Chemistry*, J.B. Garbarczyk and D.W. Jones, Oxford University Press, Oxford, **1991**, Polymorphism and the investigation of structure-property relations in organic solids.
- (3) Wöhler, F.; Liebig, J., *Am. Pharm.*, **1832**, 3, 249.
- (4) Halebian, J.K.; McCrone, W.C., *J. Pharm. Sci.*, **1969**, 58, 911-929. Pharmaceutical applications of polymorphism.
- (5) Lommerse, J.P.M.; Motherwell, W.D.S.; Ammon, H.L.; Dunitz, J.D.; Gavezzotti, A.; Hofmann, D.W.M.; Leusen, F.J.J.; Mooij, W.T.M.; Price, S.L.; Schweizer, B.; Schmidt, M.U.; Eijck, B.P.v.; Verwer, P.; Williams, D.E., *Acta Cryst.*, **2000**, B 56, 697-714. A test of crystal structure prediction of small organic molecules.
- (6) Motherwell, W.D.S., *Nova Acta Leopoldina NF*, **1999**, 79, 89-98. Crystal structure prediction and the Cambridge Structural Database.
- (7) Allen, F.H.; Davies, J.E.; Galloy, J.J.; Hohnson, O.; Kennard, O.; Macrae, C.F.; Mitchell, E.M.; Mitchell, G.F.; Smith, M.; Watson, D.G., *J. Chem. Inf. Comput. Sci.*, **1991**, 31, 187 - 204. The development of Versions 3 and 4 of the Cambridge Structural Database System.
- (8) Dunitz, J.D.; Bernstein, J., *Acc. Chem. Res.*, **1995**, 28, 193-200. Disappearing polymorphs.
- (9) Van Eijck, B.P., **2000**, Ab initio prediction of crystal structures, Plenary talk, ECM19, Nancy/France.

Chapter 1.

Molecular crystals and polymorphism

'Today, more than ever, the only practicable way to find details of particular structures or classes of structures is through the CSD for organic and organometallic structures (...). It is essential that these (...) compilations continue to be kept running smoothly and efficiently for the foreseeable future, otherwise the only people acquainted with any particular crystal structure will be the people who solved it – or the people for whom it was solved – and their friends.' ¹

This chapter gives a brief introduction into the theory of molecular crystals. This wide field of research combines very different scientific areas, such as solid-state thermodynamics, database studies and crystal engineering. One reason for the increasing interest into this group of crystal structures is the solid state phenomenon of polymorphism. The occurrence of different packing arrangements for the same molecule has proven to be a curse, though if understood and successfully exploited, it could be a blessing at the same time. But some of the basic foundations of crystallisation are still awaiting to be revealed and it is certainly not an exaggeration to say that *'...small –molecule organic crystal structures conceal the answer to some of the most challenging and fascinating questions in modern chemistry.'* ²

1.1 Molecular crystals

As a main criterion for a substance to belong to the group of molecular crystals, Kitaigorodskii ³ used the ratios between the inter- and intramolecular bond lengths. For molecular crystals, this difference is relatively large, compared to ionic crystals. Kitaigorodskii also included a few crystals formed by inorganic substances, as nitrogen and oxygen crystals, carbonyl and complex compounds of certain metals and he excluded organic salts from his definition of molecular crystals.

In the early days of organic X-ray crystallography, the main interest of studies was on the determination of the intramolecular forces: the molecular structure. The intermolecular packing patterns of *ionic* crystals had been described in the 1940s by a few simple principles as Pauling's rule of complementariness ⁴ and his set of van der Waals radii ⁵. And also for *molecular* crystals with their irregularly shaped molecules, the importance of intermolecular interactions, in particular the influence of hydrogen bonds, had been recognised at an early stage. But the importance of non-hydrogen bond close contacts was less well understood and their description very often either ignored or standardised. ² This lack of understanding can partly be attributed to the fact that the quantity of available molecular crystal structures was not large enough to draw general conclusions. It was only in the early 1960s that Kitaigorodskii put forward a method of rationalising molecular crystal structures with his theory of close packing ⁶ of molecules in crystals. His ideas reflect a subtle combination ² of different fields of chemistry, the essence of which is unfortunately often just known and cited as the '*The mutual arrangement of the molecules in a crystal is always such that the „projections“ of one molecule fit into the „hollows“ of adjacent molecules.*' ³ Early verifications ³ of this hypothesis of the avoidance of free space have been successfully carried out by the calculation of possible crystal structure packing patterns for a variety of molecules with the experimentally observed crystal structure being found among the highest density (most tightly packed) structures. As there were more close-packed crystal structures found than experimentally observed, this early work of calculating hypothetical high-density packing patterns already shows that close packing is a necessary, but not sufficient condition for crystallisation. Thermodynamics and kinetics have to be considered together for an attempt of a further understanding of the phenomenon 'solid state'.

1.2 The phenomenon of polymorphism

1.2.1 Discovery and the definition of polymorphism

Historically, the discovery of the phenomenon of polymorphism can get traced back to the end of the 18th century ⁷ when Klaproth ^{8, 9} in 1798 proposed that the minerals calcite and aragonite have the same chemical composition, CaCO_3 . Further work on this system has been reported by Thernard and Biot ¹⁰ and Mitscherlich. ¹¹ In the same paper Mitscherlich also published two different crystallographic forms for sodium phosphate. A year later, he discovered the allotropy (elements crystallising in more than one crystal form) of sulphur ¹². The first report of polymorphism in an organic compound was benzamide in 1832 by Wöhler and Liebig ¹³.

A large number of cases of polymorphism have become known, over the years, in both inorganic and organic systems. Many of them are listed in von Groth's compilation, ¹⁴ together with their physical properties, such as their melting behaviour. In 1942 Deffet ¹⁵ compiled the first documentation purely about polymorphic systems, including more than 1200 organic compounds. A later detailed collection of polymorphic crystal structures and their properties has been published by Kuhnert-Brandstätter. ¹⁶

Although polymorphism is a well-known phenomenon, it is difficult to define it accurately. A multitude of definitions has been developed, ¹⁷ often based on different experimental consequences of polymorphism. The implication of all of them is that polymorphs involve different packing arrangements of the same molecules in the solid.¹⁸ The widely accepted definition is that by McCrone. ¹⁹ He has defined a polymorph as a „solid crystalline phase of a given compound resulting from the possibility of at least two crystalline arrangements of the molecules of that compound in the solid state“ and has listed those types of solid phenomena, which are excluded from this definition. Later writers have accepted this description and have often substituted their own list of exclusions. ²⁰

The general problem in defining polymorphism arises from cases where the structure of the *molecule* is different in different polymorphic forms, such as tautomeric, zwitterionic, chiral structures and crystals with different conformers. Also the distinction between solvates, where solvent molecules form part of the crystal structure, and polymorphs is not clear cut and the often misleading term of

‘pseudopolymorphism’ has become common ²¹ particularly in the pharmaceutical industry. With the gradual discovery of the rich variety of the structures in the solid state, it was generally more and more difficult and also less sensible to maintain a narrow definition and it seems that even until today, the boundaries have not been clearly defined. ¹⁷

Despite all the work on polymorphic systems, there is still scientific dispute on the true extent of polymorphism among crystalline systems. It either being a pervasive phenomenon ²² with ‘*Every compound*’ having ‘*different polymorphic forms and the number of forms known for a given compound*’ being ‘*proportional to time and the energy spent in research on that compound.*’ ¹⁹ or restricted to a few cases, under unusual temperature or pressure conditions. Sarma and Desiraju ²³ have only identified 3.5 % of the entries in the CSD as polymorphic and Gavezzotti and Filippini ²⁴ only found 16 examples of molecules with three or more polymorphs in the CSD. This relatively small ratio is certainly related to the fact that crystallographers tend to choose the ‘best’ single crystal from a sample, rather than investigating the smaller crystallites. An extensive investigation of the thermal behaviour of organic crystals has shown that around one-third of organic substances show polymorphism under normal pressure and temperature conditions. ^{25, 26}

1.2.2 Conformational polymorphs

As pointed out before in this chapter, the *molecular* structure was the centre of attention in the early years of organic X-ray crystallography. The first report of an apparent interplay between the intra- and intermolecular forces inside a crystal, was given by Dunitz ²⁷ who observed a distortion from the expected molecular structure for tetraphenylcyclobutane. This early work exactly highlights the opportunity this phenomenon of conformational polymorphism ^{18, 28}, the existence of different molecular conformations in different polymorphic structures, offers. It provides an opportunity to study the influence of the crystal environment on the molecular conformation. ²⁹

Further examples have been published over the years, ³⁰⁻³² as the study on dichloro-benzylideneaniline, where the crystalline environment of the triclinic form stabilises the higher-intramolecular-energy planar molecular conformer. ³³

In contrast to this, the existence of non-conformational polymorphism provides a unique opportunity for the investigation of the relationship between the crystal structure and the physical and mechanical properties of a system, since the only variable between such polymorphic forms is that of structure of the crystal.

1.2.3 The importance of polymorphism

Even though the investigation and classification of all different polymorphic forms of one system for different temperatures and pressures represent a demanding goal, the study of polymorphs is not only of pure scientific interest. As the adopted crystal structure determines the properties of the material, such as density, taste, colour, conductivity, morphology, solubility, elasticity, melting point, chemical reactivity and optical properties, the determination of the structural packing arrangements within the substance is of great general interest. In particular, the *early* identification of possible polymorphs is of commercial significance for all molecular materials, including food, pharmaceuticals, pigments, electronic material and explosives. This is due to the fact that unexpected changes in the polymorphic form during manufacture and storage, *i.e.* the transformation to another structure *after* crystallisation, are likely to cause significant processing problems and compromise quality control as well as patents.

The commercial area in which the study of polymorphism has received most interest is the pharmaceutical industry because of the effect a structural transformation can have on the solubility and therefore on the bioavailability of a drug.³⁴ It has to be mentioned at this point, that Burger³⁵ suggested that the difference in solubility between polymorphs would in practice result in a significant bioavailability difference only in exceptional cases. Nevertheless, in 1998 the Abbott Laboratories withdrew its HIV drug, Ritonavir, because of the unexpected appearance of a new crystal form, which had different dissolution and absorption characteristics to the standard product.

In general, polymorphism is ubiquitous among pharmaceutical compounds³⁶⁻³⁸, and an area of enormous financial interest is the prolongation of the life of patents of drugs^{39, 40}. With the discovery of polymorphism, patents had to become more sophisticated, as illustrated in the most prominent case, Glaxo Wellcome's anti-ulcer drug, Zantac (ranitidine hydrochloride) with worldwide sales worth \$16000m⁴¹. A later discovery of an apparently second polymorph – which had in reality always been the commercial product - would have offered Glaxo Wellcome an extended patent

protection for a few more years. But a final court ruling gave Novopharm the right to produce and sell the first, originally patented but so far not commercially produced, polymorph.

Another field being dramatically influenced by the phenomenon of polymorphism is the explosives industry as a wrong polymorph can have greatly increased sensitivity to detonation.^{42, 43}

Similarly product performance is related to polymorphism, as the appearance and consumer perceived properties of products based on natural fats is very often dependent on the degree of crystallinity. The *metastable* polymorphs of fat crystals are regularly employed for the production of margarine/spreads, ice cream and chocolates⁴⁴, as they often display better dispersibility and melting behaviour at body temperature.

Organic pigments⁴⁵ are also affected, as the mode of molecular stacking clearly exerts a major influence on its colouristic properties. Although the optical properties of polymorphs have been extensively studied^{46, 47}, the specific relationship between the solid-state property, as for example the crystal color⁴⁶ and the solid state structure is not always clear. In addition, polymorphism can affect the macroscopic size and shape of crystals,⁴⁸ and hence influence the aggregation properties and dispersion performance of a pigment.

In 1961 Franken discovered that the frequency of red light from a ruby laser is exactly doubled when passed through a quartz crystal.⁴⁹ This observation of the generation of a 'second harmonic' has motivated a search for non-linear optical (NLO) materials for the application in optoelectronics.^{47, 50, 51} Most of the NLO substances are inorganic,⁵² but some organic materials show a NLO response that is several orders of magnitude higher than that observed for inorganic systems.^{53, 54} Organic molecules which contain polar groups with conjugated π -electron systems, and which crystallise in a non-centrosymmetric space group, are able to show this desirable effect. As the majority of organic substances crystallise in centrosymmetric space groups, various synthesis strategies have been developed to try to crystallise and stabilise non-centrosymmetric polymorphs with applicable properties.⁵²

And finally, even the materials in living organisms crystallise in specific polymorphic forms under selective controls, as observed in biological mineralisation.⁵⁵

Apart from all these applications, it is also important to press ahead with the work in this area to clarify the literature and to provide a good database of information. The increase of interest in polymorphism in recent years caused by pharmaceutical patent problems has brought to light how unsorted the information in this area is. There is in general no agreement for the designation of polymorphs. Numbering based on the order of the melting points would have to be changed as new polymorphic forms are discovered and a chronological order related to the date of detection would require a substantial knowledge of the history of each of the polymorphic forms. Because of these non-uniform classifications, it is unfortunately often the case that publications contain conflicting data, as the spectrum of one polymorph together with the phase transformation data of another form.¹⁷ For cimetidine, a crystal structure of an apparently novel polymorph was published⁵⁶ while the same polymorphic form had already been reported earlier⁵⁷.

The investigation of polymorphic systems, especially those with a large number of forms, provides the basis of an understanding of the intermolecular interactions and their influence on the crystal properties. This insight may help to design new molecular crystals for the use in very special applications.

1.2.4 Thermodynamic stability of polymorphs

In 1897 Ostwald⁵⁸ published his rule of stages in his ‘studies on the formation and transformation of solid phases’: *‘When leaving an unstable state, the form with the smallest free energy does not appear, rather the form which can be reached with the smallest possible loss of free energy, or the form with the next largest free energy.’* The relative equilibrium thermodynamical stability of different crystal packings at a given temperature and pressure is determined by their difference in Gibbs free energy

$$\Delta G = \Delta U + p\Delta V - T\Delta S \quad 1.1$$

with energy U , pressure p , volume V , temperature T , and entropy S .

Thus the relative stability depends on the differences in lattice energy, crystal density and entropy of the polymorphs. For cases at 0K the packing energy is the main criterion as the other two contributions can often be approximated as zero.²⁴ As the difference in density between polymorphs for small organic molecules rarely exceed a few

percent,²⁴ $p\Delta V$ is small and therefore negligible at normal pressure. At high pressures this term can be of significance, as illustrated in the example of benzene at 25 kbar.⁵⁹ The contribution of the vibrational entropy to the free enthalpy is negligible for strong covalent bonds, but as the entropy-energy compensation shows⁶⁰, it becomes important for weaker interactions. In general the influence of entropy differences to the relative thermodynamical stability of polymorphs have so far often been ignored because of the difficulty of reliable calculation.⁶¹ Nevertheless, in a recent case study for glycol and glycerol it was found that the entropy and zero-point energy give the largest contribution to free energy differences between hypothetical crystal structures, with up to 3 kJ/mol for the lowest energy structures.⁶²

Ostwald himself already reported many exceptions to his 'rule' and 100 years after his publication, at the BCA symposium in celebration of the centenary of Ostwald's rule of stages in Manchester, it is known that crystallisation does not only depend on thermodynamic factors. Each polymorph shows a unique nucleation and growth rate and in fact, little is known about the initial steps of crystal nucleation, the formation of a viable nucleus. In many cases of commercial interest the transformation process between different polymorphs is influenced by conditions like supersaturation or the solvent environment, with the simultaneous dissolution of the metastable form and growth of the stable form from solution.⁶³ In cases of conformational polymorphism, the crystallisation of the more stable crystal modification can be hindered when the molecular conformer present in the thermodynamically stable crystal is different from the most stable conformer in solution. Therefore, the growth of molecular crystals, also of metastable forms, can be controlled and inhibited by tailor-made additives⁶⁴ and by the presence of reaction by-products.⁴¹ And sometimes even multiple polymorphic forms nucleate in the same solution and mixtures of polymorphs crystallise simultaneously.^{65, 66}

1.2.5 Disappearing polymorphs

There are numerous examples⁶⁷ in the literature of polymorphs that appeared, following Ostwald's rule, at an early stage in the crystallisation process, and behaved respectably until the nucleation of a more stable form. After this, the previously obtained form, could not or just with difficulty, be crystallised again, often also

affecting distant laboratories.⁶⁸ There have been reports of successful attempts,^{69, 70} to recover a disappeared polymorph of dimethyl-benzylideneaniline by using a new laboratory.⁷¹ It is believed that any polymorph should be re-obtainable by finding the right experimental conditions.⁶⁷

As Bernstein pointed out,⁷² the idea that polymorphs can disappear is an anathema to crystal engineering and scientific principles of reproducibility.

1.3 Crystal engineering

The field of crystal engineering has been developed by structural chemists and crystallographers for the design of new materials and solid state reactions.⁷³ The term crystal engineering was initially introduced by Schmidt in the 1970s to address the problem of crystal structure prediction in the context of organic solid state photochemical dimerisation of *trans*-cinnamic acids, pioneering the topochemical approach to solid-state chemical reactions.⁷⁴ Following Lehn's analogy⁷⁵ that '*supermolecules are to molecules and the intermolecular bond what molecules are to atoms and the covalent bond*', crystal structures have been described as the '*supramolecule par excellence*'.²⁰ In this context the definition and aim of crystal engineering has been described by Desiraju in 1987⁷⁶ as *„the understanding of intermolecular interactions in the context of crystal packing and in the utilisation of such understanding in the design of new solids with desired physical and chemical properties“*.

The motifs or patterns of these interactions are described by synthons⁷⁷ and hydrogen bonds play a large role in these concepts of crystal engineering. But as the exceptional example of alloxan⁷⁸ illustrates, a molecule containing good hydrogen bond donors or acceptors may prefer a close packed crystal structure lacking hydrogen bonds over less dense structures with hydrogen bonds.

A major example for the interference of the concept of close packing in crystal engineering is the crystallisation of solid-state host guest complexes, clathrates.⁷⁹ The synthesis of these porous lattices often fails when a close packed structure⁸⁰ is realised by adopting a non-standard hydrogen bonding motif or when expected networks interpenetrate to avoid free space.⁸¹

Therefore, it is of no surprise that the rapidly growing interdisciplinary field of crystal engineering has to face the challenge that for some polymorphic systems, the same functional groups form totally different synthons in the various polymorphic forms.⁷⁷

1.4 Databases of molecular crystals and their properties

1.4.1 The Cambridge Crystallographic Database (CSD)

An enormous number of crystal structures have been solved since the introduction of X-ray diffraction into organic chemistry, unfortunately not all published. The Cambridge Crystallographic Data Centre (CCDC) keeps a record of the bibliographic, 2D chemical, and 3D structural results for published organic and organometallic compounds studied by X-ray and neutron diffraction. Its main product, the Cambridge Structural Database (CSD)⁸²⁻⁸⁶ has, together with software for search, retrieval, analysis, and display of the stored information, developed into *the* crystallographic data source with 224400 structural entries in the October 2000 release. It has become a scientific instrument for studying the empirical systematics of molecular and crystal structures.

The CSD system enables the analysis of intermolecular interactions and molecular packing arrangements by obtaining statistical data for any geometrical parameter in selected groups of structures. Correlations of different crystallographic, chemical and pharmaceutical properties as the density and the calculated packing energy were analysed by Gavezzotti and Filippini who retrieved and investigated 204 pairs of room temperature determined polymorphs.²⁴

A variety of different statistical studies show that a small number of popular space groups comprise the majority of molecular crystal structures and that these structures only contain one independent molecule in the unit cell.⁸⁷ Distributions based on early work⁸⁸ were later revised⁸⁹ introducing a statistical weight for chiral space groups. Wilson⁹⁰ was the first who actually employed the *CSD* for retrieving information on the relation between the number and kind of symmetry elements in a space group and its frequency of occurrence for organic crystal structures. He also analysed the importance of the number of molecules in the asymmetric unit.⁹¹ Many other statistical studies for organic systems followed^{92, 93} and one of the most recent

investigations ⁹⁴ showed that the nine top-ranking space groups, $P2_1/c$ (38.4 %), $P-1$ (20.1 %), $P2_12_12_1$ (10.6 %), $C2/c$ (7.4 %), $P2_1$ (5.8 %), $Pbca$ (4.3 %), $Pnma$ (1.4 %), $Pna2_1$ (1.1 %), $Pbcn$ (0.9 %), comprise 90% of all determined organic crystal structures and therefore confirmed the general outcome of previous studies. In their analysis of inorganic systems, Baur and Kastner ⁹⁵ found that there are much fewer unoccupied space groups for inorganic structures.

This dominance of a few space groups for organic molecules is frequently employed in crystal structure prediction studies to limit the number of space groups that have to be searched. Furthermore, the coordination spheres generated by the search program MOLPAK ⁹⁶ are based on common patterns found in the CSD.

1.4.2 The WebBook by the National Institute of Standards and Technology (NIST)

A very good web-based database (<http://webbook.nist.gov/>) is the experimental property collection by the U. S. National Institute of Standards and Technology (NIST). In particular the set of heats of sublimation, compiled by J. S. Chickos is very valuable for the study of molecular crystals. The heat of sublimation, the difference in energy between N molecules at infinite distance in the gas phase and N molecules having condensed, is often compared to the calculated lattice energy. But a direct comparison of these two values, although frequently used for the fitting of empirical intermolecular potential parameters, should not be done without the consideration of the inherent approximations. Sublimation enthalpies are measured at a certain temperature and calculated lattice energies correspond to 0 K, with some temperature effects being absorbed in empirical intermolecular potential parameters fitted to room temperature crystal structures. Therefore calculated lattice energies do not include contributions due to the molecular translations, vibrations or rotations at a certain temperature and they also do not consider conformational rearrangements of the molecules passing from the crystal to the gas phase. The experimental heats of sublimation include these effects and they are frequently measured at temperatures well above 300 K. Furthermore, the sublimation enthalpies also include the zero-point contribution to the lattice energy, which is not considered in the calculated lattice energies. The inherent error of this direct comparison has been estimated to about 8 kJ/mol ⁹⁷ and even up to 15 kJ/mol.⁹⁸

Unfortunately, experimental values for the heats of sublimation are normally published without any further reference to the crystal form employed, and oxalic acid seems to be the only polymorphic system for which sublimation enthalpies have been explicitly determined and are separately reported for each polymorph.⁹⁹

1.5 Tools for a theoretical comparison of molecular crystal structures and polymorphs

The theoretical study of polymorphism is fundamentally based on reliable tools with which structural similarities and differences between different polymorphic forms can be reproducibly detected and also be translated into an unambiguous, communicable mode. The following theoretical methods have been developed over the years, and they have proven to be most efficient when applied in combination with each other to allow for a more complete description of the structures. A variety of them have been employed in this thesis for the comparison of experimental polymorphs and also of hypothetical crystal structures generated in the prediction searches.

1.5.1 Visual comparison

The most figurative way of contrasting polymorphic crystal structures is by visual comparison. This is facilitated by choosing the same reference molecule and molecular reference plane for all polymorphs.²⁹ The graphical information is often analytically summarised in lists of ‘short’ intermolecular contacts.

1.5.2 Graph set theory

The graph set formalism, developed by Etter^{65, 100-104} provides a systematic method of characterising and analysing hydrogen bond patterns in crystals on the basis of topology.

The first attempt to classify networks of hydrogen bonds for crystal structures seems to have been that of Wells covering both inorganic and organic systems.¹⁰⁵ Kuleshova and Zorkii^{106, 107} recognised the correlation of these early classification schemes to the mathematical concept of graph theory. In their statistical analysis of the frequency of different patterns, as chains and layers, for polymorphic compounds, they showed that about half of their investigated crystal structures exhibited the same hydrogen-bonding patterns within all polymorphic forms of one system.

Etter ^{108, 109} adopted the symbols introduced by Kuleshova and Zorkii to create a nomenclature based on the combination of four simple motifs (chains, rings, intramolecular hydrogen-bonded patterns, and other finite patterns) to allow for a comparison of the number of hydrogen bond donors and acceptors and the number of bonds in the pattern. This encoding technique has been applied to the analysis of polymorphs, as the dimorphic system of L-glutamic acid ¹¹⁰ and the trimorphic system of iminodiacetic acid. ^{108, 111} It is generally found that the hydrogen-bond patterns differ among polymorphic forms of one compound, though the analysis must often be pursued to a higher level graph set by combinations of the basic motifs in order to recognise the distinction. ¹⁰⁹ Therefore, the graph set method also provides a way of evaluating the consequences of having different definitions of hydrogen bonds, as imposed by cutoff lengths or angles. ¹⁰⁴ The statistical analysis of functional group distributions for crystal structures with a particular graph set have resulted in the development of three hydrogen-bonding rules. ^{100, 102, 112}

Despite its limitations, as the inapplicability to interactions that cannot be considered as being of the donor-acceptor type and as defining the acceptor as a single atom, ¹¹³ the graph set theory is an efficient tool of analysing and visualising hydrogen bond patterns in crystal structures. Its use has recently been simplified by its incorporation into the CCDC software package. ^{129,130}

1.5.3 Reduced cells

In many reports of crystal structure prediction studies, the cell parameters are used to compare the experimental lattice with the structures found in the search procedure. But, although rules exist for obtaining a set of standard cell parameters for a given lattice, the conventional cell ^{114, 115}, the choice of cell parameters is not unique and the same lattice can be described by different sets of parameters. The concept of the reduced cell has become a standard tool in crystallography and a review of the different algorithms has been presented by Andrews and Bernstein ¹¹⁶. The *Bürger* reduced cell, ¹¹⁷ defined as the cell which uses as base vectors the three shortest non-coplanar vectors of the lattice, is not unique and a single lattice can be described by up to five *Bürger* reduced cells. ¹¹⁸ The unique *Niggli* reduced cell, ¹¹⁹ which has been

employed in this work, corresponds to one of these reduced cell representations and is equal to the B rger reduced cell if this is unique.

1.5.4 Similarity searches

Methods for an explicit investigation of the similarity of crystal structures have been developed. Based on a procedure by Dzyabchenko,¹²⁰ Van Eijck and Kroon developed a similarity search method for comparing cell parameters and structural fragments in different crystal structures.¹²¹ A method based on exploring the coordination sphere around a central molecule in a crystal structure¹²² has been applied in this thesis for the comparison of experimental crystal structures and also hypothetical structures generated in the prediction searches. Such an automated procedure is very efficient, especially when large numbers of crystal structures have to be compared. But for further insight into the *quality* of a possible similarity, additional comparison tools have to be considered.

1.5.5 X-ray powder diagrams

X-ray powder patterns are not only a powerful instrument for the determination of structures where X-ray single crystal diffraction is not applicable,¹²³ but as a ‘fingerprint’ of each crystal structure they also provide a further tool for comparing polymorphic forms. This method¹²⁴ has been implemented in the Polymorph Predictor program by MSI. It has to be mentioned at this point, that powder diffraction diagrams can be very similar for different polymorphic forms, as the example of terephthalic acid shows.⁴⁸

1.5.6 Hirshfeld surfaces

The molecular Hirshfeld surfaces are based on the division of electron density inside the crystal into molecular contributions by Hirshfeld’s stockholder partitioning scheme¹²⁵. Recent studies show that the 3D display of these molecular Hirshfeld surfaces offers a visual tool of comparing different polymorphic forms.¹²⁶⁻¹²⁸ For an example of a polymorphic system, oxalic acid¹²⁶, the surfaces are dramatically different, showing quite effectively the difference between packing modes utilised in the two structures. The Hirshfeld surface for the β polymorph shows evidence of strong O-

H \cdots O hydrogen bonds to form linear chains. Another polymorphic example studied by this method is p-dichlorobenzene.

1.6 Conclusions

Understanding the phenomenon of polymorphism in molecular crystal structures is a scientific challenge and the basis for the development and the production of molecular materials with reliable properties. For a more complete picture of the crystallisation process, experimental experience and theoretical knowledge have to be combined. Advances in experimental capabilities and an explicit experimental search for more polymorphic forms, not least motivated by the fact that the number of theoretical energetically feasible polymorphic structures exceeds the number of forms experimentally found, might help to reveal the true extent of polymorphism among crystal structures. For a further understanding of the intermolecular interactions present in crystalline systems and to estimate the energy differences, we need to exploit the theory of intermolecular forces present in crystalline systems (chapter 2).

References for chapter 1

- (1) Dunitz, J.D.; Gavezzotti, A., *Acc. Chem. Res.*, **1999**, 32, 677-684. Attractions and repulsions in molecular crystals: What can be learned from the crystal structures of condensed ring aromatic hydrocarbons?
- (2) Gavezzotti, A., *Crystallography Reviews*, **1998**, 7, 5-121. The crystal packing of organic molecules: challenge and fascination below 1000 Da.
- (3) Kitaigorodskii, A.I., *Molecular crystals and molecules*, **1973**, New York: Academic Press.
- (4) Pauling, L.; Delbrück, M., *Science*, **1940**, 92, 77-79. The nature of the intermolecular forces operative in biological processes.
- (5) Pauling, L., *The nature of the chemical bond*, **1940**, Ithaca, N. Y.: Cornell University Press.
- (6) Kitaigorodskii, A.I., *Organic chemical crystallography*, **1961**, New York: Consultants Bureau.
- (7) Glusker, J. P., *Crystal structure analysis for chemists and biologists*, **1994**, New York: VCH.
- (8) Mallard, E., *Ann. Mines.*, **1876**, 10, 60 - 196. Explication des phenomenes optiques anomaux qui presentent un grand nombre des substances cristallisees. (Explanations of the anomalous optical phenomena shown by many crystalline materials).
- (9) Arzruni, A., *Physikalische Chemie der Kristalle*, **1893**, Braunschweig: Vieweg.
- (10) Thernard, L.J.; Biot, J.B., *Mem. Phys. II.*, **1809**, *Soc. d' Arcueil* 2, 176 - 206. Memoire sur l'analyse comparee de l'arragonite et du carbonate de chaux rhomboidal. (On the comparative analyses of aragonite and rhombohedral calcium carbonate.).
- (11) Mitscherlich, E., *Ann. Chim. Phys.*, **1822**, 19, 350-419. Sur la relation qui existe entre la forme cristalline et les proportions chimiques. I. Memoire sure les arsenitates et les phosphates. (On the relationship between crystalline form and chimical composition. I. Note on arsenates and phosphates).
- (12) Mitscherlich, E., *Berl. Akad. Abhand.*, **1822**.
- (13) Wöhler, F.; Liebig, J., *Am. Pharm.*, **1832**, 3, 249.
- (14) Von Groth, P.H.R., *Chemische Kristallographie*, **1906-1919**, Leipzig.

- (15) Deffet, L., *Repertoire des composes organiques polymorphes*, **1942**, Liege: Dosoer.
- (16) Kuhnert-Brandstätter, M., *Thermomicroscopy in the analysis of pharmaceuticals*, **1971**, New York: Pergamon.
- (17) Threlfall, T.L., *Analyst*, **1995**, 1230, 2435-2460. Analysis of organic polymorphs.
- (18) Bernstein, J. in *Organic Crystal Chemistry*, J.B. Garbarczyk and D.W. Jones, Oxford University Press, Oxford, **1991**, 6-26. Polymorphism and the investigation of structure-property relations in organic solids.
- (19) McCrone, W.C. in *Physics and Chemistry of the Organic Solid State*, D. Fox, M.M. Labes, and A. Weissberger, Interscience, New York, **1965**, 725-767.
- (20) Dunitz, J.D., *Pure & Appl. Chem.*, **1991**, 63, 177-185. Phase transitions in molecular crystals from a chemical viewpoint.
- (21) David, R.; Giron, D., *Handbook of Powder Technology*, **1994**, 9, 193.
- (22) Haleblan, J.; McCrone, W., *J. Pharm. Sci.*, **1969**, 58, 911-929. Pharmaceutical applications of polymorphism.
- (23) Sharma, J.A.R.P.; Desiraju, G.R. in *Crystal engineering*, K.R. Seddon and M.J. Zaworothko, Kluwer, Dordrecht, **1998**, The design and application of functional solids.
- (24) Gavezzotti, A.; Filippini, G., *J. Am. Chem. Soc.*, **1995**, 117, 12299-12305. Polymorphic forms of organic crystals at room conditions: thermodynamic and structural implications.
- (25) Kuhnert-Brandstätter, M.; Riedmann, M., *Microchim. Acta*, **1987**, II, 107-120. Thermal analytical and infrared spectroscopic investigations on polymorphic organic compounds-I.
- (26) Bürger, A., *Acta Pharm. Tech.*, **1979**, Suppl. 7, 107-112. Die polymorphen Arzneistoffe des Europäischen Arzneibuches.
- (27) Dunitz, J.D., *Acta Cryst.*, **1949**, 2, 1-13. The structure of the centrosymmetric isomer of 1:2:3:4 - tetraphenylcyclobutane.
- (28) Bernstein, J. in *Organic Solid State*, G. Desiraju, Elsevier, Amsterdam, **1987**, 471-518. Conformational polymorphism.
- (29) Bernstein, J., *J. Phys. D Appl. Phys.*, **1993**, 26, B66-B76. Crystal growth, polymorphism and structure-property relationships in organic crystals.
- (30) Hagler, A.T.; Bernstein, J., *J. Am. Chem. Soc.*, **1978**, 100, 6349.

- (31) Bernstein, J.; Bar, I., *J. Phys. Chem.*, **1984**, 88, 243-248. Conformational polymorphism. 5. Crystal energetics of an isomorphic system including disorder.
- (32) Bar, I.; Bernstein, J., *Tetrahedron*, **1987**, 43, 1299-1305. Modification of crystal packing and molecular-conformation via systematic substitution.
- (33) Bernstein, J.; Hagler, A.T., *J. Am. Chem. Soc.*, **1978**, 100, 673-681. Conformational polymorphism. The influence of crystal structure on molecular conformation.
- (34) Yokoyama, T.; Umeda, T.; Kuroda, K.; Kuroda, T.; Asada, S., *Chem. Pharm. Bull.*, **1981**, 29, 194-199. Studies on drug nonequivalence. 10. Bioavailability of 6-mercaptopurine polymorphs.
- (35) Burger, A. in *Topics in Pharmaceutical Sciences*, D.D. Breimed and P. Speiser, Elsevier, **1983**, The relevance of polymorphism.
- (36) Haleblan, J.K., *J. Pharm. Sci.*, **1975**, 64, 1269-1288.
- (37) Borka, L.; Haleblan, J.K., *Acta Pharm. Jugosl.*, **1990**, 40, 71-94. Crystal polymorphism of pharmaceuticals.
- (38) Byrn, S.R., *Solid state chemistry of drugs*, **1982**, New York: Academic Press.
- (39) Cholerton, T.J., *J. Chem. Soc., Perkin Trans.*, **1984**, 2, 1761-1766. Spectroscopic studies on ranitidine - its structure and the influence of temperature and pH.
- (40) DeCamp, W.H. in *Proceedings of the 3rd International Workshop on Crystal Growth of Organic Materials.*, A.S. Myerson, D.A. Green, and P. Meenan, ACS Series, Washington DC, **1996**, 66-71. Regulatory considerations in crystallisation processes for bulk pharmaceutical industry. A reviewer's perspective.
- (41) Blagden, N.; Davey, R., *Chemistry in Britain*, **1999**, 3, 44-47. Polymorphs take shape.
- (42) Kohlbeck, J.A., *Microscope*, **1982**, 30, 249-257. Spindle stage studies on HMX and AP.
- (43) Karpowicz, R.J.; Sergio, S.T.; Brill, T.B., *I & EC Prod. Res. Dev.*, **1983**, 22, 363.
- (44) Walstra, P., *Food structure and behaviour*, **1987**, New York: Academic.
- (45) Kuhnert-Brandstaetter, M., *Microchimica Acta*, **1989**, I, 373-385. Thermal analytical and infrared spectroscopic investigations on polymorphic organic compounds - III.

- (46) Toma, P.H.; Kelley, M.P.; Borchardt, T.B.; Byrn, S.R.; Kahr, B., *Chem. Mater.*, **1994**, 6, 1317-1324. Chromoisomers and polymorphs of 9-phenylacridinium hydrogen sulfate.
- (47) Serbbutoviez, C.; Nicoud, J.-F.; Fischer, J.; Ledoux, I.; Zyss, J., *Chem. Mater.*, **1994**, 6, 1358-1368. Crystalline zwitterionic stilbazolium derivatives with large quadratic optical nonlinearities.
- (48) Davey, R.J.; Maginn, S.J.; Andrews, S.J.; Black, S.N.; Buckley, A.M.; Cottler, D.; Dempsey, P.; Plowman, R.; Rout, J.E.; Stanley, D.R.; Taylor, A., *J. Chem. Soc., Faraday Trans.*, **1994**, 90, 1003-1009. Morphology and polymorphism in molecular crystals: Terephthalic acid.
- (49) Franken, P.A.; Hill, A.E.; Peters, C.W.; Weinreich, G., *Phys. Rev. Lett.*, **1961**, 7, 118-119. Generation of optical harmonics.
- (50) Zyss, J., *J. Phys. D: Appl. Phys.*, **1993**, 26, B198-B207. Engineering new organic crystals for nonlinear optics: from molecules to oscillator.
- (51) Xu, D.; Yuan, D.-R.; Zhang, N.; Hou, W.-B.; Liu, M.-G.; Sun, S.-Y.; Jiang, M.-H., *J. Phys. D: Appl. Phys.*, **1993**, 26, B230-B235. Study of properties and structural features of some new organic and organometallic nonlinear optical crystals.
- (52) Lacroix, P.G.; Clément, R.; Nakatani, K.; Zyss, J.; Ledoux, I., *Science*, **1994**, 263, 658-660. Stilbazolium-MPS₃ nanocomposites with large second-order optical nonlinearity and permanent magnetization.
- (53) Eaton, D.F., *Science*, **1991**, 253, 281-287. Nonlinear optical materials.
- (54) Zyss, J.; Ledoux, I.; Nicoud, J.F. in *Molecular Nonlinear Optics*, J. Zyss, Academic Press, Boston, **1993**.
- (55) Addadi, L.; Weiner, S., *Angew. Chem. Int. Ed. Eng.*, **1992**, 31, 153-169. Control and design principles in biological mineralization.
- (56) Parkanyi, L.; Kalman, A.; Hegedus, B.; Harsanyi, K.; Kreidl, J., *Acta Cryst.*, **1984**, C40, 676-679. Structure of a novel and reproducible polymorph (Z) of the histamine H₂-receptor antagonist cimetidine, C₁₀H₁₆N₆S.
- (57) Shibata, M.; Kokubo, H.; Morimoto, K.; Morisaka, K.; Ishida, T.; Inoue, M., *J. Pharm. Sci.*, **1983**, 72, 1436-1442. X-ray structural studies and physicochemical properties of cimetidine polymorphism.
- (58) Ostwald, *Z. f. Phys. Chem.*, **1897**, 22, 306.

- (59) Gibson, K.D.; Scheraga, H.A., *J. Phys. Chem.*, **1995**, *99*, 3765-3773. Crystal packing without symmetry constraints. 2. Possible crystal packings of benzene obtained by energy minimization from multiple starts.
- (60) Dunitz, J.D., *Chemistry & Biology*, **1995**, *2*, 709-712. Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions.
- (61) Filippini, G.; Gramaccioli, C.M., *Acta Cryst.*, **1986**, *B42*, 605-609. Thermal motion analysis in tetraphenylmethane: a lattice-dynamical approach.
- (62) Van Eijck, B.P., **2001**, manuscript in preparation. Ab initio crystal structure predictions for flexible hydrogen-bonded molecules. Part III. Effect of lattice vibrations.
- (63) Cardew, P.T.; Davey, R.J., *Proc. Roy. Soc. London*, **1985**, *A 393*, 415-428. The kinetics of solvent-mediated phase-transformations.
- (64) Weissbuch, I.; Popovitz-Biro, R.; Leiserowitz, L.; Lahav, M. in *The Lock-and-Key Principle*, J.-P. Behr, Chichester, **1994**, 173-246. Lock-and-key processes at crystalline interfaces: Relevance to the spontaneous generation of chirality.
- (65) Etter, M.C.; Kress, R.B.; Bernstein, J.; Cash, D.J., *J. Am. Chem. Soc.*, **1984**, *106*, 6921-6927. Solid-state chemistry and structures of a new class of mixed dyes.
- (66) Aakeroy, C.B.; Nieuwenhuyzen, M.; Price, S.L., *J. Am. Chem. Soc.*, **1998**, *120*, 8986-8993. Three polymorphs of 2-Amino-5-nitropyrimidine: Experimental structures and theoretical predictions.
- (67) Dunitz, J.D.; Bernstein, J., *Acc. Chem. Res.*, **1995**, *28*, 193-200. Disappearing polymorphs.
- (68) Woodard, G.D.; McCrone, W.C., *J. Appl. Cryst.*, **1975**, *8*, 342. Unusual crystallization behaviour.
- (69) Catti, M.; Ferraris, G., *Acta Cryst.*, **1976**, *B32*, 359-369. Hydrogen bonding in the crystalline state. Structure of $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ (orthorhombic phase), and crystal chemistry of the $\text{NaH}_2\text{PO}_4 \cdot n\text{H}_2\text{O}$ series.
- (70) Czugler, M.; Kalman, A.; Kovacs, J.; Pinter, I., *Acta Cryst.*, **1981**, *B37*, 172-177. Structure of the unstable monoclinic 1,2,3,5-tetra-O-acetyl-beta-D-ribofuranose.
- (71) Bar, I.; Bernstein, J., *Acta Cryst.*, **1982**, *B38*, 121-125. Molecular conformation and electronic structure. VI. The structure of p-Methyl-N-(p-methylbenzylidene)aniline (form I).

- (72) Bernstein, J. *Disappearing and reappearing polymorphs*. in *Polymorphism in Molecular Crystals: 100 years of Ostwald's Rule*. 1997. Manchester.
- (73) Desiraju, G.R., *Angew. Chem. Int. Ed. Engl.*, **1995**, *34*, 2311-2327. Supramolecular synthons in crystal engineering - A new organic synthesis.
- (74) Schmidt, G.M.J., *Pure Appl. Chem.*, **1971**, *27*, 647-678. Photodimerization in the solid state.
- (75) Lehn, J.M., *Angew. Chem.*, **1988**, *100*, 91-116.
- (76) Desiraju, G.R. in *Organic Solid State Chemistry*, G.R. Desiraju, **1987**, 471-518.
- (77) Nangia, A.; Desiraju, G.R., *Topics in current chemistry*, **1998**, *198*, 57-95. Supramolecular synthons and pattern recognition.
- (78) Bolton, W., *Acta Cryst.*, **1964**, *17*, 147-152. Crystal structure of alloxan.
- (79) Russell, V.A.; Evans, C.C.; Li, W.; Ward, M.D., *Science*, **1997**, *276*, 575-579. Nanoporous molecular sandwiches: Pillared two-dimensional hydrogen-bonded networks with adjustable porosity.
- (80) Kolotuchin, S.V.; Fenlon, E.E.; Wilson, S.R.; Loweth, C.J.; Zimmermann, S.C., *Ang. Chem. Int. Ed. Eng.*, **1995**, *34*, 2654-2657. Self-assembly of 1,3,5-benzenetricarboxylic (trimesic) acid in the solid state.
- (81) Ermer, O., *J. Am. Chem. Soc.*, **1988**, *110*, 3747-3754. Fivefold-diamond structure of adamantane-1,3,5,7-tetracarboxylic acid.
- (82) Kennard, O.; Watson, D.G.; Allen, F.H.; Motherwell, W.S.D.; Town, W.; Rodgers, J., *Chem. Britain*, **1975**, *11*, 213-216. Crystal clear data.
- (83) Allen, F.H.; Bellard, S.; Brice, M.D.; Cartwright, C.A.; Doubleday, A.; Higgs, H.; Hummelink, T.; Hummelink-Peters, B.J.; Kennard, O.; Motherwell, W.D.S.; Rodgers, J.R.; Watson, D.G., *Acta Cryst.*, **1979**, *B35*, 2331-2339. The Cambridge Crystallographic Data Centre: Computer-based search, retrieval, analysis and display of information.
- (84) Allen, F.H.; Kennard, O.; Taylor, R., *Acc. Chem. Res.*, **1983**, *16*, 146-153. Systematic analysis of structural data as a research technique in organic chemistry.
- (85) Allen, F.H.; Davies, J.E.; Galloy, J.J.; Hohnson, O.; Kennard, O.; Macrae, C.F.; Mitchell, E.M.; Mitchell, G.F.; Smith, M.; Watson, D.G., *J. Chem. Inf. Comput. Sci.*, **1991**, *31*, 187 - 204. The development of Versions 3 and 4 of the Cambridge Structural Database System.
- (86) Allen, F.H.; Kennard, O., *Chem. Des. Autom. News*, **1993**, *8*, 1631.

- (87) Belsky, V.K.; Zorkii, P.M., *Acta Cryst.*, **1977**, A33, 1004-1006. Distribution of organic homomolecular crystals by chiral types and structural classes.
- (88) Mighell, A.D.; Himes, V.L., *Acta Cryst.*, **1983**, A39, 737-740. Space-group frequencies for organic compounds.
- (89) Donohue, J., *Acta Cryst.*, **1985**, A41, 203-204. Revised space-group frequencies for organic compounds.
- (90) Wilson, A.J.C., *Acta Cryst.*, **1988**, A44, 715-724. Space groups rare for organic structures. I. Triclinic, monoclinic and orthorhombic crystal classes.
- (91) Wilson, A.J.C., *Z. Kristallogr.*, **1991**, 197, 85-88. Space-groups rare for molecular organic structures - the arithmetic crystal class MMMP.
- (92) Padmaja, N.; Ramakumar, S.; Viswamitra, M.A., *Acta Cryst.*, **1990**, A 46, 725-730. Space-group frequencies of proteins and of organic compounds with more than one formula unit in the asymmetric unit.
- (93) Wilson, A.J.C., *Acta Cryst.*, **1993**, A49, 795-806. Space groups rare for organic structures. III. Symmorphisms and inherent molecular symmetry.
- (94) Brock, C.P.; Dunitz, J.D., *Chem. Mater.*, **1994**, 6, 1118-1127. Towards a grammar of crystal packing.
- (95) Baur, W.H.; Kassner, D., *Acta Cryst.*, **1992**, B48, 356-369. The perils of Cc: Comparing the frequencies of falsely assigned space groups with their general population.
- (96) Holden, J.R.; Du, Z.; Ammon, H.L., *J. Comp. Chem.*, **1993**, 14, 422-437. Prediction of possible crystal structures for C-, H-, N-, O- and F-containing organic compounds.
- (97) Gavezzotti, A. in *Theoretical aspects of computer modelling of the molecular solid state*, A. Gavezzotti, Wiley, Chichester, **1997**, 61-97. Energetic aspects of crystal packing: Experiment and computer simulations.
- (98) Pertsin, A.J.; Kitaigorodsky, A.I., *The atom-atom potential method*, **1987**, Berlin: Springer-Verlag.
- (99) Bernstein, J., **1998**, *personal communication*.
- (100) Etter, M.C., *J. Am. Chem. Soc.*, **1982**, 104, 1095-1096. A new role for hydrogen-bond acceptors in influencing packing patterns of carboxylic acids and amides.
- (101) Etter, M.C., *Isr. J. Chem.*, **1985**, 25, 312-319. Aggregate structures of carboxylic acids and amides.

- (102) Etter, M.C., *Acc. Chem. Res.*, **1990**, *23*, 120-126. Encoding and decoding hydrogen-bond patterns of organic compounds.
- (103) Etter, M.C.; MacDonald; Bernstein, J., *Acta Cryst.*, **1990**, *B46*, 256-262. Graph-set analysis of hydrogen-bond patterns in organic crystals.
- (104) Etter, M.C., *J. Phys. Chem.*, **1991**, *95*, 4601-4610. Hydrogen bonds as design elements in organic chemistry.
- (105) Wells, A.F., *Structural inorganic chemistry*, **1962**, Oxford: The Clarendon Press.
- (106) Kuleshova, L.N.; Zorkii, P.M., *Acta Cryst.*, **1980**, *B36*, 2113-2115. Graphical enumeration of hydrogen-bonded structures.
- (107) Zorkii, P.M.; Kuleshova, L.N., *Zh. Strukt. Khim.*, **1980**, *22*, 153.
- (108) Bernstein, J.; Etter, M.C.; MacDonald, J.C., *J. Chem. Soc. Perkin Trans. 2*, **1990**, 695-698. Decoding hydrogen-bond patterns. The case of iminodiacetic acid.
- (109) Bernstein, J.; Davis, R.E.; Shimoni, L.; Chang, N.-L., *Angew. Chem. Int. Ed. Engl.*, **1995**, *34*, 1555-1573. Patterns in hydrogen bonding: Functionality and graph set analysis in crystals.
- (110) Bernstein, J., *Acta Cryst.*, **1991**, *B47*, 1004-1010. Polymorphism of L-Glutamic Acid: Decoding the alpha-beta phase relationship via graph-set analysis.
- (111) Blagden, N.; Davey, R.J.; Liebermann, H.F.; Williams, L.; Payne, R.; Roberts, R.; Docherty, R., *J. Chem. Soc. Faraday Trans.*, **1998**, *94*, 1919. Crystal chemistry and solvent effects in polymorphic systems sulfathiazole.
- (112) Donohue, J., *J. Phys. Chem.*, **1952**, *56*, 502-510.
- (113) Subramanian, K.; Lakshmi, S.; Rajagopalan, K.; Koellner, G.; Steiner, T., *J. Mol. Struct.*, **1996**, *384*, 121. Cooperative hydrogen bond cycles involving O-H center dot center dot center dot pi and C-H center dot center dot center dot O hydrogen bonds as found in a hydrated dialkyne.
- (114) Hahn, T.E., *International tables for crystallography*, **1983**, Dordrecht: Reidel.
- (115) Le Page, Y., *J. Appl. Cryst.*, **1982**, *15*, 255-259. The derivation of the axes of the conventional unit cell from the dimensions of the B rger-reduced cell.
- (116) Andrews, L.C.; Bernstein, H.J., *Acta Cryst.*, **1988**, *A44*, 1009-1018. Lattices and reduced cells as points in 6-space and selection of Bravais lattice type by projections.
- (117) B rger, M.J., *Z. Kristallogr.*, **1957**, *109*, 42-60.

- (118) Gruber, B., *Acta Cryst.*, **1973**, A29, 433-440. The relationship between reduced cells in a general Bravais lattice.
- (119) Krivy, I.; Gruber, B., *Acta Cryst.*, **1976**, A32, 297-298. A unified algorithm for determining the reduced (Niggli) cell.
- (120) Dzyabchenko, A.V., *Acta Cryst.*, **1994**, B50, 414-425. Method of crystal structure similarity searching.
- (121) Van Eijck, B.P.; Kroon, J., *J. Comp. Chem.*, **1997**, 18, 1036-1042. Fast clustering of equivalent structures in crystal structure prediction.
- (122) Lommerse, J.P.M., *J. Appl. Cryst.*, **in preparation**, An index for crystal structure similarity.
- (123) Harris, K.D.M.; Tremayne, M., *Chem. Mater.*, **1996**, 8, 2554-2570. Crystal structure determination from powder diffraction data.
- (124) Karfunkel, H.R.; Rohde, B.; Leussen, F.J.J.; Gdanitz, R.J.; Rihs, G., *J. Comp. Chem.*, **1993**, 14, 1125-1135. Continuous similarity measure between nonoverlapping X-ray powder diagrams of different crystal modifications.
- (125) Hirshfeld, *Theor. Chim. Acta*, **1977**, 44, 129.
- (126) Spackman, M.A.; Byrom, P.G., *Chem. Phys. Lett.*, **1997**, 267, 215-220. A novel definition of a molecule in a crystal.
- (127) McKinnon, J.J.; Mitchell, A.S.; Spackman, M.A., *Chem. Eur. J.*, **1998**, 4, 2136-2141. Hirshfeld surfaces: A new tool for visualising and exploring molecular crystals.
- (128) McKinnon, J.J.; Mitchell, A.S.; Spackman, M.A., *Chem. Commun.*, **1998**, 2071-2072. Visualising intermolecular interactions in crystals: naphthalene vs. terephthalic acid.
- (129) Motherwell, W. D. S.; Shields, G. P.; Allen, F. H., *Acta Cryst.*, **1999**, B55, 1044-1056. Visualisation and characterisation of non-covalent networks in molecular crystals: automated assignment of graph set descriptors for asymmetric molecules.
- (130) Motherwell, W. D. S.; Shields, G. P.; Allen, F. H., *Acta Cryst.*, **2000**, B56, 466-473. Visualisation and characterisation of non-covalent networks in molecular crystals: automated assignment of graph set descriptors for symmetric molecules.

Chapter 2

The theory of intermolecular forces

*‘The earth’s crust may be held together mainly by ionic forces, molecules by covalent bonds, but it is weak intermolecular interactions that hold us, along with the rest of the organic world, together.’*¹

A. Gavezzotti stated in his recent review² about crystal packing of organic molecules that *“The crystal packing of pure and simple organic molecules is a vocabulary and a primer in the language of intermolecular potentials; the geometrical constitution of such crystals is easy to determine, but is to a large extent unexplained, since that language is still a cryptic one. What are the theoretical foundations of crystal packing?”* In this chapter this question will be approached by a brief introduction into the theory of intermolecular forces³ and a presentation of the state of the art of modelling molecular crystal structures.

2.1 Introduction

The forces $F(R)$ between interacting molecules can be expressed by the intermolecular potential energy $U(R)$:

$$dU(R) = -F(R)dR \quad 2.1$$

Here the interaction energy is just a function of the intermolecular separation R . But in almost all cases the intermolecular potential between two molecules will not only depend on their separation, but also on their relative orientation. Changes in the energy of the system can therefore be considered as movements on a multidimensional potential energy surface. Even for small systems, like two non-linear molecules, such an energy surface becomes very difficult to calculate. Six coordinates have to be used to describe the relative orientation of the two molecular local axes systems with respect to each other: e. g. three for the description of the separation vector R and three angles (usually the Euler angles ⁴ α, β, γ). To obtain this intermolecular potential by *ab initio* calculations is computationally very demanding as high-quality wavefunctions with extended basis sets and a high degree of electron correlation are required for each point calculated and can only be achieved for small systems like ArHF. ⁵ Many points are necessary to describe the surfaces of even fairly small systems accurately, as exemplified by a potential surface for the water dimer, which was evaluated at 20480 points ⁶. This study was feasible as density functional theory (DFT) calculations have been used, but it would have been impossible to apply a more elaborate *ab initio* method. Therefore, the development of intermolecular model potentials, the mathematical approximation of the true interaction within a system, is necessary for the description of even fairly simple systems and it is the only feasible way of calculating larger systems.

The total intermolecular interaction energy of any ensemble of N rigid molecules (e.g. the lattice energy for a molecular crystal structure or the potential energy of a liquid, as N becomes very large) relative to their energy when completely separated, can be written as

$$\begin{aligned}
U(\mathbf{R}, \mathbf{\Omega}) = & \sum_{A < B}^N U_{AB}(\mathbf{R}_{AB}, \mathbf{\Omega}_{AB}) + \sum_{A < B < C}^N U_{ABC}(\mathbf{R}_{AB}, \mathbf{R}_{AC}, \mathbf{R}_{BC}, \mathbf{\Omega}_{ABC}) \\
& + \sum_{A < B < C < D}^N U_{ABCD}(\mathbf{R}_{AB}, \mathbf{R}_{AC}, \mathbf{R}_{AD}, \mathbf{R}_{BC}, \mathbf{R}_{BD}, \mathbf{R}_{CD}, \mathbf{\Omega}_{ABCD}) + \dots
\end{aligned} \tag{2.2}$$

$U(\mathbf{R}, \mathbf{\Omega})$ is a scalar quantity and it does not depend on the orientation of the molecules in isotropic space relative to the global axis system. The first term in the intermolecular pair potential is the sum of all pair (two-body) interactions with each being a function of the separation \mathbf{R}_{AB} and relative orientation $\mathbf{\Omega}_{AB}$ of the two molecules A, B considered. Each two-body term represents the difference between the energy of a pair of molecules and their energy when completely separated. The next term (the three-body term, three-body correction) is the difference between the actual interaction energy of three molecules (A, B, C) in a given orientation and the sum of the three pair potential terms ($U_{AB} + U_{BC} + U_{AC}$). The following term is the corresponding correction to give the correct interaction energy of a quartet of molecules and so on.

As these non-additive many-body terms can only be accurately calculated for a few small systems, it is often practically assumed that they become smaller with increasing order and therefore that this series rapidly converges. A prominent example in this respect is argon. An intermolecular potential for this system⁷ needed only the addition of a three-body term to satisfactorily account for the liquid and solid state properties. In most crystal structure modelling, due to the size of the systems, only the first term in this expansion is considered. An intermolecular pair potential calculated by empirically derived pair potential parameters will have adsorbed some indefinite average over the many body forces. If the true pair potential is available and used in the calculation of the properties of a crystal, there is an error due to the omission of the non-additive many body contributions, which can significantly differ between types of molecules and crystal systems.

2.2 Classification of the forces between molecules

The intermolecular forces are commonly divided into two main types: 'long range', where the interaction energy behaves as some inverse power of \mathbf{R} , and 'short range', where the energy decreases approximately exponentially with distance. This

distinction is mainly based on whether the overlap of the charge distributions of the molecules becomes significant (short range) or can be considered as negligible (long range). The contributions to both types of intermolecular interactions, long range (electrostatic, induction, dispersion) and short range (exchange-repulsion and modifications of the long range terms arising from the overlap of the wavefunctions, as the charge-transfer, penetration and damping) forces are described below (chapter 2.2.1 – 2.2.2).

The intermolecular forces are generally relatively weak compared to the intramolecular interactions. It is therefore possible to describe these forces by Perturbation Theory.

2.2.1 Long range contributions

For molecules being separated by a relatively large distance, the overlap between their wavefunctions can be approximately ignored (the overlap is never exactly zero, but decreases exponentially with increasing distance between the molecules). Therefore all electron density can be rigorously assigned to one or the other molecule. This simplifies the mathematical approach in that sense that all terms which represent electron exchange between the molecules can be neglected and therefore the calculation of the overall wavefunction for the complex can be done without antisymmetrisation. The long range perturbation theory has first been discussed by London ⁸ and has been re-formulated by others since then.⁹⁻¹³

In the Rayleigh-Schrödinger Perturbation Theory, the infinitely separated molecules *A* and *B* with associated groundstate wavefunctions (Ψ_A^m , $m = 0$ and Ψ_B^n , $n = 0$), have Hamiltonians H_A and H_B which add up to the unperturbed Hamiltonian of the system: $H^0 = H_A + H_B$. The energy of zeroth order, U_{AB}^0 of such a system is therefore equal to the sum of the monomer energies E_A^0 and E_B^0 :

$$U_{AB}^0 = \langle \Psi_A^0 | H_A | \Psi_A^0 \rangle + \langle \Psi_B^0 | H_B | \Psi_B^0 \rangle = E_A^0 + E_B^0 \quad 2.3$$

All the important long range contributions to the intermolecular forces between molecules arise ultimately from the electrostatic (Coulombic) interaction of the charge distributions between the particles (electrons and nuclei) which make up these

molecules. The difference V (the perturbation operator) between the unperturbed Hamiltonian H^0 and the one for the perturbed system H can therefore be defined as

$$V = \frac{1}{4\pi\epsilon_0} \sum_{ij} \frac{e_i^A e_j^B}{r_{ij}} \quad 2.4$$

In this equation r_{ij} is the inter-particle separation between particles i and j in the different molecules, and the charge e_i^A refers to the i^{th} particle (electron or nucleus) associated with molecule A and e_j^B to j on B . ϵ_0 is the permittivity of vacuum. The interaction energy can be expanded in the series

$$U_{AB} = U_{AB}^0 + U_{AB}^1 + U_{AB}^2 + \dots + U_{AB}^n \quad 2.5$$

The definitions of the different contributions to the intermolecular forces, such as the electrostatic terms, the induction and the dispersion, correspond to different terms in the perturbation series expansion. The first-order perturbation represents the electrostatic energy, *i. e.* the change in energy due to the electrostatic interaction between the permanent multipole moments of the molecules. It is calculated as the expectation value of the electrostatic interaction V for the ground state ($m = n = 0$) with m, n being the order of the wavefunction

$$U_{AB}^1 = \langle \Psi_A^0 \Psi_B^0 | V | \Psi_A^0 \Psi_B^0 \rangle = U_{electrostatic} \quad 2.6$$

This classical interaction of the undistorted, non-spherical charge distributions ρ^A and ρ^B of the ground state wavefunctions in isolation of molecules A and B respectively can be calculated by an integration over the *ab initio* charge densities of the interacting molecules at each relative orientation.

$$U_{\text{electrostatic}} = \int \frac{\rho^A(\mathbf{r}_1)\rho^B(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d^3\mathbf{r}_1 d^3\mathbf{r}_2 \quad 2.7$$

The elements of the charge distributions of *A* and *B* are separated by $|\mathbf{r}_1 - \mathbf{r}_2|$ and as they are not distorted in the first-order energy correction, this contribution to the intermolecular forces is strictly pairwise additive.

This electrostatic term is zero for spherical neutral molecules (atoms), but significant for the majority of organic molecules. It is the only major contribution to the intermolecular potential that is either attractive or repulsive and is usually the interaction that persists over the longest range. Furthermore, it is the most sensitive potential contribution towards the relative orientation of the molecules and is therefore likely to dominate the orientation dependence of the intermolecular potential for polar molecules.

The second-order perturbation energy

$$U_{AB}^2 = - \sum_{\substack{\Psi_A^m \neq \Psi_A^0 \\ \Psi_B^n \neq \Psi_B^0}} \frac{\langle \Psi_A^0 \Psi_B^0 | V | \Psi_A^m \Psi_B^n \rangle \langle \Psi_A^m \Psi_B^n | V | \Psi_A^0 \Psi_B^0 \rangle}{U_{AB}^{mn} - U_{AB}^0} \\ = U_{\text{Induction}}^A + U_{\text{Induction}}^B + U_{\text{Dispersion}}^{AB} \quad 2.8$$

represents two physical effects, the induction and the dispersion interaction, and it does not contain terms with both molecules in the ground state.

The induction energy can be described as the distortion of the charge distribution on one molecule (molecular polarisability) due to the electric field of all surrounding (undistorted) molecules. In the induction energy of molecule *A*, where the field from molecule *B* polarises *A*, the distortion of *A* is described by the excited state wave functions Ψ_A^m of energy E_A^m and molecule *B* in the ground (undistorted) state

$$U_{\text{Induction}}^A = - \sum_{\Psi_A^m \neq \Psi_A^0} \frac{|\langle \Psi_A^0 \Psi_B^0 | V | \Psi_A^m \Psi_B^0 \rangle|^2}{U_A^m - U_A^0} \quad 2.9$$

The sum in this equation is over all the excited states of molecule *A*.

The induction (polarisation) energy of *B* where the field from molecule *A* polarises *B* and which has *B* in an excited state and *A* in the ground state is equal to

$$U_{Induction}^B = - \sum_{\Psi_B^n \neq \Psi_B^0} \frac{|\langle \Psi_A^0 \Psi_B^0 | V | \Psi_B^n \Psi_A^0 \rangle|^2}{U_B^n - U_B^0} \quad 2.10$$

The induction energy between ground state molecules is always attractive, as the distortions only occur to lower the energy of the pair. This highly non-pairwise additive energy term is important for the intermolecular interactions of ions, but it is often one of the smallest long range terms for neutral molecules, and zero for neutral spherical molecules (atoms).

The final term of the second order perturbation energy represents the dispersion energy, the energy lowering associated with the polarisation by instantaneous fluctuations in the charge distributions of the monomers *A* and *B*. The excited states of molecules *A* and *B* are used to describe this non-classical phenomenon.

$$U_{Dispersion}^{AB} = - \sum_{\substack{\Psi_A^m \neq \Psi_A^0 \\ \Psi_B^n \neq \Psi_B^0}} \frac{\langle \Psi_A^0 \Psi_B^0 | V | \Psi_A^m \Psi_B^n \rangle \langle \Psi_A^m \Psi_B^n | V | \Psi_A^0 \Psi_B^0 \rangle}{U_B^n - U_B^0 + (U_A^m - U_A^0)} \quad 2.11$$

For spherical neutral molecules (for example argon), which consequently have no permanent dipole or higher multipole moments, the electrostatic and induction (polarisation) contributions to the intermolecular interaction energy are absent. The only attractive interaction between such molecules at long range is the dispersion interaction, first identified by London ⁸ in the 1930s. This long range term has recently been described as the 'universal attractive glue that leads to the formation of condensed phases'. ¹⁴

2.2.2 Short range contributions

At short range, at intermolecular distances found in the solid state, the overlap of the molecular wavefunctions cannot be ignored and hence the analytical separation cannot be used. If the total wavefunction is antisymmetrised, first-order and higher electron exchange correction terms would appear in the expectation value of the energy.¹⁵ A variety of perturbation theory treatments ('exchange perturbation theories')¹⁶, which deal with intermolecular perturbation theory in the region where electron exchange cannot be neglected, are available for an estimation of the short range terms of the intermolecular potential. Symmetry-Adapted Perturbation Theory (SAPT)^{15, 17} or as a special case of the SAPT, the InterMolecular Perturbation Theory (IMPT)^{3, 18-21}, estimating the various short-range terms starting from the wavefunctions of the isolated molecule, have been developed. Whereas the long range theory provides an analytical expression for the intermolecular interaction energy, the computational requirements for the short range calculations allow only the estimation of a few points of the potential. This is not sufficient to define the full potential energy surface, but provides some means to develop and test model potentials for organic crystal structures.

Exchange-repulsion energy

This net-repulsive contribution to the intermolecular forces consists of two different components: a non-additive repulsive force, which arises from the fact that the overlapping charge densities cannot occupy the same space. This is due to the Pauli exclusion principle, which forbids electrons of the same spin to be in the same place. In addition to this dominant effect, the exchange of electrons of parallel spin between the molecules leads to a weak attractive force. For the intermolecular contacts sampled in molecular crystals, the exchange-repulsion energy is approximately additive.

Charge transfer

This contribution to the intermolecular forces results from the transfer of charge from the occupied orbitals of one molecule to the unoccupied orbitals of the other. It has an approximately exponential decay with separation and its magnitude can be seriously overestimated due to basis set superposition error (BSSE)²². Basis sets are always finite and therefore incomplete. In an *ab initio* calculation of the interaction energy

between two molecules as the difference of the energies of the interacting system and the separate monomers (supermolecule method), the basis functions of each molecule in the supermolecule become available to the other molecule. This basis set improvement for the supermolecule results in a lowering of its energy and therefore an overestimation of the interaction energy of the total system. The most popular method of correcting for the BSSE is the counterpoise correction method²³ formulated by Boys and Bernardi²⁴ which essentially uses the supermolecule basis set for all calculations, including the monomer energies.

The charge transfer effect is strongly non-additive and attractive and it can be evaluated by perturbation theory calculations.²⁵ In practice, it is usually ignored in simulations and assumed to be absorbed by the exponential terms in model potentials.

2.3 Models for calculating intermolecular interactions

2.3.1 Modelling the electrostatic contribution to the intermolecular forces

2.3.1.1 Point charge models

The simplest and traditional method of calculating the electrostatic energy for organic molecules is the atomic point charge model.^{26, 27} It is assumed that the charge is a superposition of spherical molecular charge distributions $\rho(r)$ which are approximated by a set of partial charges at the atomic sites. The charges obtained are very dependent on the method used to distribute the molecular charge density into the atomic charge contributions. Methods of dividing charge distributions into contributions on atomic sites are further discussed in 2.3.1.3. Atomic point charges can be derived directly from experiment (for example from X-ray diffraction data²⁸) or from the wavefunction by a population analysis, *i.e.* such as the Mulliken population analysis²⁹ or the natural population analysis (NPA)^{30, 31}.

Although originally not derived for calculating intermolecular forces, the highly basis-set dependent Mulliken population charges have been very popular, and their calculation is provided by virtually all semi-empirical or *ab initio* wavefunction programs. However, they basically just represent the charges of a truncated multipole series and their application for this purpose is questionable. A more recent method of reproducing the electrostatic interactions of a given *ab initio* charge distribution is to fit the charges to reproduce the correct electrostatic potential around the molecule, as calculated directly by integration over the wavefunction³²⁻³⁴.

The assumption, on which all these methods are based, that the molecular charge distribution can be accurately represented by spherical point charges, is a crude approximation, especially for atoms with non-spherical features such as lone pair and π – electrons. Wiberg and Rablen ³⁵ investigated several schemes for assigning atom charges, and concluded that: *‘the charge distribution in a molecule is much too anisotropic to be successfully modelled by _any_ single set of atom-centred charges unless only long-distance interactions are of interest. Otherwise it is necessary to include at least atomic dipole terms, and possibly higher terms as well.’*

2.3.1.2 Central multipole models

The electrostatic energy of two interacting molecules can be theoretically derived from equation 2.7 as a multipole expansion. The expansion is over the inverse of the distance between the charges in the different molecules, in terms of a vector \mathbf{R} between the centres of mass of the two molecules, and the vectors from these centres to the charges in each molecule ¹¹:

$$U_{\text{electrostatic}}(\mathbf{R}, \Omega_A, \Omega_B, \Omega_{AB}) = \frac{1}{4\pi\epsilon_0} \sum_{l_A l_B} \sum_{k_A k_B} \binom{l_A + l_B}{l_A} Q_{l_A k_A}^A Q_{l_B k_B}^B \bar{S}_{l_A l_B, l_A + l_B}^{k_A k_B} (\Omega_A, \Omega_B, \Omega_{AB}) R_{AB}^{-(l_A + l_B + 1)} \quad 2.12$$

Here $\binom{l_A + l_B}{l_A}$ is a binomial coefficient and R_{AB} is the distance between the centre of molecule A and the centre of molecule B. $Q_{l_A k_A}^A$ are the components k of the multipole moments of molecule A (defined by its local axis system with the centre of mass of the molecule often chosen as the origin) with multipole of rank l (i.e. charge ($l = 0$), dipole ($l = 1$), quadrupole ($l = 2$), octopole ($l = 3$) and hexadecapole ($l = 4$) etc.) and component k . Molecular symmetry often leads to multipoles being equal to zero and therefore simplifies these calculations. The \bar{S} functions ³⁶ describe the relative orientations of the molecular local axis systems of the interacting molecules. The electrostatic energy between the interaction of multipole l_A on molecule A and multipole l_B on molecule B decreases as $R^{-(l_A + l_B + 1)}$.

This model is not valid when the convergence spheres, which contain all the molecular charge distribution around each of the interacting molecules, overlap. If the penetration effect is considered separately, e. g. in terms of Gaussian functions as part of the short-range interaction, then the divergence sphere for a molecule is the sphere enclosing just the *nuclei*.³⁷ However, in condensed phases, like the solid state,³⁸ orientations where the spheres overlap, are very common.

However, even if the spheres do not overlap, the convergence of the central multipole expansion may be slow for certain relative orientations of even fairly spherical molecules.³⁹ Therefore a large number of terms would have to be taken into account for the representation of the true charge distribution of many organic molecules.

2.3.1.3 Distributed multipole models

A more accurate description of the entire molecular charge distribution is by a distributed multipole expansion, in which the molecule is divided into sites (e. g. atoms or small groups of atoms), each described by its own multipole moments. The difficulty with this kind of approach is, similar as for the atomic point charges, how to divide the charge distribution into contributions from each site. One of the earliest methods is Hirshfeld's 'stockholder' principle^{40, 41} in which the total molecular charge density is calculated for a hypothetical 'promolecule' as a superposition of spherical neutral atoms with non-interacting atomic charge densities. The charge density of each atomic site is determined by dividing up the total in ratio to the contribution each atom provides to the total charge density of the promolecule.

Another approach is Bader's⁴² "Atoms in Molecules (AIMs)" treatment. In this method, the charge density is partitioned by assuming that the gradient of the charge density is zero on the dividing surface ('zero-flux surface') between the atomic sites. This technique has the advantage that it is based on a fundamental definition which does not need further assumptions. Nevertheless it leads to atomic regions that are highly non-spherical which require high order multipoles to describe the charge distribution accurately.

Further methods include the Cumulative Atomic Multipole Moment (Camm) representation by Sokalski and Poirier⁴³ and similar procedures by Rein⁴⁴ and Vigne-Maeder and Claverie⁴⁵. A technique developed to optimise the convergence behaviour

is the Distributed Multipole Analysis (DMA) ^{37, 46} by Stone. It is used in this thesis and therefore described in more detail later (2.3.1.3.1).

All of these distributed multipole methods should describe the electrostatic potential outside a molecule to the same accuracy, provided that sufficient sites and multipole moments up to quadrupole are considered, as shown for the (HF)₂ dimer. ⁴⁷

2.3.1.3.1 Distributed Multipole Analysis

The Distributed Multipole Analysis (DMA) is, because of its excellent convergence behaviour, very efficient for applications to intermolecular forces. Its computational cost is negligible compared with the *ab initio* calculation of a monomer wavefunction. This method is based on the one-electron density matrix $\rho(r)$, which can be represented in terms of Gaussian type functions. For a molecule this matrix is defined as

$$\rho(r) = \sum \rho_{ab} \phi_a(r) \phi_b(r) \quad 2.13$$

with $\phi_a(r)$ and $\phi_b(r)$ being primitive Gaussian functions centred at the sites a and b respectively.

$$\phi_a(r) = f(r-a) \exp[-\alpha(r-a)^2] \quad 2.14$$

Here $f(r-a)$ is a homogeneous polynomial of degree l_a if ϕ_a is a basis function of angular momentum l_a .

The product of two Gaussian functions is another Gaussian at the overlap centre P on a line from a to b determined by the original centres a , b and corresponding exponents α , β of the original Gaussian functions.⁴⁸

$$P = \frac{\alpha a + \beta b}{(\alpha + \beta)} \quad 2.15$$

If the Gaussian primitives have angular momenta l_a and l_b , the overlap charge distribution can be represented by a finite multipole series up to the rank of the sum of

the angular momenta ($l_a + l_b$) at each overlap centre of the basis functions involved. So the overlap resulting from two s orbitals leads only to a point charge at the overlap centre, whilst any charge distribution described by an s function with a p function can be presented by a charge, dipole. The overlap of two p functions generates a charge, dipole and quadrupole. This multipole series will be located on the atom where the two Gaussian functions are. If the latter are not on the *same* atom, the contribution from the overlap centre is shifted to a different location S , usually the nearest atom site. This is done using the relationship between a point multipole at one site to an infinite series of the same and higher order multipoles at another point. 46

$$Q_{lm}(S) = \sum_{k=0}^l \sum_{q=-k}^k \binom{l+m}{k+q} \binom{l-m}{k-q}^{1/2} Q_{kq} R_{l-k,m-q}(S-P) \quad 2.16$$

It has to be guaranteed that the distance shifted is not too large, otherwise the convergence of the series is slower. Equation 2.16, in the regular spherical harmonic $R_{l-m,k-q}$, depends on the factor $|S-P|^{l-k}$ and $Q_{lm}(S)$ will therefore still be large for high values of l if the distance $|S-P|$ is large.

The electrostatic potential due to a multipole series at site S will be 46:

$$V(r) = \sum_{lm} |r-S|^{-l-1} C_{lm}(r-S) Q_{lm}(S) \quad 2.17$$

For a good convergence the potential location r has to be much further from the multipole expansion site S than any of the individual overlap populations at P contributing to it, *i. e.* $|r-S|$ has to be larger than $|S-P|$ (in equation 2.16). Therefore the DMA sites are chosen so that no overlap multipole series needs to be moved very far from its natural centre P to the DMA site S . Since the nuclei already have many multipoles associated, they are a sensible choice as expansion sites.

In general, a large number of parameters are required to describe the charge distribution, as an atomic site will be described by a point charge, three components of the dipole, five components of the quadrupole, seven components of the octopole and nine components of the hexadecapole moment. In practice, molecular symmetry and

approximate symmetry of the atomic sites will often reduce the number of significant multipoles.

A main advantage of describing the electrostatic interactions by a multipolar model, as opposed to an atomic point charge model, is that different multipole moments can be associated with well-known concepts of bonding. For example, a dipole can be seen as a representation of a lone pair on an atom and a quadrupole as an indication for π - bonding electrons.

The main limitations to the accuracy of a distributed multipole description of the electrostatic energy are the neglect of the effects of interpenetration of the charge distributions. But this can usually be absorbed by any empirical short-range exponential repulsive potential. This model for the electrostatic forces has been successfully applied for a variety of different systems, such as Van der Waals complexes ⁴⁹⁻⁵¹, DNA base pairs ^{52, 53} and in the modelling of crystal structures ⁵⁴ and it is used in this thesis.

2.3.2 Modelling the induction energy

The non-additive induction usually makes just a small contribution to the intermolecular energy and, as it is very difficult to include it easily in model potentials, its explicit treatment is therefore normally neglected. This is even the case for strongly polar systems, where the effects of induction can be quite significant. It is common to use an electrostatic model with an enhanced dipole moment to describe some form of average of the induction effects for these polar systems.

A simple and often applied approach for ionic solids is the shell model ^{55, 56} in which the valence electrons are viewed as a charged massless spherical shell attached to a point charge (the core of the ion or atom) via a harmonic spring. In a uniform field, the shell and the core will move in opposite directions, producing a dipole whose magnitude depends on the strength of the spring. The spring constant has to be chosen so that the magnitude of the generated dipole reflects the polarisability of the system.

In reality, the fields around a molecule are usually very non-uniform and it is necessary to divide the molecule into reasonably uniform regions. Calculating the induction energy by a central multipole expansion suffers from the same problems of validity and convergence as the central multipole expansion of the electrostatic energy. Especially for polyatomic systems in Van der Waals contact, the polarisability must be distributed, in order to get a realistic description. One method among others ⁵⁷⁻⁵⁹ of

distributing the polarisability amongst many sites within a molecule is the Distributed Polarisability Analysis ⁶⁰ by Stone. This model describes the flow of charge from one atom to another under the influence of an external non-uniform electrostatic field. Unfortunately, the full distributed polarisability model is too elaborate to be currently used in modelling anything but relative small systems and methods of transforming the full description into a simpler one ⁶¹ often lack transferability and accuracy. Therefore polarisability is ignored in this work.

2.3.3 Modelling the dispersion energy

This purely quantum mechanical effect between two spherical atoms (or molecules) A , B is often approximated by the application of the central multipole expansion resulting in the series

$$U_{dispersion} = - \left(\frac{C_6}{R_{AB}^6} + \frac{C_8}{R_{AB}^8} + \frac{C_{10}}{R_{AB}^{10}} + \dots \right) \quad 2.18$$

The dispersion coefficient C_n terms, often described as the correlation in the instantaneous multipolar fluctuations in the charge density of two molecules, can be evaluated by various methods ^{3, 62, 63} for small high-symmetry systems. Their numerical values increase with molecular size and polarisability of the system. The dispersion energy can also be evaluated by a distributed treatment of the polarisabilities ⁶⁴.

2.3.4 Modelling the exchange-repulsion energy

The simplest representation of the net repulsive contribution to the intermolecular interaction is the hard-sphere model in which the potential is either equal to zero or, in regions of atom-atom overlap, infinite. The chosen cutoff radii around the atoms are usually the Van der Waals radii. ^{65, 66} This crude approximation is, in combination with an accurate electrostatic model, able to account for the structures of a range of Van der Waals complexes ^{49, 50}, but does not have any predictive value in terms of intermolecular separations.

The repulsive term A/R^n (usually $n = 12$) in the widespread used Lennard-Jones potential ^{67, 68} has no theoretical justification, besides its steeply repulsive form and this potential is used for computational convenience.

The Born-Mayer potential form, $Ae^{-\alpha R}$, ⁶⁹ is based on the relationship between the repulsive force and the overlap of the wavefunctions, and therefore suggests an exponential decay of the repulsion with the distance R . This exponential form is frequently applied in model potentials for the exchange-repulsion energy of molecular crystals.

The density overlap model is based on the empirical observation that the exchange-repulsion between pairs of inert-gas atoms ⁷⁰ and between small molecules ⁷¹ is approximately proportional to the overlap between their charge densities. This density overlap between two molecules can be expressed, in an exponential form, as the sum of site-site terms. This approach has been successfully applied in modelling the exchange repulsion of complexes ⁷², molecular clusters ^{73, 74} and crystal structures ⁷⁵⁻⁷⁷. A main advantage of this method, compared to the empirical derivation of repulsion parameters, is that it does not rely on the availability of experimental data for fitting and that no transferability assumptions between atoms in different types of molecules need to be made.

2.3.5 Damping of long range potential terms

The long range potential terms, as the dispersion, the induction and the electrostatic interaction, derived by a multipole expansion, appear as a power series of R^{-n} . At very short distances, these terms will increase to infinity, and therefore give an unphysical representation of the intermolecular interaction. This singularity can be avoided by the use of damping functions ⁷⁸ in the form

$$U_{dispersion} = - \left(f_6(R) \frac{C_6}{R_{AB}^6} + f_8(R) \frac{C_8}{R_{AB}^8} + f_{10}(R) \frac{C_{10}}{R_{AB}^{10}} + \dots \right) \quad 2.19$$

where $f_6(R)$, $f_8(R)$, etc. are functions that tend to a numerical value of 1 as R approaches infinity and to a value of zero for R equal to zero.

The damping function represents the modifications of the long range terms due to the overlap of the charge distributions. It is significant at the distances found in the

Van der Waals region. Hence, such terms, usually absorbed into *empirically* derived short-range exponential exchange-repulsion potentials, will have an effect on solid-state properties.

2.4 Development of model intermolecular potentials

The accuracy of the applied intermolecular potential function is crucial for the reliability of the outcome of *any* simulation in the solid, liquid or gaseous state.

For small polyatomics, as $(\text{Cl}_2)_2$ ⁷⁹ and $(\text{H}_2\text{O})_2$, ⁸⁰ quite detailed intermolecular potentials, specific to each molecule, have been derived by summing up the anisotropic atom-atom models of the different contributions to the intermolecular interactions. These functions are usually developed to describe the interaction as accurately as possible. The water potential by Millot and Stone ⁸⁰ has further been successfully applied to calculations of the structure of the water trimer. ⁸¹

For the simulation of organic or biochemical interactions, much simpler models need to be used. They are either *empirically* derived by fitting to a range of experimental data or to accurate calculations or by simplifying accurate potential functions by omitting terms that are likely to have a negligible effect for the specific simulation. In both cases, the transferability of intermolecular atom-atom interactions between different molecules often has to be assumed.

2.4.1 Model potentials for organic crystal structures

Molecular crystal structures are very sensitive to the intermolecular potential in the region of the Van der Waals contact. The simplest realistic model for the intermolecular pair potential that can be applied to crystal structure modelling is a *6-exp* function, which is based on the assumption that the interaction between the molecules is the sum of the interactions between their constituent atoms.

$$U(R_{ik}) = \sum_{ik} U_{ik} = \sum_{ik} A_{i\kappa} \exp(-B_{i\kappa} R_{ik}) - \frac{C_{i\kappa}}{R_{ik}^6} \quad 2.20$$

Here, atom i in one molecule and atom k in the other molecule are of types ι and κ respectively. The parameters $A_{i\kappa}$, $B_{i\kappa}$ and $C_{i\kappa}$ describe the interaction between each pair of atoms. They can be empirically fitted to crystal structures and available heats of

sublimation. This potential depends only on the separation R_{ik} of the atoms in the different molecules, and the types of the atoms concerned. It can be alternatively expressed as a function of the minimum energy separation R_{lk}^0 , the well depth ϵ_{lk} and the steepness parameter λ_{lk} of each individual atom-atom interaction.

In order to reduce the number of required parameters, heteroatomic interaction parameters can be calculated by the following combining rules:

$$A_{ik} = (A_{ii}A_{kk})^{1/2}; B_{ik} = \frac{(B_{ii} + B_{kk})}{2}; C_{ik} = (C_{ii}C_{kk})^{1/2} \quad 2.21$$

This is just one of several types of combining rules⁶⁸ and although they are widely used, they do not have a strong physical justification and are known to be limited in their accuracy⁶⁴.

This type of atom-atom model potential has been pioneered by Pertsin and Kitaigorodskii⁸² and a wide variety of parameter sets have been developed empirically by fitting to experimental data. Filippini and Gavezzotti⁸³ have derived parameters for C, H, N, O, S and Cl by empirical fits to a database of 217 crystal structures and heats of sublimation (available for 122 compounds). They included hydrocarbons, oxa-, aza-, chloro-, sulfo-hydrocarbons, nitro compounds, sulfones and sulfoxides, but excluded hydrogen bonded compounds. Values of the minimum energy separation were obtained by studying the distribution of intermolecular contacts in the crystal structures. Due to the large amount of experimental data available, the well depths parameters were empirically fitted without using combining rules. The steepness parameter was averaged over all interactions. This model potential scheme was later extended to hydrogen bonded compounds⁸⁴. The electrostatic contribution to the intermolecular forces is not *explicitly* included in this potential, but partially absorbed into the empirical *6-exp* parameters. This resulted in considerable deviations of the parameters from those expected under the usual combining rules. This potential is – with some limitations – remarkably good at the reproduction of many crystal structures and their experimental heats of sublimation. The neglect of the explicit treatment of the electrostatic contributions reduces the computational expense, but provides little insight into a further understanding of the forces, which determine crystal packing.

The inclusion of point charges, as the simplest way of modelling the electrostatic contribution explicitly, has shown to be important in determining the minimum energy crystal structures even for aromatic hydrocarbons. Williams and Starr⁸⁵ used empirical charges²⁶ to improve their model potential, which gave an excellent reproduction of the quadrupole moments of the aromatic hydrocarbons studied⁸⁶.

An extensive computational scheme has been developed by Williams and coworkers for the derivation of intermolecular potential parameters for a variety of different atom types. An optimised force field (net atomic charges plus an exp-6 repulsion-dispersion potential) for perchlorohydrocarbons⁸⁷ has been derived with the repulsion-dispersion parameters for the carbon atoms being transferred from the hydrocarbon potential.⁸⁶ It has been found that the electrostatic contribution to the intermolecular potential for these structures is less dominant than for the hydrocarbon crystal structures. A similar potential type has been derived for non-hydrogen bonded oxohydrocarbons⁸⁸ with the net atomic charges being fitted to the calculated molecular electrostatic potential.³³ As the repulsion-dispersion parameters for carbon and hydrogen atoms were again transferred from hydrocarbons, the only atom type to optimise was oxygen with all oxygen atoms being assumed to have the same set of parameters. For non-hydrogen bonded azahydrocarbons⁸⁹, net atomic charges were obtained by fitting to the calculated *ab initio* molecular electrostatic potential and including, for aromatic nitrogen atoms, additional lone-pair electron potential derived site charges.⁹⁰ The repulsion-dispersion parameters for carbon and hydrogen were for an initial refinement of the nitrogen...nitrogen potential, again assumed to be transferable from the hydrocarbon study. But for a further verification of the transferability, a full set of hydrogen, carbon and nitrogen parameters were derived simultaneously from a combination of hydrocarbon and azahydrocarbon crystal structure data. The electrostatic contribution to the intermolecular potential was found to be very important for azahydrocarbons, ranging up to a maximum of 59% for nitrile crystal structures. For fluorine-fluorine intermolecular interactions,⁹¹ several potential-derived site charge models (no charge, net atomic site charge, net atomic site charge plus an extended charge site located on the extension of the C-F bond axis and a net atomic site charge model plus a bond-charge site) were tested. The best fit to the electrostatic potential has been obtained with the bond-charge site model. For the

repulsion-dispersion contribution, the carbon parameters were transferred from the work on azahydrocarbons,⁸⁹ and parameters for fluorine were derived from fits to crystal structures of perfluorocarbons. The electrostatic contribution was found to be around 10%, a similar proportion to the one obtained for hydrocarbons.

Hence, Williams has built up a library of intermolecular potential parameters (net charge model plus exp-6 repulsion-dispersion terms) for carbon, hydrogen, oxygen, nitrogen, fluorine and chlorine⁹² for organic molecules with one type of heteroatom by empirical fitting.

Since this thesis a new intermolecular force field (W99)⁹³ for hydrocarbons has been derived by assigning separate parameters for four-coordinated (alkanes and cycloalkanes) and three-coordinated (aromatics and unsaturated alkenes) carbon atoms. This was motivated by the difference in polarity of these two groups of hydrocarbons. Hydrogen atoms were given a single set of parameters. The electrostatic contribution was described by net atomic charges, and as the molecular electrostatic potential of n-alkanes and cycloalkanes cannot be well modelled by just net atomic charges^{94, 95}, methylene bisector or ring centre site charges were added where necessary. The inclusion of the different carbon types, resulted in a better reproduction of the available heats of sublimation for the studied systems. The W99 potential has recently been extended for hydrogen-bonded oxohydrocarbons.⁹⁶

As pointed out before, molecular crystal structures are very sensitive to the intermolecular potential in the region of the Van der Waals contact, and therefore the orientation dependence of the potential plays a key role in the accuracy of the structure reproduction. A study⁹⁷, using a DMA model for the electrostatic forces together with a literature-based isotropic atom-atom repulsion-dispersion potential, has been successful in reproducing the crystal structures of a variety of multifunctional molecules. The repulsion-dispersion potential (EST) applied in that study has been derived by a combination of Williams' oxohydrocarbon (oxygen)⁸⁸, azahydrocarbon (carbon, nitrogen and hydrogen bonded to carbon)⁸⁹ parameters and a parameter set for polar hydrogen atoms (H_p) derived from IMPT calculations on the formamide-formaldehyde complex.⁹⁸ Combining rules were used for the determination of the heteroatomic parameters and this potential was extended (FIT),⁵⁴ by fitting the $H_p \cdots H_p$ potential to a set of crystal structures and lattice energies to improve the

reproduction of the N-H \cdots O and N-H \cdots N intermolecular bonds and of the available heats of sublimation.

Throughout the work of this thesis, an accurate electrostatic model derived from a Distributed Multipole Analysis (DMA) combined with an empirically derived repulsion-dispersion potential has been used. In the study of the blindtest compound (chapter 5) and for modelling the polymorphs of paracetamol (chapter 6) the extensively tested FIT repulsion-dispersion potential has been applied. For the work on carboxylic acids (chapter 7) this empirical potential has been optimised by scaling the pre-exponential parameter for the polar hydrogen atoms to reproduce the hydrogen bonds found in carboxylic acids. A multipolar model for the electrostatic interactions, as opposed to an atomic point charge model, is theoretically more accurate and increases the confidence in the obtained results. The DMA has, compared to other distributed multipole models, a very good convergence behaviour and is hence a computationally very efficient way of calculating the anisotropic electrostatic intermolecular interactions accurately. The repulsion-dispersion parameters applied in this work have been empirically derived by assuming transferability between similar types of atoms in different molecules. Furthermore, no explicit treatment of the induction forces is included in these model potentials.

The crystal structures of the chlorine dimer and chlorinated hydrocarbons can only be accurately accounted for when an anisotropic atom-atom repulsion model is used.^{99, 100} This is also true for the crystal structure of s-tetrazine which displays a rare case of nitrogen...nitrogen repulsion anisotropy.^{54, 90} These examples show that the anisotropy of the repulsion can be very important for the accurate reproduction of some crystal structures and the transferability of empirically fitted isotropic parameters between different types of molecules has to be carefully assessed.

In the cause of this thesis, pioneering work, to derive *ab initio* based repulsion dispersion parameters by the density overlap model for the polymorphic system of oxalic acid,⁷⁵ an oxyboryl derivate⁷⁶ and for amide crystal structures⁷⁷ is in progress. Combined with a distributed multipole model for the electrostatic forces, these theoretically based repulsion potentials have shown to be very competitive to empirically derived potentials for crystal structure reproduction, and they have the advantage that the transferability of the potential parameters can be tested. This method

is being extended to explicitly include the theoretically derived anisotropic part of the repulsion potential.

Furthermore, a transferable *ab initio* intermolecular potential for alkanes, ether and alcohols has been developed by fitting the different contributions to the intermolecular potential separately to theoretical calculations.¹⁰¹

2.5 Conclusions

Accurate intermolecular potentials can be developed by the application of the theory of intermolecular forces to each contribution to the intermolecular interaction. These potentials with a well-defined physical meaning are not only crucial for the successful simulation of the gas, liquid and solid state and the related properties, but they are the key for a further understanding of the interactions between atoms and molecules. The successful application of theoretically derived model potentials to the condensed phase will be more and more possible in the near future due to increasing computer power. This evolution will decrease the need for empirically derived potential parameters and will foster our confidence in the accuracy of the simulation results.

References for chapter 2

- (1) Dunitz, J.D. in *Implications of Molecular and Materials Structure for New Technologies*, J.A.K. Howard, Allen, F. H., Shields, G. P., Kluwer Academic Publishers, Dordrecht, **1999**, 1-9. Into the new millennium: the present and future of crystal structure analysis.
- (2) Gavezzotti, A., *Crystallography Reviews*, **1998**, 7, 5-121. The crystal packing of organic molecules: Challenge and fascination below 1000 Da.
- (3) Stone, A.J., *Intermolecular forces*, **1996**, Oxford: Oxford University Press.
- (4) Brink, D.M.; Satchler, G.R., *Angular momentum*, **1967**, Oxford: Clarendon Press.
- (5) Van Mourik, T.; Dunning, T.H., *J. Chem. Phys.*, **1997**, 107, 2451-2462. Ab initio characterization of the structure and energetics of the ArHF complex.
- (6) Mok, D.K.W.; Handy, N.C.; Amos, R.D., *Molec. Phys.*, **1997**, 92, 667-675. A density functional water dimer potential surface.
- (7) Barker, J.A.; Fisher, R.A.; Watts, R.O., *Molec. Phys.*, **1971**, 21, 657-673. Liquid argon: Monte Carlo and molecular dynamics calculations.
- (8) London, F., *Trans. Faraday Soc.*, **1937**, 33, 8-26. The general theory of intermolecular forces.
- (9) Longuet-Higgins, H.C., *Proc. Roy. Soc.*, **1956**, A235, 537-543. The electronic states of composite systems.
- (10) Buckingham, A.D., *Adv. Chem. Phys.*, **1967**, 12, 107-143. Permanent and induced multipole moments and long-range intermolecular forces.
- (11) Stone, A.J.; Tough, R.J.A., *Chem. Phys. Lett.*, **1984**, 110, 123-129. Spherical tensor theory of long-range intermolecular forces.
- (12) Pitzer, K.S., *Adv. Chem. Phys.*, **1959**, 2, 59-83. Inter- and intramolecular forces and molecular polarizabilities.
- (13) Van Lenthe, J.H.; Van Duijneveldt-van de Rijdt, J.G.C.M.; Van Duijneveldt, F.B., *Adv. Chem. Phys.*, **1987**, 69, 521-566.
- (14) Price, S.L. in *Reviews in Computational Chemistry*, K.B. Lipkowitz and D.B. Boyd, Wiley-VCH, John Wiley and Sons, Inc., New York, **2000**, 225 - 289. Toward more accurate model intermolecular potentials for organic molecules.

- (15) Jeziorski, B.; Moszynski, R.; Szalewicz, K., *Chem. Rev.*, **1994**, *94*, 1887-1930. Perturbation theory approach to intermolecular potential energy surfaces of Van der Waals complexes.
- (16) Claverie, P. in *Intermolecular Interactions: from Diatomics to Biopolymers*, B. Pullman, Wiley, **1978**, 69-305. Elaboration of approximate formulas for the interactions between large molecules: applications in organic chemistry.
- (17) Jeziorski, B.; Kolos, W. in *Molecular Interactions*, H. Ratajczak and W.J. Orville-Thomas, Wiley, Chichester, **1982**,
- (18) Hayes, I.C.; Stone, A.J., *Molec. Phys.*, **1984**, *53*, 69-82. Matrix elements between determinantal wavefunctions of non-orthogonal orbitals.
- (19) Hayes, I.C.; Stone, A.J., *Molec. Phys.*, **1984**, *53*, 83-105. An intermolecular perturbation theory for the region of moderate overlap.
- (20) Hayes, I.C.; Hurst, G.J.B.; Stone, A.J., *Molec. Phys.*, **1984**, *53*, 107-127. Intermolecular perturbation theory. Applications to HeBe, ArHF and NeH₂.
- (21) Stone, A.J. in *Advances in biomolecular simulations*, R. Lavery, J.-L. Rivail, and J. Smith, American Institute of Physics, New York, **1991**, 3-19.
- (22) Chalasinski, G.; Gutowski, M., *Chem. Rev.*, **1988**, *88*, 943-962. Weak interactions between small systems - models for studying the nature of intermolecular forces and challenging problems for ab initio calculations.
- (23) Van Duijneveldt, F.B.; Van Duijneveldt-van deRijdt, J.G.C.M.; Van Lenthe, J.H., *Chem. Rev.*, **1994**, *94*, 1873. State-of-the-art in counterpoise theory.
- (24) Boys, S.F.; Bernardi, F., *Molec. Phys.*, **1970**, *19*, 553-566. The calculation of small molecular interactions by the difference of separate total energies. Some procedures with reduced errors.
- (25) Stone, A.J., *Chem. Phys. Lett.*, **1993**, *211*, 101-109. Computation of charge-transfer energies by perturbation theory.
- (26) Williams, D.E. in *Reviews in Computational Chemistry*, K.B. Lipkowitz and D.B. Boyd, VCH Publishers, New York, **1991**, 219-271. Net atomic charge and multipole models for the ab initio molecular electric potential.
- (27) Bachrach, S.M. in *Reviews in Computational Chemistry*, K.B. Lipkowitz and D.B. Boyd, VCH Publishers, New York, **1994**, 171-227. Population analysis and electron densities from quantum mechanics.
- (28) Spackman, M.A., *Chem. Rev.*, **1992**, *92*, 1769-1797. Molecular electric moments from X-ray diffraction data.

- (29) Mulliken, R.S., *J. Chem. Phys.*, **1955**, 83, 735.
- (30) Reed, A.E.; Weinstock, R.B.; Weinhold, F., *J. Chem. Phys.*, **1985**, 83, 735-746. Natural-population analysis.
- (31) Reed, A.E.; Weinhold, F.; Curtiss, L.A.; Pochatko, D.J., *J. Chem. Phys.*, **1986**, 84, 5687-5705. Natural bond orbital analysis of molecular interactions: theoretical studies of binary complexes of HF, H₂O, NH₃, N₂, O₂, F₂, CO and CO₂ with HF, H₂O and NH₃.
- (32) Cox, S.R.; Williams, D.E., *J. Comp. Chem.*, **1981**, 2, 304-323. Representation of the molecular electrostatic potential by a net atomic charge model.
- (33) Momany, F.A., *J. Phys. Chem.*, **1978**, 82, 592. Determination of partial atomic charges from ab initio molecular electrostatic potentials. Application to formamide, methanol, and formic acid.
- (34) Singh, U.C.; Kollman, P.A., *J. Comp. Chem.*, **1984**, 5, 129-145. An approach to computing electrostatic charges for molecules.
- (35) Wiberg, K.B.; Rablen, P.R., *J. Comp. Chem.*, **1993**, 14, 1504-1518. Comparison of atomic charges derived by different procedures.
- (36) Stone, A.J., *Molec. Phys.*, **1977**, 33, 293.
- (37) Stone, A.J.; Alderton, M., *Molec. Phys.*, **1985**, 56, 1047-1064. Distributed Multipole Analysis - methods and applications.
- (38) Mulder, F.; Huiszoon, C., *Molec. Phys.*, **1977**, 34, 1215.
- (39) Stone, A.J.; Price, S.L., *J. Phys. Chem.*, **1988**, 92, 3325-3335. Some new ideas in the theory of intermolecular forces - anisotropic atom atom potentials.
- (40) Hirshfeld, F.L., *Theor. Chim. Acta*, **1977**, 44, 129.
- (41) Ritchie, J.P., *J. Am. Chem. Soc.*, **1985**, 107, 1829-1837. Electron-density distribution analysis for nitromethane, nitromethide, and nitramide.
- (42) Bader, R.F.W., *Acc. Chem. Res.*, **1985**, 18, 9-15. Atoms in molecules.
- (43) Sokalski, W.A.; Poirier, R.A., *Chem. Phys. Lett.*, **1983**, 98, 86-92. Cumulative atomic multipole representation of the molecular charge distribution and its basis set dependence.
- (44) Rein, R., *Adv. Quantum Chem.*, **1973**, 7, 335-396. Physical properties and interactions of polyatomic molecules: with applications to molecular recognition in biology.

- (45) Vigne-Maeder, F.; Claverie, P., *J. Chem. Phys.*, **1988**, 88, 4934-4948. The exact multicentre multipolar part of a molecular charge distribution and its simplified representations.
- (46) Stone, A.J., *Chem. Phys. Lett.*, **1981**, 83, 233-239. Distributed Multipole Analysis; or how to describe a molecular charge distribution.
- (47) Spackman, M.A., *J. Chem. Phys.*, **1986**, 85, 6587-6601. A simple quantitative model of hydrogen-bonding.
- (48) Boys, S.F., *Proc. R. Soc. A*, **1950**, 200, 542-554. Electronic wave functions. I. A general method of calculation for the stationary states of any molecular system.
- (49) Buckingham, A.D.; Fowler, P.W., *J. Chem. Phys.*, **1983**, 79, 6426-6428. Do electrostatic interactions predict structures of Van der Waals molecules?
- (50) Buckingham, A.D.; Fowler, P.W., *Canad. J. Chem.*, **1985**, 63, 2018-2025. A model for the geometries of Van der Waals complexes.
- (51) Hurst, G.J.B.; Fowler, P.W.; Stone, A.J.; Buckingham, A.D., *Int. J. Quantum Chem.*, **1986**, 29, 1223-1239. Intermolecular forces in Van der Waals dimers.
- (52) Price, S.L.; Lo Celso, F.; Treichel, J.A.; Goodfellow, J.M.; Umrana, Y., *J. Chem. Soc., Faraday Trans.*, **1993**, 89, 3407-3417. What base pairings can occur in DNA-a distributed multipole study of the electrostatic interactions between normal and alkylated nucleic-acid bases.
- (53) Medhi, C.; Mitchell, J.B.O.; Price, S.L.; Tabor, A.B., *Biopolymers*, **2000**, 52, 84-93. Electrostatic factors in DNA intercalation.
- (54) Coombes, D.S.; Price, S.L.; Willock, D.J.; Leslie, M., *J. Phys. Chem.*, **1996**, 100, 7352-7360. Role of electrostatic interactions in determining the crystal structures of polar organic molecules. A distributed multipole study.
- (55) Dick, B.G.; Overhauser, A.W., *Phys. Rev.*, **1958**, 112, 90-103. Theory of the dielectric constants of alkali halide crystals.
- (56) Cochran, W., *Proc. Roy. Soc.*, **1959**, A 253, 260-276. Theory of the lattice vibrations of germanium.
- (57) Jansen, G.; Haettig, C.; Hess, B.A.; Angyan, J.G., *Mol. Phys.*, **1996**, 88, 69-92. Intermolecular interaction energies by topologically partitioned electric properties .1. Electrostatic and induction energies in one-centre and multicentre multipole expansions.
- (58) Karlström, G., *Theor. Chim. Acta*, **1982**, 60, 535-541. Local polarizabilities in molecules, based on ab initio Hartree-Fock calculations.

- (59) Celebi, N.; Angyan, J.G.; Dehez, F.; Millot, C.; Chipot, C., *J. Chem. Phys.*, **2000**, *112*, 2709-2717. Distributed polarizabilities derived from induction energies: A finite perturbation approach.
- (60) Stone, A.J., *Molec. Phys.*, **1985**, *56*, 1065-1082. Distributed polarizabilities.
- (61) Le Sueur, C.R.L.; Stone, A.J., *J. Mol. Phys.*, **1994**, *83*, 293-307. Localisation methods for distributed polarizabilities.
- (62) Slater, J.C.; Kirkwood, J.G., *Phys. Rev.*, **1931**, *37*, 682.
- (63) Kumar, A.; Meath, W.J., *Mol. Phys.*, **1992**, *75*, 311-324. Dipole oscillator strength properties and dispersion energies for acetylene and benzene.
- (64) Stone, A.J.; Tong, C.S., *J. Comp. Chem.*, **1994**, *15*, 1377-1392. Anisotropy of atom-atom repulsion.
- (65) Pauling, L., *The nature of the chemical bond*, **1960**, Cornell University Press.
- (66) Bondi, A., *J. Phys. Chem.*, **1964**, *68*, 441-451. Van der Waals volumes and radii.
- (67) Lennard-Jones, J.E., *Proc. Roy. Soc.*, **1924**, *A106*, 441, 463.
- (68) Maitland, G.C.; Rigby, M.; Smith, E.B.; Wakeham, W.A., *Intermolecular forces: their origin and determination*, **1981**, Oxford: Clarendon Press.
- (69) Born, M.; Mayer, J.E., *Z. Phys.*, **1932**, *75*, 1.
- (70) Kita, S.; Noda, K.; Inouye, H., *J. Chem. Phys.*, **1976**, *64*, 3446-3449. Repulsion potentials for Cl⁻-R and Br⁻-R (R=He, Ne and Ar) derived from beam experiments.
- (71) Wheatley, R.J.; Price, S.L., *Molec. Phys.*, **1990**, *69*, 507-533. An overlap model for estimating the anisotropy of repulsion.
- (72) Wheatley, R.J.; Hutson, J.M., *Molec. Phys.*, **1995**, *84*, 879-898. A systematic model potential for Li⁺ - H₂O.
- (73) Wheatley, R.J., *Molec. Phys.*, **1996**, *87*, 1083. The solvation of sodium ions in water clusters: intermolecular potentials for Na⁺ - H₂O and H₂O - H₂O.
- (74) Nobeli, I.; Price, S.L.; Wheatley, R.J., *Molec. Phys.*, **1998**, *95*, 525-537. Use of molecular overlap to predict intermolecular repulsion in N...H-O hydrogen bonds.
- (75) Nobeli, I.; Price, S.L., *J. Phys. Chem. A*, **1999**, *103*, 6448-6457. A non-empirical intermolecular potential for oxalic acid crystal structures.

- (76) Tsui, H.H.Y.; Price, S.L., *CrystEngComm*, **1999**, *7*, A non-empirical method of determining atom-atom repulsion parameters: application to crystal structure prediction of an oxyboryl derivative.
- (77) Mitchell, J.B.O.; Price, S.L., *J. Phys. Chem. A*, **2000**, *104*, 10958-10971. A systematic nonempirical method of deriving model intermolecular potentials for organic molecules: Application to amides.
- (78) Knowles, P.J.; Meath, W.J., *Molec. Phys.*, **1987**, *60*, 1143-1158. A separable method for the calculation of dispersion and induction energy damping functions with applications to the dimers arising from He, Ne and HF.
- (79) Wheatley, R.J.; Price, S.L., *Mol. Phys.*, **1990**, *71*, 1381-1404. A systematic intermolecular potential method applied to chlorine.
- (80) Millot, C.; Stone, A.J., *Mol. Phys.*, **1992**, *77*, 439-462. Towards an accurate intermolecular potential for water.
- (81) Gregory, J.K.; Clary, D.C., *J. Phys. Chem.*, **1996**, *103*, 8924-8930. Three-body effects on molecular properties in the water trimer.
- (82) Kitaigorodsky, A.I., *Molecular crystals and molecules*, **1973**, New York: Academic Press.
- (83) Filippini, G.; Gavezzotti, A., *Acta Cryst.*, **1993**, *B49*, 868-880. Empirical intermolecular potentials for organic crystals: the '6-exp' approximation revisited.
- (84) Gavezzotti, A.; Filippini, G., *J. Phys. Chem.*, **1994**, *98*, 4831-4837. Geometry of the intermolecular X-H...Y (X, Y = N, O) hydrogen bond and the calibration of empirical hydrogen-bond potentials.
- (85) Williams, D.E.; Starr, T.L., *J. Comp. Chem.*, **1977**, *1*, 173.
- (86) Price, S.L., *Chem. Phys. Lett.*, **1985**, *114*, 359-364. A distributed multipole analysis of the charge-densities of some aromatic hydrocarbons.
- (87) Hsu, L.-Y.; Williams, D.E., *Acta Cryst.*, **1980**, *A36*, 277-281. Intermolecular potential-function models for crystalline perchlorohydrocarbons.
- (88) Cox, S.R.; Hsu, L.-Y.; Williams, D.E., *Acta Cryst.*, **1981**, *A37*, 293-301. Nonbonded potential function models for crystalline oxohydrocarbons.
- (89) Williams, D.E.; Cox, S.R., *Acta Cryst.*, **1984**, *B40*, 404. Nonbonded potentials for azahydrocarbons: The importance of the coulombic interaction.

- (90) Williams, D.E.; Weller, R.R., *J. Am. Chem. Soc.*, **1983**, *105*, 4143-4148. Lone-pair electronic effects on the calculated ab initio SCF-MO electric potential and the crystal structures of azabenzenes.
- (91) Williams, D.E.; Houpt, D.J., *Acta Cryst.*, **1986**, *B42*, 286-295. Fluorine nonbonded potential parameters derived from crystalline perfluorocarbons.
- (92) Williams, D.E.; Gao, D., *Inorg. Chem.*, **1997**, *36*, 782-788. Effects of molecular electric potential and anisotropic atomic repulsion in the dichlorine dimer and crystalline chlorine.
- (93) Williams, D.E., *J. Molec. Struct.*, **1999**, *485-486*, 321-347. Improved intermolecular force field for crystalline hydrocarbons containing four- or three-coordinated carbon.
- (94) Williams, D.E., *J. Comp. Chem.*, **1994**, *15*, 719-732. Failure of net atomic charge models to represent the Van-der-Waals envelope electric-potential of n-alkanes.
- (95) Williams, D.E.; Abraha, A., *J. Comp. Chem.*, **1999**, *20*, 579-585. Site charge models for molecular electrostatic potentials of cycloalkanes and tetrahedrane.
- (96) Williams, D.E., *J. Comp. Chem.*, **2001**, *22*, 1-20. Improved intermolecular force field for crystalline oxohydrocarbons including O-H...O hydrogen bonding.
- (97) Coombes, D.S., *Philosophical Magazine B*, **1996**, *73*, 117-125. Deriving intermolecular potentials for predicting the crystal structures of polar molecules.
- (98) Mitchell, J.B.O.; Price, S.L., *J. Comp. Chem.*, **1990**, *11*, 1217-1233. The nature of the N-H O=C hydrogen-bond - an intermolecular perturbation theory study of the formamide formaldehyde complex.
- (99) Price, S.L.; Stone, A.J., *Molec. Phys.*, **1982**, *47*, 1457-1470. The anisotropy of the Cl₂-Cl₂ pair potential as shown by the crystal structure - evidence for intermolecular bonding or lone pair effects.
- (100) Price, S.L.; Stone, A.J.; Lucas, J.; Rowland, R.S.; Thornley, A.E., *J. Am. Chem. Soc.*, **1994**, *116*, 4910-4918. The nature of Cl...Cl intermolecular interactions.
- (101) Mooij, M.T.W.; Van Duijneveldt, F.B.; Van Duijneveldt-van de Rijdt, J.G.C.M.; Van Eijck, B.P., *J. Phys. Chem. A*, **1999**, *103*, 9872-9882. Transferable ab initio intermolecular potentials. 1. Derivation from methanol dimer and trimer calculations.

Chapter 3

Prediction of molecular crystal structures

*'...the supramolecular Everest of predicting the crystal structure of any given molecule is as challenging and formidable a task today as the total synthesis of vitamin B12, ginkgolide or palytoxin were in earlier times for organic synthesis.'*¹

In this chapter 'crystal structure prediction of small organic molecules' is reviewed. The question if 'a reliable prediction of the structure(s) actually adopted' still belongs to 'the realm of wishful thinking'² is approached by a brief discussion of the definition and the scientific and commercial importance of crystal structure prediction, followed by an overview of current methods and prediction programs and by a synopsis of published studies. Pioneering work to simulate the different stages involved in the nucleation process is briefly discussed.

3.1 Introduction

The expression ‘crystal structure prediction’ is not used unambiguously in the literature. This is not only caused, as for any scientific term, by an evolution with time, but it is the result of a problematic fusion³ of the terminology used in crystal structure prediction and crystal engineering.⁴ The ultimate goal of crystal structure prediction could be summarised as the unfailing reliable determination of the full solid state phase diagram of any organic compound using just its atomic connectivity. This exclusion of any experimental data is currently – possibly not unproblematically - described as ‘*ab initio*’ prediction. A prediction study would not only include the unequivocal detection of the most stable and all possible metastable polymorphs at a given temperature (or pressure) but also the determination of the phase which will actually appear under given crystallisation conditions.² The prediction should be feasible *a priori*, i.e. prior to the synthesis of the molecule.

Such a prediction would be highly desirable to any molecular material industry as different polymorphic forms of a given compound generally have significantly different solid state properties (chapter 1.2.3) and a basic requirement for a theoretical investigation of these properties is the precise knowledge of the crystal structure. The scientific investigations towards this ultimate goal will certainly reveal some of the fundamental processes that determine the crystallisation of organic molecular crystal structures. Furthermore, the methodology developed in crystal structure prediction can also be applied to determine crystal structures from powder diffraction data (chapter 3.6).

Different opinions can be found in the scientific literature on whether crystal structures are predictable. The endlessly cited Nature editorial by John Maddox from 1988⁵ where he indicated that ‘*One of the continuing scandals in the physical sciences is that it remains in general impossible to predict the structure of even the simplest crystalline solids from a knowledge of their chemical composition*’ actually referred mainly to inorganic solids. In 1990, Hawthorne confirmed⁶ ‘*Thus rigorous general solution to the question of crystal structure prediction may not be forthcoming in the near future.*’ Concerning molecular crystals, Fagan and Ward⁷ stated in 1992 that ‘*Unfortunately the molecular interactions in a lattice become too complicated for any human or computer to handle*’. Two years later Gavezzotti⁸ dissected the global problem into smaller challenges and discussed the related achievements and quests. In

spring 1999, the ability of currently available methods to predict crystal structures was tested by the first crystal structure prediction blind test for small organic molecules organised by the Cambridge Crystallographic Data Centre (CCDC).⁹ Eleven participants were invited to submit three crystal structures for each of the three experimentally determined but so far unpublished compounds. Although there were some successes, none of the participants predicted all structures correctly (For the study of compound I and the overall results of this blindtest see chapter 5).

The majority of current crystal structure prediction programs (Table 3.1) are based on lattice energy minimisation, with the assumption that the global lattice energy minimum found in the search corresponds to the stable polymorph and other low lattice energy minima to metastable forms. The number of low energy structures often seems to depend on the thoroughness of the search¹⁰ and the particular ordering of these structures is heavily dependent on the force field employed. For large irregularly shaped molecules, in contrast to small spherical molecules, there seems to be a better chance that the experimental structure corresponds to the global minimum.³ But generally, the recognition of absolute stability is often impossible.¹¹ A partial success criterion could therefore be defined as the ranking of the experimental structure as one of the most favourable candidate structures. Mooij suggested this process is more appropriately described as crystal structure generation.¹²

3.2 Crystal structure prediction based on lattice energy minimisation

There are four major elements to any crystal structure prediction program based on lattice energy minimisation: the input molecular model which may allow for conformational flexibility, a method for searching for crude initial crystal structures, a potential energy function and an efficient means of evaluating this function for the generated structures to locate lattice energy minima.

The molecular structure is often – as it would be for *a priori* crystal structure prediction – represented by an *ab initio* optimised rather than an experimentally determined model to not bias the search towards one of the polymorphic forms. Furthermore, the structures of negligibly flexible molecules are often treated as rigid bodies. If flexibility of the model is included, it is usually controlled by rotation of rigid fragments about selected bonds, as in the programs CRYSCA and MPA (the programs are explained in detail below). In some cases, as for UPACK and the Polymorph

Predictor, parts of the molecular model can be relaxed in the crystal structure at some point near the calculated minimum energy packing. An investigation of the influence of the molecular model on the lattice energy minima was published on an organo-metallic system.¹³ A further investigation for a molecular crystal, paracetamol, has been carried out in this thesis (chapter 6).

The generation of crude crystal structures can either be done systematically (ICE9, MOLPAK, MPA, PROMET, UPACK, method by Hammond *et al.*, method by Perlstein) or randomly (CRYSCA, MPA, MDCP, UPACK, PMC, Polymorph Predictor, method by Shoda, method by Scheraga, method by Chin). The efficiency and general applicability varies for each of the particular implementations, which are explained in detail below.

Lattice energy minimisation is an important and generally the most time consuming step of the crystal structure prediction process. Many crude structures can be generated rapidly, but all must be minimised to obtain low energy crystal structures. Several minimisation algorithms¹⁴ are available, as steepest descent (CRYSCA, MDCP), conjugate gradient (UPACK), simulated annealing (Polymorph Predictor, method by Chin), molecular dynamics (MDCP) or even a combination of different methods (MPA, UPACK) to improve speed. The minimisation program DMAREL employed in this thesis is further discussed in chapter 4.

The lattice energy of polar crystals, *i.e.* crystals where the unit cell has a nonzero dipole moment, depends on the external shape of the crystal and therefore can in principle not be calculated without the consideration of the macroscopic crystal shape.¹⁵⁻¹⁷ It has been suggested¹⁵ that correction terms to the lattice energy can be omitted in crystal structure prediction. as it is very likely that the crystal will adopt an energetically optimal shape (a needle, with the dipole moment directed along the needle axis, or a platelet with the dipole moment in its plane) or that the charges on the crystal surface are counterbalanced by external charges.

At different stages of a crystal structure prediction, it may be necessary to reduce the number of structures under consideration by eliminating similar arrangements. The general problem and possible clustering methods have been discussed in 1.5.4 -1.5.5.

3.2.1 Implementations of crystal structure prediction methods

A variety of different crystal structure prediction programs have been developed and several reviews about these methods and their applications have been published.¹⁸⁻²⁰ A summary of each approach is given in table 3.1 together with the number of molecular crystal structures studied.

The method by Perlstein is based on the work of Scaringe^{21, 22} and Scaringe and Perez²³ to predict the one- and two-dimensional structure of rigid organic molecules in crystals. Using Kitaigorodskii's aufbau principle²⁴, Perlstein's algorithm packs molecules into one-dimensional stable aggregates²⁵⁻²⁷ by applying the symmetry operations of translation, glide plane, screw axis and inversion. This is followed by the packing of each of these aggregates into layers.²⁸ Scaringe showed that seven types of layers occur most frequently²² and the combination of these types of layers yields crystal structures in the most frequently occurring space groups.^{29, 30} Perlstein employed a Monte Carlo procedure to search for low energy aggregates with the orientation of the molecules and the repeat distances as variables. The molecules are treated as rigid bodies with the exception of dihedral angles and simulated annealing is used to evaluate the lattice energy.

This approach is similar to the systematic search program Promet³¹ which builds up dimers and chains of two to four rigid molecules by applying the same symmetry operators as Perlstein. The energetically most promising clusters are then used to generate three-dimensional trial crystal structures by applying translational symmetry. These approximate crystal structures are then refined by rigid body lattice energy minimisation.³² A sometimes unfavourable consequence of this aufbau style is that it cannot reach crystal structures in which no particularly stable substructures are present.

The program MOLPAK (Molecular Packing)³³ has been developed to search systematically for densely packed structures. A rigid central molecule in a variable orientation is surrounded by a coordination sphere consisting of fourteen molecules applying a repulsion-only potential function. Different types of common coordination spheres, representing different crystal symmetries with one molecule in the asymmetric unit, were obtained from an analysis of 242 crystal structures containing nitrogen, oxygen and fluorine retrieved from the CSD. This analysis also established fourteen as

the most probable number of molecules in a coordination sphere, *i.e.* molecules which are in contact or close to van der Waals contact with the central molecule. The orientation of the central molecule is systematically altered in 10° increments, about three Eulerian axes over a 180° range, the densest structures are refined to 2°, and the procedure is repeated to ultimately produce a three-dimensional map of minimum unit cell volume as a function of the central molecule's orientation. The most densely packed arrangements, usually 25 to 50 of each coordination group, are then refined by lattice energy minimisation using the program WMIN³⁴ or, as in this thesis, in combination with DMAREL (chapter 4)³⁵, a program which enables the use of a DMA electrostatic model in the crystal structure calculation. MOLPAK currently allows for the search in 23 coordination groups covering the space groups *P1*, *P1̄*, *P2₁*, *P2₁/c*, *C2*, *C2/c*, *P2₁2₁2₁*, *Pca2₁*, *Pna2₁*, *Pbca*.

A systematic grid search program, ICE9,³⁶ generates close packed crystal structures by applying full space group symmetry to a given number of independent rigid molecules in a unit cell, followed by an initial energy minimisation with a (6-12) Lennard-Jones pair potential. The structures are then refined by a further energy minimisation, using a cutoff radius and a molecular multipole expansion for the electrostatic interactions.

A further systematic grid-based search method has been developed by Hammond *et al.*³⁷ It has been extensively used in combination with experimental X-ray diffraction data,³⁸ but also work purely based on lattice energy criteria has been published.³⁹

The UPACK (Utrecht crystal PACKer) program allows for a systematic grid search^{40, 41} or a random search⁴² and it has been extended to handle more than one molecule in the asymmetric unit.⁴² The program was originally developed for the specific problem of predicting crystal structures of pyranoses, flexible molecules which form hydrogen-bonded structures. Therefore UPACK is able to treat the conformational freedom of hydroxyl hydrogen atoms automatically during the generation of trial structures. For the clustering of generated structures two different algorithms have been developed.^{43, 44} In a comparative test the random search method was found to be approximately equally as efficient as the systematic grid search. An advantage of the random search is that the number of equivalent structures obtained is an indication of

the completeness of the search. Furthermore it has been observed that the number of equivalent structures decreases with increasing energy and that the deepest minima in the potential energy landscape are broader than others. It has been suggested that this might be a reason that random searches are often successful, although they might not be complete.⁴²

The program MPA (Molecular Packing Analysis)⁴⁵ also allows for a systematic⁴⁶ and a random⁴⁷ search. In the latter approach a Monte Carlo procedure is used to randomly orient a given number of independent molecules in an expanded trial unit cell without imposing any space group symmetry. A combination^{47, 48} of rigid body minimisation techniques is then used to minimise the crude crystal structures. The program mpg (molecular packing-group method)^{46, 49, 50}, an extension of the MPA method, takes one molecule in the unit cell as independent and the others related by assumed space group symmetry. Furthermore it reduces the number of possible space groups by considering the relationship between the assumed number of molecules in the asymmetric unit and in the unit cell.

In the random search program CRYSCA (CRYStal structure CAlculation)⁵¹⁻⁵³ a steepest descent algorithm is used to minimise symmetry-constrained crude crystal structures allowing for intramolecular rotation of rigid fragments about selected bonds. This method has been applied to a variety of organo-metallic compounds and more recently to pigments.

A further random search program PMC (Packing of Molecules in Crystals) has been developed by Dzyabchenko.⁵⁴

The only commercial crystal structure prediction program is the Polymorph Predictor,^{19, 55-58} a component of the Cerius2 software by Molecular Simulations (MSI) Limited. It employs a Monte Carlo-Simulated Annealing (MC-SA)⁵⁹ technique to randomly search for crude packing arrangements of rigid molecular structures. A new trial structure is accepted or rejected according to the Metropolis algorithm⁶⁰ by comparing the energy of a new structure with that of the last accepted structure. A structure with a lower energy than the former one will always be accepted and one with a higher energy has a probability of being accepted dependent on the energy difference and the current 'temperature' of the simulation. Hence energy barriers can be overcome, which is especially important in the beginning of the simulation. A search finishes when

the simulation temperature is so low that a certain number of structures are rejected and the crystal is 'frozen'. A clustering procedure based on interatomic distances is followed by a lattice energy minimisation of the remaining crude structures. This symmetry-constrained minimisation allows for molecular flexibility. Finally the created structures are clustered again to remove duplicate structures.

A very similar approach by Chin *et al.*⁶¹ builds on the former procedure, but adds constraints of selecting structures with 'reasonable' intermolecular hydrogen bond geometries. A requirement of the method is the prior identification of hydrogen bond donors and acceptors in the molecular structure. This approach is referred to as Hydrogen-Bond Biased Simulated Annealing Monte Carlo (HBB-SAMC). In the lattice energy minimisation stage intramolecular degrees of freedom are also considered.

A further random search method, MDCP (Molecular Dynamics for Crystal Packing)⁶² uses constant pressure molecular dynamics (MD) to search for crude structures which are, after a clustering process, subjected to a steepest-descent rigid body lattice energy minimisation.

A random search technique based on potential energy smoothing methods has been developed by Scheraga *et al.*^{16, 63, 64} The potential energy surface is transformed until only a few traces of the deepest minima of the original surface remain. These deepest minima are recovered in a reversing procedure, in which the deformation is gradually removed.

3.3 Prediction methods based on statistical distributions of CSD data

A few methods have recently been developed based on the statistical distribution of chosen subsets of molecular structures retrieved from the CSD. As the majority of potential energy functions are derived by empirical fitting to experimental data, this approach is not clearly distinct from methods based on lattice energy minimisation. Unfortunately, the CSD does not store any systematic information about crystallisation conditions, as crystallisation method or solvent, although these conditions often determine which polymorph crystallises. A systematic consideration of this so far neglected information would certainly be an important contribution towards this subject. In general prediction methods based on CSD statistical data have so far not been tested as widely as the conventional lattice energy procedures.

In the program FlexCryst⁶⁵⁻⁶⁷ the inverse Boltzmann equation is applied to draw the relationship between the frequency of statistical distributions of observed interatomic distances and energies to derive a scoring function in form of an atom-atom pair potential. Crystal structures are constructed systematically by translating clusters of molecules on a grid and these structures are ranked, without any further minimisation, according to the scoring function. The use of an atom-atom pair potential function to evaluate energies and the lack of structure minimisation makes the method very fast, possibly at the expense of accuracy. The method has mainly been tested on structures in the space groups $P1$ and $P\bar{1}$ and for these cases it has been very successful in finding the experimental structure among the lowest energy structures, but just moderate success was reported for locating the observed structure as the global minimum.

Rancel⁶⁸, as FlexCryst, uses distributions of intermolecular distance data from the CSD, but the statistical data is derived from a limited number of manually selected structures. These distributions of distances are then divided into different zones and employed to derive a cost function which is minimised by the genetic algorithm (GA) method of differential evolution.⁶⁹ In this procedure, penalty functions are employed to handle cases with closest contacts smaller than the distances observed in the CSD data or when expected interactions do not occur in the search.

The program PackStar⁷⁰ takes, in addition to the distribution of intermolecular distances, the directionality of interactions into account. It is built on the IsoStar library⁷¹ of intermolecular interactions and employs the stored scatterplots to derive propensity vector maps of interactions and their directionalities. A cost function based on these propensity maps is minimised by simulated annealing to maximise the fit with the CSD data.

3.4 Studies based on lattice energy minimisation

A variety of different molecular crystal structures have been investigated by crystal structure prediction methods based on lattice energy minimisation. Table 3.2 gives a summary of most of the studies published from 1984 to the last quarter of 2000, excluding the results of the CCDC crystal structure prediction blindtest (chapter 5) and the investigations done for this thesis. The molecular crystals are sorted into groups (alcohols, carboxylic acids, drugs, hydrocarbons, molecular magnets, pigments, polyalcohols, pyranoses, other and hydrates) and the systems in each group are listed by

publication year. The characteristics and results of each study are briefly described in the table. The significance of the fact that the experimental structure has been found in the search among the low energy minima (a), or even as the global minimum (c) decreases for cases where the search has just been performed in the experimental space group (d). In 93 out of 228 searches for a particular polymorphic form, the study has successfully found the experimental structure among the low energy minima, though not as the global minimum. In 72 cases the polymorphic form corresponded to the global search minimum. Sometimes authors report that the experimental structure has been found in the search, but no further information about other low lattice energy minima and their relative energies is provided (b). This has been observed for 21 search studies. This list does not include any studies merely aiming at structure solution from powder diffraction data (chapter 3.6), but it comprises two studies where Rietveld refinement has been used to search for a subset of the known polymorphic forms (e) and seven investigations where this technique has been employed *in addition to* crystal structure prediction (+e). Eight cases have been published where unindexed powder diagrams are employed to compare search structures to the experimentally observed forms (f). One could assume that unsuccessful studies are less likely to be published. This synopsis shows that for 32 out of 228 studies, the employed procedure was reported to have failed to locate the experimental structure of at least one of the polymorphic forms (g). For three of these 'failures' (n-hexane, isoiridomyrmecin and urea) more recent re-investigations were able to locate the experimental structure. The question if crystal structures are predictable should best be investigated by *a priori* studies. For seven polymorphic forms the prediction has been made in advance of any experimental investigation (k).

3.5 Additional approaches towards solving the prediction problem

The search for lattice energy minima is only a first step towards the prediction of small organic crystals. Locating a structure at or near the global minimum is only a necessary, but not sufficient condition for a form to be found experimentally as nucleation is a kinetic process. Current experimental techniques are so far not able to provide any information about the early stages of crystal formation of organic molecules within a solution. Pioneering computational work has been done by Gavezzotti who studied the evolution with time of different nuclei of tetrolic acid in carbon tetrachloride solution by Molecular Dynamics (MD) simulation.⁷² Tetrolic acid is known to

crystallise in two forms, a chain and a cyclic dimer. Gavezzotti observed both motifs next to other clusters in the solution, with the catemer developing from dimer structures by rupture of one of the two hydrogen bonds in so-called ‘catemer-jumps’. The observation of small clusters in the solution of carboxylic acids is also confirmed by *experimental* work.⁷³ As a variety of different arrangements has been found in the simulation, the appearance of a motif in the solution seems to be a necessary but not sufficient condition for its growth into a crystal structure.

In a similar study for acetic acid⁷⁴ similar types of clusters of solute molecules and ‘catemer-jumps’ were observed. As a hypothesis based on the character of the Molecular Dynamics (MD) simulation, the formation of a liquidlike state in which order occurs at a later stage, has been postulated as a first nucleation step. The formation of such a state seems to be more controlled by molecular diffusion and solvent reorganisation rather than long-range interactions of the solute molecules. Unfortunately, the study cannot reveal information on further stages of the nucleation process, as for example the subsequent formation of the chain structure found in the experimental form.

A simulation on 2-pyridone⁷⁵ in a carbon tetrachloride solution leads to a variety of clusters and also in this case ‘catemer-jumps’ can be observed. The hereby created catemer arrangement can be found in the final crystal structure.

The simulation of an entire nucleation process is currently not feasible due to timescale limitations. But the study of the *melting* behaviour of a molecular crystal is possible by a simulated heating procedure. In the case study of the melting of acetic acid⁷⁶ the lattice disaggregates at a certain temperature and the system evolves quickly into an isotropic liquid with fluctuating hydrogen bonds. No selective features, as persistent hydrogen bonds could be observed and the author suggested therefore that the formation of such intermolecular bonds in crystal *nucleation* is part of a later stage.

3.6 Structure determination from powder data

If a suitable - in size and quality - single crystal of a compound is available, standard single-crystal X-ray diffraction⁷⁷ or neutron techniques can be applied to solve the structure. However, for yet unknown reasons, crystals of a reasonable size cannot be grown for all molecules. X-ray powder diagrams can be obtained from powders with crystal sizes below 1 μm ⁷⁸, but they contain the structural information

compressed into a one-dimensional space. This leads to a considerable overlap of peaks in the diagrams and makes it more difficult to extract the intensities of individual diffraction maxima. If a powder diagram can be indexed, *i.e.* the positions of the peaks can be determined accurately, the possible cell parameters and space groups can be assigned. In this case, conventional approaches, as the Patterson method or direct methods, can be applied to try to solve the structure.⁷⁹ These methods are routinely used in single-crystal diffraction and require the extraction of intensities of individual reflections. This limits their use in powder diffraction to small molecules. A different type of approach, based on the search techniques derived in crystal structure prediction, is the theoretical derivation of structural models and the comparison of their powder diagrams with the experimental diagram. The advantage of this method is that the powder diagrams are compared as a whole using a profile R-factor and therefore no individual diffraction intensities have to be extracted. The generation of trial structures for this purpose has been realised by different methods, as random Monte Carlo sampling⁸⁰ or the application of genetic algorithms.⁸¹⁻⁸⁴ If the pattern cannot be indexed, these latter techniques could still be applied, but a complete crystal structure prediction study would have to be performed.⁸⁵⁻⁸⁸ In all cases, the promising calculated structures have to be subjected to a refinement process, as Rietfeld refinement.⁸⁹

3.7 Conclusions

The prediction of molecular crystal structures has so far been mainly approached by methods based on lattice energy minimisation. As nucleation is a kinetically driven process, for a structure to be at or near the global minimum in the search is a necessary but not sufficient criterion for it to be found experimentally. A literature survey of crystal structure prediction studies gives an overview of the characteristics and results of each investigation. Pioneering simulation work to reveal some insight into the first stages of nucleation and also prediction methods based on the distribution of experimental data retrieved from the CSD are promising starting points to try out new ideas in this challenging scientific area. The methodologies applied in this work are further described in the following chapter.

search program	search type	minimisation of crude structures	intramolecular flexibility	fitness function	coulombic interactions included	Ewald summation	clustering	application of symmetry	No of published investigated structures
Chin	random	yes	yes	force field	yes	yes	yes	yes	3
CRYSCA	random	yes	yes	force field	yes	no	no	yes	1
Hammond	systematic	yes	yes	force field	no	no	no	yes	2
ICE9	systematic	yes	no	force field	molecular multipole moments	no	no	yes	15
MDCP	random (MD)	yes	no	force field	yes	yes	no	no	4
MOLPAK	systematic (density)	yes	no	force field	minimisation	no	no	yes	10 (WMIN) 7 (DMAREL)
MPA	random (MD) or systematic	yes	no	force field	yes	yes	yes	no	4
mpg	random (MD) or systematic	yes	no	force field	yes	yes	yes	yes	11
Perlstein	random	yes	yes	force field	yes	no	no	yes	3
PMC	random	no	no	force field	yes	no	yes	yes	1
Polymorph Predictor	random (MC)	yes	yes	force field	yes	yes	yes	yes	24
PROMET	systematic (cluster)	yes	no	force field	in final stages only	yes	yes	yes	23
Scheraga	random	yes	no	force field	yes	yes	no	yes	1
UPACK	systematic or random	yes	yes	force field	yes	no	yes	yes	40 (systematic) 11 (random)
FlexCryst	systematic	no	no	scoring function	no	no	yes	yes	226
PackStar	GA	yes	no	cost function	no	no	yes	yes	in preparation
Rancel	GA	yes	no	cost function	no	no	no	yes	6

Table 3.1 Methods based on lattice energy minimisation, plus methods based on the statistical distribution of CSD data.

The main characteristics of each method, as the search type, information on the minimisation procedure (if applied) and the clustering process are listed together with information about the total number of investigated molecular crystals (only published cases).

molecule	name	No of known polymorphs	space group	Z' (Z)	program	year	reference	comment
Alcohols								
<chem>HO-CH3</chem>	methanol	3	α $P2_12_12_1$ β disordered γ not determined	1 (4)	- Polymorph Predictor + ab initio derived ff	1999	90	for α : a
<chem>H3C-CH2OH</chem>	ethanol	1	Pc $P2_1/c$ (high pressure)	2 (4) 1 (4)	- Polymorph Predictor + ab initio derived ff - UPACK (random)	1999 2000	90 42	for both: a for both: a
Carboxylic acids								
	terephthalic acid	2	$P\bar{1}$ $P\bar{1}$	0.5 (1) 0.5 (1)	- Polymorph Predictor	1996	91	for both: g, +e
	tetrolic acid	2	α $P\bar{1}$ (dimer) β $P2_1$ (chain)	1 (2) 1 (2)	- PROMET UPACK (systematic)	1997	72	for both: a, d
	acetic acid	2	I $Pna2_1$ (chain) II $P2_1/c$ (dimer, highpressure)	1 (4)	- Polymorph Predictor - Polymorph Predictor	1998 1998	92 44	for both: a, +e for both: m, a, f
	bromoacetic acid	2	I $P2_1/c$ (dimer) II $Pccn$ (dimer)	1 (4) 1 (8)	- Polymorph Predictor	1998	92	for both: g
	chloroacetic acid	2	α $P2_1/c$ (tetramer) β $P2_1/c$ (dimer)	2 (8) 1 (4)	- Polymorph Predictor	1998	92	for both: a
	fluoroacetic acid	1	$P2_1/c$ (dimer)	1 (4)	- Polymorph Predictor	1998	92	a
Drugs								
	aspirin	1	$P2_1/c$	1 (4)	- PROMET - Polymorph Predictor	1995 1999	93 94	c a
	estrone	3	$P2_1$ $P2_12_12_1$ $P2_12_12_1$	2 (4) 1 (4) 1 (4)	- Polymorph Predictor	1996	91	for all: b
	paracetamol	3	I $P2_1/c$ II $Pbca$ III not determined	1 (4) 1 (8)	- Polymorph Predictor	1996	19	for I: c for II: a

Table 3.2 (continued)

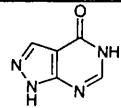
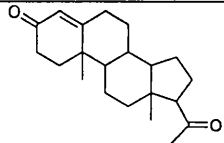
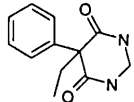

	allopurinol	1	$P2_1/c$	1 (4)	- MOLPAK/DMAREL	1997	95	c
	progesterone	2	$\alpha P2_12_12_1$ $\beta P2_12_12_1$	1 (4) 1 (4)	- Polymorph Predictor	1999	96	for α : c, d for β : a, d
	primidone	2	A $P2_1/c$ B $Pbca$	1 (4) 1 (8)	- Polymorph Predictor	1999	96	for A: a, d, +e for B: a, d
Hydrocarbons								
	benzene	5	I $Pbca$ III $P2_1/c$ (high pressure) II not determined IV not determined	0.5 (4) 0.5 (2)	- PMC - MPA - MDCP - ICE9 - MPA - Polymorph Predictor - UPACK (systematic)	1984 - 1989 1994- 1995 1995 1995 1996 1996 1996 1996 1998	97-101 102, 103 62 104 36 47 105 91 106	I: c, III: a II: a, k low pressure: I: c, III: a high pressure: <1.4GPa: III: c >1.4GPa I: c I: c, III: a I: c, d III: a, d for I: a for I: b for III: b for II: c, k (0.5 GPa < p < 1.0 GPa) (special investigation for form II) I: b, III: b I: c, III: a II: f, k

Table 3.2 (continued)

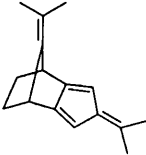
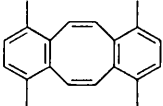
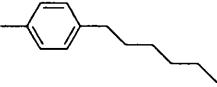
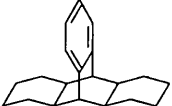
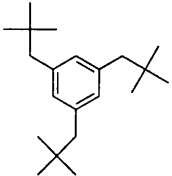
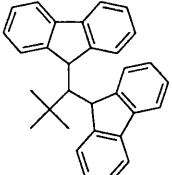
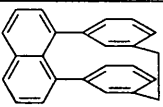
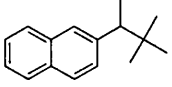
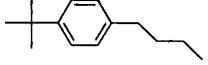
	4,5,6,7-Tetrahydro-2,8-diisopropylidene-4,7-methano-2H-indene	1	$P\bar{I}$	1 (2)	- PROMET	1991	31	a, d
	1,4,7,10-Tetramethyldibenzo (a,e) cyclo-octene	1	$P\bar{I}$	1 (2)	- PROMET	1991	31	a
	1-Methyl-4-hexylbenzene	1	$P\bar{I}$	1 (2)	- PROMET	1991	31	a, d
	cis, cis-Dodecahydrotritycene	1	$P\bar{I}$	1 (2)	- PROMET	1991	31	a, d
	1,3,5-Trineopentylbenzene	1	$P\bar{I}$	1 (2)	- PROMET	1991	31	g, d
	9-t-Butyl-9 (9-fluorenyl) fluorene	1	$P\bar{I}$	1 (2)	- PROMET	1991	31	g, d
	(2.0.0)(1,3) Benzeno (1,8) naphthaleno (1,3) benzenophane	1	$P\bar{I}$	1 (2)	- PROMET	1991	31	g, d
	2,2-Dimethyl-3-(2-naphyl)-butane	1	$P2_1$	1 (2)	- PROMET	1991	31	g, d
	1-tert-Butyl-4-n-butylbenzene	1	Pc	1 (2)	- PROMET	1991	31	g, d

Table 3.2 (continued)

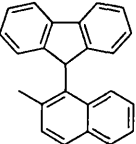
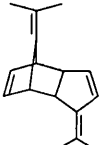
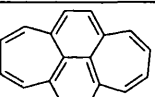
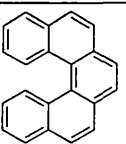
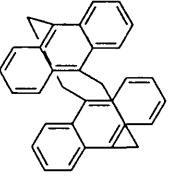
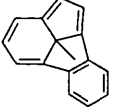
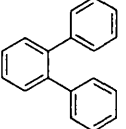
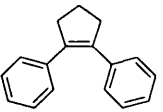
	9-(2-Methyl-1-naphthyl)-fluorene	1	$P2_1/n$	1 (4)	- PROMET	1991	31	g, d
	9,10-Di-isopropylidene-tricyclo (4.3.0.1 ^{2,5}) deca-3,7-diene	1	$P2_12_12_1$	1 (4)	- PROMET	1991	31	c
	Dicyclohepta(de,ij)naphthalene	1	$P2_1/c$	1 (4)	- PROMET	1991	31	c
	3,4-5,6-Dibenzophenanthrene	2	$P2_1/c$ $C2/c$	1 (4) 1.5 (12)	- PROMET	1991	31	for $P2_1/c$ g, d
	Bi (anthracene-9,10-dimethylene)	1	$P2_1/c$	0.5 (2)	- PROMET	1991	31	g, d
	9c-Methyl-9cH-cyclopenta(jk)fluorene	1	$P2_12_12_1$	1 (4)	- PROMET	1991	31	g, d
	1,2-Diphenylbenzene	1	$P2_12_12_1$	1 (4)	- PROMET	1991	31	a, d
	1,2-Diphenylcyclopentene	1	$P2_1$	1 (2)	- PROMET	1991	31	g

Table 3.2 (continued)

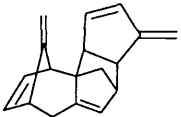

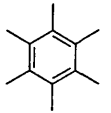
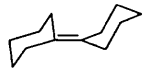
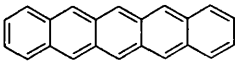
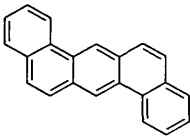
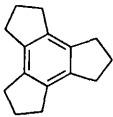
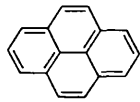


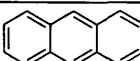
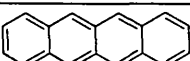
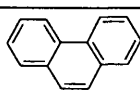
					- PROMET	1991	31	k
	ethene	1	$P2_1/n$	0.5 (2)	- Polymorph Predictor	1992	55	b
	hexamethylbenzene	1	$P\bar{1}$	0.5 (1)	- Polymorph Predictor	1992	55	b
	bicyclohexylidene	1	$P\bar{1}$	0.5 (1)	- ICE9	1996	36	g
	pentacene	1	$P\bar{1}$	1 (2)	- ICE9	1996	36	g
	1:2:5:6-dibenzanthracene	2	A $P2_1$ B $Pcab$	1 (2) 0.5 (4)	- ICE9	1996	36	for A: k for B: g
	trindan	1	$P2_1/c$	1 (4)	- ICE9	1996	36	g
	pyrene	1	$P2_1/a$	1 (4)	- ICE9	1996	36	a
	durene	1	$P2_1/a$	0.5 (2)	- ICE9	1996	36	a
	naphthalene	1	$P2_1/a$	0.5 (2)	- ICE9	1996	36	a
	anthracene	1	$P2_1/a$	0.5 (2)	- ICE9	1996	36	a
	tetracene	1	$P\bar{1}$	1 (2)	- ICE9	1996	36	g
	phenanthrene	2	$P2_1$ $P2_1/c$	1 (2) 1 (4)	- ICE9	1996	36	for $P2_1$: g

Table 3.2 (continued)

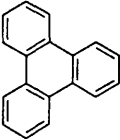
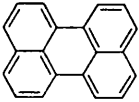






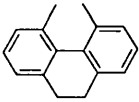
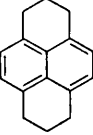

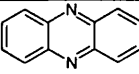
	triphenylene	1	$P2_12_12_1$	1 (4)	- ICE9	1996	36	a
	perylene	2	$\alpha P2_1/a$ $\beta P2_1/c$ (not fully determined)	1 (4) 0.5 (2)	- ICE9 - Hammond	1996 1999	36 39	for α : a for α : c for β : k
	n-hexane	1	$P\bar{1}$	0.5 (1)	- ICE9 - mpg	1996 1999	36 107	g b
	n-octane	1	$P\bar{1}$	0.5 (1)	- ICE9 - mpg	1996 1999	36 107	a b
	n-pentane	1	$Pbcn$	0.5 (4)	- mpg	1999	107	b
	adamantane	1	$P\bar{4}2_1/c$	0.25 (2)	- mpg	1999	107	b
	iceane	1	$P6_3/m$	(2)	- mpg	1999	107	b
	Pentaspiro[2.0.0.2.0.2.0.0.2.0]tridecane	1	$Pbcn$	0.5 (2)	- mpg	1999	107	b
	4,5 Dimethyl-9,10-dihydrophenanthrene	1	$P\bar{1}$	1 (2)	- mpg	1999	46	c
	sym-hexahdropyrene	1	$P2_1/c$	0.5 (2)	- mpg	1999	46	c
	1,2-dihydrocyclobutabenzene	1	$P\bar{1}$	1 (2)	- mpg	1999	46	c
	phenazine	2	$\alpha P2_1/n$ $\beta P2_1/n$ (not determined)	0.5 (2) 1 (4)	- Hammond	1999	39	for α : c for β : k

Table 3.2 (continued)

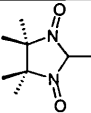
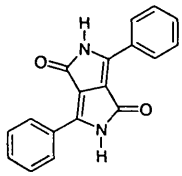
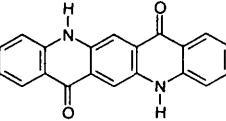
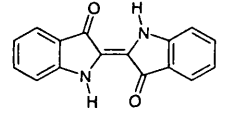
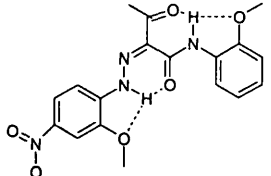
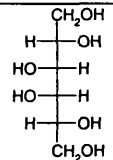
Molecular magnets								
	4,4,5,5-tetramethyl-4,5-dihydro-1H-imidazol-1-oxyl 3-oxid (2-hydronitronitroxide radical) HNN	2	α : $P2_1/c$ β : $P2_1/c$	1 (4) 4 (16)	PROMET	1999	85	for α : c, f
Pigments								
	pigment red	1	$P-1$	0.5 (1)	- Polymorph Predictor	1993	108	a, d,
	quinacridone	3	α (not determined) β (not determined) γ $P2_1/c$	0.5 (2)	- Polymorph Predictor	1996	91	for γ : b for α , β : e
	indigo	2	A $P2_1/c$ B $P2_1/c$	0.5 (2) 0.5 (2)	-MOLPAK/DMAREL	1998	109	for A: c for B: a
	pigment yellow 74	1	$P-1$	1 (2)	-CRYSCA	1999	78	c
Polyalcohols								
	Galactitol	1	$P2_1/c$	1 (4)	- UPACK (systematic)	1999	41	a, d

Table 3.2 (continued)

$ \begin{array}{c} \text{CH}_2\text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{CH}_2\text{OH} \end{array} $	D-Glucero-L-galacto-heptitol	1	$P2_1$	1 (2)	- UPACK (systematic)	1999	41	c, d
$ \begin{array}{c} \text{CH}_2\text{OH} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{CH}_2\text{OH} \end{array} $	D-Altritol	1	$P2_1$	1 (2)	- UPACK (systematic)	1999	41	a, d
$ \begin{array}{c} \text{CH}_2\text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{CH}_2\text{OH} \end{array} $	D-Glycero-L-gulo-heptitol	1	$P2_12_12$	1 (4)	- UPACK (systematic)	1999	41	c, d
$ \begin{array}{c} \text{CH}_2\text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{CH}_2\text{OH} \end{array} $	meso-Glycero-allo-heptitol	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	a, d
$ \begin{array}{c} \text{OH} \\ \\ \text{HO} - \text{C} - \text{OH} \\ \\ \text{HO} - \text{C} - \text{OH} \\ \\ \text{OH} - \text{C} - \text{OH} \\ \\ \text{OH} \end{array} $	Muco-inositol 1,2,4,5,6-Cyclohexane-hexol	1	$P2_1/c$	1 (4)	- UPACK (systematic)	1999	41	c, d
$ \begin{array}{c} \text{CH}_2\text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{CH}_2\text{OH} \end{array} $	Ribitol	1	$P2_1/c$	1 (4)	- UPACK (systematic)	1999	41	c, d
$ \begin{array}{c} \text{CH}_2\text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{CH}_2\text{OH} \end{array} $	meso-D-Glycero-L-ido-heptitol	1	$P\bar{1}$	1 (2)	- UPACK (systematic)	1999	41	a, d

Table 3.2 (continued)

$ \begin{array}{c} \text{CH}_2\text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{CH}_2\text{OH} \end{array} $	meso-D-Glycero-L-altro-heptitol	1	<i>Pbca</i>	1 (8)	- UPACK (systematic)	1999	41	d, g
$ \begin{array}{c} \text{CH}_2\text{OH} \\ \\ \text{CHOH} \\ \\ \text{CHOH} \\ \\ \text{CHOH} \\ \\ \text{CHOH} \\ \\ \text{CHOH} \\ \\ \text{CH}_2\text{OH} \end{array} $	D, L-Glycero-D, L-galacto-heptitol	1	<i>Cc</i>	1 (4)	- UPACK (systematic)	1999	41	c, d
$ \begin{array}{c} \text{CH}_2\text{OH} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{CH}_2\text{OH} \end{array} $	D-Iditol	1	<i>P2₁</i>	1 (2)	- UPACK (systematic)	1999	41	c, d
$ \begin{array}{c} \text{CH}_2\text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{CH}_2\text{OH} \end{array} $	D-Glucitol	1	<i>P2₁2₁2₁</i>	1 (4)	- UPACK (systematic)	1999	41	c, d
$ \begin{array}{c} \text{CH}_2\text{OH} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{CH}_2\text{OH} \end{array} $	D-Mannitol	2	β : <i>P2₁2₁2₁</i> k: <i>P2₁2₁2₁</i>	1 (4) 1 (4)	- UPACK (systematic)	1999	41	for β : c, d for k: a, d
$ \begin{array}{c} \text{CH}_2\text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{CH}_2\text{OH} \end{array} $	Allitol	1	<i>P2₁/c</i>	0.5 (2)	- UPACK (systematic)	1999	41	c, d
$ \begin{array}{c} \text{CH}_2\text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{CH}_2\text{OH} \end{array} $	DL-Arabinitol	1	<i>Pna2₁</i>	1 (4)	- UPACK (systematic)	1999	41	c, d

Table 3.2 (continued)

$ \begin{array}{c} \text{CH}_2\text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{CH}_2\text{OH} \end{array} $	D,L-Mannitol	1	$Pna2_1$	1 (4)	- UPACK (systematic)	1999	41	c, d
$ \begin{array}{c} \text{OH} \\ \\ \text{HO} - \text{C} - \text{OH} \\ \\ \text{HO} - \text{C} - \text{OH} \\ \\ \text{OH} \end{array} $	epi-Inositol	1	$P2_1/c$	1 (4)	- UPACK (systematic)	1999	41	c, d
$ \begin{array}{c} \text{CH}_2\text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{CH}_2\text{OH} \end{array} $	D-Glycero-L-allo-heptitol	1	$P2_12_12$	1 (4)	- UPACK (systematic)	1999	41	c, d
$ \begin{array}{c} \text{OH} \\ \\ \text{HO} - \text{C} - \text{OH} \\ \\ \text{HO} - \text{C} - \text{OH} \\ \\ \text{OH} \end{array} $	L-chiro-Inositol	1	$P2_1$	1 (2)	- UPACK (systematic)	1999	41	a, d
$ \begin{array}{c} \text{CH}_2\text{OH} \\ \\ \text{CHOH} \\ \\ \text{CHOH} \\ \\ \text{CHOH} \\ \\ \text{CHOH} \\ \\ \text{CHOH} \\ \\ \text{CH}_2\text{OH} \end{array} $	D,L-Iditol	1	$P2_1/c$	1 (4)	- UPACK (systematic)	1999	41	d, g
$ \begin{array}{c} \text{CH}_2\text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{CH}_2\text{OH} \end{array} $	D-Glycero-D-manno-heptitol	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	c, d
$ \begin{array}{c} \text{CH}_2\text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{HO} - \text{C} - \text{H} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{CH}_2\text{OH} \end{array} $	Xylitol	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	c, d

Table 3.2 (continued)

	neo-Inositol	1	$P\bar{1}$	0.5 (1)	- UPACK (systematic)	1999	41	c, d
	glycol	1	$P2_12_12_1$	1 (4)	- UPACK (systematic) + <i>ab initio</i> derived ff	2000	110	a
	glycerol	1	$P2_12_12_1$	1 (4)	- UPACK (systematic) + <i>ab initio</i> derived ff	2000	110	c
Pyranoses								
	α -D-Galactose	1	$P2_12_12_1$	1 (4)	- UPACK (systematic) - UPACK (systematic)	1995 1999	40 41	c, d a
	α -D-Glucose	1	$P2_12_12_1$	1 (4)	- UPACK (systematic) - UPACK (systematic)	1995 1999	40 41	a, d a
	α -D-Talose	1	$P2_12_12_1$	(4)	- UPACK (systematic) - UPACK (systematic)	1995 1999	40 41	c, d a
	β -D-Allose	1	$P2_12_12_1$	1 (4)	- UPACK (systematic) - UPACK (systematic)	1995 1999	40 41	a, d a
	β -D-Galactose	1	$P2_12_12_1$	1 (4)	- UPACK (systematic) - UPACK (systematic)	1995 1999	40 41	a, d a
	β -D-Glucose	1	$P2_12_12_1$	1 (4)	- UPACK (systematic) - UPACK (systematic)	1995 1999	40 41	a, d a
	Methyl- α -D-altropyranoside	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	a

Table 3.2 (continued)

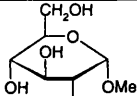
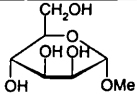
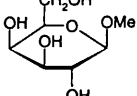
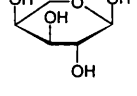
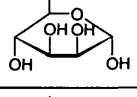
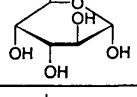
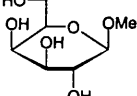
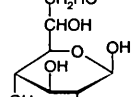
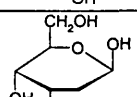
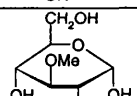
	Methyl- α -D-glucopyranoside	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	a
	Methyl- α -D-mannopyranoside	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	c
	Methyl- β -D-galactopyranoside	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	c
	β -L-Arabinose	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	a
	α -D-Manno-2-heptulose	1	$P2_1$	1 (2)	- UPACK (systematic)	1999	41	a
	α -L-Fucose	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	c
	Methyl-7-deoxy-L-glycero- β -D-galacto-heptopyranoside	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	a
	β -Glucoheptose	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	c
	2-Deoxy- β -D-arabino-hexapyranose	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	a
	3-o-Methyl- α -D-glucopyranose	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	c

Table 3.2 (continued)

	1-Deoxy-1-methyl- α -D-glucopyranose	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	a
	1,5-Anhydro-D-glucitol	1	$P2_1$	1 (2)	- UPACK (systematic)	1999	41	a
	2-Desoxy- β -D-Lyxohexose	1	$P2_1$	1 (2)	- UPACK (systematic)	1999	41	a
	β -D-Fructopyranose	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	c
	β -L-Lyxopyranose	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	a
	Methyl- α -L-arabinoside	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	a
	Methyl- β -L-arabinopyranoside	1	$P2_1$	1 (4)	- UPACK (systematic)	1999	41	a
	Methyl- β -D-arabinopyranoside	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	c
	Methyl- β -D-ribopyranoside	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	a
	Methyl- α -D-lyxopyranoside	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	a

Table 3.2 (continued)

	Methyl-6-deoxy-α-D-idopyranoside	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	c
	Methyl-β-xyloside	1	$P2_1$	1 (2)	- UPACK (systematic)	1999	41	a
	α-L-Xylopyranose	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	a
	Methyl-α-L-fucopyranoside	1	$P2_1$	1 (2)	- UPACK (systematic)	1999	41	a
	Methyl-α-L-rhamnopyranoside (Methyl-6-deoxy-α-L-mannopyranoside)	1	$P2_12_12_1$	1 (4)	- UPACK (systematic)	1999	41	g
	α-D-Mannose	1	$P2_12_12_1$	2 (8)	- UPACK (random)	2000	42	g
Other								
	azo-bis(isobutyronitrile)	1	$P\bar{1}$	0.5 (1)	- Polymorph Predictor	1992	56	c, d
	4,8-dimethoxy-3,7-diazatricyclo(4.2.2.2-2,5)- dodeca-3,7,9,11-tetraene	1	$P\bar{1}$	0.5 (1)	- Polymorph Predictor	1992	56	c, d
	cyclo-L-alanyl-L-alanyl		$P1$	1 (1)	- Polymorph Predictor - Polymorph Predictor	1992 1993	56 [112]	a, d a
	cyclo-bis(dehydro-alanyl) 3,6-dimethylene-piperazine-2,5-dione	1	$P2_1$	1 (2)	- Polymorph Predictor	1992- 1993	56, 57, 111	a

Table 3.2 (continued)

	isoiridomyrmecin	1	$P2_1$	0.5 (1)	- Polymorph Predictor - Polymorph Predictor - Perlstein	1992 1992 1999	56 57 29	g c, d c, d
	2,4,6-Trinitro-N-methyldiphenylamine	1	$P2_1/c$	1 (4)	- MOLPAK/WMIN	1993	33	c
	Methyl- β -D-glucopyranosidetetranitrate	1	$P2_12_12_1$	1 (4)	- MOLPAK/WMIN	1993	33	c
	cis-1,3,5,7-Tetranitro-1,3,5,7-tetra-azadecalin	1	$P2_1$	1 (2)	- MOLPAK/WMIN	1993	33	c
	2,3-Dimethyl-2,3-dinitrobutane	1	$P\bar{1}$	1 (2)	- MOLPAK/WMIN	1993	33	a
	1,4-Dinitro-cyclo-octatetraene	1	$Pna2_1$	1 (4)	- MOLPAK/WMIN	1993	33	b
	2-Nitro-6,7,8,9-tetrahydronaphtho(2,1-b)furan	1	$P\bar{1}$	1 (2)	- MOLPAK/WMIN	1993	33	b
	2-Methoxy-N-(4-nitrobenzylidene)-5-pyridylamine	2	$P2_1/c$ $P2_1$	1 (4) 1 (2)	- MOLPAK/WMIN	1993	33	for $P2_1/c$: b
	Dinitropenamethylenetetramine	1	$P2_1/c$	1 (4)	- MOLPAK/WMIN	1993	33	b

Table 3.2 (continued)

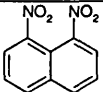
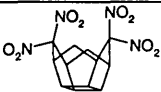

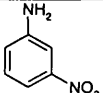

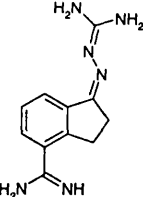
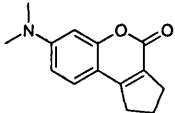

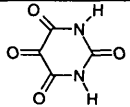
	1,8-dinitronaphthalene	2	$P2_12_12_1$ form II I2/a	1 (4) 1 (8)	- MOLPAK/WMIN - PROMET	1993 1996	33	for $P2_12_12_1$: a
	1,1,5,5-tetranitro- [4]peristylane	1	$P2_1$	1 (2)	- MOLPAK/WMIN	1994	112	c, +e
CO_2		2	Pa3 Cmca (high pressure)		- MDCP	1995	113	for Pa3: c for Cmca: c (at high pressure)
	pyrimidine	1	$Pna2_1$	1 (4)	- MDCP	1995	62	g
	m-nitroaniline	1	$Pbc2_1$	1 (4)	- MPA	1995	114	c, d, f
	1,2-dimethoxyethane	1	$C2/c$	0.5 (4)	- MDCP	1995	115	for $C2/c$: g
	4-Amidino-indanone guanyl-hydrazone	2	A $P2_1/c$ B $P\bar{1}$	1 (4) 1 (2)	- Polymorph Predictor	1996	86	for A: a for B: d, c, +e
	7- Dimethylaminocyclopenta[c]coumarin	2	$Pbca$ $Pna2_1$	1 (8) 1 (4)	- PROMET	1996	116	for $Pbca$: c for $Pna2_1$: a
	urea	1	$P\bar{4} 2_1m$	0.25 (2)	- Polymorph Predictor - MPA - mpg	1996 1996 1999	91 47 50	g b c
	alloxan	1	$P4_12_12$	0.5 (4)	- MOLPAK/DMAREL	1997	117	g in molecular space group

Table 3.2 (continued)


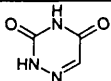
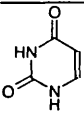
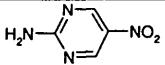
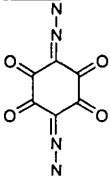
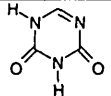
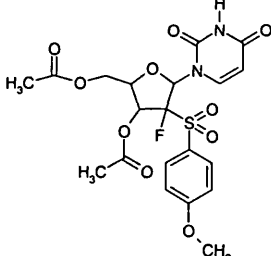
	2-Pyridone	1	$P2_12_12_1$	1 (4)	- PROMET	1997	75	c
	6-azauracil	1	$P2_12_12_1$	1 (4)	- MOLPAK/DMAREL	1997	95	a
	uracil	1	$P2_1/a$	1 (4)	-MOLPAK/DMAREL	1997	95	c
Cl_2	chlorine		Cmca		- MPA	1997	118	a
	2-amino-5-nitropyrimidine	3	I $P2_1/c$ II $P2_1/n$ III $Pccn$	1 (4) 1 (4) 1 (4)	- MOLPAK/DMAREL	1998	119	for I, II: c
	3,6-Bis(diazo)cyclohexanetetrone	1	$P2_1/c$	0.5 (2)	- mpg	1999	50	b
	5-azauracil	1	$Pbca$	1 (8)	- MOLPAK/DMAREL	1999	120	c
	3', 5'-di-O-acetyl-(2'S)-deoxy-2'-fluoro-2'-((4-methoxyphenyl)sulfinyl)) uridine	1	PI	1 (1)	- Perlstein	1999	29	c, d

Table 3.2 (continued)

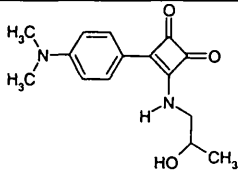

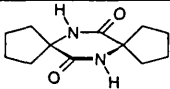
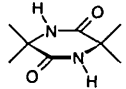
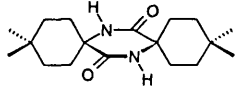
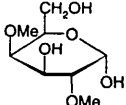
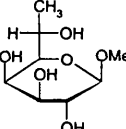
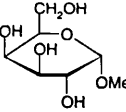
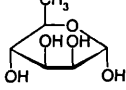
	(-)-1-(4-dimethylaminophenyl)-2-(2-hydroxypropylamino)cyclobutene-3,4-dione	1	$P1$	1 (1)	- Perlstein	1999	29	c, d
	1,4-dioxane	2	I $P2_1/c$ II $P2_1/c$	0.5 (2) 0.5 (2)	- UPACK (systematic) + <i>ab initio</i> derived ff	1999	90	for I, II: a
	3,6-(cyclooctamethylene)-2,5-diketopiperazine	1	$P\bar{1}$	0.5 (1)	- Chin	1999	61	c, f
	3,6-(tetramethyl)-2,5-diketopiperazine	1	$P\bar{1}$	0.5 (1)	- Chin	1999	61	c, f
	3,6-(4,4-dimethylcyclohexane)-2,5-diketopiperazine	1	$P\bar{1}$	0.5 (1)	- Chin	1999	61	g, f
Hydrates								
	2,4-Di-O-methyl-α-D-galactopyranose monohydrate	1	$P2_12_12_1$	1 (4)	- UPACK (random)	2000	42	a
	Methyl-7-deoxy-D-glycero-β-D-galactohexopyranoside monohydrate	1	$C2$	1 (4)	- UPACK (random)	2000	42	a
	Methyl-α-D-galactopyranoside monohydrate	1	$P2_12_12_1$	1 (4)	- UPACK (random)	2000	42	a
	α-L-Rhamnose monohydrate	1	$P2_1$	1 (2)	- UPACK (random)	2000	42	a

Table 3.2 (continued)

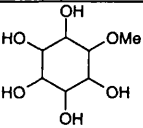
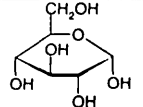
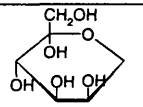
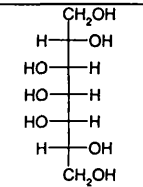
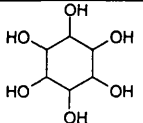
	(+) -Ononitol monohydrate	1	$P1$	1 (1)	- UPACK (random)	2000	42	a
	α -D-Glucose monohydrate	1	$P2_1$	1 (2)	- UPACK (random)	2000	42	a
	D-manno-3-heptulose monohydrate	1	$P2_12_12_1$	1 (4)	- UPACK (random)	2000	42	a
	meso-D-Glycero-L-altro-heptitol monohydrate	1	$C2/c$	1 (8)	- UPACK (random)	2000	42	g
	myo-Inositol dihydrate	1	$P2_1/a$	1 (4)	- UPACK (random)	2000	42	g

Table 3.2 Summary of published crystal structure prediction studies.

The molecular crystals are sorted into groups (alcohols, carboxylic acids, drugs, hydrocarbons, molecular magnets, pigments, polyalcohols, pyranoses, other and hydrates) and the systems in each group are listed by publication year. The comments include the following point:

- a The experimental structure has been found in the search among the low energy minima, but not as the global minimum.
- b The experimental structure has been found in the search, but no further information has been provided about the structures and relative energies of other low energy minima.
- c The experimental structure has been found as the global minimum in the search.
- d The search has just been performed in the experimental space group.
- e Rietfeld refinement has been used (+e in addition to *ab initio* crystal structure prediction)
- f Comparison of unindexed powder patterns.
- g The experimental structure has not been found in the search.
- k The prediction has been made in advance of an experimental investigation.

References for chapter 3

- (1) Nangia, A.; Desiraju, G.R., *Topics in current chemistry*, **1998**, 198, 57-95. Supramolecular synthons and pattern recognition.
- (2) Dunitz, J.D. in *Implications of Molecular and Materials Structure for New Technologies*, J.A.K. Howard, F.H. Allen, and G.P. Shields, Kluwer Academic Publishers, Dordrecht, **1999**, 175-184. Weak interactions in molecular crystals.
- (3) Gavezzotti, A. in *Crystal engineering: from molecules to crystals to materials*, D. Braga, F. Grepioni, and A.G. Orpen, Kluwer, Dordrecht, **1999**, Molecular packing in liquids, solutions and crystals: Acetic acid as a test case.
- (4) Desiraju, G.R., *Science*, **1997**, 278, 404-405. Crystal gazing: Structure prediction and polymorphism.
- (5) Maddox, J., *Nature*, **1988**, 335, 201. Crystals from 1st principles.
- (6) Hawthorne, F.C., *Nature*, **1990**, 345, 297. Crystals from 1st principles.
- (7) Fagan, P.J.; Ward, M.D., *Scientific American*, **1992**, 7, 28 -43. Building molecular crystals.
- (8) Gavezzotti, A., *Acc. Chem. Res.*, **1994**, 27, 309-314. Are crystal structures predictable?
- (9) Lommerse, J.P.M.; Motherwell, W.D.S.; Ammon, H.L.; Dunitz, J.D.; Gavezzotti, A.; Hofmann, D.W.M.; Leusen, F.J.J.; Mooij, W.T.M.; Price, S.L.; Schweizer, B.; Schmidt, M.U.; Van Eijck, B.P.; Verwer, P.; Williams, D.E., *Acta Cryst.*, **2000**, B 56, 697-714. A test of crystal structure prediction of small organic molecules.
- (10) Van Eijck, B.P., **2000**, Ab initio prediction of crystal structures, Plenary talk, ECM19, Nancy/France.
- (11) Gavezzotti, A.; Filippini, G., *Chem. Comm.*, **1998**, Self-organization of small organic molecules in liquids, solutions and crystals: static and dynamic calculations.
- (12) Mooij, W.T.M., *Ab initio prediction of crystal structures, PhD thesis, Utrecht University, Utrecht, The Netherlands*, **2000**.
- (13) Breu, J.; Domel, H.; Norrby, P.-O., *Eur. J. Inorg. Chem.*, **2000**, 2409-2419. Chiral recognition among trisdiimine - metal complexes, 7 racemic compound Formation versus conglomerate formation with (M(bpy)₃(PF₆)₂(M=Ni, Zn, Ru); Lattice energy minimisations and implications for structure prediction.

- (14) Press, W.H.; Flannery, B.P.; Teukolsky, S.A.; Vetterling, W.T., *Numerical recipes: the art of scientific computing*, **1989**, Cambridge: Cambridge University Press.
- (15) Van Eijck, B.P.; Kroon, J., *J. Phys. Chem. B*, **1997**, *101*, 1096-1100. Coulomb energy of polar crystals.
- (16) Pillardy, J.; Wawak, R.J.; Arnautova, Y.A.; Czaplewski, C.; Scheraga, H.A., *J. Am. Chem. Soc.*, **2000**, *122*, 907-921. Crystal structure prediction by global optimisation as a tool for evaluating potentials: Role of the dipole moment correction term in successful predictions.
- (17) Van Eijck, B.P.; Kroon, J., *J. Phys. Chem. B*, **2000**, *104*, 8089. Comment on "Crystal structure prediction by global optimization as a tool for evaluating potentials: Role of the dipole moment correction term in successful predictions".
- (18) Docherty, R.; Jones, W. in *Organic molecular solids: Properties and applications*, W. Jones, CRC Press, Boca Raton, Florida, **1997**, Chapter 5.
- (19) Verwer, P.; Leusen, F.J.J. in *Rev. Comp. Chem.*, K.B. Lipkowitz and D.B. Boyd, Wiley-VCH, John Wiley and Sons, Inc., New York, **1998**, 327-365. Computer simulation to predict possible crystal polymorphs.
- (20) Gdanitz, R.J., *Current opinion in solid state & materials science*, **1998**, *3*, 414-418. Ab initio prediction of molecular crystal structures.
- (21) Scaringe, R.P. . in *45th annual meeting electron microscopy Society of America*. 1987. San Francisco: San Francisco Press.
- (22) Scaringe, R.P. in *Electron Crystallography of Organic Molecules*, J.R. Fryer and D.L. Dorset, Kluwer, Dordrecht, **1990**, 85-113. A theoretical technique for layer structure prediction.
- (23) Scaringe, R.P.; Perez, S., *J. Phys. C: Solid State Phys.*, **1987**, *91*, 2394-2403. A novel method for calculating the structure of small-molecule chains on polymeric templates.
- (24) Kitaigorodskii, A.I., *Organic chemical crystallography*, **1961**, New York.
- (25) Perlstein, J., *J. Am. Chem. Soc.*, **1992**, *114*, 1955-1963. Molecular Self-Assemblies: Monte Carlo prediction for the structure of the one-dimensional translation aggregate.
- (26) Perlstein, J., *J. Am. Chem. Soc.*, **1994**, *116*, 455-470. Molecular self-assemblies. 2. A computational method for the prediction of the structure of one-

dimensional screw, glide, and inversion molecular aggregates and implications for the packing of molecules in monolayers and crystals.

- (27) Perlstein, J., *Chem. Mater.*, **1994**, *6*, 319-326. Molecular self-assemblies. 3. Quantitative predictions for the packing geometry of perylenedicarboximide translation aggregates and the effects of flexible end groups. Implications for monolayers and three-dimensional crystal structure predictions.
- (28) Perlstein, J., *J. Am. Chem. Soc.*, **1994**, *116*, 11420-11432. Molecular self-assemblies. 4. Kitaigorodskii's aufbau principle for quantitatively predicting the packing geometry of semiflexible organic molecules in translation monolayer aggregates.
- (29) Perlstein, J. in *Crystal engineering: from molecules to crystals to materials*, D. Braga, F. Grepioni, and A.G. Orpen, Kluwer, Dordrecht, **1999**, Introduction to packing patterns and packing energetics of crystalline self-assembled structures. Predicting crystal structures using Kitaigorodskii's aufbau principle.
- (30) Abdallah, D.; Bachman, R.E.; Perlstein, J.; Weiss, R., *J. Phys. Chem. B*, **1999**, *103*, 9269-9278. Crystal structures of symmetrical tetra-n-Alkyl ammonium and phosphonium halides. Dissection of competing interactions leading to 'biradial' and 'tetraradial' shapes.
- (31) Gavezzotti, A., *J. Am. Chem. Soc.*, **1991**, *113*, 4622-4629. Generation of possible crystal structures from the molecular structure for low-polarity organic compounds.
- (32) Gavezzotti, A., *Minopac - a new lattice energy minimiser*, Milano, 1999.
- (33) Holden, J.R.; Du, Z.; Ammon, H.L., *J. Comp. Chem.*, **1993**, *14*, 422-437. Prediction of possible crystal structures for C-, H-, N-, O- and F-containing organic compounds.
- (34) Busing, W.R., *Oak Ridge National Laboratory, Oak Ridge, TN 37831. E-mail: busingwr@ornl.gov*, **1981**, WMIN, A computer program to model molecules and crystals in terms of potential energy functions. Oak Ridge National Report - 5747.
- (35) Willock, D.J.; Price, S.L.; Leslie, M.; Catlow, C.R., *J. Comp. Chem.*, **1995**, *16*, 628-647. The relaxation of molecular crystal structures using a distributed multipole electrostatic model.

- (36) Chaka, A.M.; Zaniwski, R.; Youngs, W.; Tessier, C.; Klopman, G., *Acta Cryst.*, **1996**, *B52*, 165-183. Predicting the crystal structure of organic molecular materials.
- (37) Hammond, R.B.; Roberts, K.J.; Docherty, R.; Edmondson, M., *J. Phys. Chem. B*, **1997**, *101*, 6532-6536. Computationally assisted structure determination for molecular materials from X-ray powder diffraction data.
- (38) Hammond, R.B.; Roberts, K.J.; Docherty, R.; Edmonson, M.; Gairns, R., *J. Chem. Soc., Perkin Trans. 2*, **1996**, 1527-1528. X-Form metal-free phthalocyanine: crystal structure determination using a combination of high-resolution X-ray powder diffraction and molecular modelling techniques.
- (39) Hammond, R.B.; Roberts, K.J.; Smith, E.D.L.; Docherty, R., *J. Phys. Chem. B*, **1999**, *103*, 7762-7770. Application of a computational systematic search strategy to study polymorphism in phenazine and perylene.
- (40) Van Eijck, B.P.; Mooij, W.T.; Kroon, J., *Acta Cryst.*, **1995**, *B51*, 99-103. Attempted prediction of the crystal structures of six monosaccharides.
- (41) Van Eijck, B.P.; Kroon, J., *J. Comp. Chem.*, **1999**, *20*, 799-812. UPACK program package for crystal structure prediction: Force fields and crystal structure generation for small carbohydrate molecules.
- (42) Van Eijck, B.P.; Kroon, J., *Acta Cryst.*, **2000**, *B56*, 535-542. Structure predictions allowing more than one molecule in the asymmetric unit.
- (43) Van Eijck, B.P.; Kroon, J., *J. Comp. Chem.*, **1997**, *18*, 1036-1042. Fast clustering of equivalent structures in crystal structure prediction.
- (44) Mooij, W.T.M.; Van Eijck, B.P.; Price, S.L.; Verwer, P.; Kroon, J., *J. Comp. Chem.*, **1998**, *19*, 459 - 474. Crystal structure predictions for acetic acid.
- (45) Williams, D.E., *Program MPA, Molecular Packing Analysis*, Copyright by D. E. Williams, University of Louisville, Louisville, Kentucky, 1996.
- (46) Williams, D.E. in *Crystal engineering: from molecules to crystals to materials*, D. Braga, F. Grepioni, and A.G. Orpen, Kluwer, Dordrecht, **1999**, Theoretical prediction of crystal structures of rigid organic molecules.
- (47) Williams, D.E., *Acta Cryst.*, **1996**, *A 52*, 326-328. Ab initio molecular packing analysis.
- (48) Williams, D.E., *Chem. Phys. Lett.*, **1992**, *192*, 538-543. OREMWA prediction of the structure of benzene clusters: transition from subsidiary to global energy minima.

- (49) Williams, D.E., *Program MPA/mpg, Molecular Packing Analysis/Molecular Packing Graphics*, Department of Chemistry, University of Louisville, Louisville, Kentucky, 1999.
- (50) Williams, D.E.; Gao, D.Q., *Acta Cryst.*, **1999**, A55, 621-627. Molecular packing groups and ab initio crystal structure prediction.
- (51) Schmidt, M.U., *Kristallberechnungen metallorganischer Molekülverbindungen*, **1995**, Technische Universität Aachen, Aachen, Germany: Verlag Shaker.
- (52) Schmidt, M.U.; Englert, U., *J. Chem. Soc., Dalton Trans.*, **1996**, *Dalton Trans.*, 2077-2082. Prediction of crystal structures.
- (53) Schmidt, M.U.; Kalkhof, H., *CRYSCA, Program for crystal structure calculations of flexible molecules*, Clariant GmbH, Frankfurt am Main, Germany, 1997.
- (54) Dzyabchenko, A.V.; Pivina, T.S.; Arnautova, E.A., *J. Mol. Struc.*, **1996**, 378, 67-82. Prediction of structure and density for organic nitramines.
- (55) Gdanitz, R.J., *Chem. Phys. Lett.*, **1992**, 190, 391-396. Prediction of molecular crystal structures by Monte Carlo simulated annealing without reference to diffraction data.
- (56) Karfunkel, H.R.; Gdanitz, R.J., *J. Comp. Chem.*, **1992**, 13, 1171-1183. Ab initio prediction of possible crystal structures for general organic molecules.
- (57) Karfunkel, H.R.; Leusen, F.J.J., *speedup*, **1992**, 6, 43-50. Practical aspects of predicting possible crystal structures on the basis of molecular information only.
- (58) Gdanitz, R.J.; Karfunkel, H.R.; Leusen, F.J.J., *J. Mol. Graphics*, **1993**, 11-12, The prediction of yet-unknown molecular crystal structures by solving the packing problem.
- (59) Kalos, M.H.; Whitlock, P.A., *Monte Carlo methods*, **1986**, New York: John Wiley & Sons.
- (60) Metropolis, N.; Rosenbluth, A.W.; Rosenbluth, M.N.; Teller, A.H.; Teller, E., *J. Chem. Phys.*, **1953**, 21, 1087-1092. Equation of state calculations by fast computing machines.
- (61) Chin, D.N.; Palmore, T.R.; Whitesides, G.M., *J. Am. Chem. Soc.*, **1999**, 121, 2115-2122. Predicting crystalline packing arrangements of molecules that form hydrogen-bonded tapes.

- (62) Tajima, N.; Tanaka, T.; Arikawa, T.; Sakurai, T.; Teramae, S.; Hirano, T., *Bull. Chem. Soc. Jpn.*, **1995**, 68, 519-527. A heuristic molecular-dynamics approach for the prediction of a molecular crystal structure.
- (63) Wawak, R.J.; Gibson, K.D.; Liwo, A.; Scheraga, H.A., *Proceedings of the national academy of sciences of the united states of america*, **1996**, 93, 1743-1746. Theoretical prediction of a crystal structure.
- (64) Wawak, R.J.; Pillardy, J.; Liwo, A.; Gibson, K.D.; Scheraga, H.A., *J. Phys. Chem. A*, **1998**, 102, 2904-2918. Diffusion equation and distance scaling methods of global optimization: Applications to crystal structure prediction.
- (65) Hofmann, D.W.; Lengauer, T., *Acta Cryst.*, **1997**, A 53, 225-235. A discrete algorithm for crystal structure prediction of organic molecules.
- (66) Hofmann, D.W.M.; Lengauer, T., *J. Mol. Mod.*, **1998**, 4, 132-144. Crystal structure prediction based on statistical potentials.
- (67) Hofmann, D.W.M.; Lengauer, T., *J. Mol. Struct.*, **1999**, 474, 13-23. Prediction of crystal structures of organic molecules.
- (68) Motherwell, W.D.S., *Nova Acta Leopoldina NF*, **1999**, 79, 89-98. Crystal structure prediction and the Cambridge Structural Database.
- (69) Storn, R.; Price, K., *Dr Dobb's Journal*, **1997**, April, 18-24. Differential Evolution.
- (70) Lommerse, J.P.M., **2001**, manuscript in preparation. PackStar.
- (71) Bruno, I.J.; Cole, J.C.; Lommerse, J.P.M.; Rowland, R.S.; Taylor, R.; Verdonk, M.L., *J. Comput.-Aided Mol. Des.*, **1997**, 11, 525-537.
- (72) Gavezzotti, A.; Filippini, G.; Kroon, J.; Van Eijck, B.P.; Klewinghaus, P., *Chem. Eur. J.*, **1997**, 3, 893-899. The crystal polymorphism of tetrolic acid: A molecular dynamics study of precursors in solution, and a crystal structure generation.
- (73) Mochizuki, S.; Usui, Y.; Wakisaka, A., *J. Chem. Soc., Faraday Trans.*, **1998**, 94, 547-552. Acid-base interaction from the viewpoint of molecular clustering - effects of solvent, $pK(a)$ and size of alkyl group.
- (74) Gavezzotti, A., *Chem. Eur. J.*, **1999**, 5, 567-576. Molecular aggregation in a carbon tetrachloride solution: A molecular dynamics study with a view to crystal nucleation.
- (75) Gavezzotti, A., *Faraday Discussions*, **1997**, 106, 63-77. Computer simulations of organic solids and their liquid-state precursors.

- (76) Gavezzotti, A., *J. Mol. Struct.*, **1999**, 485-486, 485-499. A molecular dynamics view of some kinetic and structural aspects of melting in the acetic acid crystal.
- (77) Dunitz, J.D., *X-ray analysis and the structure of organic molecules*, **1995**, Basel: Verlag Helvetica Chimica Acta.
- (78) Schmidt, M.U. in *Crystal engineering: from molecules to crystals to materials*, D. Braga, F. Grepioni, and A.G. Orpen, Kluwer, Dordrecht, **1999**, Energy minimization as a tool for crystal structure determination of industrial pigments.
- (79) Harris, K.D.M.; Tremayne, M., *Chemistry of Materials*, **1996**, 8, 2554-2570. Crystal structure determination from powder diffraction data.
- (80) Harris, K.D.M.; Tremayne, M.; Lightfoot, P.; Bruce, P.G., *J. Am. Chem. Soc.*, **1994**, 116, 3542-3547. Crystal structure determination from powder diffraction data by Monte Carlo methods.
- (81) Harris, K.D.M.; Johnston, R.L.; Kariuki, B.M., *Acta Cryst.*, **1998**, A54, 632-645. The genetic algorithm: Foundations and applications in structure solution from powder diffraction data.
- (82) David, W.I.F.; Shankland, K.; Shankland, N., *Chem. Comm.*, **1998**, 931-932. Routine determination of molecular crystal structures from powder diffraction data.
- (83) Kariuki, B.M.; Serrano-Gonzalez, H.; Johnston, R.L.; Harris, K.D.M., *Chem. Phys. Lett.*, **1997**, 280, 189-195. The application of a genetic algorithm for solving crystal structures from powder diffraction data.
- (84) Kariuki, B.M.; Calcagno, P.; Harris, K.D.M.; Philip, D.; Johnston, R.L., *Angew. Chem. Int. Ed. Engl.*, **1999**, 38, 831-835. Evolving opportunities in structure solution from powder diffraction data - crystal structure determination of a molecular system with 12 variable torsion angles.
- (85) Filippini, G.; Gavezzotti, A.; Novoa, J.J., *Acta Cryst.*, **1999**, B55, 543. Modelling the crystal structure of the 2-hydronitronitroxide radical (HNN): observed and computer-generated polymorphs.
- (86) Karfunkel, H.R.; Wu, Z.J.; Burkhard, A.; Rihs, G.; Sinnreich, D.; Bürger, H.M.; Stanek, J., *Acta Cryst.*, **1996**, B 52, 555-561. Crystal packing calculations and rietveld refinement in elucidating the crystal structures of two modifications of 4-amidinoindanone guanylhyazone.

- (87) Karfunkel, H.; Wilts, H.; Hao, Z.; Iqbal, A.; Mizuguchi, J.; Wu, Z., *Acta Cryst.*, **1999**, *B55*, 1075-1089. Local similarity in organic crystals and the non-uniqueness of X-ray powder patterns.
- (88) Schmidt, M.U.; Dinnebier, R.E., *J. Appl. Cryst.*, **1999**, *32*, 178. Combination of energy minimizations and rigid body Rietfeld refinement: The structure of 2,5-dihydroxy-benzo(de)benzo(4,5)imidazo(2,1-a)isoquinolin-7-one.
- (89) Rietfeld, H.M., *J. Appl. Cryst.*, **1969**, *2*, 65. A profile refinement method for nuclear and magnetic structures.
- (90) Mooij, W.T.M., Van Eijck, B.P., and Kroon, J., *J. Phys. Chem. A*, **1999**, *103*, 9883-9890. Transferable ab initio intermolecular potentials. 2. Validation and application to crystal structure prediction.
- (91) Leusen, F.J.J., *Journal of Crystal Growth*, **1996**, *166*, 900-903. Ab initio prediction of polymorphs.
- (92) Payne, R.S., Roberts, R.J., Rowe, R.C., and Docherty, R., *J. Comp. Chem.*, **1998**, *19*, 1-20. Generation of crystal structures of acetic acid and its halogenated analogs.
- (93) Gavezzotti, A. and Filipinni, G., *J. Am. Chem. Soc.*, **1995**, *117*, 12299-12305. Polymorphic forms of organic crystals at room conditions: Thermodynamic and structural implications.
- (94) Payne, R.S., Rowe, R.C., Roberts, R.J., Charlton, M.H., and Docherty, R., *J Comp Chem*, **1999**, *20*, 262-273. Potential polymorphs of aspirin.
- (95) Price, S.L. and Wibley, K.S., *J. Phys. Chem. A*, **1997**, *101*, 2198-2206. Predictions of crystal packings for uracil, 6-azauracil, and allopurinol: The interplay between hydrogen bonding and close packing.
- (96) Payne, R.S., Roberts, R.J., Rowe, R.C., and Docherty, R., *International Journal of Pharmaceutics*, **1999**, *177*, 231-245. Examples of successful crystal structure prediction: polymorphs of primidone and progesterone.
- (97) Dzyabchenko, A.V., *J. Struct. Chem.*, **1984**, *25*, 416-420. Theoretical structures of crystalline benzene: The search for a global minimum of the lattice energy in four space groups.
- (98) Dzyabchenko, A.V., *J. Struct. Chem.*, **1984**, *25*, 559-563. Theoretical structures of crystalline benzene: II. Verification of atom-atom potentials minima in *Pbca*, *P21/c*, *P-1*, *C2/c*.

- (99) Dzyabchenko, A.V. and Bazilevskii, M.V., *J. Struct. Chem.*, **1985**, 26, 553-558. Theoretical structure of crystalline benzene: III. Hydrostatic pressure effect.
- (100) Dzyabchenko, A.V., *J. Struct. Chem.*, **1987**, 28, 862-869. Theoretical structures of crystalline benzene. VI. Global search in the bisystemic structural class.
- (101) Dzyabchenko, A.V., *Sov. Phys. Crystallogr.*, **1989**, 34, 131-133. Emergence of symmetry in optimization of the packing of molecules.
- (102) Shoda, T., Yamahara, K., Okazaki, K., and Williams, D.E., *J. Mol. Struct. (Theochem)*, **1994**, 119, 321-334. Molecular packing analysis - prediction of experimental crystal structures of benzene starting from unreasonable initial structures.
- (103) Shoda, T., Yamahara, K., Okazaki, K., and Williams, D.E., *J. Mol. Struct. (Theochem.)*, **1995**, 333, 267-274. Molecular packing analysis of benzene crystals. Part 2. Prediction of experimental crystal structure polymorphs at low and high pressure.
- (104) Gibson, K.D. and Scheraga, H.A., *J. Phys. Chem.*, **1995**, 99, 3765-3773. Crystal Packing without Symmetry Constraints. 2. Possible crystal packings of benzene obtained by energy minimization from multiple starts.
- (105) Thiéry, M.M. and Rérat, C., *J. Chem. Phys.*, **1996**, 104, 9079-9089. High pressure solid phases of benzene. III Molecular packing analysis of the crystalline structures of C₆H₆.
- (106) Van Eijck, B.P., Spek, A.L., Mooij, W.T.M., and Kroon, J., *Acta Cryst.*, **1998**, B 54, 291-299. Hypothetical crystal structures of benzene at 0 and 30 kbar.
- (107) Gao, D.Q., Williams, D.E., *Acta Cryst.*, **1999**, A 55, 221-227. Molecular packing groups and ab initio crystal structure prediction.
- (108) Karfunkel, H.R., Rohde, B., Leusen, F.J.J., Gdanitz, R.J., and Rihs, G., *J. Comp. Chem.*, **1993**, 14, 1125. Continuous similarity measure between nonoverlapping X-ray powder diagrams of different crystal modifications.
- (109) Price, S.L. and Beyer, T., *Transactions ACA*, **1998**, 33, 23-31. Progress and problems in the computer prediction of molecular crystal structures and polymorphism.
- (110) Mooij, W.T.M., Van Eijck, B.P., and Kroon, J., *J. Am. Chem. Soc.*, **2000**, 122, 3500-3505. Ab initio crystal structure predictions for flexible hydrogen-bonded molecules.

- (111) Karfunkel, H.R., Leusen, F.J.J., and Gdanitz, R.J., *J. Comp.-Aided Mat. Design*, **1993**, *1*, 177. The ab initio prediction of yet unknown molecular crystal structures by solving the crystal packing problem.
- (112) Ammon, H.L., Du, Z., Holden, J.R., and Paquette, L.A., *Acta Cryst.*, **1994**, *B50*, 216-220. Structure of 1,1,5,5-tetranitro-(4)peristylane. Structure solution from molecular packing analysis.
- (113) Hirano, T., Tsuzuki, S., Tanabe, K., and Tajima, N., *Chem. Lett.*, **1995**, 1073-1074. Totally ab initio prediction of the structures of CO₂ molecular crystal.
- (114) Shoda, T. and Williams, D.E., *J. Mol. Struct. (Theochem)*, **1995**, *357*, 1-8. Molecular packing analysis. Part 3. The prediction of m-nitroaniline crystal structure.
- (115) Arikawa, T., Tajima, N., Tsuzuki, S., Tanabe, K., and Hirano, T., *Theochem-Journal of Molecular Structure*, **1995**, *339*, 115-124. A possible crystal structure of 1,2-dimethoxyethane-prediction based on a lattice variable molecular dynamics.
- (116) Gavezzotti, A., *Acta Cryst.*, **1996**, *B52*, 201-208. Polymorphism of 7-dimethylaminocyclopenta(c)coumarin: Packing analysis and generation of trial crystal structures.
- (117) Coombes, D.S., Nagi, G.K., and Price, S.L., *Chem. Phys. Lett.*, **1997**, *265*, 532-537. On the lack of hydrogen bonds in the crystal structure of alloxan.
- (118) Williams, D.E. and Gao, D., *Inorg. Chem.*, **1997**, *36*, 782-788. Effects of molecular electric potential and anisotropic atomic repulsion in the dichlorine dimer and crystalline chlorine.
- (119) Aakeroy, C.B., Nieuwenhuyzen, M., and Price, S.L., *J. Am. Chem. Soc.*, **1998**, *120*, 8986-8993. Three polymorphs of 2-amino-5-nitropyrimidine: Experimental structures and theoretical predictions.
- (120) Potter, B.P., Palmer, R.A., Withnall, R., Chowdhry, B.Z., and Price, S.L., *J. Mol. Struct.*, **1999**, *485-486*, 349-361. Aza analogues of nucleic acid bases: experimental determination and computational prediction of the crystal structure of anhydrous 5-azauracil

Chapter 4.

Calculation of crystal structures and their properties

*'The molecules on the screen are often so 'real' that one can easily forget that one is dealing with models, not with reality.'*¹

As described in the previous three chapters, lattice energy is a key component of crystal structure prediction. The following summarises the methodology of the lattice energy minimisation program DMAREL employed in this work. It also describes the calculation of properties used in this thesis. This includes the theoretical determination of the elastic constants from the energy second derivatives of the potential energy hypersurface, as implemented in DMAREL, and the calculation of the Bravais-Friedel-Donnay-Harker (BFDH) and attachment energy (AE) morphologies by the MSI program Cerius²3.5.

4.1 The lattice energy minimisation program DMAREL

The energy minimisation of the crude hypothetical crystal structures is an important and usually the most time consuming part of crystal structure prediction procedures. The choice of the intermolecular potential determines to a large extent the speed of the minimisation, and accuracy is often compromised to reduce the calculation time. The program DMAREL ² allows for the application of distributed multipole moments as a highly accurate representation of the electrostatic part of the intermolecular potential. The atomic multipoles (up to hexadecapoles) are obtained from a DMA of an *ab initio* wavefunction of the isolated rigid molecule by the program CADPAC. ³ The repulsion/dispersion part is represented by an isotropic atom-atom intermolecular potential for all work in this thesis. Recent development also allows for the use of an anisotropic repulsion/dispersion intermolecular potential (chapter 2) within DMAREL.

For the calculations on experimental crystal structures determined by X-ray diffraction, the bond lengths to hydrogen in the molecular structure are usually normalised to standard neutron diffraction values.⁴

A unit cell of N molecules is described by $6N + 6$ variables, comprising six coordinates per molecule and six strain elements describing the unit cell. Three of the six coordinates represent the positions of the centres of mass of each molecule in the unit cell and the remaining three describe the orientation of each molecule relative to the global axis system. The values of these $6N + 6$ variables are represented by a generalised vector \mathbf{r} in coordinate space. The lattice energy $U(\mathbf{r})$ is expanded in a Taylor expansion to second order

$$U(\mathbf{r} + \delta\mathbf{r}) = U(\mathbf{r}) + \mathbf{G}\delta\mathbf{r} + \frac{1}{2}\delta\mathbf{r}\mathbf{W}\delta\mathbf{r} \quad 4.1$$

with $U(\mathbf{r})$ the lattice energy at the stationary point, \mathbf{G} the vector of first derivatives and \mathbf{W} the matrix of second derivatives. At the minimum in the lattice energy the first derivative of $U(\mathbf{r})$ with respect to \mathbf{r} is equal to zero. To find these locations on the potential energy surface a pseudo Newton-Raphson minimisation process is employed which involves choosing a search direction and then estimating the best distance to

move in that direction along the potential energy surface. The maximum step size in this process can be chosen by the user.

The energy of the crystal lattice is calculated by a pairwise additive approximation of atom-atom interactions. As each atom in a crystal system has an infinite number of intermolecular interactions, the calculation of the lattice energy is often done by using a cutoff radius of typically 15 Å for the intermolecular separation r and therefore neglecting all interactions between atoms separated by a larger distance than this predefined cutoff. The systematic error inherent in this approximation is minor for repulsion/dispersion intermolecular interactions as the magnitude of this contribution to the intermolecular energy decreases exponentially and with r^{-6} , respectively. The convergence of the electrostatic energy is slower as it decreases with r^{-1} . Therefore the contributions involving charge-charge, charge-dipole and dipole-dipole interactions are evaluated using an Ewald summation technique.⁵ In this method, part of the summations are calculated in reciprocal space which leads to a much faster convergence. The higher multipole interactions are calculated using a whole molecule cutoff. For the lattice energy calculations, relationships between sites can be used to reduce the number of necessary lattice sums which need to be calculated.

The minimisation of all crystal structure prediction searches in this thesis are done without symmetry constraints, *i.e.* in space group *P1*. During the progress of this thesis, DMAREL has been extended so that the original space group symmetry can be constrained. If the minimisation process does not yield in a true minimised structure, but a saddle point, the symmetry can be relaxed in the direction of the most negative eigenvalue of the Hessian (the matrix *W* in equation 4.1) to allow for a minimisation in a subgroup of the original symmetry. This has been applied to the lowest energy structures of paracetamol (chapter 6) and of the carboxylic acids (chapter 7).

4.2 Mechanical stability of crystal structures

The mechanical stability of molecular crystal structures determines the application qualities of many industrially important materials in production processes. An example is the tablet formation of pharmaceuticals^{6, 7} which is hindered by the reversible phenomenon of elastic deformation observed in soft crystals and supported by the irreversible process of brittle fracture and plastic deformation. For crystalline arrangements with a high stiffness in all three dimensions, tablets often cannot be

compressed directly and therefore binders (as starches or gelatin) need to be employed in the production process. This leads to a disadvantageous increase in tablet size, production time and cost and therefore polymorphic forms with good compression properties are highly desirable (for the case study of paracetamol see chapter 6).

Mechanical stress and strain in crystal structures does not only arise due to a mechanical load, but also during any solid-state transformation, as observed for phase transitions or chemical reaction.⁸

The elastic properties of a crystal can be described by its set of elastic constants. In contrast to ionic or metallic crystals, elastic constants of just a few molecular crystals have been determined experimentally. This is caused by measurement problems due to a lack of sufficiently large single crystals and the weakness of the intermolecular interactions found in organic solids. The fact that the majority of molecular systems crystallises in low energy space groups, increases the number of unique elements in the elastic matrix, therefore more measurements are needed to completely characterise the mechanical properties of a crystal. A theoretical derivation is possible from full lattice dynamic calculations⁹⁻¹³, numerical methods by applying specific stresses and strains on a crystal and calculating the response¹⁴ and analytic expressions directly relating the energy second derivatives of the potential energy hypersurface to the elastic constant matrix.¹⁵⁻¹⁷

The first three diagonal elements (C_{11} , C_{22} , C_{33}) in the elastic stiffness matrix (6×6) describe the axial strain towards a stress in the x, y, z direction, respectively. The shear stiffness constants (C_{44} , C_{55} , C_{66}) are very important for the compaction properties of a crystal as they describe the shearing behaviour.

Elastic properties are largely determined by the forces between molecules, and hence the choice of the intermolecular potential to locate a stationary point on the energy surface, corresponding to a stress-free crystal structure, is a key factor in the accuracy of the calculations. Recently the analytical method for calculating the elastic matrix from the second energy derivatives for rigid organic crystals at 0 K has been incorporated into DMAREL and the use of anisotropic electrostatic models in the calculation of elastic constants could be tested.^{18, 19}

For application purposes the bulk mechanical properties, *i.e.*, the elastic properties of a polycrystalline aggregate rather than of the single crystal, are important. For the calculation of these properties it is often assumed that the crystals are randomly

oriented in the polycrystal, and that therefore the mechanical properties of the aggregate can be approximated as isotropic. These bulk moduli are usually determined from the single crystal properties by two different averaging procedures by Voigt ²⁰ assuming a uniform strain throughout the crystal and Reuss ²¹ assuming a uniform stress. These averages determine the lower and upper boundaries, respectively, of the experimental bulk moduli, provided there is no preferential orientation in the aggregate.

Elastic constants calculations have been employed in this thesis to investigate the difference in elastic properties of the two experimental polymorphs of paracetamol (chapter 6) and as a measure for the mechanical stability of the hypothetical prediction search structures of paracetamol (chapter 6) and the carboxylic acids (chapter 7).

4.3 Morphology calculations

The morphology of a crystal summarises the complete set of its faces. These can be further characterised by their habit, which comprises the relative sizes and shapes of all faces. The habit of a crystal is determined by the crystal structure and can therefore be different for different polymorphs. It is also influenced by external factors, as temperature, supersaturation, solvent or impurity concentrations. ²² The habit of a crystalline product influences its behaviour during production, storage and usage and therefore the shape of a crystal has to be controlled in all application industries.

Each face of a crystal is quantitatively described by a set of Miller indices (hkl) ²³ and its appearance is determined by the relative growth rate of the different faces. Faces with a low growth velocity have a higher morphological importance, *i.e.* their area is relatively large and their relative frequency of occurrence is relatively high. ²²

The theoretical prediction of the morphology of a crystal from the knowledge of its crystal structure has been attempted by a small number of different methods. In this chapter, just the Bravais-Friedel-Donnay-Harker (BFDH) and the attachment energy (AE) model are introduced as they are employed in this thesis and most widely used for organic crystals.

4.3.1 Bravais-Friedel-Donnay-Harker (BFDH) model

One of the earliest attempts to empirically relate the habit of a crystal to its microscopic structure was by Bravais ²⁴ who observed that the morphological importance, the relative extension of a particular face, is proportional to its inter-planar

spacing d_{hkl} . Therefore the ease of the growth of a face is reflected by its thickness. This observation was confirmed by further studies done by Friedel ²⁵ and this approach was later refined by Donnay and Harker ^{26, 27} who took submultiples of d_{hkl} into account by relating the translational symmetry operators of the crystal, as screw axis and glide planes to the faces. In the BFDH model the Donnay-Harker methodology is employed to generate a list of possible growth faces with their relative growth rates estimated by the Bravais-Friedel rules. This method is purely based on geometrical calculations and does not take any energetical characteristics of the system into account. It is very useful for identifying possible growth faces, but the determination of their relative growth rates is limited especially for systems with anisotropic intermolecular interactions such as organic molecules.

4.3.2 Attachment energy (AE) model

The attachment energy model takes the energetics of a crystal into account by calculating the attachment energy, *i.e.* the energy per molecule released on the attachment of a growth layer to a growing crystal surface. ²⁸ The attachment energy is the difference between the lattice energy of a crystal and the energy of the growth slice of thickness d_{hkl} . ²² Hartman and Perdok ²⁹⁻³² observed that the growth rate of a face is directly proportional to its attachment energy. Therefore faces with a high attachment energy grow relatively fast and have a low morphological importance. The list of possible faces for which the attachment energies are calculated is usually generated by the BFDH model. As the attachment energy model is based on energetical criteria, it is more reliable for predicting the relative sizes and shapes of the different surfaces. Hartman and Bennema ³³ observed that for crystals grown from low supersaturations, the assumed proportionality between the growth rate of a face and the attachment energy is a valid approximation for the flat F-faces (morphologically dominant faces) and further work confirming this observation has been published. ³⁴⁻³⁶

Both methods, the BFDH and the attachment energy model do not take any temperature, solvent or impurity effects into account which can have a fundamental influence on the morphology of the experimental crystal. Therefore calculated morphologies are by these methods most likely to predict vapour grown crystals.

Several hypotheses of the influence of the solvent on crystal growth have been proposed and the 'true' mechanism of solvent interaction is still a matter of debate. ^{22,}

37-40 The growth rates and crystal habits can already be modified by a very small quantity of solvent or impurity²² by either inhibition or increase in the growth rate.⁴¹,⁴² Therefore, in cases where the attachment energy calculated and the observed habit forms differ, this is often attributed to solvent and impurity effects.⁴³,⁴⁴

In this thesis the BFDH and the attachment energy model, as implemented in the MSI Cerius software, have been used to estimate the morphological properties of the hypothetical structures generated in the prediction search procedures (chapter 6 and 7). This estimate has been employed to discuss the growth rates of the lowest energy crystal structures and to link this to their likelihood of being observed experimentally as possible polymorphs.

The morphologies for the paracetamol structures were calculated using the program Cerius²3.5 with the Dreiding-2.21 force field⁴⁵ with equilibrium charges,⁴⁶ following the reasonable success of this method in predicting the experimental morphologies grown from supersaturated IMS solution.⁴⁷

For the carboxylic acids morphology calculations, the empirically derived force field by Hagler *et al.*⁴⁸⁻⁵⁰ has been applied.

References for chapter 4

- (1) Dunitz, J.D. in *Implications of Molecular and Materials Structure for New Technologies*, J.A.K. Howard, F.H. Allen, and G.P. Shields, Kluwer Academic Publishers, Dordrecht, **1999**, 175-184. Weak interactions in molecular crystals.
- (2) Willock, D.J.; Price, S.L.; Leslie, M.; Catlow, C.R., *J. Comp. Chem.*, **1995**, *16*, 628-647. The relaxation of molecular crystal structures using a distributed multipole electrostatic model.
- (3) Amos, R.D.; Alberts, I.L.; Andrews, J.S.; Colwell, S.M.; Handy, N.C.; Jayatilaka, D.; Knowles, P.J.; Kobayashi, R.; Laidig, K.E.; Laming, G.; Lee, A.M.; Maslen, P.E.; Murray, C.W.; Rice, J.E.; Simandiras, E.D.; Stone, A.J.; Su, M.D.; Tozer, D.J., *CADPAC6: The Cambridge Analytic Derivatives Package*, Issue 6: Cambridge University, 1995.
- (4) Allen, F.H.; Kennard, O.; Watson, D.G.; Brammer, L.; Orpen, A.G.R., *J. Chem. Soc., Perkin Trans.*, **1987**, *2*, S1 - S9. Tables of bond lengths determined by X-ray and neutron diffraction. Part 1. Bond lengths in organic compounds.
- (5) Ewald, P.P., *Ann. Physik*, **1921**, *64*, 253-265. Die Berechnung optischer und elektrostatischer Gitterpotentiale.
- (6) Payne, R.S.; Roberts, R.J.; Rowe, R.C.; McPartlin, M.; Bashal, A., *Int. J. Pharm.*, **1996**, *145*, 165-173. The mechanical properties of the two forms of primidone predicted from their crystal structures.
- (7) Roberts, R.J.; Payne, R.S.; Rowe, R.C., *Eur. J. Pharm. Sci.*, **2000**, *9*, 277-283. Mechanical property predictions for polymorphs of sulphathiazole and carbamazepine.
- (8) Boldyreva, E. in *Implications of molecular and materials structure for new technologies*, J.A.K. Howard, F.H. Allen, and G.P. Shields, Kluwer, Dordrecht, **1999**, 151-174. Solid-state reactivity and implications for catalytic processes.
- (9) Drabble, J.R.; Husain, A.H.M., *J. Phys. C.: Solid State Phys.*, **1980**, *13*, 1377.
- (10) Brose, K.-H.; Eckhardt, C., *J. Chem. Phys. Lett.*, **1986**, *125*, 235-240. Calculation of elastic constants of TCNQ crystals using atom-atom potentials.
- (11) He, H.X.; Welberry, T.R., *J. Phys. Chem. Solids*, **1988**, *49*, 421-424. Calculation of elastic constants for crystalline acenaphthylene, C₁₂H₈, using semiempirical atom-atom potentials.

- (12) Michalski, D.; Swanson, D.R.; Eckhardt, C.J., *J. Phys. Chem.*, **1996**, *100*, 9506-9511. Elasticity, bulk modulus, and mode Gruneisen parameters of the HPTB molecular crystal: Computational investigation of a clathrate precursor.
- (13) Michalski, D.; Eckhardt, C.J., *J. Phys. Chem. B*, **1997**, *101*, 9690-9694. Computational determination of the elastic properties of the alpha-phenazine crystal.
- (14) Pavlides, P.; Pugh, D.; Roberts, K.J., *Acta Cryst.*, **1991**, *A47*, 846-850. Elastic-tensor atom-atom potential calculations for molecular crystals C_6H_6 AND $CO(NH_2)_2$.
- (15) Walmsley, S.H., *J. Chem. Phys.*, **1968**, *48*, 1438.
- (16) Walmsley, S.H. in *Lattice dynamics and intermolecular forces*, L.V. Corso, Academic Press, New York, **1975**, Basic theory of lattice dynamics of molecular crystals.
- (17) Sharp, N.D.; Walmsley, S.H., *Chem. Phys. Lett.*, **1994**, *222*, 546-550. The method of homogeneous deformations of atomic and molecular crystals. General formulation and rotational invariances.
- (18) Day, G.M.; Price, S.L.; Leslie, M., *Crystal Growth & Design*, **2001**, *1*, 13-27. Elastic constant calculations for molecular organic crystals.
- (19) Day, G.M.; Price, S.L. in *Handbook of elastic properties of solids, liquids and gases*, Levy, Bass, and Stern, Academic Press, **2001**, 3-50. Properties of crystalline organic molecules.
- (20) Voigt, W., *Lehrbuch der Kristallphysik*, **1928**, Leipzig: Teubner.
- (21) Reuss, A., *Z. Angew. Math. Mech.*, **1929**, *9*, 55.
- (22) Berkovitch-Yellin, Z., *J. Am. Chem. Soc.*, **1985**, *107*, 8239-8253. Toward an ab initio derivation of crystal morphology.
- (23) Miller, W.H., *A treatise on crystallography*, **1839**, Cambridge: Pitt Press.
- (24) Bravais, A., *Etudes crystallographiques*, **1866**, Paris: Gauthier-Villars.
- (25) Friedel, G., *Bull. Soc. Fr. Mineral.*, **1907**, *30*, 326. Etudes sur la loi de Bravais.
- (26) Donnay, J.D.H.; Harker, D., *Am. Mineral.*, **1937**, *22*, 446-467. A new law of crystal morphology extending the law of Bravais.
- (27) Donnay, J.D.H.; Donnay, G.C.R., *Acad. Sci. Paris*, **1961**, 252, 908.
- (28) Docherty, R.; Clydesdale, G.; Roberts, K.J.; Bennema, P., *J. Phys. D: Appl. Phys.*, **1991**, *24*, 89-99. Application of Bravais-Friedel-Donnay-Harker,

attachment energy and Ising models to predicting and understanding the morphology of molecular crystals.

- (29) Hartman, P.; Perdok, W.G., *Acta Cryst.*, **1955**, 8, 49-52. On the relations between structure and morphology of crystals. I.
- (30) Hartman, P.; Perdok, W.G., *Acta Cryst.*, **1955**, 8, 521-524. On the relations between structure and morphology of crystals. II.
- (31) Hartman, P.; Perdok, W.G., *Acta Cryst.*, **1955**, 8, 525-529. On the relations between structure and morphology of crystals. III.
- (32) Hartman, P. in *Crystal Growth: An Introduction*, P. Hartman, North Holland, Amsterdam, **1975**, 367-402.
- (33) Hartman, P.; Bennema, P., *Journal of Crystal Growth*, **1980**, 49, 145-156. The attachment energy as a habit controlling factor.
- (34) Tassoni, D.; Riquet, J.P.; Durand, F., *Acta Cryst.*, **1980**, A 36, 420-428. Morphologie théorique du composé Al_3Ni et comparaison avec les formes observées.
- (35) Hartman, P., *J. Cryst. Growth*, **1980**, 49, 157-165.
- (36) Hamer, R.; Tassoni, D.; Riquet, J.P.; Durand, F., *J. Cryst. Growth*, **1981**, 51, 493-501. Morphology of the intermetallic compound Al_2Cu .
- (37) Wells, A.F., *Philos. Mag.*, **1946**, 37, 184.
- (38) Bennema, P. in *Crystal Growth: An Introduction*, P. Hartman, North-Holland, Amsterdam, **1973**, 274.
- (39) Davey, R.J.; Milisavljevic, B.; Bourne, J.R., *J. Phys. Chem.*, **1988**, 92, 2032-2036. Solvent interactions at crystal-surfaces-the kinetic story of alpha-resorcinol.
- (40) Weissbuch, I.; Popoviti-biro, R.; Lahav, M.; Leiserowitz, L., *Acta Cryst.*, **1995**, B 51, 115-148. Understanding and control of nucleation, growth, habit, dissolution and structure of 2-dimensional and 3-dimensional crystals using tailor-made auxiliaries.
- (41) Davey, R.J., *J. Cryst. Growth*, **1976**, 34, 109-119.
- (42) Gilmer, G.H., *Science*, **1980**, 208, 355-363.
- (43) Hartman, P., *J. Cryst. Growth*, **1980**, 49, 166-170.
- (44) Visser, R.A.; Bennema, P., *Neth. Milk Dairy J.*, **1983**, 37, 109-137. Interpretation of the morphology of alpha-lactose hydrate.

- (45) Mayo, S.L.; Olafson, B.D.; Goddard III, W.A., *J. Phys. Chem.*, **1990**, *94*, 8897-8909. DREIDING: A generic force field for molecular simulations.
- (46) Rappe, A.K.; Goddard III, W.A., *J. Phys. Chem.*, **1991**, *95*, 3358-3363. Charge equilibration for molecular-dynamics simulations.
- (47) Nichols, G.; Frampton, C.S., *Journal of Pharmaceutical Sciences*, **1998**, *87*, 684-693. Physicochemical Characterization of the orthorhombic polymorph of paracetamol crystallized from solution. Atomic coordinates by private communication.
- (48) Lifson, S.; Hagler, A.T.; Dauber, P., *J. Am. Chem. Soc.*, **1979**, *101*, 5111-5121. Consistent force field studies of intermolecular forces in hydrogen-bonded crystals. 1. Carboxylic acids, amides, and the C=O...H-hydrogen bonds.
- (49) Hagler, A.T.; Lifson, S.; Dauber, P., *J. Am. Chem. Soc.*, **1979**, *101*, 5122-5130.
- (50) Lifson, S.; Hagler, A.T.; Dauber, P., *J. Am. Chem. Soc.*, **1979**, *101*, 5131-5141. Consistent force field studies of intermolecular forces in hydrogen-bonded crystals. 2. A benchmark for the objective comparison of alternative force fields.

Chapter 5. Cambridge Crystallographic Data Centre (CCDC) crystal structure prediction workshop 1999

*'... A comprehensive and unified test run using a convenient molecule and comparing the performances of all these packages would be desirable, but is very difficult to organize (ideally, this would be a good task for some committee of the International Union of Crystallography).'*¹

In this chapter the international crystal structure prediction workshop 1999 for small organic molecules organised by the Cambridge Crystallographic Data Centre (CCDC) will be introduced and the study of blindtest compound I, 3-oxabicyclo(3.2.0)hepta-1,4-diene, will be discussed in detail. The studies for compound II and III have been carried out by SL Price and HHY Tsui / SL Price respectively and only the geometry optimisations and general conclusions for these two structures will be presented here. It has been the first workshop of this kind for molecular crystals and the results of it have been published.²

5.1 The blindtest challenge

'... Enclosed you will find three structural diagrams of molecules. The crystal structures of these compounds have been solved in different laboratories, but are not published yet. They have been chosen by an independent crystallographer from a combined list of about 20 structural diagrams. All compounds have been crystallised in one of the common space groups and have one molecule only in the asymmetric unit ($Z' = 1$). This is the only information we have available. It is now up to you to predict or propose the cell dimensions and the complete crystal packing of the experimental structure....' This was the invitation to eleven selected international groups in the beginning of 1999 for the first crystal structure prediction workshop organised by Dr Sam Motherwell and Dr Jos Lommerse from the Cambridge Crystallographic Data Centre (CCDC). Although the idea of predicting crystal structures of small organic molecules *ab initio*, i.e. just from the knowledge of the molecular connectivity, has been on the scientific platform for about ten years, it is still, as shown in chapter 3 of this thesis, far from being done routinely. This comparative study was the first of its kind for molecular crystals and it has, without doubt, become a benchmark in the history of crystal structure prediction.

For each of the three compounds (table 5.1) different obstacles had to be overcome in the prediction process. The first and the third molecule do not have any hydrogen donor atoms and therefore cannot form conventional hydrogen bonds in the crystal. The non-directional intermolecular interactions between these molecules are likely to allow for structural shifts by sliding of the molecules relative to each other without large changes in the lattice energy. Such a rearrangement is generally reinforced by *planar* molecular structures and an approximate planarity is very likely for molecule I and can possibly be expected for molecule III. Therefore, especially for compound I, as it is relatively small and lacking a particular shape, a large number of low lattice energy crystal structures can be expected as an outcome of the prediction. Furthermore, all the molecules contain atoms (ether-oxygen, sulphur and boron, respectively) for which intermolecular repulsion-dispersion potential parameters in conjunction with an accurate electrostatic potential derived by a Distributed Multipole Analysis (DMA) had not been tested prior to this prediction workshop. In addition, the *empirical* derivation and validation of such repulsion-dispersion parameters for boron would have been difficult due to the lack of experimental crystal structure data for boron compounds in the CSD.

5.2 Geometry optimisation of the molecular structures for all compounds

As already discussed in chapter 3, one of the main assumptions of each crystal structure prediction study, is the molecular model. As the lattice energy minimisation program DMAREL assumes the molecular structure to be rigid, an accurate determination of the molecular model is vital for the ongoing prediction procedure. Therefore, the molecular geometries have been calculated for each of the three molecules by a high level (MP2 6-31G**) *ab initio* optimisation using the program package Gaussian98³. These calculations were performed on a Digital 8400 computer (Columbus) provided by the National Supercomputing Centre. The starting geometries for these optimisations were derived from bond lengths and angles of similar structures found in the CSD, version December 1998 (table 5.2). For I and II, two test molecules were selected and for III, due to lack of experimental data, only one suitable test molecule was found. In cases where more than one value for any one parameter was available, an average was taken. For all three molecules, no constraints were set to any bonding parameters during the geometry optimisation.

For compound I the optimisation resulted in one unique planar structure. The calculation of molecule II gave two low energy structures differing in the orientation of the hydrogen atom of the hydroxyl group with an intramolecular energy difference between these two conformers of around 2 kJ/mol. The optimisation of compound III proved to be a difficult case because of the size of the molecule and its flexibility. The energy convergence was very slow as the orientation of the plane of the benzene ring in respect to the rest of the molecule kept changing in an oscillating manner without large energy changes. The calculation was terminated and the lowest energy structure obtained was taken as a starting point for the further procedure. In this structure, the benzene ring is out of plane in respect to the rest of the molecule by around 19.4°. Other molecular structures with slightly smaller or larger torsion angles are slightly less stable and the corresponding planar structure is just 0.9 kJ/mol less stable. This low energy barrier between the twisted and planar form, implies that the molecular conformation adopted in the crystal will be determined by the packing forces.

Comparing the bond parameters of these *ab initio* calculated structures for all three molecules with comparable experimental data of similar structures found in the CSD (table 5.2) does not reveal any significant discrepancies. The largest difference, disregarding the hydrogen bond lengths, is less than 0.03 Å for one of the carbon – oxygen and less than 0.04 Å for one of the carbon – carbon bonds in molecule I.

5.3 Testing the intermolecular potential for compound I

To confirm that a chosen intermolecular potential is suitable for the use in a prediction study of a particular molecule, it is necessary to test the successful reproduction of its experimental crystal structure, if known, or of a similar structure retrieved from the CSD with the chosen intermolecular energy function. For compound I, the experimental crystal structures of furane (FURANE10) and (2.2) (3,4) furanophane (WEYXON) (Table 5.2) were lattice energy minimised by DMAREL. The energy function for these calculations consisted of a highly accurate electrostatic model derived by a distributed multipole analysis (DMA) of a 6-31G** MP2 wavefunction by the program CADPAC ⁴ and two different empirically derived repulsion-dispersion parameter sets, FIT (Chapter 2.4) and FIT/HA. The parameter set FIT, as already discussed in chapter 2, has been successfully tested in conjunction with the accurate distributed multipolar electrostatic model, on a wide range of crystal structures. But as these molecules did not contain ether-oxygen atoms, a further repulsion-dispersion parameter set (FIT/HA, Table 5.4) has been derived for compound I by combining the Williams parameters ⁵ for carbon and hydrogen with ether oxygen parameters developed by Halgren *et al.* ⁶ using the usual combining rules.

The results (table 5.4) of these calculations revealed that for both systems the electrostatic contribution to the intermolecular potential is minor (< 25%). Therefore the quality of the structure reproduction is even more dependent on the suitability of the chosen repulsion-dispersion potential. Comparing the root mean square (RMS %) errors for the cell dimensions for both test molecules, the parameter set FIT (FURANE10: 7.2 % RMS, WEYXON: 2.6% RMS) gave better results than FIT/HA (FURANE10: 10.2 % RMS, WEYXON: 6.2% RMS), although neither set was very satisfactory for furane. The main problem for the reproduction of this crystal structure seems to be a structural shift along the *c*-direction of around 11% (FIT) and 14% (FIT/HA). This displacement does not change the general orientation of the furane molecules in the crystal with respect to each other, but it has some influence on the relative intermolecular contacts (table 5.4). This behaviour could probably be related to the planarity of the furane molecule and the non-directionality of the intermolecular contacts.

5.4 The attempt to predict the crystal structure of compound I

As the result of the potential test, the repulsion-dispersion parameter set FIT was selected for the further prediction procedure as the better choice of the two available empirical potential sets. The MOLPAK search covered the common organic molecular space groups $P1$, $P\bar{1}$, $P2_1$, $P2_1/c$, $P2_12_12_1$, $Pna2_1$, $Pca2_1$, $Pbca$, $C2/c$ and generated 25 high density starting structures in each of the 20 coordination groups, a total of 500 crude hypothetical structures. The lattice energy of each of these structures was minimised by DMAREL and the assumed model intermolecular potential.

As expected, this search found a large number of low lattice energy minima (black symbols in figure 5.2) similar in energy and density. A further analysis of the lowest 15 structures, within an energy window of 2.5 kJ/mol, revealed no obvious clustering of these structures in respect of similar coordination types, same intermolecular contacts or general similarity. As this molecular structure does not form conventional hydrogen bonds in the crystal, graph set analysis cannot be applied. Since the structures all appeared distinct and as we wanted to be consistent with the low-lattice-energy-criterion of our method, the minima with the lowest lattice energy in three space groups were chosen as the three allowed submissions (structures AF20, AI19 and CA25 in table 5.5). We noted that we had low confidence in having found the experimental structures (Appendix 5.1).

5.5 Comparison with the experimental structure and discussion

Two polymorphic forms ⁷ were experimentally determined for compound I. Both were crystallised *in situ* by a laser technique at 235 K from the neat sample, which is a liquid at room temperature. By cooling of the sample, the data for polymorph 1 ($Pbca$) was collected at 200K and the one of form 2 ($P2_1/c$) at 116 K. Collecting high angle data for the low temperature crystal was not successful in the first approach as the cooling system broke down and the sample melted. Therefore this part of the study had been explicitly repeated and finally low temperature high angle data was collected successfully. But thereafter a crystallisation of the high temperature structure, despite many attempts with 'clean' laboratory material, over a time of about two weeks, was impossible. Further studies would probably prove the $Pbca$ structure to be another representative of the family of disappearing polymorphs ^{8, 9}. Both crystal structures have a melting point of 242 K.

Although our guesses included the space group $P2_1/c$, none of the three selected structures is close in packing to either of the two experimental polymorphs (table 5.5). The best prediction could have aimed at finding the same structure as the one obtained from minimising the experimental structure using the *ab initio* optimised molecular structure and the same intermolecular potential. Therefore each of the three main components of this study, the intermolecular potential, the molecular model and the search method could contribute to this failure.

Minimising the experimental polymorphs, without normalisation of the bonds to hydrogen, with the same potential as employed in the prediction, to test the reproduction capability of the energy function, gave satisfying results for the low energy *Pbca* structure and a significant shift for the high energy structure $P2_1/c$ form (table 5.5 and figure 5.3). This shift may have a similar origin as the one observed in the reproduction test for furane (figure 5.1). Therefore it is quite obvious that the intermolecular potential has problems at reproducing at least one of the two polymorphic experimental structures. This might be related to the fact that Williams' empirical potential parameters for hydrocarbons have been derived with a foreshortened C-H bondlength. It is unfortunately not obvious from his publication whether this foreshortening is applied to X-ray or standard neutron values.⁵

Significant differences between the *ab initio* optimised and experimental molecular structures were only found in the lengths of bonds to hydrogen, which are elongated by about 0.07 – 0.12 Angstroms in the *ab initio* calculated model compared to the reported X-ray positions. A further MOLPAK search for the experimental space group with the experimental molecular structure, without normalisation of the bonds to hydrogen, and the same potential parameter set, showed that the *Pbca* polymorph would have been found as *the* lowest structure (CC10) of its space group (table 5.5). This shows the considerable sensitivity of this search method to the molecular model. As just the *experimental* space group had been investigated in this post analysis, it was not a final proof that form 1 would have been one of the lowest candidates in total. In the case of the second polymorph ($P2_1/c$) just a similar structure was found in the MOLPAK run with the experimental X-ray structure and it was *not* one of the lowest structures in the $P2_1/c$ space group.

The experimental structures cannot be directly compared to the calculated structures found by the initial MOLPAK search as their molecular structures are different. Therefore, the optimised molecular structure was pasted into the two

experimental crystal structures (Figure 5.2) which were then lattice energy minimised with the same intermolecular potential. This changed their lattice energy and decreased their density considerably. It also revealed that the structure of the second polymorph ($P2_1/c$) had already been sampled in the first search as one of the higher energy candidates (AM25) (table 5.5).

In general the results for compound I suggests that the empirical repulsion-dispersion potential employed in this study is not appropriate for the reproduction of the crystal structures of this type of molecular structure. Pioneering work on theoretically derived repulsion-dispersion potentials (chapter 2.4) has shown that these novel potentials are very competitive to empirically derived potentials. Furthermore the upper limit of 25 high density search structures for each coordination group generated in the MOLPAK procedure does not seem to be sufficient. The second polymorph had been generated as rank 25 in coordination group AM and therefore had just been included the DMAREL minimisation stage, and the $Pbca$ polymorph has not been found at all. Hence our failure to predict the polymorphs is due to the sensitivity of the search method to the molecular model and an inadequate repulsion/dispersion potential.

5.6 Comparison with the results of other groups and conclusions

The comparison of the submitted structures was performed in respect of different criteria: Visual inspection, the molecular environment of a chosen central molecule ¹⁰, space group and cell dimensions (Niggli reduced cell ¹¹). Taking these criteria into account, the first polymorph ($Pbca$) was reported by four of the eleven different participants as one of their first three prediction results. In contrast to this, the second polymorph ($P2_1/c$) was not reported by any of these groups as their choice, though Mooij found it as 5th rank and our study located it as the 16th rank, well within the energy of possible polymorphism.

The $Pbca$ structure was successfully found by B. Schweitzer (Switzerland) / J. D. Dunitz (Switzerland) / A. Gavezzotti (Italy) using the program PROMET ¹²: ‘... $Pbca$ has the best lattice energy. However, as we find many structures (>20) within an energy window of 2 – 3 kJ/mol and as the relative order of these structures varies with assumed atomic charges (...) and since the calculations refer to molecules at rest, we have no basis of confidence in our prediction.’ Also B. P. van Eijck (The Netherlands) ‘The best in the OPLS force field, but not the first choice because it is bad in the Williams’

potential” using the program UPACK ^{13, 14} and different force fields (OPLS potential ¹⁵ and Williams’ potential) was successful in predicting this structure. Nevertheless, B. P. van Eijck himself was not too satisfied with the result as he would have liked to include molecular translations and rotations (in terms of the fractional coordinates) in the evaluation process. P. Verwer (The Netherlands), testing the MSI Polymorph Predictor ¹⁶⁻¹⁸, found the correct structure: *“Lowest energy structure. The packing seems reasonable. Similar packings are found with slightly higher energy. This polymorph can probably be obtained under the right crystallisation conditions. Confidence level (1-10) is 5 by energy.”* D. E. Williams (USA) ^{19, 20} also successfully reported the *Pbca* structure. The knowledge based search method did not succeed in finding these two polymorphs.

The crystal structure of compound II was found by one program (MSI Polymorph Predictor) as the 5th lowest energy rank and UPACK located compound III correctly, as the 2nd lowest energy structure. HHY Tsui and SL Price derived a theoretically based repulsion-dispersion potential (chapter 2) for their work on compound III. ²¹ Due to the uncertainty in the geometry of the molecular structure the rigid molecule search did not generate the experimental structure as the global minimum. In the post analysis the experimental structure was reproduced with very minor errors in the cell parameters which is certainly a proof of the high accuracy of the intermolecular potential.

A senior scientist had been against the holding of this crystal structure prediction blindtest, either it being a total success, which would declare the problem to be solved, or a complete failure which would discourage people to carry on. None of these two extreme options came true. As even the successes were accompanied beforehand by comments of low confidence and some of the failures by a far greater level of certainty, predicting molecular crystal structures is not yet fully established. The organisational effort of setting up such a comparative study is enormous, but the results provide a fair objective tool of comparing the performance of different search methods. This is certainly unique and much needed as the discussion of the literature survey in chapter 3 already indicated. The 2nd CCDC crystal structure prediction blindtest is going to be completed in May 2001.


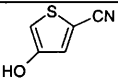
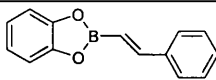
Compound	Molecular connectivity	Name	No of known forms	Space group	Z' (Z)	Ref-erence
I		3-oxabicyclo (3.2.0) hepta-1,4-diene	2	1 <i>Pbca</i> 2 <i>P2₁/c</i>	1 (8) 1 (4)	7
II		4-hydroxy-2-thiophenecarbonitrile	1	<i>P2₁/n</i>	1 (4)	22
III		2-(2-phenylethenyl)-1,3,2-benzodioxaborole	1	<i>P2₁/c</i>	1 (4)	23

Table 5.1 The compounds of the CCDC crystal structure prediction blindtest.


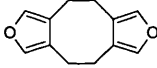
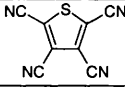
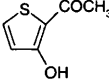
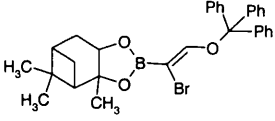
Compound	Test molecule from CSD	Name	Space group	Z' (Z)	Reference
I		furane (FURANE10)	<i>P4₁2₁2</i>	0.5 (4)	24
		(2.2) (3,4) furanophane, (WEYXON)	<i>Pbca</i>	0.5 (4)	25
II		tetracyanothiophene (TCTHPH)	<i>Pa</i>	1 (2)	26
		2-acetyl-3-hydroxylthiophene (TISMAL)	<i>P2₁/c</i>	1 (4)	27
III		(S)-pinanediol (S) - ((trityloxy)ethyl) boronate (ZANNUX)	<i>P2₁2₁2₁</i>	1 (4)	28

Table 5.2 Test molecules retrieved from the CSD (Version December 1998) for the determination of starting parameters for the *ab initio* optimisation of the molecular structures of the blindtest compounds.

Interaction		A		B / Å ⁻¹	C	
Atoms	Labels	eV	kJ/mol		eV Å ⁶	kJ/mol Å ⁶
C C	CA-CA	3832.1	369743.0	3.60	25.286949	2439.8
C H	CA-HY	689.5	66529.6	3.67	5.978972	576.9
H H	HY-HY	124.1	11971.0	3.74	1.413698	136.4
H O	HY-OX	543.9	52479.5	3.85	4.057452	391.5
		(1405.2)	(135575.8)	(3.93)	(2.643612)	(255.1)
O O	OX-OX	2384.5	230064.1	3.96	11.645288	1123.6
		(15913.9)	(1535445)	(4.13)	(4.943548)	(477.0)
C O	CA-OX	3022.9	291658.3	3.78	17.160239	1655.7
		(7809.2)	(753472)	(3.86)	(11.180664)	(1078.8)

Table 5.3: Repulsion-dispersion potential parameters, FIT (FIT/HA).

	FURANE10			WEYXON		
	Expt	FIT	FIT/HA	Expt	FIT	FIT/HA
Cell vectors / Å (Δ / %) #						
a	5.690	5.485 (3.6)	5.584 (1.9)	6.960	7.250 (-4.2)	7.693 (-10.5)
b	5.690	5.485 (3.6)	5.584 (1.9)	12.321	12.203 (1.0)	12.411 (-0.7)
c	11.920	13.270 (-11.3)	14.002 (-17.5)	11.132	10.968 (1.5)	10.912 (2.0)
Cell angles / °						
α	90.0	90.0	90.0	90.0	90.0	90.0
β	90.0	90.0	90.0	90.0	90.0	90.0
γ	90.0	90.0	90.0	90.0	90.0	90.0
Close intermolecular contacts						
H...H	2.9	2.8	2.9	2.5	2.5	2.5
H...O	2.5	2.6	2.9	2.6	2.9	3.0
H...C	2.7	2.9	3.0	2.7	2.9	2.9
C...O	3.4	3.6	3.9	3.5	3.6	3.7
C...C	3.7	3.9	3.9	3.7	3.8	4.0
Volume / cm ³ (Δ / %) #	385.9	399.2 (-3.4)	436.7 (-13.2)	954.6	970.3 (-1.7)	1041.8 (-9.1)
Cell density / g/cm ³ (Δ / %) #	1.2	1.1 (3.3)	1.0 (11.6)	1.3	1.3 (1.7)	1.2 (8.4)
Energy / kJ/mol	-	-38.4	-30.0	-	-99.5	-79.7
Energy electrostatic / %	-	24.0	24.7	-	15.3	17.8
RMS / %	-	7.2	10.2	-	2.6	6.2

Δ = 100*(expt – calculated)/expt

Table 5.4 Comparison of the performance of the repulsion-dispersion potentials FIT and FIT/HA in the reproduction of the experimental structures of the two test molecules furane (FURANE10) and (2.2) (3,4) furanophane (WEYXON).

Molecular Structure	Space group	Close intermolecular contacts / Å						Energy/ kJ/mol	Reduced cell lengths / Å			Cell angles / °		
		H...H	H...O	H...C	C...O	O...O	C...C		a	b	c	α	β	γ
Form I	expt	2.6	2.7	2.8	3.3	4.5	3.5	-57.5	5.309	12.648	14.544	90.0	90.0	90.0
Min	expt	2.5	2.9	2.9	3.3	4.4	3.6	-58.9	5.295	12.709	14.594	90.0	90.0	90.0
Min	opt	2.5	2.7	2.9	3.4	4.5	3.7	-54.9	5.362	12.893	14.892	90.0	90.0	90.0
Form II	expt	2.7	2.7	2.8	3.3	3.6	3.4	-52.9	4.954	9.679	9.845	90.0	90.0	90.6
Min	expt	2.7	2.7	2.8	3.5	3.4	3.6	-56.2	4.753	9.985	10.584	90.0	90.1	90.0
Min	opt	2.7	2.6	2.8	3.4	3.5	3.6	-52.7	4.838	10.118	10.721	90.0	91.1	90.0
AF20	opt	2.6	2.7	2.8	3.3	4.5	3.5	-55.6	5.440	6.085	7.811	90.0	90.0	97.4
AI19	opt	2.5	2.7	2.9	3.4	3.5	3.6	-55.0	5.420	8.707	10.954	102.2	90.0	90.0
CA25	opt	2.5	2.7	3.1	3.6	4.7	3.5	-54.4	5.432	6.324	8.061	83.0	74.4	74.9
AM25	opt	2.7	2.6	2.8	3.4	3.5	3.6	-54.4	4.838	10.118	10.721	90.0	91.1	90.0
CC10	expt	2.7	2.7	2.9	3.3	4.4	3.6	-58.9	5.295	12.743	14.593	90.0	90.0	90.0

Table 5.5 Results of the blindtest prediction and post analysis for compound I in reduced cells.

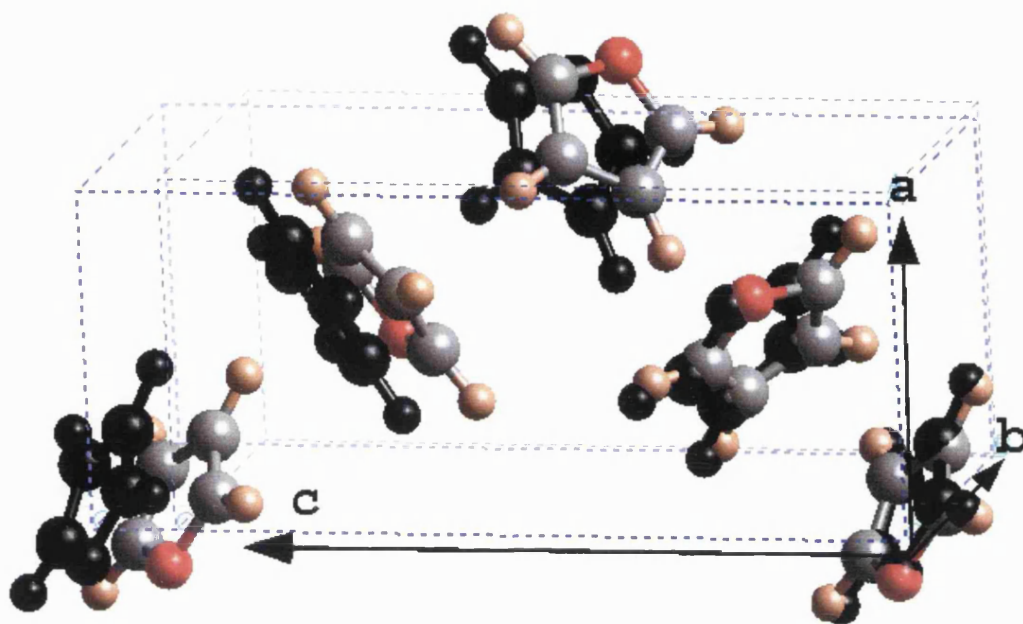


Figure 5.1 Overlay of the experimental and lattice energy minimised (FIT)(black) crystal structure of furane (FURANE10).

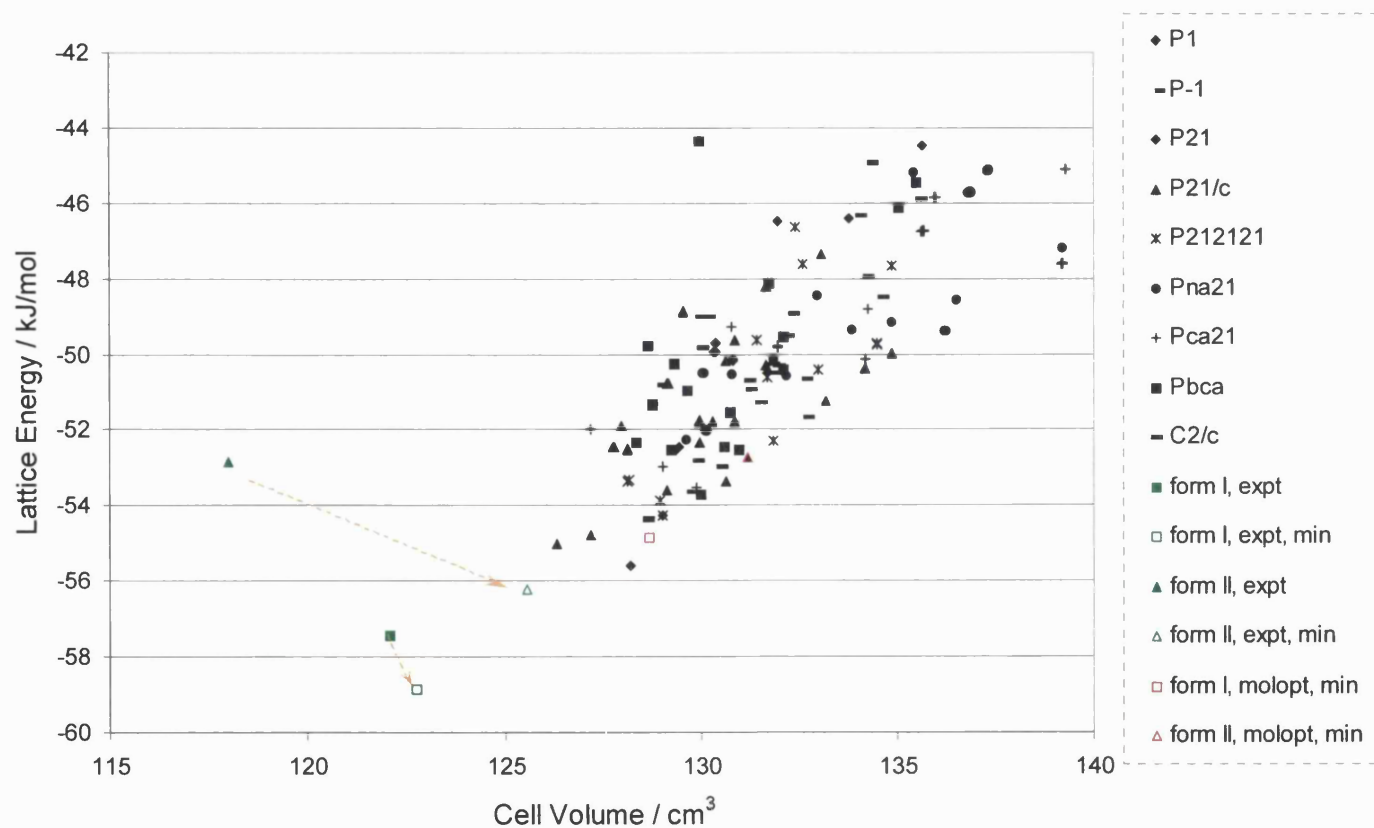


Figure 5.2

The minima in lattice energy (black symbols) for blindtest compound I found by the MOLPAK/DMAREL search. Superimposed are the points corresponding to the experimental structures, the minimisations of the experimental structures and the minimisations of the experimental structure with the *ab initio* optimised molecular structure (opt).

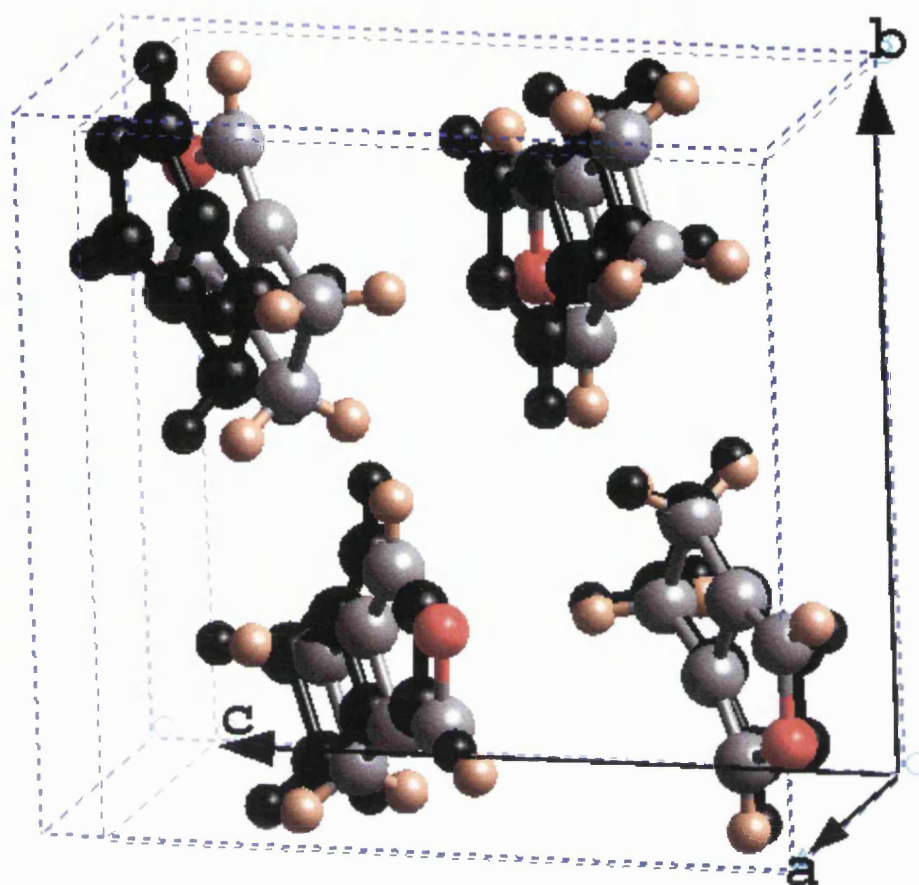


Figure 5.3 Overlay of the experimental and minimised (FIT) (black) crystal structure of the 2nd polymorph ($P2_1/c$) of compound I.

Appendix 5.1

Fractional coordinates of the three submitted crystal structures for the first blindtest molecule.

TITL Price I 1

SPACEGROUP P21

CELL 6.084870 7.811194 5.439376 90.000000 97.390740 90.000000

ATOM O1	0.012106	0.142573	-0.290221
ATOM C1	-0.066207	0.011825	-0.448315
ATOM C2	-0.239419	-0.063058	-0.353740
ATOM C5	-0.268695	0.023427	-0.132023
ATOM C6	-0.112871	0.149321	-0.096041
ATOM H6	-0.064839	0.247691	0.038633
ATOM H1	0.020683	-0.004183	-0.606919
ATOM C3	-0.430128	-0.185465	-0.332506
ATOM C4	-0.462977	-0.088796	-0.084815
ATOM H2	-0.569929	-0.175123	-0.478602
ATOM H3	-0.384483	-0.319937	-0.308016
ATOM H4	-0.436156	-0.167963	0.081305
ATOM H5	-0.621484	-0.023217	-0.089600

ENERGY -55.6 kJ/mol (AF20)

COMMENT Structure with the lowest lattice energy,

COMMENT next lowest structure (2nd choice) is 0.6 kJ/mol less stable.

COMMENT Structure has a large intermolecular O...O distance (4.5 Angstroms),

COMMENT which is also existent in two CSD structures

COMMENT (FURANE10, WEYXON) with similar molecular structures.

TITL Price I 2

SPACEGROUP P21/c

CELL 8.707496 5.420314 12.467059 90.000000 120.831677 90.000000

ATOM O1	-0.412563	-0.095107	-0.289745
ATOM C1	-0.404989	-0.271837	-0.366496
ATOM C2	-0.289703	-0.189810	-0.402086
ATOM C5	-0.224690	0.043050	-0.346009
ATOM C6	-0.301411	0.098194	-0.277308
ATOM H6	-0.297735	0.246700	-0.218785
ATOM H1	-0.487449	-0.431351	-0.382323
ATOM C3	-0.181274	-0.187865	-0.466291
ATOM C4	-0.108465	0.072362	-0.403484
ATOM H2	-0.079200	-0.330547	-0.436071
ATOM H3	-0.258943	-0.184524	-0.568093
ATOM H4	-0.144499	0.224414	-0.469392
ATOM H5	0.035219	0.078029	-0.337498

ENERGY -55.0 kJ/mol (AI19)

COMMENT Structure with the densest packing.

COMMENT Short intermolecular O...O distance (3.5 Angstroms).

TITL Price I 3

SPACEGROUP P-1

CELL 6.323677 5.431903 8.060828 74.403375 96.988952 105.117685

ATOM O1 -0.318035 -0.158540 -0.345715

ATOM C1 -0.205690 0.102247 -0.389999

ATOM C2 -0.004043 0.122402 -0.305523

ATOM C5 0.008342 -0.131733 -0.206789

ATOM C6 -0.185856 -0.301773 -0.233107

ATOM H6 -0.258778 -0.508885 -0.190529

ATOM H1 -0.295034 0.231417 -0.478003

ATOM C3 0.226681 0.248586 -0.246542

ATOM C4 0.240647 -0.035607 -0.136290

ATOM H2 0.343478 0.334094 -0.347527

ATOM H3 0.237769 0.390494 -0.170684

ATOM H4 0.259850 -0.056005 0.002658

ATOM H5 0.365359 -0.112194 -0.174304

ENERGY -54.4 kJ/mol (CA25)

COMMENT Short intermolecular O...O distance (3.6 Angstroms).

COMMENT Very similar 'planar dimer' motif (oxygen atoms head-to-head)

COMMENT observed here is also found in a slightly lower in energy P21/c

COMMENT structure.

COMMENT There are many structures that are only marginally less favourable

COMMENT than these four structures. Hence, we cannot be very confident that

COMMENT the observed structure is one of these four.

References for chapter 5

- (1) Gavezzotti, A., *Crystallography Reviews*, **1998**, 7, 5-121. The crystal packing of organic molecules: challenge and fascination below 1000 Da.
- (2) Lommerse, J.P.M.; Motherwell, W.D.S.; Ammon, H.L.; Dunitz, J.D.; Gavezzotti, A.; Hofmann, D.W.M.; Leusen, F.J.J.; Mooij, W.T.M.; Price, S.L.; Schweizer, B.; Schmidt, M.U.; Van Eijck, B.P.; Verwer, P.; Williams, D.E., *Acta Cryst.*, **2000**, B 56, 697-714. A test of crystal structure prediction of small organic molecules.
- (3) Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Zakrzewski, V.G.; Montgomery, J.A.; Stratmann, R.E.; Burant, J.C.; Dapprich, S.; Millam, J.M.; Daniels, A.D.; Kudin, K.N.; Strain, M.C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, A.; Ayala, P.Y.; Cui, Q.; Morokuma, K.; Malick, D.K.; Rabuck, A.D.; Raghavachari, K.; Foresman, J.B.; Cioslowski, J.; Ortiz, J.V.; Stefanov, B.B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R.L.; Fox, D.J.; Keith, T.; Al-Laham, M.A.; Peng, C.Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P.M.V.; Johnson, B.G.; Chen, W.; Wong, M.W.; Andres, J.L.; Head-Gordon, M.; Replogle, E.S.; Pople, J.A., *Gaussian 98, Revision A.6*, Gaussian, Inc., Pittsburgh PA, 1998.
- (4) Amos, R.D.; Alberts, I.L.; Andrews, J.S.; Colwell, S.M.; Handy, N.C.; Jayatilaka, D.; Knowles, P.J.; Kobayashi, R.; Laidig, K.E.; Laming, G.; Lee, A.M.; Maslen, P.E.; Murray, C.W.; Rice, J.E.; Simandiras, E.D.; Stone, A.J.; Su, M.D.; Tozer, D.J., *CADPAC6: The Cambridge Analytic Derivatives Package*, Issue 6: Cambridge University, 1995.
- (5) Williams, D.E.; Starr, T.L., *J. Comp. Chem.*, **1977**, 1, 173.
- (6) Halgren, *J. Am. Chem. Soc.*, **1992**, 114, 7827-7843. Representation of van der Waals (vdW) interactions in molecular mechanics force-fields-potential form, combination rules, and vdW parameters.
- (7) Boese, R.; Garbarczyk, J., **1998**, unpublished results.
- (8) Dunitz, J.D.; Bernstein, J., *Acc. Chem. Res.*, **1995**, 28, 193-200. Disappearing polymorphs.

- (9) Bernstein, J.; Henck, J.O., *Mater. Research Bull.*, **1998**, *SS*, 119-128. Disappearing and reappearing polymorphs - An anathema to crystal engineering?
- (10) Lommerse, J.P.M., *J. Appl. Cryst.*, **2001**, *in preparation*.
- (11) Krivy, I.; Gruber, B., *Acta Cryst.*, **1976**, *A32*, 297-298. A unified algorithm for determining the reduced (Niggli) cell.
- (12) Gavezzotti, A., *PROMET 5.3, A program for the generation of possible crystal structures from the molecular structure of organic compounds*, University of Milano (Italy), 1997.
- (13) Van Eijck, B.P.; Mooij, W.T.; Kroon, J., *Acta Cryst.*, **1995**, *B51*, 99-103. Attempted prediction of the crystal structures of six monosaccharides.
- (14) Van Eijck, B.P.; Kroon, J., *J. Comp. Chem.*, **1999**, *20*, 799-812. UPACK program package for crystal structure prediction: Force fields and crystal structure generation for small carbohydrate molecules.
- (15) Jorgensen, W.L.; Madura, J.D.; Swenson, C.J., *J. Am. Chem. Soc.*, **1984**, *106*, 6638-6646. Optimized intermolecular potential functions for liquid hydrocarbons.
- (16) Karfunkel, H.R.; Gdanitz, R.J., *J. Comp. Chem.*, **1992**, *13*, 1171-1183. Ab initio prediction of possible crystal structures for general organic molecules.
- (17) Leusen, F.J.J., *Journal of Crystal Growth*, **1996**, *166*, 900-903. Ab initio prediction of polymorphs.
- (18) Leusen, F.J.J. *Industrial applications of polymorph prediction*. in *5th World Congress on Chemical Engineering*. 1996. San Diego, USA.
- (19) Williams, D.E., *Program MPA, Molecular Packing Analysis*, Copyright by D. E. Williams, University of Louisville, Louisville, Kentucky, 1996.
- (20) Williams, D.E. in *Crystal engineering: from molecules to crystals to materials*, D. Braga, F. Grepioni, and A.G. Orpen, Kluwer, Dordrecht, **1999**, Theoretical prediction of crystal structures of rigid organic molecules.
- (21) Tsui, H.H.Y.; Price, S.L., *CrystEngComm*, **1999**, *7*, A non-empirical method of determining atom-atom repulsion parameters: application to crystal structure prediction of an oxyboryl derivative.
- (22) Blake, A.J.; Clark, B.A.J.; Gierens, H.; Gould, R.O.; Hunter, G.A.; McNab, H.; Morrow, M.; Sommerville, C.C., *Acta Cryst.*, **1999**, *B55*, 963-974.

Intramolecular and intermolecular geometry of thiophenes with oxygen-containing substituents.

- (23) Clegg, W.; Scott, A.J., **1998**, unpublished results.
- (24) Fourme, R., *Acta Cryst.*, **1972**, B 28, 2984-2991. Etude par diffraction X des structures cristallines du furanne à la pression atmosphérique.
- (25) Pollmann, M.; Mullen, K., *J. Am. Chem. Soc.*, **1994**, 116, 2318-2323. Semiflexible ribbon-type structures via repetitive diels-alder cycloaddition - cage formation versus polymerization.
- (26) Rychnovsky, V.; Britton, D., *Acta Cryst.*, **1968**, B 24, 725-730. The crystal structure of tetracyanothiophene.
- (27) Danielsen, J., *Acta Chem. Scand.*, **1969**, 23, 2031.
- (28) Ho, O.C.; Soundararajan, R.; Lu, J.; Matteson, D.S.; Wang, Z.; Chen, X.; Wei, M.; Willett, R.D., *Organometallics*, **1995**, 14, 2855-2860. ((trityloxy)methyl)boronic esters.

Chapter 6. Polymorphs of paracetamol – sensitivity of the lattice energy minima to experimental variations in the molecular structure, crystal structure prediction, morphology and mechanical properties

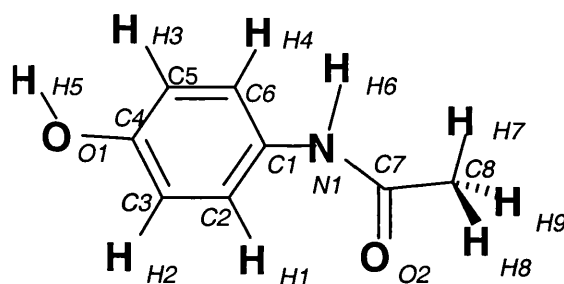
In this chapter, an extensive theoretical study on the polymorphic drug paracetamol (acetaminophen, N-(4-hydroxyphenyl)acetamide) is presented. As a variety of crystal structure determinations (X-ray and neutron) at different temperatures are available for both polymorphic forms it was feasible to evaluate how much variation in the temperature and accuracy of the experimental structure affect the calculated lattice energy minima. The small variations in the molecular structure produce variations of a few percent in the cell lengths and a few kJ mol^{-1} of the lattice energy in the rigid molecular structure minima. This is comparable with previous estimates based on the thermal expansion of this type of organic crystal structure between 0 K and room temperature.

A systematic crystal structure prediction search has been performed and both experimental forms were successfully found as lowest energy forms, reproducing the order of thermodynamical stability. To reduce the number of hypothetical low energy crystal structures that can be considered as potential polymorphs, the elastic properties and vapour growth morphologies have been estimated. These inexpensive calculations give reasonable agreement with the available experimental data for the known polymorphs. Some of the hypothetical structures are predicted to have a low growth rate and plate-like morphology and so are unlikely to be observed. Another is only marginally mechanically stable. Thus this first consideration of such properties in a crystal structure prediction appears to reduce the number of predicted polymorphs whilst leaving a few candidates for the uncharacterised form.

A shorter version of this chapter has been published. ^{1, 2}

6.1 Introduction

Paracetamol (Scheme 6.1) in the solid form is a white, odourless crystalline powder with a bitter taste. It was launched in the UK as a drug in 1956. In 1963, it was added to the British Pharmacopoeia and its popularity as an over-the-counter analgesic increased rapidly. It has now grown to be the most widely accepted antipyretic (fever suppressant) and analgesic (pain killer) in the world, used by over 30 million people in the UK every year with an annual production of about 3 billion tablets.³



Scheme 6.1: Molecular structure of paracetamol

Paracetamol is known to be polymorphic with two reported metastable forms: one orthorhombic form with better tableting properties than the stable monoclinic form, and another, which remains uncharacterised. The transition from the orthorhombic to the monoclinic form has been investigated by X-ray powder diffraction pattern studies as a function of temperature ^{4, 5} The third polymorphic form, form III, has been mentioned by Bürger ⁶, and has been recently recrystallised by Di Martino *et al.* ⁷ by keeping melted paracetamol for several hours between a slide and a cover slip at 54 °C. Besides large transparent crystals (identified as form II), fine crystals appeared in the melted mass which transformed at 70 °C into form II again. This new third polymorphic form has been found in both studies to be too unstable for an experimental investigation of its crystal structure. A search for the third polymorph by slow evaporation of an aqueous solution of commercial paracetamol actually produced a coupled dimer oxidation product. ⁸ The molecular structures of the two entities in the dimer, except for the torsion angle between the ring and the amide chain, are similar to the one in pure paracetamol and a water molecule is incorporated into the crystal structure formed by the coupled dimer.

The crystal structures of the two experimentally investigated forms (form I and II, figure 6.1) have been determined by X-ray and neutron diffraction at different temperatures (table 6.1). The earliest X-ray determinations are by Haisa *et al.* for form II (HXACAN) ⁹ in 1974 and for form I (HXACAN01) ¹⁰ in 1976. Further X-ray determinations for form I are by Welton ¹¹ in 1988, Naumov *et al.* (HXACAN04) ¹² in 1998 and for form II by Bürger ⁶ in 1982 and by Kuhnert Brandstätter *et al.* ¹³ in 1990. More recent investigations are by Frampton *et al.* ¹⁴ at 123K (CSFI_LT) and at 298 K (CSFI_RT) for form I and for form II at 123 K (CSFII_LT, also denoted Form II Expt_{low}) and at 298 K (CSFII_RT, also denoted Form II Expt_{high}). Form I was determined by single crystal pulsed neutron diffraction at 20 K (CCW_20 K, also denoted Form I Expt_{low}), 50 K (CCW_50 K), 80 K (CCW_80 K), 150 K (CCW_150 K), 200 K (CCW_200 K), 250 K (CCW_250 K) and at 330 K (CCW_330 K, also denoted Form I Expt_{high}) by Wilson. ^{15, 16} This neutron study over a wide range of temperatures has revealed a large amplitude librational motion of the methyl group at high temperatures.

The available variety of determinations at different temperatures for both polymorphs allows a careful assessment of the errors involved in the use of crystal structure modelling methods neglecting the influence of temperature effects. It is generally assumed that the main cause for the difference between the experimental crystal structure and the lattice energy minimum is the approximation in the model intermolecular potential. However, there are many other approximations involved in this comparison and one major is the model used for the molecule. Even for molecules that most chemists would consider rigid, the molecular structure in the crystal will be slightly affected by the environment, which differs between polymorphs. Although the extent of these variations has been studied ¹⁷ for organic molecular crystals, their effect on the corresponding lattice energy minima has not. This study therefore provides a concrete example of the level of significance of the accuracy with which crystal structures can be reproduced by lattice energy minimisation. It complements the estimates that differences of 2 kcal/mol ¹⁸ or 3-4 kcal/mol ¹⁹ between the minimised lattice energy and the experimental heat of sublimation are not a cause for concern (chapter 1.4.2), based on the theoretical approximations, as well as experimental errors. Hence we compare here the lattice energy minima obtained using an *ab initio* optimised molecular structure with those using the experimental molecular structures to access the

variation in the predicted lattice energies and crystal structures with the experimental determination. This serves as an estimate of the uncertainties involved for similar molecules where only one experimental structure is available.

The tablet formation properties differ between the two determined polymorphs due to the differences in their crystal structures.¹⁴ The commercial form I requires binders (gelatin, PVP, starch or starch derivatives) for tablet production. As this is not only costly in material and production time, but also results in an enlarged pill size for the same quantity of drug, the metastable polymorphs of paracetamol are of particular industrial interest. The crystal structure of form I has pleated sheets stacked along the *b* axis (figure 6.1a), making it relatively stiff, resulting in poor compression properties²⁰. Although improved compression properties can be obtained by varying the habit of the crystallites²¹⁻²³, such a production process has problems in completely eliminating the residues of the required solvent²⁴. The crystal structure of the orthorhombic form II (figure 6.1b) has parallel hydrogen-bonded sheets along the *c* axis, giving slip planes which allow plastic deformation at higher pressure and a lower elastic recovery during decompression.¹⁴ Thus the direct compression of the orthorhombic form II into tablets has been investigated for potential industrial use.¹⁴ It has been found that a possible transition of form II to I in tablets during storage should be without consequences on the bioavailability.²⁴

The morphology of paracetamol crystals also affects the industrial processing. Form I crystals grown from aqueous solution²⁵ have an elongated prismatic morphology when grown at low supersaturation, but for relatively high supersaturation the faces are equally developed, and so solvent molecules play a role in determining the habit.²⁶ The morphology of both forms when grown from supersaturated industrially methylated spirits (IMS) solution has been determined¹⁴ and compared with predictions from the attachment energy model, which is most appropriate for crystals grown from vapour. For the monoclinic form, the experimental prismatic to plate-like morphology is well predicted¹⁴ by the attachment energy model using the Dreiding-2.21 force-field²⁷. However, the same model underestimates the *c* axis elongation of the prismatic habit for form II. This was attributed to the effect of the solvent and to the rapidity with which the orthorhombic crystal had been precipitated.

Paracetamol is, because of its polymorphic behaviour and the so far uncharacterised form, an ideal candidate for a test for crystal structure prediction. A preceding study on paracetamol has already been undertaken with the MSI Polymorph Predictor ²⁸, using the Dreiding-2.21 force field ²⁷ in combination with semiempirical MNDO-ESP atomic charges ²⁹. This found both experimentally known crystal structures in the correct stability order, and a third potential polymorph in the $P2_12_12_1$ space group. After recalculation with a different force-field it was concluded ²⁸ that this unknown structure was too unstable to exist. It is unfortunate that in the documentation of this study no information is given about the existence of low energy competing structural candidates created by the search procedure. A previous documentation by the same author ³⁰ suggests that the space group reported in the literature for the metastable orthorhombic polymorph is incorrect as the simulations identify a structure in the space group $Pca2_1$.

Thus, a further aim of this study is to investigate the ability of rigid molecule lattice modelling to provide a genuine prediction of the known polymorphs of paracetamol and their industrially important properties. We consider also the morphology and mechanical properties of the hypothetical structures, for the first time in a crystal structure prediction study. This is to establish whether the properties provide any discrimination between the energetically plausible structures, either by suggesting that they are unlikely to be found, or to show that they would have desirable properties if they could be obtained experimentally. The crystallographic details of possible candidates for form III may prove useful in characterising this elusive structure if, for example, powder diffraction diagrams could be obtained experimentally.

6.2 Method

For fourteen selected experimental crystal structures (X-ray and neutron) of paracetamol, covering the two polymorphic forms and a large range of temperatures, a symmetry-constrained lattice energy minimisation was performed by DMAREL, ³¹ using both the corresponding experimental molecular structure, and a model gas phase structure. The Hessian of the second derivatives at each lattice energy minimum was examined to confirm that it was a true minimum (*i.e.* all the eigenvalues are positive). For all experimental structures determined by X-ray diffraction the bonds to hydrogen were adjusted to a standard length ³² of 1.08 Å for C–H, 1.02 Å for O–H and 1.01 Å for

N–H along the experimental bond direction. For neutron data, the original hydrogen bond lengths were taken.

The intermolecular potential consisted of a highly accurate DMA electrostatic model derived of the MP2/6–31G** wavefunction of the isolated molecular structure and the empirical repulsion-dispersion FIT potential parameter set (chapter 2).

A molecular model was generated without reference to an experimental crystal structure by geometry optimisation of a high quality *ab initio* SCF/6–31G** wavefunction, obtained using the program suite CADPAC.³³ This ‘prediction’ *ab initio* molecular model was incorporated into each experimental crystal structure so that the centres of mass of the prediction and experimental molecular structure coincided, and their axes systems were parallel. The molecular axis system was defined with x along C4 to C1, and the xy plane being parallel to the plane containing C1, C4 and C2, so essentially the phenyl ring was located in the same position. The minimum obtained by this process (of minimising the experimental crystal structure using the same rigid molecular structure as would have been used in the search procedure) is the closest that a crystal structure prediction method based on rigid lattice energy minimisation could ‘predict’ the experimental crystal structure.

For the search for possible crystal structures the MOLPAK/DMAREL^{31, 34} program suite was applied. Low energy structures which could have a surface energy correction term,³⁵⁻³⁸ were identified after the minimisation. For the lowest energy structures within an energy range of about 10 kJ/mol, the Hessian of the second derivatives was examined for negative eigenvalues and the elastic constant matrix³⁹ was then calculated for the stress-free crystal by DMAREL.

The morphologies of the experimental and various hypothetical crystal structures of paracetamol were calculated using the Bravais-Friedel-Donnay-Harker (BFDH)⁴⁰ and the attachment energy model⁴¹ by the program⁴² Cerius²3.5 and the Dreiding-2.21 force field²⁷ with equilibrium charges²⁹.

6.3 Results

6.3.1 Analysis of the molecular structures

The variation in the molecular structures between the different determinations of the monoclinic form was known to be small for the non-hydrogenic atoms. An overall similarity quantification,⁴³ based on superimposing the conformations and co-

ordination environment of the non-hydrogenic atoms onto the low temperature structure CCW_20 K, gave a root-mean-square deviation from 0.003 up to 0.012 Å for the conformations of the molecules, and from 0.013 up to 0.12 Å for the superposition of a co-ordination group of 10 molecules, with the higher errors corresponding to higher temperatures. The *ab initio* molecular structure is contrasted with the experimentally determined molecular structures from each crystal structure in table 6.1. This presents the bond lengths and angles of the experimental and *ab initio* optimised molecular structures, for all bond lengths and angles where any experimental structure deviates from the *ab initio* structure by more than 1% and all torsion angles where there is a difference of more than 3°. These deviations can be visualised in figure 6.2, where the *ab initio* optimised structure is overlaid with the lowest temperature experimental structures for each form. The differences in the molecular framework (the bond lengths and angles not involving hydrogen atoms) are mainly too small to be in table 6.1. Thus, there is no significant difference in the overall dimensions of the frame, as measured by the separation of the hydroxyl oxygen and the amide carbonyl (O1...O2) or methyl carbon (O1...C8) atoms.

The X-ray bond lengths to hydrogen are all quite severely underestimated (table 6.1), due to the systematic error resulting from the X-ray determination locating electron density, ⁴⁴ which is displaced into the bond for hydrogen nuclei. This systematic error¹⁸ is accounted for in the lattice energy calculations, where the hydrogen atoms are repositioned by elongating the bonds to standard neutron values. The *ab initio* N–H and O–H bond lengths are somewhat shorter than both these standard values and the neutron determinations in paracetamol.

Most of the crystal structures have the hydroxyl group tilted out of the plane of the aromatic ring by around 17° in form I and 22° in form II (with the exception of CSFII_RT), whereas it is coplanar in the *ab initio* structure. This suggests that these are genuine distortions of the molecular structure by the hydrogen bonding in the crystal, which appear to be similar in the two forms. In both polymorphs the hydrogen bonds have the same graph sets ⁴⁵ and quite a degree of similarity, although the sequence of the different types of hydrogen bond (N(H)...O(H) or O(H)...OC) is different and there is a considerable difference in the angles between the planes of the hydrogen bonded molecules.

The intramolecular energy differences between the molecules in the gas and crystalline phases, due to the rotation of the hydroxyl group, are estimated to be less than 1.6 kJ mol^{-1} . This is the difference in the total MP2 energy between the *ab initio* optimised molecular structure with a torsion angle of 0° and the same structure with the hydroxyl proton repositioned to give a torsion angle of 20° . This estimate is consistent with the nature of the atoms involved and the observation that torsion angles associated with an intramolecular energy of more than 1 kcal mol^{-1} appear¹⁷ to be very unusual in crystal structures. The errors in this estimate of the intramolecular energy difference arise from the *ab initio* method and the assumption that all other geometric parameters are constant. Although the *ab initio* optimised structure is, as expected, the most stable at the SCF level at which it was optimised (by 16 kJ mol^{-1} relative to CCW_20 K), the inclusion of electron correlation makes it less stable by 0.8 kJ mol^{-1} . This is another example⁴⁶ of an experimental crystal structure providing a more accurate estimate of the gas phase structure than SCF optimisation, by the criterion of the variation principle. We also note that the justified concerns⁴⁷ about calculating the relative stability of conformers using *ab initio* methods on molecules of fixed geometry⁴⁸ are well illustrated by paracetamol. The energies of the experimental molecular structures, relative to the most stable (CCW_20 K), range from 1.1 kJ mol^{-1} for CCW_50 K to over 100 kJ mol^{-1} for CCW_300 K. This is consistent with the great sensitivity of fixed geometry calculations to slight errors in the experimental bondlengths and angles, and the uncertainties in the positioning of the hydrogen atoms increasing with thermal motion.

There is also considerable variation in the methyl group geometry. The considerable deviations from a tetrahedral methyl group observed in HXACAN01 and CSFI_RT are primarily an artefact of the inaccuracy of X-ray located protons.⁴⁴ For the more accurately located methyl protons, the variation in torsion angles (table 6.1) is consistent shown with their large thermal ellipsoids. As Wilson deduced from the thermal parameter analysis,¹⁶ there is considerable libration of this group with a residual zero-point root mean squared librational amplitude of 11° , and a barrier height for rotation of 2.2 kJ mol^{-1} . Thus, the substantial libration of the methyl group, above 80 K, results in an apparent shortening of the C8–H bondlengths in the neutron structures.

6.3.2 Analysis of the experimental temperature effects

The temperature dependence of the lattice constants of the monoclinic form I of paracetamol are shown in figure 6.3 and table 6.2. The reduced cell parameters in table 6.2 correspond to the $P2_1/n$ setting for form I and the $Pabc$ setting for form II. There is considerable variation in the three thermal expansion coefficients. There is a small expansion coefficient of $\alpha = 1 \times 10^{-5} \text{ deg}^{-1}$ along a , which is approximately the direction of the hydrogen bonds (figure 6.4a). There is a moderate expansion along c of $\alpha = 4 \times 10^{-5} \text{ deg}^{-1}$, corresponding to changes in the relative tilt of the molecules in the zig-zag layer and the protrusion of the methyl groups. By far the largest expansion of $\alpha = 9 \times 10^{-5} \text{ deg}^{-1}$ is in the b direction along which the pleated sheets are stacked. The more limited data on form II show quite small expansions in the bc plane of the hydrogen bonded sheets, with one side being essentially temperature independent, but considerable expansion of a which measures the separation of the sheets (figure 6.4b).

6.3.3 Analysis of the lattice energy minima

All the lattice energy minima are quite close to their corresponding experimental crystal structures (table 6.2). The maximum % change in a cell length is 3.6%, and the r.m.s. error over the three cell lengths is less than 2.3%. The lowest r.m.s. % error for both forms can be found for structure determinations in the region around $150 < T < 298 \text{ K}$ for form I and $T = 295 \text{ K}$ for form II, *i.e.* approximately room temperature. This probably reflects the fact that the repulsion–dispersion parameters were fitted to room temperature crystal structures. The variations in the minimised cell lengths with experimental structure reflect the changes in the molecular structures with temperature only, as the simulation is effectively 0 K. As Fig. 3 shows, this results in a significant but erratic variation in the cell lengths of the lattice energy minima with the temperature of the experimental determination. In particular, the b and c parameters of form I can be overestimated or underestimated by lattice energy minimisation, depending on the crystal structure used.

A more detailed examination of the lattice energy minima (table 6.2) shows that only the a parameter of form I is consistently predicted to have significant deviation from experiment, albeit only being over-estimated by 1.6 to 3.5%. The intermolecular $\text{N(H)} \cdots \text{O(H)}$ hydrogen bond lengths are well modelled ($1.2 < \Delta < 1.8\%$), but the difference between the experimental and calculated $\text{O(–H)} \cdots \text{O(C)}$ intermolecular bonds

is much higher ($6.1 < \Delta < 8.6\%$). This is consistent with these hydrogen bonds having a significant component along a for form I and in the bc plane of form II. It has been observed that this FIT potential gave a comparable elongation of the $\text{CO}\cdots\text{H}-\text{O}$ bonds found in carboxylic acids.⁴⁹ Hence, the use of distinct repulsion parameters for the two polar hydrogens, $\text{H}_\text{N}-\text{N}$ and $\text{H}_\text{O}-\text{O}$, may lead to an improvement in the model potential.

In the prediction model, when the *ab initio* optimised structure is pasted into any of the experimental structures for a given form, the lattice energy minima are identical. This implies that the crystal structures for each form correspond to the same basin of attraction for the minimisation, as using the same intermolecular potential and molecular model produces the same lattice energy minimum from each crystal structure. The prediction model minima differ from the experimental lattice energy minima in having a larger cell volume. In the monoclinic form I, the zig-zag sheets have expanded along the non-hydrogen bonded direction c , which reduces the protrusion of the methyl groups so that they stack with a smaller b (figure 6.4a). The sheets of the orthorhombic form are expanded (figure 6.4b). These changes are nevertheless quite small and consistent with the enforced methyl conformation and planarity of the hydroxyl group. The $\text{O}(-\text{H})\cdots\text{O}$ distance is about 0.2 \AA larger in the prediction model than in the experimental reproductions.

The calculations agree with the available energetic information on this system, *i.e.* that the monoclinic form I is the more stable. The various experimental structures give lattice energy minima estimates that vary with the experimental molecular model, from between 118 and 120 kJ mol^{-1} for form I, and between 112 and $117.5 \text{ kJ mol}^{-1}$ for form II. In both cases, the prediction model gives a significantly less stable intermolecular lattice energy, by about 10 kJ mol^{-1} . The prediction model estimates that the orthorhombic form II is 3.6 kJ mol^{-1} less stable than the monoclinic form I, whereas the experimental estimates would be 7.2 kJ mol^{-1} using the 298 K crystal structures, or 1 kJ mol^{-1} for the 123 K crystal structures.

We note from table 6.2 that the initial lattice energies show a somewhat larger range of values than the minimised lattice energies, and also that the polymorphic energy difference, for comparable determinations, such as the 123 or 295 K sets, predicts the wrong order of stability. This emphasises that lattice energy estimates derived from minimised crystal structures are more reliable than those derived from the initial lattice energies at the experimental structures. This is because small errors in the

structure or repulsive potential can significantly destabilise the lattice since the repulsive wall is exponential.

6.3.4 Crystal structure prediction search

The lattice energy minima found in the MOLPAK search using the *ab initio* molecular model and the same intermolecular potential are shown in figure 6.5. The low energy structures were compared in detail, using visual inspections and their reduced cells⁵⁰, with the resulting clustering of very similar minima being confirmed by a similarity index analysis based on the coordination environment.⁴³ The fourteen distinct crystal structures with a lattice energy within 10 kJ/mol of the global minimum, are listed in table 6.3, designated by the MOLPAK coordination group and number.

The search readily found the monoclinic stable form as the global minimum (table 6.3). The match of the structure designated AM30 with the minimum obtained from any of the experimental determinations of form I, using the rigid *ab initio* molecular model and the same model potential, was exact. An extremely close match for the corresponding minimum for the metastable orthorhombic form was also found as CB47 after an extended search in the *Pbca* space group, using two sets of initial orientations of the probe molecule.

The hypothetical structures in table 6.3 often involve different types of hydrogen bonding motifs. The majority of these structures consist of stacks of hydrogen-bonded chains (AI22, AQ6, AQ14, CC8, AI16, AM4). Four structures form chains interlinked by OH \cdots OH bonds (CC19, AK6, AK22, AK4). There is one sheet structure (CB9) and another structure (AY8) has a three-dimensional hydrogen-bonding network. Despite this variety, all these structures are within the small energy of plausible polymorphism, and show only a 6% variation in density.

6.3.5 Elastic constants and morphology calculations

The calculated elastic properties of the lowest temperature crystal structures of both known forms of paracetamol are given in table 6.4. The structural differences in the two polymorphic forms lead to distinct differences in the mechanical properties. In form I, the pleated sheets have N-H \cdots O(H)-C hydrogen bonds approximately along the *a* axis, interlinked by OH \cdots O=C hydrogen bonds in the *c* direction (figure 6.1). This results in larger elastic constants in this plane ($C_{11} \approx C_{33} \approx 21$ GPa) than in the *b*

direction along which the pleated sheets are stacked ($C_{22} = 14$ GPa). All three shear elements of the stiffness matrix have relatively large values for molecular crystals ($C_{ii} > 5$ GPa, $i = 4, 5, 6$) describing resistance to shearing along the axial planes.

In contrast, the orthorhombic form II shows considerable anisotropy in both the axial ($C_{11} : C_{22} : C_{33} = 4.4 : 2.2 : 1$) and shear ($C_{44} : C_{55} : C_{66} = 4.7 : 1 : 15.3$) diagonal elements of the elastic tensor. The large axial matrix element along the a axis correlates with the $\text{O-H}\cdots\text{O}=\text{C}$ hydrogen bonds being predominantly along this axis. There is also significant intermolecular stiffness in the b direction due to the $\text{N-H}\cdots\text{O(H)-C}$ hydrogen bonding which complete the ab sheets. These hydrogen-bonded sheets are separated by approximately 3.7 \AA along the c direction, and the weakness of the intermolecular forces between the sheets is reflected by a small value of C_{33} . The smallest shearing element for form II is C_{55} representing the facile slippage of the ab plane along the a axis which produces the desirable tabletability.

Quantitative experimental information on the elastic properties of paracetamol is limited. Values for the Young's modulus of a compaction of monoclinic paracetamol ^{51, 52} vary by about 40%. The more recently determined higher value of 11.7 GPa is bounded by the calculated Reuss ⁵³ ($E_R = 11.5$ GPa) and Voigt ⁵⁴ ($E_V = 14.9$ GPa) averages. These two averages assume a macroscopically homogeneous aggregate, which may not be valid as the habit of single crystals of monoclinic paracetamol is found to be prismatic to plate-like. ¹⁰ In the limit of idealised stacked plates, the Voigt average is more appropriate. ⁵⁵ However, all the calculated elastic constants are expected to be too high (of order of 40% by comparison with other hydrogen bonded molecular crystals ³⁹) because of the neglect of thermal softening. Thus, the results are in reasonable agreement with the known elastic properties of paracetamol, and qualitatively explain the different tabletability of the two forms.

The elastic properties of the two lowest energy crystal structures found in the search are also shown in table 6.4, and for the other hypothetical structures in table 6.5. The elastic properties calculated from the experimental structures are well reproduced by the corresponding structures found in the search (table 6.4), apart from the underestimate of C_{11} for form II which can be attributed to the differences in the hydroxyl proton position in the two rigid molecular models. The predicted bulk properties of the known and hypothetical structures of paracetamol do not vary greatly, consistent with the small variation in the density. However, the anisotropy of the

stiffness matrix varies considerably, as expected from the differences in the hydrogen-bonding networks.

All the elastic stiffness matrix elements C_{ij} are positive definite, satisfying the Born criteria for mechanical stability. However, one of the hypothetical structures, CC8, has C_{55} of only 0.2 GPa, only just satisfying the Born criterion and indicating that the structure is even more easily deformed than the orthorhombic form II.

The predicted morphologies for the experimental and low energy structures found in the search are shown in figure 6.6. The minor differences between the experimental structures and the corresponding search minima have a small effect on the predicted morphologies. As previously found ¹⁴, the attachment energy predictions are in reasonable agreement with the experimental morphology of mature crystals of form I grown from IMS solution, but are insufficiently elongated along c for form II, although the observed faces and chunky prismatic shape are correctly predicted. Many of the hypothetical low energy structures are also predicted to have fairly equant prismatic habits. There are qualitative differences between the attachment energy and the BFDH model predictions for AM30 and AY8, as found for the experimental structures ¹⁴, resulting from variations in relative areas of the main faces, and sometimes additional small faces. However the two methods always predicted similar morphologies.

Figure 6.6 shows that the morphologies of four of the hypothetical structures (AK6, CC19, AK22 and AK4) have distinct platy shapes. The distinction between these extreme plate-like morphologies and the other approximately isodimensional morphologies is so marked that it is unlikely to be changed by the use of a more realistic force-field or method of predicting morphologies.

The differences in morphology are quantified in table 6.6, which gives the ratio of distances between the centre of the furthest and nearest faces (the aspect ratio) for both the attachment energy and BFDH morphologies. Using the assumption that the growth rate of a face is proportional to the magnitude of its attachment energy, the slowest growing face is tabulated for each crystal structure, along with its relative area and attachment energy. This is usually the morphologically dominant face, though if another slow growing face occurs more frequently, this may have a larger relative surface area in the crystal. The attachment energies for any such faces, and for any faces which have a larger inter-planar spacing than the slowest growing face, and so would be larger in the BFDH model morphology, are also given in table 6.6. This shows that such

faces only have moderately more attractive attachment energies and hence faster growth rates, than the slowest growing faces.

Since all the attachment energies in table 6.6 are calculated for crystal structures of the same molecule, using the same intermolecular potential, we can use them to estimate the relative growth rates of the different crystal structures. The attachment energies for the slowest growing faces are largest for the experimental polymorphs, their corresponding structures found in the search, AY8 and AQ14. The hypothetical crystal structures which are predicted to form thin plates have attachment energies for the slowest growing, morphologically dominant faces, which are less than a fifth of those for the experimentally observed forms. Hence the thin plate crystals will grow considerably more slow than the experimental and other more isodimensional morphologies.

6.4 Discussion

The comparison of the eleven experimental molecular structures within form I and three within form II show that the paracetamol molecule in its crystalline phases is as well represented by a rigid *ab initio* optimised structure as can be reasonably expected for a typical organic molecule. There are variations with crystal structure determination and temperature, but these are mainly in the positions of the hydrogen atoms and reflect environmental distortions, dynamical motions and the limitations of X-ray crystallography ⁴⁴ which are much as would be expected, and are only quantified through the extensive neutron studies. However, lattice energy minimisation with the same model intermolecular potential overestimates the *b* cell length by 1.7% in CCW_20 K and underestimates it by 2.7% in CCW_330 K. Hence a variation of over 4% in the minimised *b* cell length of form I comes from changing the experimental determination and therefore the molecular structure. The *c* cell length similarly can be under- or over-estimated using the same model potential, though with smaller but more erratic variations. The more limited data for form II also suggest that the differences between the lattice energy minimum and corresponding experimental structure are too small and erratic to reflect inadequacies in the model potential.

Thus, the comparison of the paracetamol lattice energy minima suggests that differences of around 3% between the minimised and experimental structure may arise from approximations other than inadequacies in the model potential. The cell parameters of the experimental structures for form I vary by about 0.5% for *a*, 2.3% for

b and 1.3% for *c* between 20 K and room temperature. These thermal expansions are approximately similar to the larger differences between each experimental structure and the corresponding lattice energy minimum for *b* and *c*. Thus this comparison of lattice energy minima of paracetamol produces the same conclusion as the typical thermal expansion argument, namely that a few % in the cell parameters is satisfactory agreement between lattice energy minima and experimental structures for neutral organic molecules.

The reduced cell parameter *a* of the monoclinic structure is consistently overestimated, (though only by 1.6 to 3.2%), so that it could be interpreted as revealing a deficiency in the model potential, probably in the CO \cdots H–O interactions. However, this conclusion could only be reached from having a range of determinations over two polymorphs, and because it is reasonable for hydrogens bonded to oxygen to have different repulsion–dispersion parameters from those bonded to nitrogen. The results from comparing any one lattice-energy minimum and the corresponding crystal structure would suggest that the potential was adequate within the limitations of lattice energy minimisation. Thus, although the model potential used in this study is far from definitively accurate, and the theory of intermolecular forces suggests improvements in the repulsion–dispersion and neglected contributions, such improvements cannot be made through empirical fitting using lattice energy minimisation.

The systematic search for lattice energy minima has found the known stable polymorph as the global minimum. Therefore it would have been “predicted” with moderate confidence as the most likely crystal structure for paracetamol. However, there are at least a dozen crystal structures which are within 3.6 to 10 kJ/mol of the most stable form, and therefore within the energy range of possible polymorphism. The energies separating these energetically feasible crystal structures are so small that improvements in the model intermolecular potential, changes in the molecular model, including intramolecular distortion, or entropy effects would probably change the relative energies. Hence, it is probably fortuitous that the well-characterised metastable polymorph has the second lowest lattice energy. Thus, there are various other hypothetical structures that appear as likely to be observed as the known metastable polymorph, on the lattice energy criterion used in crystal structure prediction. Hence, lattice energy minimisation is not sufficient to predict the polymorphism of paracetamol, and has produced a larger number of hypothetical structures than the three suggested by the extensive experimental studies of this important compound. What

other factors can be used to decide which of the energetically feasible crystal structures will be observed?

The effects of temperature need to be considered. The two structures CB9 and CB47 (form II) are very similar and only differ by a translation of the hydrogen bonded sheets, producing a different stacking of the methyl groups. The two minima were located from similar starting structures, suggesting that the potential energy surface between the two minima is such that thermal motion would easily produce the transformation of CB9 to the more stable CB47. This is born out by the small resistance to shear of the hydrogen bonded sheets shown in table 4 and 5. Thus the less stable CB9 is unlikely to be observed, whereas CB47 is form II within the limitations of the modelling.

This and other transformations from the hypothetical structures to more stable minima might be observed by performing Molecular Dynamics simulations, as this ‘shake up’ has been reported to significantly reduce the number of minima in the case of acetic acid.⁵⁶ For example, inspection of the crystal structures, hydrogen bonding motifs and space groups suggests transformations between AQ6 and AQ14 might be quite facile. However, given that this type of inspection would suggest a similar relationship between the known polymorphs of indigo⁵⁷ and terephthalic acid⁵⁸, it is clear that more elaborate dynamical studies are required to determine whether the barrier is sufficient for both the similar crystal structures to be observed. We have considered other, readily calculated, properties of the energetically feasible hypothetical structures to determine whether they are kinetically and mechanically plausible polymorphs.

The mechanical stability of the growing crystallite may also be a factor in determining which polymorphs are observed. Very soft crystallites are likely to be distorted by the stresses experienced during crystallisation. Although all lattice energy minima must satisfy the Born criteria for mechanical stability, those which have a very small diagonal shear elastic constant only just meet this criterion. Such hypothetical structures are less likely to be sufficiently mechanically stable to grow suitable crystals for structure determination. The fact that a small shear constant ($C_{55} \approx 0.7$ GPa) is calculated for the observed form II of paracetamol, but that it is an unusually deformable hydrogen-bonded molecular crystal, shows that a lower boundary for mechanical stability is less than 0.7 GPa. It seems unlikely that CC8, with a shear

elastic constant of 0.2 GPa is likely to be sufficiently mechanically stable to grow in competition with forms of similar thermodynamic stability.

Kinetic factors certainly have a major influence on which polymorphs are formed. Using attachment energy calculations, the observed forms are two of the four fastest growing crystal structures of paracetamol. Although it is physically reasonable that the fastest growing of the approximately isoenergetic crystal structures should be the observed forms, and our results are consistent with this hypothesis, this remarkable success has to be viewed with caution because of the limitations of the attachment energy model (chapter 4). However, if we take the attachment energy of the slowest growing, morphologically important face as a measure of the speed of growth of a crystal, then if several low energy crystal structures of paracetamol had nucleated and were growing, the observed forms are predicted to produce the largest crystals.

This argument can only compare the low energy forms that nucleate and grow under the same crystallisation conditions, and so we cannot make quantitative comparisons of the likelihood of observing crystal structures with reasonable growth rates. However, four of the hypothetical structures have particularly small attachment energies for their slowest growing face, which in all cases dominates their thin plate-like morphology. It has been observed that difficulties in growing crystals of a size suitable for X-ray diffraction are generally associated with thin plate-like or fine needle morphologies.⁵⁹ Therefore, it seems unlikely that polymorphs of such habits will be found when structures with more equant habits are equally thermodynamically favourable.

Thus, consideration of simple models of mechanical properties and crystal growth rates can reduce the number of hypothetical structures that appear to be possible polymorphs, as summarised in table 7. Undoubtedly, the relative thermodynamic stability of these structures will depend on temperature and pressure. The question arises as to whether the elusive third polymorph can be expected to be amongst the remaining four to six structures. Since the third polymorph has only been observed⁷ under the cover slip of a microscope slide, it may be a pressure stabilised form and hence would not be found even by a more exhaustive search for minima in the lattice energy under conditions of zero pressure.

6.5 Conclusions

The accuracy with which lattice energy minimisation can predict crystal structures using an *ab initio* molecular structure is qualitatively excellent, but the crystal structures are not reproduced quantitatively as well as if the actual experimental molecular structures had been used. In the case of paracetamol, the only systematic difference between the experimental and the *ab initio* structures that is likely to be a genuine effect of the crystal packing is a small out-of-plane rotation of the hydroxyl group. Thus, this is a system where the rigid molecule approximation in a genuine crystal structure prediction should be fairly good. This is apparently the case, as the prediction model gives cell lengths that are generally within the range of the experimental determinations and reproductions, and only slightly outside this range in the directions most affected by the intramolecular rotation of the hydroxyl group.

The differences in the rigid methyl group model appear to have a fairly minor but observable effect on the crystal structure modelling of paracetamol. This is similar to the effect of positioning of the methyl groups in theophylline, where a full 60° rotation of the methyl group left the lattice energy remarkably unaffected.⁶⁰ Hence, it seems that the uncertainties associated in using a rigid model for the methyl hydrogen atoms may be quite small, but this is very dependent on the crystal packing. This is fortunate, as it would be very demanding to model the dynamical motion of methyl groups⁶¹ beyond correcting for the apparent bond shrinkage caused by librational motion. Recent rigid-body lattice energy calculations on PF₆ salts of metal complexes showed that correcting the P–F bondlength for librational motion had a significant effect on calculated lattice energies.⁴⁷ The standard methodology adopted here, of correcting X-ray determined proton positions to give standard bondlengths, does provide a rigid-body correction for the foreshortening caused by the librational motion. However, the neutron molecular structures were not corrected for this effect, and the results indicate that this might have had a small effect on the lattice energy minima for the higher temperature structures.

Although the changes in the molecular and crystal structures are not atypical for the range of temperatures and types of determination considered, the sensitivity of calculated lattice energies to the positions of the hydrogen atoms does result in some variation in the predicted lattice energies. The differences of order of 1 kcal mol⁻¹ (i.e. 2 kJ mol⁻¹ for form I, 5 kJ mol⁻¹ for form II) are small relative to the total lattice energy.

Even the lattice energy estimates using the gas phase molecular structure are not unreasonable compared with typical experimental errors in determining of heats of sublimation and the other errors implicit in comparing the two quantities. Unfortunately these differences are highly significant compared with the estimated energy difference between the two polymorphs.

Thus, even for good quality experimental crystal structures of rigid molecules, discrepancies in lattice energy minima of a few percent in the cell dimensions and a kcal mol⁻¹ or so in the lattice energy could arise from variations in the experimental molecular structure. Thus static minimisation estimates of the energy differences between polymorphs, based on experimental molecular and crystal structures, have associated errors that are generally comparable to the actual energy difference.

The systematic search for the rigid body lattice energy minima of paracetamol correctly identifies the stable crystal structure as the global minimum. A variety of energetically feasible alternative structures are predicted, including the well-characterised orthorhombic form. Consideration of the predicted morphologies and growth rates, using the attachment energy model, can be used to eliminate some of the hypothetical structures, as they are growing as thin plates so slowly that they are unlikely to be found in competition with equally thermodynamically stable structures.

Calculation of the elastic tensors of the experimental polymorphs quantifies the microscopic reason for the very different tabletabilities of the two forms. A few of the hypothetical structures (AI22 and AI6 in $P2_1/c$ and AQ14 in $P2_12_12_1$) also have the low resistance to shear that produces the desirable direct tableting of the known metastable orthorhombic form. However, this resistance to shear is so low for one hypothetical structure (CC8) that it is unlikely to be found.

The proposed additional criteria of considering the growth rate, as practically and approximately computed through the attachment energy, and the mechanical stability, through eliminating structures that are only just stable according to the calculated elastic constants, is promising. Taken in conjunction with the lattice energy, this approach favors the experimentally known structures of paracetamol, and reduces the number of hypothetical structures somewhat, though still leaving more than have been detected experimentally in this much-studied pharmaceutical.

This method has been applied to carboxylic acids and the results are discussed in the following chapter.

ab initio values	Form II, X-ray			form I, X-ray				form I, neutron							
	HXACAN	CSFII LT	CSFII RT	HXACAN01	HXACAN04	CSFI RT	CSFI LT	CCW, 20K	CCW, 50K	CCW, 80K	CCW, 150K	CCW, 200K	CCW, 250K	CCW, 300K	
	Δ / %	Δ / %	Δ / %	Δ / %	Δ / %	Δ / %	Δ / %	Δ / %	Δ / %	Δ / %	Δ / %	Δ / %	Δ / %	Δ / %	
Bondlengths/Å															
C2 C3	1.381	0.3	0.7	-0.2	-0.1	0.5	0.0	0.5	1.2	1.2	1.5	1.4	2.0	1.8	1.6
C5 C6	1.385	0.4	0.1	-0.4	0.6	0.0	-0.6	-0.1	0.4	0.6	0.6	0.4	1.0	0.7	1.2
C4 O1	1.354	2.0	1.8	1.6	1.7	1.6	1.3	1.2	0.7	0.9	-0.1	0.7	0.8	0.5	0.8
N1 C7	1.360	-1.3	-1.1	-1.7	-1.4	-1.0	-1.5	-1.1	-1.1	-1.0	-1.2	-1.0	-0.9	-0.8	-1.6
C7 O2	1.198	2.1	3.2	3.1	2.8	3.2	2.8	3.2	3.1	3.3	3.1	2.7	2.3	2.4	2.1
C7 C8	1.515	-0.3	-0.5	-0.7	-0.4	-0.4	-0.7	-0.7	-0.6	-0.5	-0.1	0.3	0.7	1.2	0.8
C2 H2	1.069	-11.2	-8.5	-13.0	-13.2	-12.8	-11.8	-10.6	0.7	0.5	1.2	0.2	-0.9	-0.1	-1.2
C3 H3	1.074	-12.7	-10.7	-13.5	-8.9	-10.3	-9.7	-8.3	0.2	0.7	-0.7	1.0	1.5	1.0	0.5
O1 H1	0.942	-12.7	-4.7	-13.0	-6.1	-2.1	-2.5	-1.3	4.9	4.7	4.9	3.3	3.9	3.6	2.3
C5 H5	1.077	-10.8	-10.8	-13.7	-11.8	-10.4	-10.7	-8.4	0.4	0.4	-0.4	-0.3	0.1	0.2	-0.1
C6 H6	1.077	-12.3	-7.7	-13.7	-10.9	-9.1	-11.6	-9.4	0.3	0.5	-1.2	-1.9	-1.3	-2.5	-1.9
N1 H4	0.993	-20.1	-6.7	-13.4	-9.0	-6.9	-8.2	-7.2	1.3	1.2	2.2	1.0	1.7	0.1	-1.5
C8 H7	1.084	-14.8	-6.8	-11.4	-19.0	-13.4	-14.2	-10.5	-0.9	-1.4	-6.6	-10.2	-11.5	-13.3	-14.7
C8 H8	1.08	-18.8	-8.9	-11.1	-12.5	-13.1	-10.3	-9.5	-0.1	0.3	-1.1	-1.5	-4.2	-2.8	-5.8
C8 H9	1.086	-11.1	-12.1	-11.6	-1.1	-13.5	-11.1	-12.2	0.3	0.0	-0.3	-4.0	-3.9	-5.0	-9.6
Angles/°															
O1 C4 C5	123	-1.6	-1.6	-1.6	0.0	-0.8	0.0	-0.8	-0.8	-0.8	-1.6	-0.8	-0.8	-0.8	-1.6
C6 C1 N1	118	-1.7	-1.7	0.0	-0.9	-0.9	-0.8	-0.8	-1.7	-1.7	-1.7	-1.7	-1.7	-1.7	-1.7
C1 N1 C7	129	0.8	0.0	0.8	-0.8	-0.8	-0.8	-0.8	-0.8	-0.8	-0.8	-0.8	-0.8	-1.6	-0.8
O2 C7 C8	121	1.7	1.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
H1 O1 C4	111	-0.9	-2.7	-1.8	-2.7	0.0	-0.9	0.0	0.9	0.9	0.9	1.8	0.9	1.8	1.8
C7 C8 H7	113	0.0	-0.9	-3.5	-0.9	-2.7	0.9	0.9	0.9	-0.9	-0.9	0.0	0.0	0.9	-0.9
C7 C8 H8	108	2.8	-1.9	0.9	0.9	0.9	1.9	2.8	1.9	1.9	1.9	2.8	0.0	0.9	2.8
C7 C8 H9	109	-1.8	0.9	0.0	0.0	0.0	0.0	0.9	0.0	0.0	-0.9	0.0	0.9	-1.8	0.9
Overall/Å															
O1C8	7.918	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
O1O2	6.506	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Torsions/°															
H1O1C4C5	0	-23	22	-1	-19	17	16	17	-18	-17	-19	-18	-17	-17	-17
H1O1C4C3	180	158	-159	-180	161	-163	-164	-164	163	163	162	163	162	163	163
O1C4C5H5	0	-3	4	3	1	-1	-3	-1	2	2	0	1	0	2	1
O1C4C3H3	0	2	-2	-2	-6	1	2	1	-3	-3	-3	-3	-3	-3	-4
C6C1N1H4	12	15	-19	-16	13	-15	-18	-16	15	15	15	15	15	16	16
N1C7C8H7	-33	-29	22	0	6	17	-1	11	-14	-14	-12	-14	-10	-11	-23
N1C7C8H8	-154	-159	-100	-120	-146	-103	-106	-106	-138	-139	-141	-143	-146	-147	-153
N1C7C8H9	88	84	147	120	110	137	147	139	106	105	105	104	104	104	94

Table 6.1 Differences between the experimental and the *ab initio* optimised molecular structure of paracetamol.

For the bond lengths, angles and for the overall dimension, % errors with $\Delta = 100(\text{expt} - \text{ab initio})/\text{ab initio}$ and for the torsions the absolute values in degrees are listed. Parameters for which the comparison of the *ab initio* form with all experimental structures yields $\Delta < 1\%$ or torsional differences of $\delta < 3^\circ$ are omitted. (atom labels are shown in Scheme 6.1.)

	T/K	Origin. Space group	Method (R factor)	Reduced cell parameters /Å (error in min Δ/%)				Cell Vol /Å ³	Initial lattice energy / kJ/mol	Min lattice energy / kJ/mol	r.m.s. % Error	Hydrogen bonds	
				a	b	c	β/°					N(H)···O /Å	O(H)···O /Å
FORM I													
CCW_20K	20	<i>P2₁/a</i>	Neutron	7.073 (3.2)	9.166 (1.7)	11.546 (0.1)	98.1 (2.0)	741.16 (4.5)	-111.1	-118.3	2.2	2.90 (0.9)	2.65 (8.2)
CCW_50K	50	<i>P2₁/a</i>	Neutron	7.073 (3.0)	9.173 (1.6)	11.572 (0.1)	98.0 (2.1)	743.54 (4.2)	-111.6	-118.6	2.1	2.90 (0.8)	2.65 (8.0)
CCW_80K	80	<i>P2₁/a</i>	Neutron	7.077 (2.7)	9.173 (1.5)	11.574 (0.1)	97.9 (2.3)	744.23 (3.7)	-113.0	-119.6	1.9	2.90 (0.7)	2.65 (7.4)
CSFI_LT	123	<i>P2₁/n</i>	X-ray	7.094 (3.5)	9.232 (0.5)	11.620 (-0.3)	97.8 (2.2)	753.94 (3.2)	-111.2	-118.4	2.0	2.91 (0.5)	2.66 (8.4)
CCW_150K	150	<i>P2₁/a</i>	Neutron	7.079 (2.4)	9.240 (0.0)	11.628 (0.3)	97.8 (2.6)	753.61 (2.1)	-113.7	-119.7	1.6	2.91 (0.0)	2.65 (7.3)
HXACAN04	150	<i>P2₁/n</i>	X-ray (0.055)	7.094 (3.4)	9.262 (0.3)	11.657 (-0.7)	97.7 (2.4)	759.09 (2.3)	-110.6	-118.3	2.0	2.91 (0.3)	2.66 (8.6)
CCW_200K	200	<i>P2₁/a</i>	Neutron	7.084 (2.1)	9.278 (0.0)	11.664 (0.0)	97.7 (2.8)	759.77 (1.3)	-114.1	-119.9	1.7	2.91 (-0.1)	2.66 (6.7)
CCW_250K	250	<i>P2₁/a</i>	Neutron	7.086 (2.0)	9.315 (-0.5)	11.676 (0.0)	97.6 (2.9)	763.89 (0.7)	-113.8	-119.6	1.7	2.92 (-0.5)	2.66 (6.7)
HXACAN01	295	<i>P2₁/a</i>	X-ray (0.072)	7.100 (2.7)	9.400 (-0.4)	11.721 (-0.7)	97.1 (2.8)	776.27 (0.9)	-110.9	-118.3	1.9	2.93 (-0.5)	2.66 (7.8)
CSFI_RT	298	<i>P2₁/n</i>	X-ray	7.106 (3.0)	9.382 (-0.7)	11.704 (-1.1)	97.4 (2.8)	773.88 (0.5)	-112.9	-119.4	1.9	2.93 (-0.6)	2.67 (7.4)
CCW_330K	330	<i>P2₁/a</i>	Neutron	7.085 (1.6)	9.370 (-2.7)	11.706 (1.1)	97.5 (3.6)	770.51 (-1.0)	-115.2	-120.4	2.1	2.93 (-1.2)	2.66 (6.1)
Prediction model				7.278	8.944	12.119	100.0	776.99		-110.1		2.91	2.88
FORM II													
CSFII_LT	123	<i>Pbca</i>	X-ray	7.212 (3.4)	11.777 (0.8)	17.166 (1.8)	90 (0.0)	1458.08 (6.1)	-111.9	-117.4	2.2	2.94 (0.3)	2.71 (6.9)
HXACAN	295	<i>Pcab</i>	X-ray (0.077)	7.393 (0.2)	11.805 (1.2)	17.164 (1.6)	90 (0.0)	1497.98 (3.0)	-112.0	-116.2	1.1	2.97 (-0.5)	2.73 (6.9)
CSFII_RT	298	<i>Pbca</i>	X-ray	7.405 (-0.1)	11.831 (0.6)	17.156 (2.5)	90 (0.0)	1502.99 (2.9)	-108.2	-112.2	1.5	2.98 (1.8)	2.73 (6.9)
Prediction model				7.384	12.076	17.263	90	1539.25		-106.5		2.92	2.95

Table 6.2 Experimental crystal structures of paracetamol and the changes in parameters on lattice energy minimisation.

 $\Delta/\% = 100(\text{min-expt})/\text{expt}$

	Space group	a/Å	b/Å	Reduced cell parameters				Hydrogen bonds ^a			U _{latt} / kJ mol ⁻¹	ΔU _{latt} ^b / kJ mol ⁻¹	ρ / g cm ⁻³	
				c/Å	α/°	β/°	γ/°	O-(H)···O=C / Å (°)	N-(H)···O-H / Å (°)	O-(H)···O-H / Å (°)				
Form I														
Expt _{low}	<i>P2₁/a</i>	7.073	9.166	11.546	90.0	98.1	90.0	2.65 (165)	2.90 (164)	-	-	-111.1	-	1.35
Expt _{high}	<i>P2₁/a</i>	7.085	9.370	11.706	90.0	97.5	90.0	2.66 (167)	2.93 (163)	-	-	-115.2	-	1.30
Min _{exptlow}	<i>P2₁/a</i>	7.297	9.326	11.554	90.0	100.0	90.0	2.86 (163)	2.92 (160)	-	-	-118.3	-	1.30
Min _{expthigh}	<i>P2₁/a</i>	7.201	9.120	11.829	90.0	100.9	90.0	2.82 (161)	2.90 (160)	-	-	-120.4	-	1.32
Min _{opt}	<i>P2₁/a</i>	7.278	8.944	12.120	90.0	100.0	90.0	2.88 (155)	2.91 (164)	-	-	-110.1	0	1.29
AM30	<i>P2₁/c</i>	7.277	8.944	12.119	90.0	100.0	90.0	2.88 (155)	2.91 (164)	-	-	-110.1	0	1.29
Form II														
Expt _{low}	<i>Pbca</i>	7.212	11.777	17.166	90.0	90.0	90.0	2.71 (171)	2.94 (164)	-	-	-111.9	-	1.38
Expt _{high}	<i>Pbca</i>	7.405	11.831	17.156	90.0	90.0	90.0	2.73 (144)	2.98 (160)	-	-	-108.2	-	1.34
Min _{exptlow}	<i>Pbca</i>	7.454	11.877	17.471	90.0	90.0	90.0	2.90 (172)	2.95 (161)	-	-	-117.4	-	1.30
Min _{expthigh}	<i>Pbca</i>	7.398	11.897	17.578	90.0	90.0	90.0	2.92 (156)	3.03 (157)	3.71 (95)	-	-112.2	-	1.30
Min _{opt}	<i>Pbca</i>	7.384	12.076	17.263	90.0	90.0	90.0	2.95 (146)	2.92 (153)	3.69 (99)	-	-106.5	3.6	1.30
CB47	<i>Pbca</i>	7.382	12.086	17.248	90.0	90.0	90.0	2.95 (145)	2.93 (153)	3.69 (99)	-	-106.5	3.6	1.30
Hypothetical structures which have not been experimentally characterised														
AI22	<i>P2₁/c</i>	6.749	8.248	13.540	96.8	90.0	90.0	2.80 (166)	-	-	-	-105.6	4.5	1.34
AY8 ^c	<i>Pca2₁</i>	4.522	10.692	15.863	90.0	90.0	90.0	2.91 (163)	3.07 (168)	-	-	-105.4	4.7	1.31
CC19 ^d	<i>P2₁/c</i>	5.073	9.886	31.121	97.8	90.0	90.0	-	-	2.92 (143) 2.87(146)	3.38 (136) 3.54(134)	-103.3	6.8	1.30
AQ6	<i>P2₁2₁2₁</i>	6.571	7.344	16.281	90.0	90.0	90.0	2.80 (163)	-	-	-	-102.8	7.3	1.28
AK6	<i>P2₁/c</i>	5.072	9.648	16.050	100.9	90.0	90.0	-	-	2.90 (143)	3.49 (135)	-101.9	8.2	1.30
AQ14	<i>P2₁2₁2₁</i>	6.776	7.071	15.834	90.0	90.0	90.0	2.80 (174)	3.98 (139)	-	-	-101.9	8.2	1.32
CB9	<i>Pbca</i>	7.266	12.207	17.432	90.0	90.0	90.0	3.06 (158)	2.96 (146)	-	3.96 (93)	-101.8	8.3	1.30
CC8	<i>Pbca</i>	6.848	13.491	16.664	90.0	90.0	90.0	2.79 (170)	-	-	-	-101.0	9.1	1.30
AK22	<i>P2₁/c</i>	5.071	9.861	15.834	93.3	90.0	90.0	-	-	2.89 (144)	3.43 (136)	-100.2	9.9	1.27
AI16	<i>P2₁/c</i>	7.553	8.112	12.716	90.0	103.5	90.0	3.19 (114)	3.04 (179)	-	-	-100.2	9.9	1.33
AM4	<i>P2₁/c</i>	5.937	7.590	17.071	90.0	99.3	90.0	2.96 (153)	-	3.84 (112)	4.00 (145)	-100.1	10.0	1.32
AK4	<i>P2₁/c</i>	5.293	8.080	19.034	101.6	90.0	90.0	-	-	2.91 (153)	3.34 (144)	-100.0	10.1	1.26

Table 6.3 Low energy crystal structures of paracetamol

The lowest (Expt_{low}) and highest (Expt_{high}) temperature experimental structures of paracetamol (20 K and 330 K neutron structures for form I ¹⁶, and 123 K and 298 K X-ray structures for form II ¹⁴) are contrasted with lattice energy minima calculated using the same intermolecular potential. All minimisations used the *ab initio* optimised molecular structure, except Min_{exptlow} and Min_{expthigh}, where the experimental molecular structures were used (X-ray bondlengths to hydrogen atoms were standardised ³²). The minima obtained with the *ab initio* structure Min_{opt} were the same for all experimental determinations of each form. The lowest energy crystal structures found in the MOLPAK/DMAREL search are designated by the MOLPAK coordination group and number, and the space group is that of the conventional setting of each structure. The Niggli reduced cell parameters ⁵⁰ are tabulated here to aid comparison.

^aThe hydrogen bonds were characterised using PLUTO ⁶², tabulating the N/O···O distance (<4 Å) and N/O-H···O angle (>90°), and each hydrogen bond has one occurrence per molecule. ^bΔU_{latt} is the energy above the global minimum found in the search. ^cThis structure has a total dipole moment of 3.1 eÅ per unit cell along *b*, which may produce a small destabilising dipole correction term.

^dThis P2₁/c Z'=2 structure was found as a true minimum from a Pbca starting structure.

Optical axes	Form I $P2_1/c$			Form II $Pbca$		
	$x // a^*, z // c, y \perp x, z$			$x // a, y // b, z // c$		
	Min _{exptlow}	Min _{opt}	AM30	Min _{exptlow}	Min _{opt}	CB47
C_{11} / GPa	21.2	22.1	22.1	53.2	30.3	30.0
C_{22} / GPa	14.4	12.0	12.0	26.3	22.7	22.7
C_{33} / GPa	21.6	20.1	20.4	12.0	12.7	12.5
C_{44} / GPa	5.5	5.1	5.3	3.5	3.0	3.0
C_{55} / GPa	5.9	6.1	6.0	0.7	0.8	0.8
C_{66} / GPa	8.9	5.9	5.8	11.3	7.6	7.6
C_{12} / GPa	12.3	11.5	11.4	21.2	14.5	14.5
C_{13} / GPa	10.5	12.1	12.0	3.6	4.3	4.4
C_{23} / GPa	11.0	10.3	10.4	8.3	8.8	8.7
C_{15} / GPa	-0.4	0.1	0.1			
C_{25} / GPa	1.4	2.0	2.0			
C_{35} / GPa	2.8	2.8	2.9			
C_{46} / GPa	2.3	2.4	2.4			
Bulk modulus	12.9	11.0	11.0	10.8	10.6	10.6
/ GPa	13.9	13.6	13.6	17.5	13.4	13.4
Shear modulus	4.3	3.6	3.6	2.4	2.3	2.3
/ GPa	5.6	4.8	4.8	7.0	4.8	4.8
Young's modulus	11.5	9.7	9.7	6.7	6.6	6.6
/ GPa	14.9	12.9	12.9	18.5	12.9	12.9
Experimental Young's modulus for Form I ^{51, 52} 8.4 GPa, 11.7 GPa						

Table 6.4 The mechanical properties for the lowest temperature forms of paracetamol and the closest structures found in the search with the *ab initio* molecular structure.

The bulk mechanical properties of a macroscopically homogeneous aggregate have been estimated by both the Reuss and Voigt averages, and given in the format Reuss above Voigt.

	AI22	AY8	CC19	AQ6	AK6	AQ14	CB9	CC8	AK22	AI16	AM4	AK4
C ₁₁	13.8	17.9	10.7	10.0	14.2	12.3	36.9	11.1	14.8	11.7	15.8	18.2
C ₂₂	16.2	11.3	14.8	22.8	14.6	26.8	16.7	29.3	12.7	16.9	12.4	18.7
C ₃₃	41.1	19.7	14.4	11.9	11.0	13.1	10.8	13.1	14.4	23.2	41.3	9.5
C ₄₄	4.8	4.2	6.9	7.2	4.3	10.0	1.8	1.3	3.3	12.5	2.4	4.2
C ₅₅	2.1	5.5	3.5	3.2	7.0	1.8	0.6	0.2	6.9	1.8	5.9	3.8
C ₆₆	0.9	3.1	3.4	2.5	3.3	0.9	6.1	3.2	3.3	1.0	5.0	3.5
C ₁₂	8.2	10.5	8.1	8.3	7.6	4.6	10.8	11.2	7.6	10.0	10.0	6.5
C ₁₃	6.9	9.9	8.8	8.4	10.8	7.2	2.3	7.0	12.5	9.5	9.1	10.2
C ₂₃	16.8	10.5	9.8	9.2	8.4	10.2	8.9	7.6	6.4	8.4	4.7	8.8
C ₁₅	-0.01		1.0		2.8				-3.4	-0.04	-0.4	6.0
C ₂₅	0.4		-3.1		-1.3				1.8	1.5	-1.2	-1.6
C ₃₅	3.9		1.1		1.8				-5.5	-1.3	3.1	1.9
C ₄₆	2.0		0.1		0.4				-0.1	0.9	-0.2	-0.5
K _R	11.5	11.1	9.0	9.4	10.0	9.5	9.1	9.4	9.3	11.1	10.8	7.7
K _V	15.0	12.3	10.4	10.7	10.4	10.7	12.0	11.7	10.5	12.0	13.0	10.8

Table 6.5 Calculated elastic properties (in GPa) for the energetically feasible, hypothetical structures of paracetamol.

In all calculations the *ab initio* optimised molecular model is used. The optical axes are defined with x parallel to **a**, y to **b** and z to **c**, for the orthorhombic crystals (AY8, AQ6, AQ14, CB9, CC8) and z parallel to **c**, x parallel to **a*** and y perpendicular to xz for the monoclinic crystals (AI22, CC19, AK6, AK22, AI6, AM4, AK4). The bounds on the bulk modulus of a homogeneous aggregate of paracetamol are estimated by the Reuss (K_R) and Voigt (K_V) averages.

	Aspect ratio		Slowest growing faces for AE model (frequency of occurrence)	Attachment Energy / kJ mol ⁻¹	% ^a Area of face
	BFDH	AE			
Form I, Expt _{low}	1.4	1.9	110 (4)	-12.8	48.5
Form I, Min _{opt} & AM30	1.4	1.7	110 (4)	-14.7	43.2
Form II, Expt _{low}	1.8	1.7	002 (2)	-13.6	21.4
			<i>111</i> (8)	-17.5	32.1
			*200 (2)	-19.1	
Form II, Min _{opt} & CB47	1.8	1.4	020 (2)	-15.2	16.5
			<i>111</i> (8)	-15.6	42.3
			*200 (2)	-15.7	
AI22	1.5	2.8	100 (2)	-7.2	38.7
			*102 (2)	-19.0	
AY8	2.1	2.2	200 (2)	-14.1	33.2
CC19	7.7	13.9	001 (2)	-1.5	85.1
AQ6	1.7	1.5	110 (4)	-12.3	35.5
AK6	4.0	6.4	100 (2)	-3.1	71.0
AQ14	1.7	1.7	101 (4)	-13.7	34.7
			*020 (2)	-19.8	
CB9	1.8	1.8	200 (2)	-9.9	30.0
CC8	1.7	2.2	002 (2)	-8.1	33.9
			*020 (2)	-14.1	
AK22	4.0	7.7	100 (2)	-2.7	74.6
AI16	1.5	3.3	100 (2)	-6.0	47.3
AM4	1.9	1.8	002 (2)	-11.1	32.2
AK4	5.1	11.2	100 (2)	-1.7	81.7

Table 6.6 Growth rate and morphology predictions for the known and hypothetical forms of paracetamol.

The morphologically dominant, slowest growing face, in the attachment energy (AE) prediction of the morphology is given in normal type, along with corresponding AE. For crystals where another face provides a higher percentage of the total surface area, because of a higher frequency of occurrence, this face is given in italics. The aspect ratio for the AE predicted morphology is contrasted with that using the BFDH model, and where the BFDH model predicts a different slowest growing face, this is given after *.

^a100×(total surface area of slowest growing face)/total surface area

Crystal Structure, Space Group	Lattice Energy / kJ mol ⁻¹	Smallest C _{ii} Elastic Constant / GPa	Attachment Energy of slowest growing face / kJ mol ⁻¹	Description of hydrogen bonded motif
AM30, <i>P2₁/c</i>	-110.1	C ₄₄ =5.3	AE(110) = -15.5	KNOWN , zig-zag sheets hydrogen bonded along <i>a</i> , stacked along <i>b</i>
CB47, <i>Pbca</i>	-106.5	C ₅₅ =0.8	AE(020) = -15.9	KNOWN , flat sheets (slip planes) in <i>ab</i> plane, stacked along <i>c</i>
AI22, <i>P2₁/c</i>	-105.6	C ₆₆ =0.9	AE(110) = -8.2	Soft zig-zag chains along <i>c</i> , approx. parallel stacking along <i>a</i>
AY8, <i>Pca2₁</i>	-105.4	C ₆₆ =3.1	AE(200) = -14.4	3D hydrogen bonding network
CC19, <i>P2₁/c</i>	-103.3	C ₆₆ =3.4	<i>AE(001) = -1.3</i>	UNLIKELY , chains interlinked by OH...OH bonds along <i>a</i>
AQ6, <i>P2₁2₁2₁</i>	-102.8	C ₆₆ =2.5	AE(110) = -13.0	Soft zig-zag chains along <i>b</i> , packed approx. parallel (slightly tilted) along <i>a</i>
AK6, <i>P2₁/c</i>	-101.9	C ₆₆ =3.3	<i>AE(100) = -2.9</i>	UNLIKELY , chains, interlinked by OH...OH bonds along <i>b</i>
AQ14, <i>P2₁2₁2₁</i>	-101.8	C ₆₆ =0.9	AE(110) = -15.8	Soft zig-zag chains along <i>b</i> , approx. parallel (slightly tilted) stacking along <i>a</i>
CB9, <i>Pbca</i>	-101.4	C ₅₅ =0.6	AE(200) = -10.0	UNLIKELY , flat sheets (slip planes) in <i>ab</i> plane, stacked along <i>c</i> , facile transformation to form II
CC8, <i>Pbca</i>	-101.0	C ₅₅ =0.2	AE(002) = -9.0	UNLIKELY , soft zig-zag chains (buckled) along <i>b</i> , strong methyl-methyl group interaction
AK22, <i>P2₁/c</i>	-100.2	C ₄₄ =C ₆₆ =3.3	<i>AE(100) = -2.3</i>	UNLIKELY , chains of dimers, interlinked by OH...OH bonds
AI16, <i>P2₁/c</i>	-100.2	C ₆₆ =1.0	AE(100) = -6.9	Soft zig-zag chains along <i>c</i> , approx. parallel (slightly tilted) stacking along <i>a</i>
AM4, <i>P2₁/c</i>	-100.1	C ₄₄ =2.4	AE(002) = -11.5	Buckled chains along <i>c</i> , stacked along <i>b</i>
AK4, <i>P2₁/c</i>	-100.0	C ₆₆ =3.5	<i>AE(100) = -1.4</i>	UNLIKELY , chains interlinked by OH...OH bonds along <i>b</i>

Table 6.7 Summary of properties of the lowest energy crystal structures of paracetamol found in the lattice energy search.
Properties which make the structure unlikely to be an observed polymorph are in *italic*.

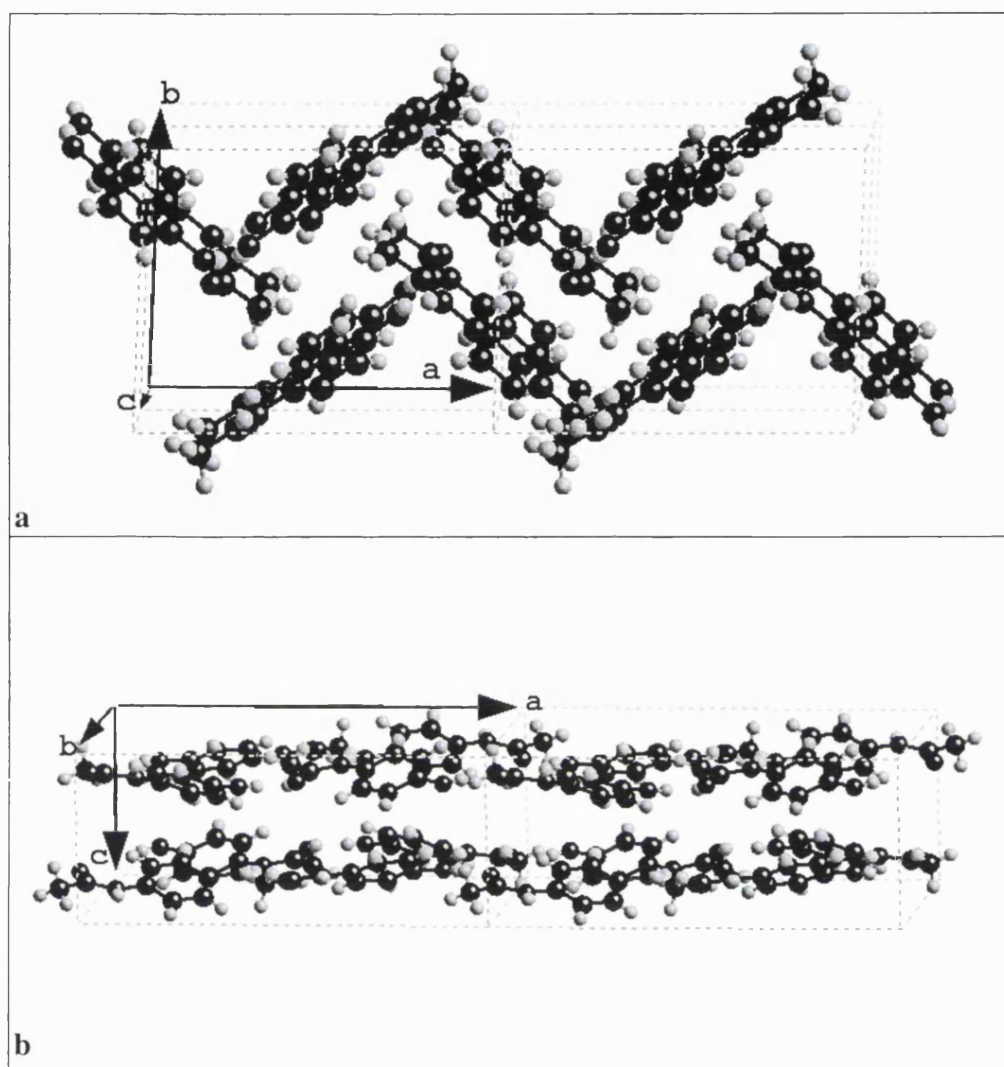


Figure 6.1 The crystal structures of paracetamol in (a) form I (determined at 20 K)¹⁶ and (b) form II (123 K)¹⁵.

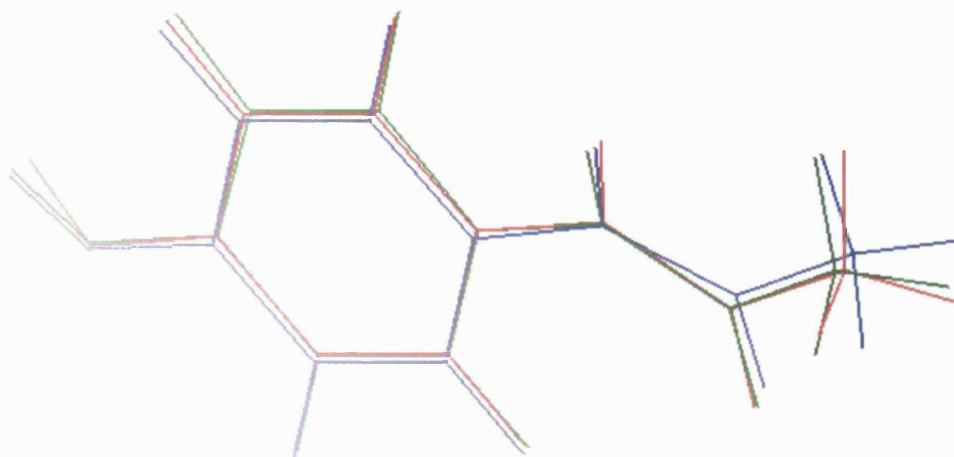


Figure 6.2 An overlay of experimental and *ab initio* molecular structures of paracetamol.
 Blue: the lowest temperature (20 K) neutron determination of form I (CCW_20K).
 Green: the lowest temperature (123 K) determination (X-ray) of form II (CSFII_LT).
 Red: the *ab initio* optimised structure.

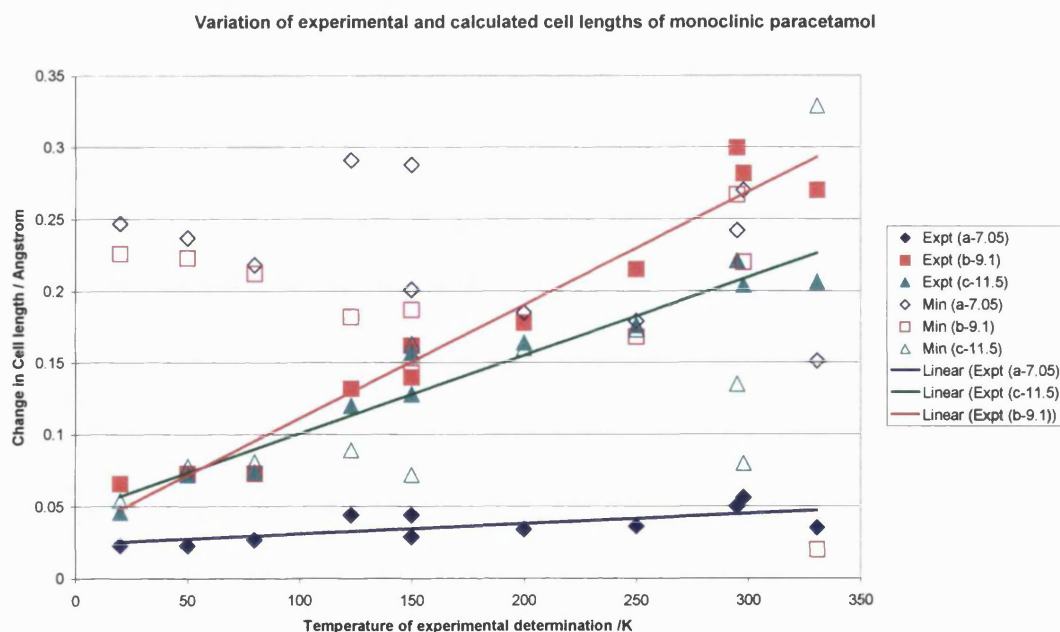


Figure 6.3 Variation of the experimental and lattice energy minimised cell lengths of form I with temperature of determination. Constant offsets of 7.05 Å in *a*, 9.1 Å in *b* and 11.5 Å in *c* have been used.

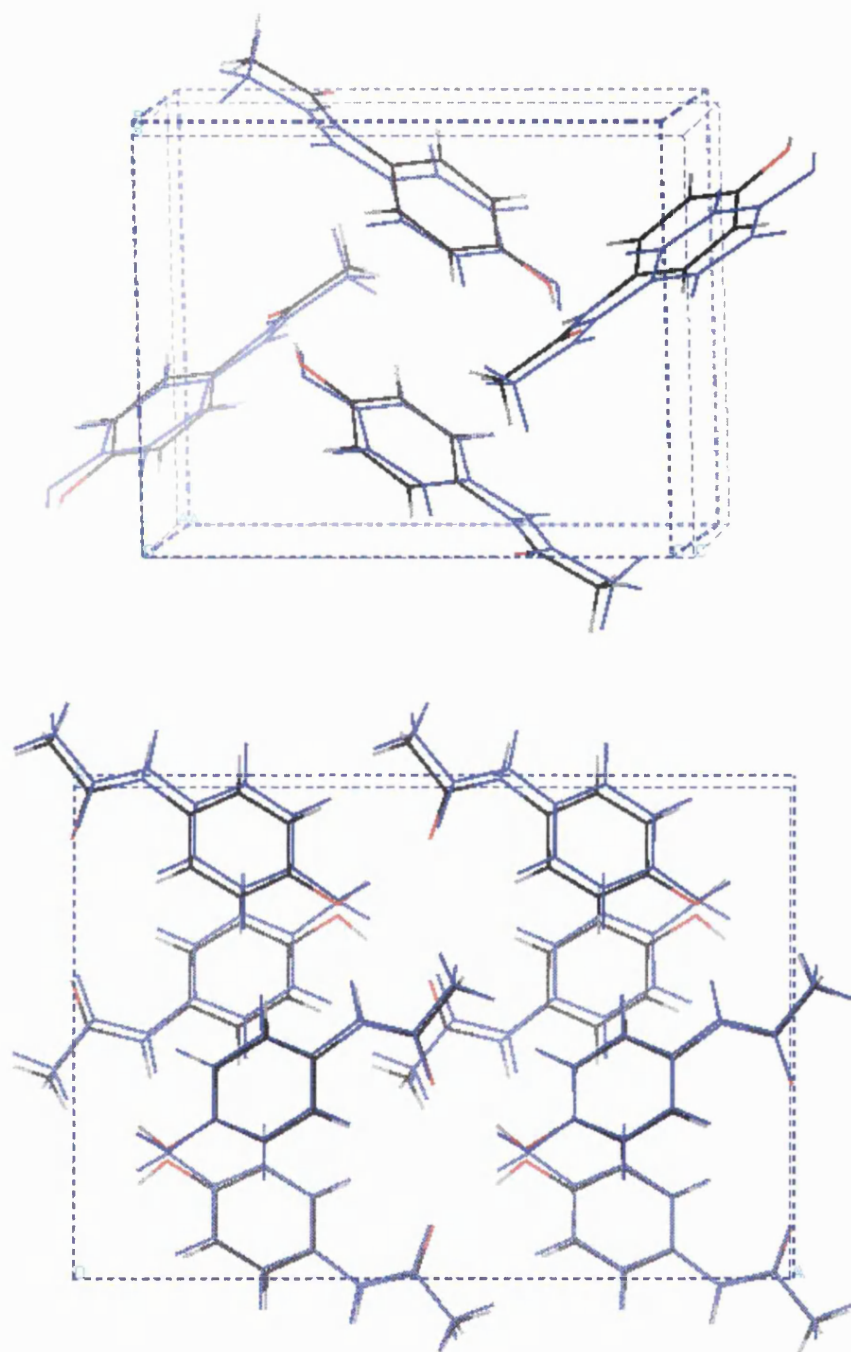


Figure 6.4 Overlay of the experimental (123K) and prediction model (blue) crystal structures.

a) Form I in $P2_1/n$, using CSFI_LT

b) Form II using CSFII_LT in $Pabc$. In the reduced cell setting, the layers are in the bc plane ($b < c$), and this diagram is viewed down a .

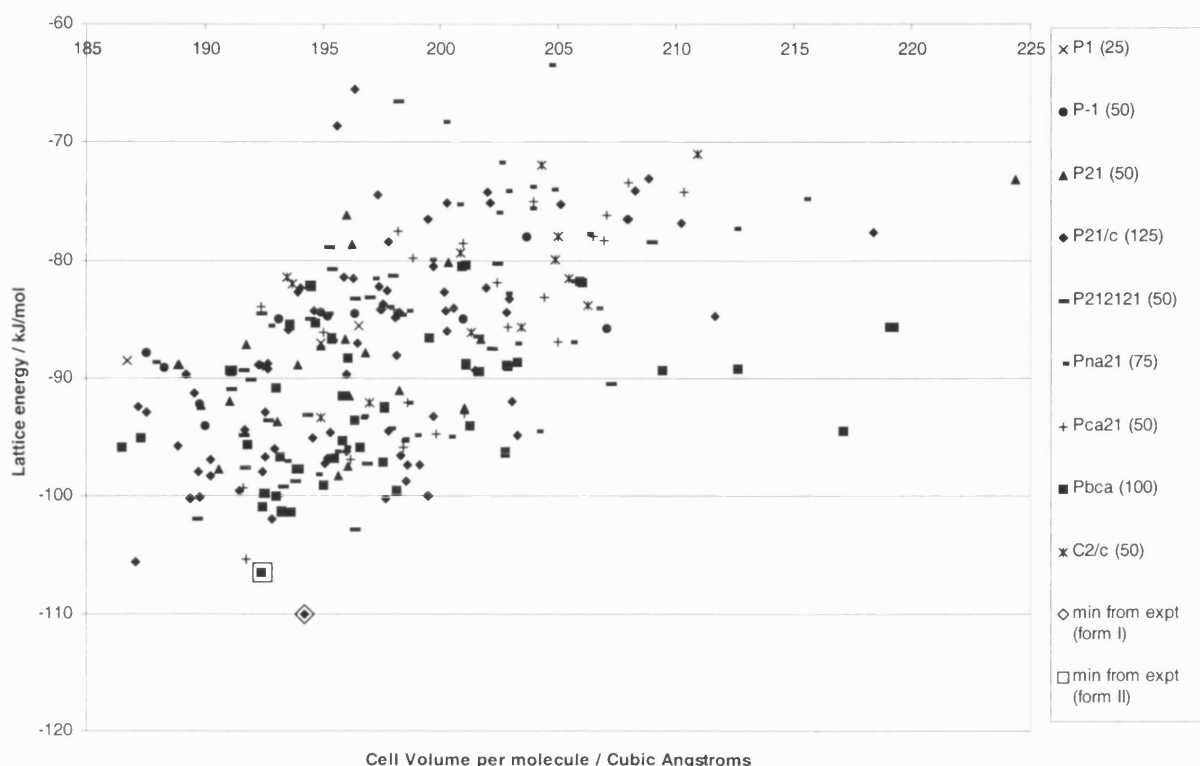


Figure 6.5. The minima in the lattice energy of paracetamol found by the MOLPAK/DMAREL search.

The symbol denotes the space group of the MOLPAK starting structure, and, in the legend, is followed by the number of minimisations performed for that space group. Two sets of minimisations were performed in *Pbca*. The initial search using the molecular axis system with the xy plane defined by C1, C4, C2, resulted in CB9 as the lowest *Pbca* structure, whereas a further search, using C1, C7, O2 to define the plane, found CB47 which is even more similar to form II. This unusual sensitivity to the initial probe orientation may arise from the nature of the potential energy surface for this type of sheet structure.


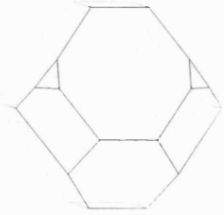
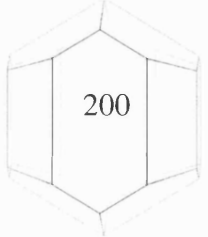
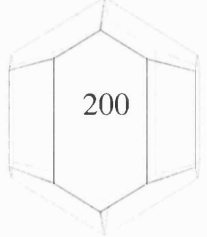
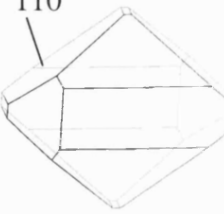
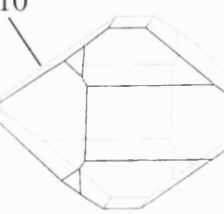
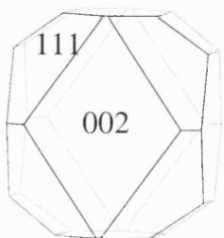
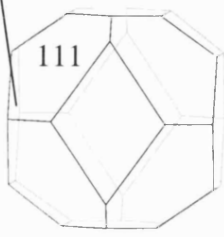
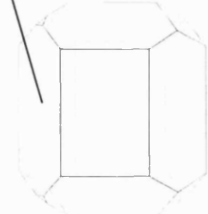
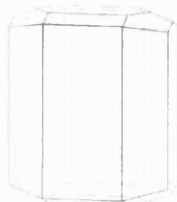
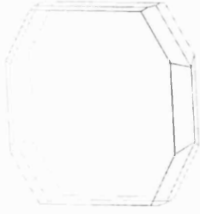
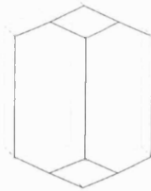
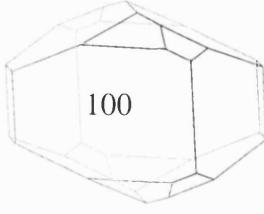
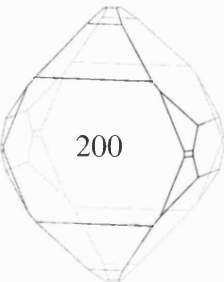
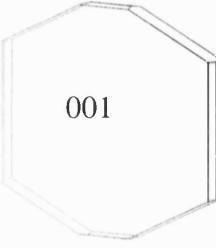
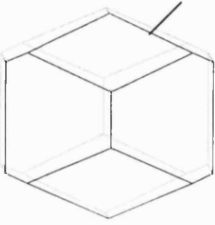
	AM30 form I (expt _{low}) (=form I min _{opt})		CB47 form II (expt _{low}) (=form II min _{opt})	
BFDH				
AE				
	AI22	AY8	CC19	AQ6
BFDH				
AE				

Figure 6.6 (continued)

	AK6	AQ14	CB9	CC8
BFDH				
AE				
	AK22	AI16	AM4	AK4
BFDH				
AE				

Figure 6.6 BFDH and attachment energy (AE) predicted morphologies for the known and hypothetical crystal structures of paracetamol. The slowest growing faces are labelled.

Appendix 6.1

Fractional coordinates of the lowest energy structures of paracetamol found by the MOLPAK/DMAREL search.

```
-----
TITLE formI Minopt, P21/a, E_latt = -110.1 kJ/mol
CELL 13.011999 8.943599 7.278221 90.000000 113.456300 90.000000
SYMM X, Y, Z
SYMM -X+0.5, Y+0.5, -Z
SYMM -X, -Y, -Z
SYMM X+0.5, -Y+0.5, Z
ATOM C1 0.930827 0.349233 0.849290
ATOM C2 1.003332 0.347144 0.750924
ATOM C3 1.099598 0.261647 0.824943
ATOM C4 1.126075 0.177575 0.997794
ATOM C5 1.054451 0.179272 1.096087
ATOM C6 0.957513 0.264269 1.021326
ATOM C7 0.777541 0.505773 0.607064
ATOM C8 0.674887 0.593131 0.592720
ATOM H1 0.983905 0.410981 0.617059
ATOM H2 1.155621 0.259716 0.749297
ATOM H3 1.073259 0.114507 1.230382
ATOM H4 0.902368 0.263581 1.099242
ATOM H5 1.232059 0.044290 1.180691
ATOM H6 0.797176 0.444552 0.884618
ATOM H7 0.631931 0.543223 0.678194
ATOM H8 0.618358 0.600518 0.436915
ATOM H9 0.700334 0.705545 0.648719
ATOM N1 0.831883 0.435930 0.785828
ATOM O1 1.222025 0.096896 1.062575
ATOM O2 0.807851 0.503650 0.471679
END
```

```
-----
TITLE AM30, P21/c, E_latt = -110.1 kJ/mol
CELL 12.119 8.944 7.278 90.00 80.03 90.00
SYMM X, Y, Z
SYMM -X+0.5, Y+0.5, -Z+0.5
SYMM -X, -Y, -Z
SYMM X+0.5, -Y+0.5, Z+0.5
ATOM C1 0.930790 0.349254 0.918439
ATOM C2 1.003347 0.347130 0.747610
ATOM C3 1.099584 0.261572 0.725240
ATOM C4 1.126091 0.177517 0.871663
ATOM C5 1.054460 0.179281 1.041646
ATOM C6 0.957514 0.264284 1.063808
ATOM C7 0.777527 0.505799 0.829550
ATOM C8 0.674882 0.593148 0.917785
ATOM H1 0.983888 0.410928 0.633137
ATOM H2 1.155559 0.259685 0.593573
ATOM H3 1.073288 0.114552 1.157113
ATOM H4 0.902407 0.263661 1.196944
ATOM H5 1.232022 0.044292 0.948624
ATOM H6 0.797178 0.444631 1.087429
ATOM H7 0.631951 0.543258 1.046220
ATOM H8 0.618355 0.600504 0.818579
ATOM H9 0.700385 0.705580 0.948368
ATOM N1 0.831858 0.435983 0.953870
ATOM O1 1.222024 0.096864 0.840543
ATOM O2 0.807843 0.503637 0.663821
END
```

```
-----
TITLE formII Minopt, Pbca, E_latt = -106.5 kJ/mol
CELL 17.263060 12.075615 7.383896 90.000000 90.000000 90.000000
SYMM X, Y, Z
```

SYMM -X+0.5, -Y, Z+0.5
SYMM -X, Y+0.5, -Z+0.5
SYMM X+0.5, -Y+0.5, -Z
SYMM -X, -Y, -Z
SYMM X+0.5, Y, -Z+0.5
SYMM X, -Y+0.5, Z+0.5
SYMM -X+0.5, Y+0.5, Z

ATOM C1	0.830993	0.056987	0.212399
ATOM C2	0.831515	0.162275	0.289507
ATOM C3	0.764491	0.224613	0.293873
ATOM C4	0.695828	0.183725	0.223118
ATOM C5	0.694954	0.079273	0.146491
ATOM C6	0.762263	0.016923	0.141030
ATOM C7	0.972166	0.013284	0.235517
ATOM C8	1.027295	-0.083960	0.221942
ATOM H1	0.884069	0.195137	0.344115
ATOM H2	0.764731	0.305883	0.353105
ATOM H3	0.642290	0.045980	0.090325
ATOM H4	0.760618	-0.064040	0.079899
ATOM H5	0.588876	0.213439	0.181166
ATOM H6	0.885658	-0.091019	0.172239
ATOM H7	1.008948	-0.145256	0.123590
ATOM H8	1.083979	-0.053020	0.186049
ATOM H9	1.031218	-0.123735	0.353515
ATOM N1	0.896604	-0.012968	0.205740
ATOM O1	0.632188	0.249002	0.232653
ATOM O2	0.994723	0.104140	0.274026

END

TITLE CB47, Pbca, E_{latt} = -106.5 kJ/mol

CELL 17.248591 12.086164 7.382187 90.000000 90.000000 90.000000

SYMM X, Y, Z
SYMM 0.5-X, -Y, Z+0.5
SYMM -X, Y+0.5, 0.5-Z
SYMM X+0.5, 0.5-Y, -Z
SYMM -X, -Y, -Z
SYMM X+0.5, Y, 0.5-Z
SYMM X, 0.5-Y, Z+0.5
SYMM 0.5-X, Y+0.5, Z

ATOM C1	0.168839	0.942902	0.288305
ATOM C2	0.168328	0.837665	0.211355
ATOM C3	0.235391	0.775311	0.207483
ATOM C4	0.304055	0.816124	0.278566
ATOM C5	0.304904	0.920533	0.355034
ATOM C6	0.237565	0.982917	0.359994
ATOM C7	0.027592	0.986722	0.264312
ATOM C8	-0.027537	1.083948	0.277413
ATOM H1	0.115774	0.804849	0.156504
ATOM H2	0.235164	0.694061	0.148381
ATOM H3	0.357576	0.953771	0.411454
ATOM H4	0.239191	1.063836	0.421016
ATOM H5	0.411032	0.786340	0.321189
ATOM H6	0.114162	1.090876	0.327911
ATOM H7	-0.009251	1.145230	0.375790
ATOM H8	-0.084320	1.053120	0.313051
ATOM H9	-0.031317	1.123612	0.145728
ATOM N1	0.103201	1.012886	0.294486
ATOM O1	0.367716	0.750822	0.269506
ATOM O2	0.005008	0.895941	0.225830

END

TITLE AI22, P21/c, E_{latt} = -105.6 kJ/mol

CELL 8.248119 6.749419 16.669219 90.000000 126.237189 90.000000

SYMM X, Y, Z
SYMM -X, Y+0.5, -Z+0.5
SYMM -X, -Y, -Z
SYMM X, -Y+0.5, Z+0.5

```
ATOM C1 0.734490 0.922200 0.871973
ATOM C2 0.782771 0.731408 0.911507
ATOM C3 0.821752 0.584713 0.867281
ATOM C4 0.812715 0.623912 0.783060
ATOM C5 0.764727 0.812967 0.743539
ATOM C6 0.726689 0.960495 0.788077
ATOM C7 0.719942 1.090543 1.001146
ATOM C8 0.646135 1.280846 1.017751
ATOM H1 0.791009 0.698985 0.976682
ATOM H2 0.859200 0.437585 0.897763
ATOM H3 0.757294 0.846743 0.678263
ATOM H4 0.690681 1.107220 0.756609
ATOM H5 0.841207 0.515322 0.686540
ATOM H6 0.628163 1.198073 0.868058
ATOM H7 0.657841 1.407642 0.981427
ATOM H8 0.732752 1.307522 1.096820
ATOM H9 0.489540 1.262630 0.988707
ATOM N1 0.688539 1.079477 0.911396
ATOM O1 0.851676 0.473107 0.743339
ATOM O2 0.794325 0.961369 1.062285
END
```

```
-----
TITLE AY8, Pca21, E_latt = -105.4 kJ/mol
CELL 15.863353 4.522143 10.691577 90.000000 90.000000 90.000000
SYMM X, Y, Z
SYMM 0.5-X, Y, Z+0.5
SYMM X+0.5, -Y, Z
SYMM -X, -Y, Z+0.5
ATOM C1 0.675009 0.113188 0.965707
ATOM C2 0.652514 -0.046448 1.071973
ATOM C3 0.710025 -0.233618 1.127952
ATOM C4 0.790460 -0.267383 1.079466
ATOM C5 0.813127 -0.109584 0.973960
ATOM C6 0.755719 0.079738 0.918117
ATOM C7 0.544173 0.414980 0.940216
ATOM C8 0.500068 0.608413 0.844798
ATOM H1 0.590794 -0.022159 1.110752
ATOM H2 0.692802 -0.356830 1.210006
ATOM H3 0.875479 -0.131823 0.934821
ATOM H4 0.774665 0.202132 0.836370
ATOM H5 0.896300 -0.462261 1.098154
ATOM H6 0.639231 0.369485 0.817563
ATOM H7 0.543700 0.731446 0.786485
ATOM H8 0.459175 0.760067 0.893830
ATOM H9 0.461635 0.469432 0.784761
ATOM N1 0.619554 0.304735 0.901292
ATOM O1 0.843565 -0.455555 1.138548
ATOM O2 0.512619 0.361474 1.039554
END
```

```
-----
TITLE CC19, P21/c (Z'=2), E_latt = -103.3 kJ/mol
CELL 9.886174 5.073661 31.121307 90.000000 82.229344 90.000000
SYMM X, Y, Z
SYMM -X + 0.5, Y + 0.5, -Z
SYMM -X, -Y, -Z
SYMM X + 0.5, -Y + 0.5, Z
ATOM C1 0.111556 0.058038 0.884202
ATOM C2 0.027766 -0.161262 0.882272
ATOM C3 -0.079153 -0.206345 0.914792
ATOM C4 -0.105422 -0.035047 0.949632
ATOM C5 -0.022563 0.182874 0.951711
ATOM C6 0.085226 0.227546 0.919161
ATOM C7 0.282515 -0.033076 0.819082
ATOM C8 0.395383 0.102786 0.789402
ATOM H1 0.046906 -0.295670 0.855693
ATOM H2 -0.143882 -0.375345 0.913413
ATOM H3 -0.040903 0.318302 0.978491
```

```
ATOM H4  0.149175  0.397516  0.921501
ATOM H5 -0.221141  0.038019  1.002352
ATOM H6  0.259384  0.301190  0.853872
ATOM H7  0.450214  0.247226  0.806042
ATOM H8  0.465635 -0.045440  0.774882
ATOM H9  0.351130  0.202272  0.763673
ATOM N1  0.221325  0.121351  0.851782
ATOM O1 -0.212221 -0.090329  0.980272
ATOM O2  0.249856 -0.255725  0.812742
ATOM C1  0.486165  0.094967  0.385469
ATOM C2  0.567075 -0.130751  0.380720
ATOM C3  0.645412 -0.199573  0.412451
ATOM C4  0.645588 -0.046057  0.449311
ATOM C5  0.565448  0.178131  0.454190
ATOM C6  0.486261  0.246643  0.422419
ATOM C7  0.368711  0.047954  0.319608
ATOM C8  0.283874  0.206585  0.291818
ATOM H1  0.567768 -0.251658  0.352550
ATOM H2  0.707832 -0.373582  0.408901
ATOM H3  0.563815  0.300037  0.482570
ATOM H4  0.423820  0.421131  0.426949
ATOM H5  0.717983 -0.006512  0.502661
ATOM H6  0.369108  0.364097  0.358429
ATOM H7  0.217892  0.348802  0.310287
ATOM H8  0.222965  0.071838  0.275597
ATOM H9  0.351938  0.312124  0.267329
ATOM N1  0.405268  0.182189  0.354059
ATOM O1  0.725195 -0.124250  0.479051
ATOM O2  0.402202 -0.174294  0.310479
END
```

```
-----
TITLE AQ6, P212121, E_latt = -102.8 kJ/mol
CELL 7.343724 16.280771 6.570977 90.000000 90.000000 90.000000
SYMM X, Y, Z
SYMM 0.5-X, -Y, Z+0.5
SYMM -X, Y+0.5, 0.5-Z
SYMM X+0.5, 0.5-Y, -Z
ATOM C1  0.911951  0.720715  0.906340
ATOM C2  0.984767  0.694741  0.721476
ATOM C3  1.013192  0.750545  0.566414
ATOM C4  0.969043  0.832756  0.591111
ATOM C5  0.896740  0.858858  0.774237
ATOM C6  0.869164  0.803030  0.930311
ATOM C7  0.933406  0.588837  1.099661
ATOM C8  0.862843  0.548976  1.292399
ATOM H1  1.019992  0.631609  0.700214
ATOM H2  1.069419  0.730727  0.423739
ATOM H3  0.861851  0.922458  0.796924
ATOM H4  0.813520  0.824540  1.072308
ATOM H5  0.964911  0.938218  0.465125
ATOM H6  0.799225  0.691247  1.180953
ATOM H7  0.844265  0.592320  1.415987
ATOM H8  0.957667  0.501732  1.338135
ATOM H9  0.732474  0.520219  1.260222
ATOM N1  0.875214  0.667352  1.071003
ATOM O1  0.999510  0.884048  0.432584
ATOM O2  1.028329  0.552077  0.982613
END
```

```
-----
TITLE AK6, P21/c, E_latt = -101.9 kJ/mol
CELL 16.050382 5.072115 9.647982 90.000000 79.143922 90.000000
SYMM X, Y, Z
SYMM -X, Y+0.5, 0.5-Z
SYMM -X, -Y, -Z
SYMM X, 0.5-Y, Z+0.5
ATOM C1  0.769671  0.921483  1.012862
ATOM C2  0.763579  1.144555  1.098450
```

```
ATOM C3 0.828091 1.205618 1.169709
ATOM C4 0.899301 1.046796 1.158003
ATOM C5 0.905695 0.825233 1.073206
ATOM C6 0.841292 0.764482 1.001053
ATOM C7 0.638271 0.982780 0.912598
ATOM C8 0.580899 0.831769 0.832775
ATOM H1 0.709147 1.269466 1.108187
ATOM H2 0.823585 1.377626 1.235677
ATOM H3 0.960602 0.699294 1.062367
ATOM H4 0.847606 0.591960 0.934816
ATOM H5 1.005353 0.996573 1.216084
ATOM H6 0.712391 0.661086 0.899194
ATOM H7 0.614938 0.689621 0.759247
ATOM H8 0.548790 0.971372 0.776555
ATOM H9 0.533605 0.727050 0.908130
ATOM N1 0.705861 0.842190 0.938718
ATOM O1 0.959940 1.117495 1.231182
ATOM O2 0.622627 1.205006 0.951114
END
```

```
-----
TITLE AQ14, P212121, E_latt = -101.9 kJ/mol
CELL 7.070802 15.834421 6.776105 90.000000 90.000000 90.000000
SYMM X, Y, Z
SYMM 0.5-X, -Y, Z+0.5
SYMM -X, Y+0.5, 0.5-Z
SYMM X+0.5, 0.5-Y, -Z
ATOM C1 0.015736 0.666860 0.058064
ATOM C2 0.090650 0.633307 0.231397
ATOM C3 0.120573 0.684711 0.392941
ATOM C4 0.078122 0.770143 0.385650
ATOM C5 0.003829 0.803826 0.214051
ATOM C6 -0.027314 0.752217 0.052190
ATOM C7 -0.017616 0.533611 -0.136473
ATOM C8 -0.050650 0.503814 -0.346147
ATOM H1 0.123493 0.567470 0.239653
ATOM H2 0.178096 0.659053 0.526813
ATOM H3 -0.030615 0.869952 0.205219
ATOM H4 -0.086226 0.779372 -0.079924
ATOM H5 0.078310 0.873876 0.528278
ATOM H6 -0.042416 0.652697 -0.235363
ATOM H7 -0.141680 0.545839 -0.429573
ATOM H8 -0.112418 0.441484 -0.341121
ATOM H9 0.084652 0.499745 -0.421343
ATOM N1 -0.015683 0.618979 -0.114947
ATOM O1 0.111766 0.816982 0.549042
ATOM O2 0.008912 0.484502 -0.004727
END
```

```
-----
TITLE CB9, Pbca, E_latt = -101.8 kJ/mol
CELL 17.431956 12.207169 7.265817 90.000000 90.000000 90.000000
SYMM X, Y, Z
SYMM -X+0.5, -Y, Z+0.5
SYMM -X, Y+0.5, -Z+0.5
SYMM X+0.5, -Y+0.5, -Z
SYMM -X, -Y, -Z
SYMM X+0.5, Y, -Z+0.5
SYMM X, -Y+0.5, Z+0.5
SYMM -X+0.5, Y+0.5, Z
ATOM C1 0.177268 0.197062 0.776489
ATOM C2 0.172513 0.090786 0.707670
ATOM C3 0.238027 0.028186 0.689015
ATOM C4 0.309282 0.069775 0.737027
ATOM C5 0.314380 0.175247 0.805345
ATOM C6 0.248626 0.237861 0.825214
ATOM C7 0.037436 0.241441 0.790484
ATOM C8 -0.015834 0.338643 0.812112
ATOM H1 0.117937 0.057430 0.670567
```

```
ATOM H2  0.234500 -0.053809  0.636119
ATOM H3  0.369068  0.209105  0.843928
ATOM H4  0.253595  0.319643  0.879630
ATOM H5  0.416540  0.040217  0.750702
ATOM H6  0.125955  0.345283  0.820639
ATOM H7  0.007421  0.401868  0.900830
ATOM H8 -0.069793  0.309911  0.867284
ATOM H9 -0.026413  0.374123  0.677406
ATOM N1  0.113305  0.267326  0.796479
ATOM O1  0.371135  0.004107  0.714358
ATOM O2  0.013022  0.150829  0.765688
END
```

TITLE CC8, Pbc_a, E_{latt} = -101.0 kJ/mol

CELL 6.848235 16.664305 13.490591 90.000000 90.000000 90.000000

SYMM X, Y, Z

SYMM 0.5-X, -Y, Z+0.5

SYMM -X, Y+0.5, 0.5-Z

SYMM X+0.5, 0.5-Y, -Z

SYMM -X, -Y, -Z

SYMM X+0.5, Y, 0.5-Z

SYMM X, 0.5-Y, Z+0.5

SYMM 0.5-X, Y+0.5, Z

ATOM C1 0.086197 0.051411 0.869104

ATOM C2 0.274066 0.026402 0.893961

ATOM C3 0.418435 0.082424 0.912033

ATOM C4 0.379804 0.163878 0.905269

ATOM C5 0.193698 0.188976 0.880507

ATOM C6 0.048454 0.132875 0.862988

ATOM C7 -0.079601 -0.082147 0.865111

ATOM C8 -0.266822 -0.121036 0.829382

ATOM H1 0.306030 -0.036164 0.899801

ATOM H2 0.563325 0.063265 0.931231

ATOM H3 0.160343 0.252060 0.875089

ATOM H4 -0.096024 0.153628 0.844392

ATOM H5 0.486609 0.269091 0.916721

ATOM H6 -0.185276 0.023288 0.816856

ATOM H7 -0.391886 -0.081308 0.833813

ATOM H8 -0.293384 -0.174020 0.873459

ATOM H9 -0.248166 -0.139172 0.752641

ATOM N1 -0.068630 -0.001889 0.847500

ATOM O1 0.528295 0.215398 0.923397

ATOM O2 0.047548 -0.120878 0.903073

END

TITLE AK22, P2₁/c, E_{latt} = -100.2 kJ/mol

CELL 15.834393 5.071468 9.861235 90.000000 86.684102 90.000000

SYMM X, Y, Z

SYMM -X, Y+0.5, 0.5-Z

SYMM -X, -Y, -Z

SYMM X, 0.5-Y, Z+0.5

ATOM C1 0.725814 0.068435 0.916750

ATOM C2 0.731278 -0.152157 0.999936

ATOM C3 0.668045 -0.204235 1.098661

ATOM C4 0.598692 -0.038496 1.117184

ATOM C5 0.592911 0.180671 1.035010

ATOM C6 0.656062 0.232245 0.935441

ATOM C7 0.853648 -0.009675 0.761336

ATOM C8 0.910529 0.132891 0.656469

ATOM H1 0.784195 -0.282416 0.986630

ATOM H2 0.671977 -0.374431 1.162876

ATOM H3 0.539525 0.311800 1.047389

ATOM H4 0.650316 0.403263 0.871803

ATOM H5 0.495052 0.024260 1.220337

ATOM H6 0.783053 0.319163 0.777808

ATOM H7 0.876926 0.276123 0.598201

ATOM H8 0.940115 -0.011786 0.588772

ATOM H9 0.959754 0.235199 0.707938
ATOM N1 0.788478 0.138703 0.815362
ATOM O1 0.539221 -0.100704 1.216235
ATOM O2 0.867751 -0.232028 0.794317
END

TITLE AI16, P21/c, E_latt = -100.2 kJ/mol
CELL 7.553258 8.112605 12.715831 90.000000 103.524632 90.000000
SYMM X, Y, Z
SYMM -X, Y+0.5, 0.5-Z
SYMM -X, -Y, -Z
SYMM X, 0.5-Y, Z+0.5
ATOM C1 0.762527 0.929735 0.116932
ATOM C2 0.701144 0.935724 0.211967
ATOM C3 0.723450 0.801057 0.280229
ATOM C4 0.805617 0.658163 0.255413
ATOM C5 0.866695 0.651474 0.161090
ATOM C6 0.845458 0.786848 0.092872
ATOM C7 0.696381 1.220431 0.054846
ATOM C8 0.681301 1.326746 -0.044788
ATOM H1 0.637960 1.045228 0.232416
ATOM H2 0.676328 0.805329 0.353604
ATOM H3 0.931380 0.541832 0.140382
ATOM H4 0.894846 0.779734 0.020132
ATOM H5 0.879538 0.439819 0.301066
ATOM H6 0.766436 1.031821 -0.029399
ATOM H7 0.774844 1.290298 -0.093426
ATOM H8 0.706821 1.453368 -0.019227
ATOM H9 0.543434 1.318068 -0.094300
ATOM N1 0.741771 1.060717 0.041669
ATOM O1 0.822229 0.530464 0.325578
ATOM O2 0.665633 1.274858 0.136389
END

TITLE AM4, P21/c, E_latt = -100.1 kJ/mol
CELL 5.937262 7.590126 17.070748 90.000000 80.714464 90.000000
SYMM X, Y, Z
SYMM -X, Y+0.5, -Z+0.5
SYMM -X, -Y, -Z
SYMM X, -Y+0.5, Z+0.5
ATOM C1 0.222943 0.353032 0.789656
ATOM C2 0.439002 0.276580 0.775251
ATOM C3 0.558008 0.247661 0.837258
ATOM C4 0.466036 0.294482 0.914200
ATOM C5 0.251675 0.370493 0.928885
ATOM C6 0.131557 0.398660 0.866810
ATOM C7 0.123265 0.331522 0.653209
ATOM C8 -0.047355 0.402965 0.604149
ATOM H1 0.512020 0.239214 0.716230
ATOM H2 0.724509 0.188537 0.826293
ATOM H3 0.176725 0.407545 0.988241
ATOM H4 -0.035678 0.457040 0.879392
ATOM H5 0.514510 0.299921 1.021921
ATOM H6 -0.042011 0.467854 0.745847
ATOM H7 -0.214395 0.426093 0.638861
ATOM H8 -0.061653 0.310337 0.557174
ATOM H9 0.016469 0.526798 0.577475
ATOM N1 0.091964 0.390460 0.729469
ATOM O1 0.592318 0.263120 0.972062
ATOM O2 0.274585 0.235004 0.624981
END

TITLE AK4, P21/c, E_latt = -100.0 kJ/mol
CELL 19.034544 5.293694 8.080315 90.000000 101.559109 90.000000
SYMM X, Y, Z
SYMM -X, Y+0.5, 0.5-Z
SYMM -X, -Y, -Z

SYMM X, 0.5-Y, Z+0.5
ATOM C1 0.764407 0.992856 0.927930
ATOM C2 0.785262 1.197039 1.035439
ATOM C3 0.853493 1.205598 1.133846
ATOM C4 0.902088 1.012074 1.128495
ATOM C5 0.881638 0.808896 1.022141
ATOM C6 0.813380 0.800878 0.922532
ATOM C7 0.643040 1.144914 0.785569
ATOM C8 0.574206 1.048428 0.675971
ATOM H1 0.748564 1.348176 1.040635
ATOM H2 0.869706 1.362890 1.216728
ATOM H3 0.918573 0.656836 1.015618
ATOM H4 0.798542 0.641853 0.839730
ATOM H5 0.996483 0.890998 1.215695
ATOM H6 0.684283 0.797224 0.775150
ATOM H7 0.583262 0.900670 0.589635
ATOM H8 0.546947 1.205412 0.605447
ATOM H9 0.540068 0.973483 0.756945
ATOM N1 0.695315 0.967055 0.825716
ATOM O1 0.967861 1.031871 1.228889
ATOM O2 0.648693 1.359085 0.834354
END

References for chapter 6

- (1) Beyer, T.; Price, S.L., *CrystEngComm* (<http://www.rsc.org/is/journals/current/crystengcomm/cecpub.htm>), **2000**, 34. The errors in lattice energy minimisation studies: sensitivity to experimental variations in the molecular structure of paracetamol.
- (2) Beyer, T.; Day, G.M.; Price, S.L., *J. Am. Chem. Soc.*, **2001**, in press. The prediction, morphology and mechanical properties of the polymorphs of paracetamol.
- (3) *Pharm. J.*, **1996**, 256. 40 years of paracetamol.
- (4) Lennartz, P., *Untersuchungen zu speziellen Eigenschaften und zur inneren Struktur von Mintabletten aus Paracetamol und spruehgetrockneter Laktose*, in *Fachbereich Chemie*. 1998, Dissertation, Universität Hamburg: Hamburg.
- (5) Nürnberg, E.; Hopp, A., *Pharm. Ind.*, **1982**, 44, 1081. Galenisch relevante Untersuchungen von Paracetamol. 1. Modell - Kristallisationsversuche im Roentgen-Diffractometer.
- (6) Bürger, A., *Acta Pharm. Technol.*, **1982**, 28, 1-20. Zur Interpretation von Polymorphie-Untersuchungen.
- (7) Di Martino, P.; Conflant, P.; Drache, M.; Huvenne, J.-P.; Guyot-Hermann, A.-M., *Journal of Thermal Analysis*, **1997**, 48, 447-458. Preparation and physical characterization of forms II and III of paracetamol.
- (8) Clegg, W.; Teat, S.J.; Prasad, K.V.R.; Ristic, R.I.; Sheen, D.B.; Shepherd, E.A.; Sherwood, J.N.; Vrcelj, R.M., *Acta Cryst.*, **1998**, C54, 1881-1882. An oxidatively coupled dimer of paracetamol.
- (9) Haisa, M.; Kashino, S.; Maeda, H., *Acta Cryst.*, **1974**, B30, 2510-2512. The orthorhombic form of p-hydroxyacetanilide.
- (10) Haisa, M.; Kashino, S.; Kawai, R.; Maeda, H., *Acta Cryst.*, **1976**, B32, 1283-1285. The monoclinic form of p-hydroxyacetanilide.
- (11) Welton, J.M.; McCarthy, G.J., *Powder Diffraction*, **1988**, 3, 102-103. X-ray powder data for acetaminophen.
- (12) Naumov, D.Y.; Vasilchenko, M.A.; Howard, J.A.K., *Acta Cryst.*, **1998**, C54, 653-655. The monoclinic form of acetaminophen at 150K.
- (13) Kuhnert-Branstatter, M.; Geiler, M.; Wurian, I., *Sci. Pharm.*, **1990**, 48, 250.
- (14) Nichols, G.; Frampton, C.S., *Journal of Pharmaceutical Sciences*, **1998**, 87, 684-693. Physicochemical characterization of the orthorhombic polymorph of

paracetamol crystallized from solution. Atomic coordinates by private communication.

- (15) Wilson, C.C., *Journal of Molecular Structure*, **1997**, 405, 207-217. Neutron diffraction of p-hydroxyacetanilide (paracetamol): libration or disorder of the methyl group at 100K.
- (16) Wilson, C.C., *Chem. Phys. Lett.*, **1997**, 280, 531-534. Zero point motion of the librating methyl group in p-hydroxyacetanilide.
- (17) Allen, F.H.; Harris, S.E.; Taylor, R., *Journal of Computer-Aided Molecular Design*, **1996**, 10, 247-254. Comparison of conformer distributions in the crystalline state with conformational energies calculated by ab initio techniques.
- (18) Gavezzotti, A. in *Theoretical Aspects and Computer Modeling of the Molecular Solid State*, A. Gavezzotti, John Wiley & Sons, Chichester, **1997**, 97. Crystal symmetry and molecular recognition.
- (19) Pertsin, A.J.; Kitaigorodskii, A.I., *The atom-atom potential method. Applications to organic molecular solids*, **1987**, Berlin: Springer-Verlag.
- (20) Joiris, E.; Di Martino, P.; Berneron, C.; Guyot-Hermann, A.-M.; Guyot, J.-C., *Pharmaceutical Research*, **1998**, 15, 1122-1130. Compression behaviour of orthorhombic paracetamol.
- (21) Fachaux, J.M.; Guyot-Hermann, A.-M.; Guyot, J.-C.; Conflant, P.; Drache, M.; Huvenne, J.P.; Bouche, R., *Int. J. of Pharm.*, **1993**, 99, 99-107. Compression ability improvement by solvation/desolvation process: application to paracetamol for direct compression.
- (22) Fachaux, J.-M.; Guyot-Hermann, A.-M.; Guyot, J.-C.; Conflant, P.; Drache, M.; Veessler, S.; Boistelle, R., *Powder Technology*, **1995**, 82, 123-128. Pure paracetamol for direct compression Part I. Development of sintered-like crystals of paracetamol.
- (23) Fachaux, J.-M.; Guyot-Hermann, A.-M.; Guyot, J.-C.; Conflant, P.; Drache, M.; Veessler, S.; Boistelle, R., *Powder Technology*, **1995**, 82, 129-133. Pure paracetamol for direct compression Part II. Study of the physicochemical and mechanical properties of sintered-like crystals of paracetamol.
- (24) Di Martino, P.; Guyot-Hermann, A.-M.; Conflant, P.; Drache, M.; Guyot, J.-C., *International Journal of Pharmaceutics*, **1996**, 128, 1-8. A new pure paracetamol for direct compression: the orthorhombic form.

- (25) Shekunov, B.Y.; Grant, D.J.W., *J. Phys. Chem. B*, **1997**, *101*, 3973-3979. In situ optical interferometric studies of the growth and dissolution behavior of paracetamol (acetaminophen) 1. Growth kinetics.
- (26) Shekunov, B.Y.; Aulton, M.E.; AdamaAcquah, R.W.; Grant, D.J.W., *J. Chem. Soc., Faraday Trans.*, **1996**, *92*, 439-444. Effect of temperature on crystal growth and crystal properties of paracetamol.
- (27) Mayo, S.L.; Olafson, B.D.; Goddard III, W.A., *J. Phys. Chem.*, **1990**, *94*, 8897-8909. DREIDING: A generic force field for molecular simulations.
- (28) Verwer, P.; Leusen, F.J.J. in *Rev. Comp. Chem.*, K.B. Lipkowitz and D.B. Boyd, Wiley-VCH, John Wiley and Sons, Inc., New York, **1998**, 327-365. Computer simulation to predict possible crystal polymorphs.
- (29) Rappe, A.K.; Goddard III, W.A., *J. Phys. Chem.*, **1991**, *95*, 3358-3363. Charge equilibration for molecular-dynamics simulations.
- (30) Leusen, F.J.J. . in *5th World Congress on Chemical Engineering*. 1996. San Diego.
- (31) Willock, D.J.; Price, S.L.; Leslie, M.; Catlow, C.R.A., *J. Comp. Chem.*, **1995**, *16*, 628-647. The relaxation of molecular crystal structures using a distributed multipole electrostatic model.
- (32) Allen, F.H.; Kennard, O.; Watson, D.G.; Brammer, L.; Orpen, A.G.R.; Taylor, R., *J. Chem. Soc., Perkin Trans.*, **1987**, *2*, S1 - S9. Tables of bond lengths determined by X-ray and neutron diffraction. Part 1. Bond lengths in organic compounds.
- (33) Amos, R.D.; Alberts, I.L.; Andrews, J.S.; Colwell, S.M.; Handy, N.C.; Jayatilaka, D.; Knowles, P.J.; Kobayashi, R.; Laidig, K.E.; Laming, G.; Lee, A.M.; Maslen, P.E.; Murray, C.W.; Rice, J.E.; Simandiras, E.D.; Stone, A.J.; Su, M.D.; Tozer, D.J., *CADPAC6: The Cambridge Analytic Derivatives Package*, Issue 6: Cambridge University, 1995,
- (34) Holden, J.R.; Du, Z.; Ammon, H.L., *J. Comp. Chem.*, **1993**, *14*, 422-437. Prediction of possible crystal-structures for C-containing, H-containing, N-containing, O-containing and F-containing organic-compounds.
- (35) Pillardy, J.; Wawak, R.J.; Arnautova, Y.A.; Czaplewski, C.; Scheraga, H.A., *J. Am. Chem. Soc.*, **2000**, *122*, 907-921. Crystal structure prediction by global optimization as a tool for evaluating potentials: Role of the dipole moment correction term in successful predictions.

- (36) van Eijck, B.P.; Kroon, J., *J. Phys. Chem. B*, **1997**, *101*, 1096-1100. Coulomb energy of polar crystals.
- (37) van Eijck, B.P.; Kroon, J., *Journal of Physical Chemistry B*, **2000**, *104*, 8089-8089. Comment on "Crystal structure prediction by global optimization as a tool for evaluating potentials: Role of the dipole moment correction term in successful predictions".
- (38) Wedemeyer, W.J.; Arnautova, Y.A.; Pillardy, J.; Wawak, R.J.; Czaplewski, C.; Scheraga, H.A., *Journal of Physical Chemistry B*, **2000**, *104*, 8090-8092. Reply to "Comment on 'Crystal structure prediction by global optimization as a tool for evaluating potentials: Role of the dipole moment correction term in successful predictions'" by B. P. van Eijck and J. Kroon.
- (39) Day, G.M.; Price, S.L.; Leslie, M., *Crystal Growth and Design*, **2001**, *1*, 13-27. Calculating elastic constants of molecular organic crystals - sensitivity to the model potential.
- (40) Donnay, J.D.H.; Harker, D., *Am. Mineral.*, **1937**, *22*, 446.
- (41) Berkovitch-Yellin, Z., *J. Am. Chem. Soc.*, **1985**, *107*, 8239. Toward an ab initio derivation of crystal morphology.
- (42) Cerius2, Version 3.5: Molecular Simulations Inc., 1997,
- (43) Lommerse, J.P.M., *J. Applied. Cryst.*, **in preparation**, An index for crystal structure similarity.
- (44) Speakman, J.C. in *Molecular Structure by Diffraction Methods*, G.A. Sim and L.E. Sutton, The Chemical Society, London, **1973**, 201-231. Neutron diffraction.
- (45) Etter, M.C., *Acc. Chem. Res.*, **1990**, *23*, 120-126. Encoding and decoding hydrogen-bond patterns of organic compounds.
- (46) Aakeroy, C.B.; Nieuwenhuyzen, M.; Price, S.L., *Journal of the American Chemical Society*, **1998**, *120*, 8986-8993. Three polymorphs of 2-amino-5-nitropyrimidine: Experimental structures and theoretical predictions.
- (47) Breu, J.; Domel, H.; Norrby, P.-O., *Eur. J. Inorg. Chem.*, **2000**, *2000*, Chiral recognition among trisdiimine-metal complexes, 7. Racemic compound formation versus conglomerate formation with $[M(\text{bpy})_3](\text{PF}_6)_2$ ($M=\text{Ni}$, Zn , Ru): Lattice energy minimisations and implications for structure prediction.
- (48) Buttar, D.; Charlton, M.H.; Docherty, R.; Starbuck, J., *Journal of the Chemical Society-Perkin Transactions 2*, **1998**, 763-772. Theoretical investigations of conformational aspects of polymorphism. Part 1: O-acetamidobenzamide.

- (49) Beyer, T.; Price, S.L., *J. Phys. Chem. B*, **2000**, *104*, 2647-2655. Dimer or catemer? Low-energy crystal packings for small carboxylic acids.
- (50) Krivy, I.; Gruber, B., *Acta Cryst.*, **1976**, *A32*, 297-298. A unified algorithm for determining the reduced (Niggli) cell.
- (51) Duncan-Hewitt, W.C.; Weatherly, G.C., *J. Mat. Sci. Lett.*, **1989**, *8*, 1350-1352. Evaluating the hardness, Young's modulus and fracture toughness of some pharmaceutical crystals using microindentation techniques.
- (52) Rowe, R.C.; Roberts, R.J. in *Advances in Pharmaceutical Sciences*, D. Ganderton, T. Jones, and J. McGinty, Academic Press, London, **1995**, 1-62. The mechanical properties of powders.
- (53) Reuss, A., *Z. Angew. Math. Mech.*, **1929**, *9*, 49-58. Berechnung der Fließsgrenze von Mischkristallen auf Grund der Plastizitätsbedingung für Einkristalle.
- (54) Voigt, W., *Lehrbuch der Kristallphysik*, **1928**, Leipzig: Teubner.
- (55) Volkov, S.D.; Klinskikh, N.A., *Dokl. Akad. Nauk. SSSR*, **1963**, *146*, 565-568.
- (56) Mooij, W.T.M.; van Eijck, B.P.; Price, S.L.; Verwer, P.; Kroon, J., *Journal of Computational Chemistry*, **1998**, *19*, 459-474. Crystal structure predictions for acetic acid.
- (57) Price, S.L.; Beyer, T., *Transactions ACA*, **1998**, *33*, 23-31. Progress and problems in the computer prediction of molecular crystal structures and polymorphism.
- (58) Davey, R.J.; Maginn, S.J.; Andrews, S.J.; Black, S.N.; Buckley, A.M.; Cottier, D.; Dempsey, P.; Plowman, R.; Rout, J.E.; Stanley, D.R.; Taylor, A., *J. Chem. Soc., Faraday Trans.*, **1994**, *90*, 1003-1009. Morphology and polymorphism in molecular crystals - terephthalic acid.
- (59) Harding, M.M., *J. Synchrotron Rad.*, **1996**, *3*, 250-259. Recording diffraction data for structure determination for very small crystals.
- (60) Smith, E.D.; Hammond, R.B.; Jones, M.J.; Roberts, K.J.; Mitchell, J.B.O.; Price, S.L.; Harris, R.K.; Apperley, D.C.; Cherryman, J.C.; Docherty, R., *J. Phys. Chem. B*, **submitted**, The determination of the crystal structure of anhydrous theophylline by X-ray powder diffraction with a systematic search algorithm, lattice energy calculations ¹³C and ¹⁵N solid state NMR: A question of polymorphism in a given unit cell.
- (61) Johnson, M.R.; Prager, M.; Grimm, H.; Neumann, M.A.; Kearley, G.J.; Wilson, C.C., *Chemical Physics*, **1999**, *244*, 49-66. Methyl group dynamics in

paracetamol and acetanilide: probing the static properties of intermolecular hydrogen bonds formed by peptide groups.

- (62) Motherwell, W.D.S., *PLUTO (RPLUTO from www.ccdc.cam.ac.uk)*, . 1998, CCDC: Cambridge.

Chapter 7. Dimer or catemer? Calculations of low energy packings, morphology and mechanical properties for small carboxylic acids

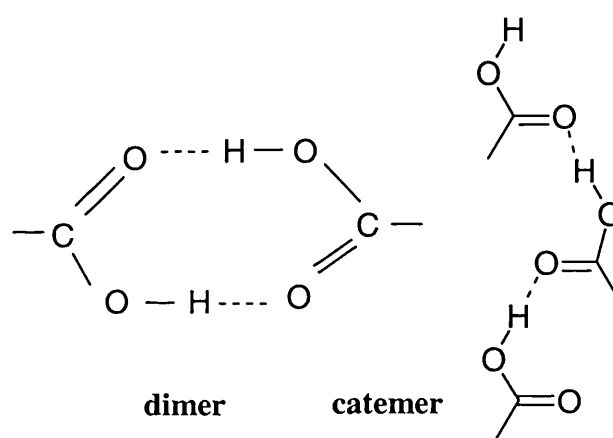
The ability of a search for minima in the lattice energy to predict the crystal structures of three carboxylic acids was examined. In the intermolecular potential applied in this study, the DMA electrostatic model was combined with an empirical atom-atom repulsion-dispersion potential with the polar hydrogen atom parameters specifically adjusted to reproduce a range of carboxylic acids. The catemer structure of formic acid, the dimer-based structure of benzoic acid and both the catemer and dimer polymorphs of tetrolic acid were found reasonably accurately, close to the global minimum in the lattice energy. However, as in the recently highlighted case of acetic acid, there are a variety of energetically competitive structures, with both the catemer and dimer motifs, for formic and tetrolic acid. In contrast, all low energy structures for benzoic acid were based on the dimer motif.

To reduce the number of hypothetical low energy structures that can be considered as potential polymorphs the elastic properties and vapour growth morphologies have been estimated. In contrast to the paracetamol study (chapter 6) these additional criteria do not eliminate many hypothetical structures of the small carboxylic acids as mechanically weak or with relatively slow growth.

The first part of this study has been published.¹

7.1 Introduction

For directing the synthesis of new molecular solids, it is very important to understand the relationship between the functional groups in a molecule and the crystal structure it forms (chapter 1.3). Carboxylic acids are so commonly observed to crystallise forming a dimer motif (scheme 7.1) that this is regarded as a synthon in crystal engineering.² However, there are exceptions where the crystal structure does not include the dimer motif, such as formic, acetic and β -tetrolic acid, in which cases the catemer motif (scheme 7.1) is observed.



Scheme 7.1: Hydrogen bonding motifs observed in carboxylic acids.

A recent statistical study³ of crystal structures in the Cambridge Structural Database (CSD) showed, that the dimer motif occurs for 95% of monocarboxylic, and for 85% of dicarboxylic acids, when there are no other hydrogen bonding groups present. In a non-restricted dataset, where competing functional groups are present, the carboxylic acid dimer has only a 33% probability of being observed. This is consistent with an earlier study,⁴ which concluded that the cyclic dimer predominates in monofunctional acids, but it appears only in 34% of the polyfunctional organics that have the potential to form this particular pattern.⁵ The obvious explanation for the reduced frequency of the carboxylic acid dimer motif in polyfunctional organics is the competition between different hydrogen bond donors and acceptors. For example,^{2, 6} 3-hydroxybenzoic acid, 2,5-dihydroxybenzoic acid and 2,6-dihydroxybenzoic acid can form completely different packing patterns by involving the competing hydrogen bond donor and acceptor groups. However, when the carboxylic acid is the only non-

hydrocarbon group, the crystal structure will almost inevitably be dominated by hydrogen bonding interactions between the acid groups.⁷

The main argument advanced for the preponderance of carboxylic acid dimer over catemer motifs in the crystal structures of $R\text{-CO}_2\text{H}$ is that the bulk of the R substituent will often prevent the formation of the catemer structure.⁸ In cases where R is sufficiently small that it does not obstruct the optimum packing of either motif, such as formic and acetic acid, then the preferred hydrogen bonding motif is presumed to reflect energetic differences between the two hydrogen bonding geometries. An early molecular packing analysis⁹ of acetic acid compared the lattice energies of the catemer and hypothetical dimer structures and concluded that they were practically equal. Detailed studies of the electrostatic interactions between formic acid molecules in chains related by different crystal symmetry elements¹⁰ showed that all hypothetical alternative chain arrangements were at least 5 kcal/mol less stable than the observed catemer structure (but no dimer structures were considered). Studies with an empirical Lennard-Jones plus point charge intermolecular potential¹¹ concluded that the electrostatic interactions stabilise the catemer motif relative to the cyclic dimer motif for small monocarboxylic acids, but that the catemer motif is destabilised by a large R group.

The assumption that differences in the intermolecular interaction energies of the hydrogen-bonding motifs in isolation are the determining factor was challenged by two simultaneous attempts to evaluate different methods of predicting the crystal structure of acetic acid.^{12, 13} These searches for lattice energy minima, using a range of model potentials, found a large number of dimer and catemer based hypothetical crystal structures of acetic acid within a small energy range of the global minimum in the lattice energy. This implies that any small energy differences between the catemer and dimer motif in isolation can be readily compensated for by the other interactions in the crystal.

In this study, we extended the acetic acid work by finding the low energy hypothetical crystal structures of other carboxylic acids and considered the extent to which the differences in lattice energy determine which crystal structures and hydrogen-bonding motifs are adopted by this group of molecules. We investigated formic acid, which adopts a catemer structure, benzoic acid, which forms a dimer structure, and tetrolic acid, which has two known polymorphs, one a dimer and the other a catemer.

We also considered the morphology and mechanical properties of the hypothetical structures to establish whether the properties provide any discrimination between the energetically plausible structures.

The derivation of optimised intermolecular potential functions for the strongly hydrogen bonded carboxylic acids has been a challenge for the last 30 years. Pioneering work in this area, ¹⁴⁻¹⁶ based on semi-empirical calculations, resulted in large discrepancies between the calculated and observed structural data and in overestimated lattice energies. The intermolecular potential for formic acid by Minicozzi and Stroot ¹⁷ was criticised ^{18, 19} for containing too many empirical parameters derived from too little experimental data. A fully empirically derived interatomic potential function ¹⁹ for carboxylic acids reported very promising results for the structural reproduction and reasonable lattice energies. A non-empirical intermolecular potential, ²⁰ based on fitting atom-centred functions to *ab initio* interaction energies, gave competitive results compared to this empirical potential. Unfortunately, the results published for both types of parameters ^{19, 20} have later been found not to be reproducible. ¹⁸ Moreover, two empirically derived force fields ^{21, 22} for carboxylic acids have been based on potential parameters sequentially developed for intermolecular interactions in alkanes, alkenes and finally for the study of the amide hydrogen bond in amide crystals. ^{23, 24} These consistent force fields consist of partial atomic charges, obtained from population analysis, for the electrostatic contribution and a Lennard-Jones type functional form (12-6 or 9-6) for the repulsion-dispersion term. Both potentials successfully reproduce the experimental structural parameters and observed sublimation energies of the majority of a large set of mono- and dicarboxylic acids. ¹¹ A study, based on the experimental electron density distribution for carboxylic acids, used the atomic charge, dipole and quadrupole moments, derived by Hirshfeld's partitioning scheme ²⁵, to describe the electrostatic contribution of a variety of different carboxylic acids. For the repulsion dispersion terms a modified form of the Hagler Lifson potential ^{21, 26} has been derived. This potential has been successfully applied in the analyse of various characteristic properties of carboxylic acids.

Our calculations are performed using a model potential which includes an *ab initio* based DMA model for the dominant electrostatic interactions, to provide the best currently available estimates of the relative lattice energies. The repulsion-dispersion

interactions are represented by an empirical 6-exp potential. This requires additional parameters for the polar hydrogen atoms in carboxylic acids, which are derived, and the total potential is validated.

The issue of the sensitivity of the calculated lattice energy to the geometry of the molecule, and the implications of the differences in the apparent molecular structure in different phases, are also explored in this study.

7.2 Method

7.2.1 Choice of small carboxylic acids for investigation

Three small carboxylic acids, one adopting a catemer, the other a dimer, and the third being polymorphic have been chosen, where the molecular structure can be assumed to be rigid (table 7.1). The molecules of formic acid ²⁷ are hydrogen-bonded in roughly planar chains (catemer motif). The hydrogen-bonded chains are tightly packed in layers forming the space group $Pna2_1$. This hydrogen bond pattern can be specified by graph set analysis ^{7, 28} as C(4). The structure of benzoic acid ²⁹ consists of centrosymmetric dimers, in which two monomer units are linked by a pair of hydrogen bonds between their carboxyl groups, in the expected graph set $R_2^2(8)$. The crystal is formed in the space group $P2_1/c$ with four molecules in the unit cell. The molecules in the triclinic α -form ³⁰ of tetrolic acid form hydrogen-bonded centrosymmetric dimers through the carboxylic group (graph set $R_2^2(8)$), which are arranged in planar layers, forming the spacegroup $P\bar{1}$. The monoclinic β -form ³⁰ consists of planar chains (graph set C(4)), propagating along the twofold screw axis through a single O-H \cdots O hydrogen bond between the pairs of molecules (catemer motif), forming the space group $P2_1$. Both polymorphs have been crystallised from *n*-pentane solution and the β -form converts at 56-58 °C into the α -form. The reverse transformation could not be observed and it is thought ⁸ that the α -form is metastable at low temperatures.

7.2.2 Development of the intermolecular potential

The intermolecular potential employed in this study consists of an *ab initio*-based DMA model for the electrostatic contribution derived of the MP2 6-31G** wavefunction of the isolated molecules by the program suite CADPAC ³¹ and an

empirical isotropic atom-atom repulsion-dispersion potential set. The parameters for the interactions involving the polar hydrogen atoms, H_O , were empirically derived from the parameter set for the polar hydrogens H_N in $N-H_N \cdots N$ and $N-H_N \cdots O=C$ hydrogen bonds in the potential set FIT (chapter 2). The pre-exponential polar-hydrogen repulsion parameters $A_{H_N H_N}$ were scaled to optimise the reproduction of the investigated carboxylic acid crystal structures, thereby empirically taking into account the differences in the character of the polar hydrogens. Therefore a scale factor s ($A_{H_O H_O} = s \times A_{H_N H_N}$) has been introduced with $1 \geq s \geq 0.2$. For validating this new potential parameter, the crystal structures (table 7.1) of acetic acid,³² the polymorphic forms of fumaric acid^{33, 34} and oxalic acid³⁵ have also been considered. In this part of the study, the experimental molecular structures from X-ray data were used, with the bonds to hydrogen adjusted to a standard length³⁶ of 1.08 Å for C-H and 1.02 Å for O-H along the experimental bond direction. For the α -polymorph of tetrolic acid, the location of the hydrogen atoms was not determined experimentally and had to be calculated using the program PLUTO.³⁷

The lattice energies were calculated and minimised using the program DMAREL.³⁸

7.2.3 Crystal structure prediction searches

A high quality *ab initio* MP2 6-31G** model for the molecular structures of all investigated systems, obtained by the program suite CADPAC,³¹ was used in the search. Hypothetical crystal structures for the three investigated molecules were generated by the search program MOLPAK.³⁹ For each molecule, the 24 possible coordination types for $Z'=1$ in the space groups $P1$, $P\bar{1}$, $P2_1$, $P2_1/c$, $P2_12_12_1$, $Pna2_1$, $Pca2_1$, $Pbca$, $C2/c$, $C2$ and Cc , were considered.

The hypothetical crystal structures were then compared to the experimentally determined structures. Therefore, the *ab initio* optimised molecular structures have been pasted into the experimental crystal structures and the systems have been relaxed using the same intermolecular potential (table 7.3). The cell parameters, space group, lattice energy and the simulated powder diffraction patterns were compared.

The morphologies of the experimental and the lowest energy hypothetical crystal structures were calculated for all investigated systems using the Bravais-Friedel-

Donnay-Harker (BFDH) ⁴⁰ and the attachment energy model ⁴¹ by the program ⁴² Cerius²3.5 and the empirically derived force field by Hagler *et al.*^{21, 22, 26}

7.3 Results

7.3.1 Optimisation of the empirical repulsion potential parameters

Scaling the pre-exponential polar hydrogen repulsion parameters reduced the hydrogen bond lengths and improved the overall reproduction of the crystal parameters. The optimum reproduction for most of the structures, with the exception of β -oxalic acid, is achieved for a medium value of the scale factor, i.e. the curve runs through a minimum r.m.s. percentage error in middle range values of s (figure 7.1a). For the intermolecular hydrogen bond lengths, an optimum is achieved for all systems for a scale factor between 0.35 and 0.6 (figure 7.1b).

A scaling factor of 0.45 for the pre-exponential polar hydrogen repulsion parameter gives a reasonable reproduction of most of the crystal structures, hydrogen bond lengths and lattice energies (table 7.2). Thus, the parameters in table 7.3 were used in the crystal structure energy calculations. The exceptions which were not reproduced well were the dimer motifs of α -tetrolic acid and β -oxalic acid, with r.m.s percentage errors of 5.9 % and 11.4 % respectively. Other empirical model potentials fitted to carboxylic acids have also encountered problems in reproducing the crystal structure of β -oxalic acid, including an empirical exp-6 potential, ⁴ fitted to a large database of mono- and di-carboxylic acids; the Lennard-Jones plus point charge potential of Hagler *et al.* ²¹; and the potential set by Smit derived by fitting to *ab initio* data. ²⁰ Berkovitch-Yellin *et al.* ¹⁰ showed that the offset between the chains in the β -polymorph is primarily determined by van der Waals forces, despite the polar nature of the carboxyl groups. Therefore the problems in reproducing this crystal structure are in the repulsion-dispersion potential, probably mainly in the C \cdots O interactions. ⁴ A non-empirical potential specifically derived for oxalic acid, ⁴³ which does not assume that the two different types of oxygen atoms have the same repulsion-dispersion parameters, and does not use any combining rules to derive heterogeneous parameters, gives a reasonable reproduction of both the polymorphs of oxalic acid. Hence the poor reproduction of the chain offset in β -oxalic acid in our transferable potential is not a cause for concern in this study.

The reproduction of the structure of α -tetrolic acid is very sensitive to the molecular geometry, and is acceptably reproduced with the *ab initio* optimised molecular structure (see next section). Hence, this DMA + exp-6 potential (table 7.3) has the advantage of being an accurate model for the dominant electrostatic interaction, and has been validated against experimental crystal structures as well as other carboxylic acid potentials have been.

7.3.2 Effect of intramolecular structure variations

The molecular structures from the experimental crystal structures and the *ab initio* optimised molecular models are compared in table 7.4. The structural parameters that are not directly involved in the carboxylic acid group have very similar experimental and *ab initio* values, with a maximum difference of less than 4%. All optimised structures have the expected planarity and linearity of the molecular fragments, whereas there are slight deviations in some of the experimental structures. There is similar agreement between the *ab initio* and X-ray values of the bond parameters of the acid groups (*i.e.* a difference of less than 3.5%), with the notable exceptions of the acid groups which are in the eight-membered ring dimer motif. For these dimer structures, the experimental C-OH and the C=O bonds are quite similar in length, and so the differences from the *ab initio* values are much larger (benzoic acid 8.1% and 3.8% and α -tetrolic acid 7.2% and 2.6% for the C-O and C=O bonds respectively). This shortening of the bond lengths in the dimer motif structures is an artifact of crystal structure determination, reflecting a state of disorder, either static or dynamic.⁸ Gavezzotti,⁴ in his analysis of intermolecular hydrogen bonds in carboxylic acids, has chosen a difference of 0.05 Å between the C-OH /C=O bond lengths as a minimum threshold to differentiate between disordered and non-disordered structures. Hydrogen bond dynamics in carboxylic acid dimers have been studied experimentally, using a variety of spectroscopic techniques.⁴⁴⁻⁴⁸ This shows that our gas phase *ab initio* model for the molecular structure differs from the apparent X-ray molecular structure in the crystal because of proton disorder in the eight-membered ring, in addition to usual limitations in comparing *ab initio* and crystal structures.

A comparison of the minimised crystal structures using the experimental molecular structure (table 7.2), and the *ab initio* optimised molecular structure (table 7.5), shows that changing the molecular geometry has a fairly small effect on the

minimised crystal structure (compared to the errors inherent in comparing static minima with room temperature crystal structures). The only exception is α -tetrolic acid, where the reproduction of the crystal structure is much better using the optimised molecular structure (an r.m.s. percentage error of 2.9 %, in comparison of 5.9 % using the experimental molecular structure), which may be due to errors in the assumed direction of the C-H bonds in the 'experimental' molecular structure.

7.3.3 Generation of hypothetical crystal structures

The minima produced by the MOLPAK/DMAREL search are compared with the minimised experimental structures of each molecule in figure 7.2. For all carboxylic acids, the search procedure has found a reasonable match to the minimum corresponding to the experimental structure, calculated using the same *ab initio* optimised molecular structure and intermolecular potential. The cell lattice parameters, energies and hydrogen bond lengths for the minimisations starting from the experimental structures are compared with the closest hypothetical structures found in the search in table 7.5. The r.m.s. percentage error in the reduced cell lengths is 1.9 % for formic acid, 2.6 % for benzoic acid, 0.0 % for α -tetrolic acid and 5.7 % for β -tetrolic acid. This agreement is also apparent when the crystal structures and the simulated powder diffraction diagrams for these crystal structures (figure 7.2) are compared.

In all cases the minimised experimental structure and the corresponding structure found in the search are among the lowest energy structures (figure 7.3). For benzoic acid and β -tetrolic acid, the minimised experimental structure gives the global minimum in the lattice energy. The maximum deviation of the minimised experimental structure from the global minimum was less than 5 kJ/mol, found for the metastable α -polymorph of tetrolic acid. Thus the search has reproduced the experimental order of stability for tetrolic acid.⁸ Although there is a close correspondence between the minimised experimental structures and the corresponding packing patterns found in the search, the structural differences produce a small difference in calculated lattice energy, of less than 2 kJ/mol. In the case of benzoic acid, these structures are separated in energy from other, very different hypothetical crystal structures (figure 7.3b). For formic acid, four structures found in the search have a lower lattice energy than the search structure which corresponds to the experimental structure (figure 7.3a). There is

a multitude of hypothetical low energy structures of tetrolic acid in the energy range defined by the two known polymorphs (figure 7.3c).

An investigation of the low energy hypothetical crystal structures (table 7.6) reveals a variety of different closely packed structures with both catemer and dimer motifs for all but benzoic acid. For benzoic acid, the studied system with the largest substituent, the catemer with the lowest energy is found 8 kJ/mol above the experimental structure found in the search (the global minimum). There are a variety of different packings of the benzoic acid dimer motif that are more favourable in energy than any catemer, but the observed structure is the most dense and energetically favourable.

For formic acid, the structures within 1 kcal/mol of the global minimum include the experimental catemer (scheme 7.1) in various packings, a hypothetical dimer motif structure, a chain motif incorporating an intermolecular torsion angle of 140° and chains formed through the hydroxyl groups. For tetrolic acid, the low energy hypothetical structures contain catemer and dimer motifs in approximately equal numbers.

7.3.4 Elastic constants and morphology calculations

The calculated elastic properties for the minimised experimental and for all the energetically feasible, hypothetical structures of formic, benzoic and tetrolic acid are given in tables 7.1-7.3. The elastic properties calculated from the experimental structures are well reproduced for α -tetrolic acid by the corresponding structures found in the search, but just reasonably approximated for the other molecules. This shows that there are small differences between these structures, for all investigated polymorphs, but α -tetrolic acid. Furthermore, the elastic constants are very sensitive to variations in the 2nd derivative matrix. Although the crystal structures of neighbouring minima on the potential energy surface might be quite similar, their Hessian matrices are very likely to be different.

All the elastic stiffness matrix elements C_{ij} are positive definite for all hypothetical structures, satisfying the Born criteria for mechanical stability. However, three hypothetical structures for tetrolic acid, AZ8, BD12 and BD25 have at least one matrix element $C_{ij} \leq 0.5$, with $C_{66} = 0.03$ for BD12. This indicates that these structures are easily deformed.

The predicted morphologies for the experimental and the low energy structures found in the search are shown in figures 7.8.1-7.8.3. In general the geometry based BFDH and AE morphologies do predict some variations in shape for all investigated carboxylic acids. The calculated AE morphologies for the experimental structures, with the exception of α -tetrolic acid, show an approximately plate-like shape. For benzoic acid an experimentally determined morphology has been published ⁴⁹ and this is well reproduced by the attachment energy calculations. For α -tetrolic acid, the search found an exact match. The structural differences between the experimental structures and the corresponding search minima for formic, benzoic and β -tetrolic acid have only a small effect on the predicted AE morphologies and the overall shape is well reproduced. The sensitivity of these properties to small variations in the molecular structure is therefore minor, compared to the differences obtained when morphologies of different crystal structures are compared.

The differences in morphology are quantified in tables 7.8.1-7.8.3. The values for the aspect ratio, the distance between the centre of the furthest and nearest faces are similar for both methods. The slowest growing face is tabulated for each crystal structure, along with its relative area and attachment energy. This is usually the most dominant face, though if another slow growing face occurs more frequently, this may have a larger relative surface area in the crystal. The attachment energy for any such faces, and for any faces which have a larger inter-planar spacing than the slowest growing face, and so would be larger in the BFDH model morphology, are also given in table 7.6. All of such faces have more attractive attachment energies and hence a faster growth rate, than the slowest growing faces.

The comparison of the attachment energies calculated for crystal structures for the same carboxylic acid molecule, using the same intermolecular potential, does not reveal any large differences in attachment energies for the experimental and hypothetical structures. Although the experimental structures do not have the largest attachment energy for their slowest growing faces, the differences to the maximum value are so small for all carboxylic acids compared to the errors inherent in these calculations, that no conclusions can be drawn about their relative growth rates.

7.4 Discussion

7.4.1 Catemer versus dimer?

Our results for formic and tetrolic acid, along with recent studies on acetic acid ^{12, 13} show that the dimer and catemer structures, and a variety of other motifs, have very similar low lattice energies. Thus, there is no great intrinsic energy difference between the catemer and dimer motifs for the smaller acids. Any small differences can easily be compensated for by the variations in the lattice energy contribution from the packing of these motifs. The lack of low energy catemer structures for benzoic acid seems to support Leiserowitz's argument ⁸ that the catemer motif can only be formed by molecules with small substituents on the acid group, as the steric bulk of the aromatic ring prevents catemer formation. Nevertheless, we did find a distorted catemer structure within only 8.0 kJ/mol of the minimum. The packing in this structure consists of non-planar chains with relatively large intermolecular O...O distances of 2.75 Å compared to 2.63 Å in the experimental dimer motif. The existence of a chain suggests that the steric strain put on the catemer motif by the phenyl substituent is relatively small. This is consistent with the chains of dimers being experimentally observed for some substituted benzoic acids, ⁵⁰ and the chain structure of trimellitic anhydride, ⁵¹ where the acid chain motif is stabilised by the interaction between aromatic hydrogen atoms and the anhydride carbonyl.

Thus, the prevalence of the dimer motif in crystal structures of carboxylic acids does not imply that this motif, in itself, is significantly more stable than alternatives. The crystal packing and steric interactions of the other functional groups will play a major role in determining the relative energetics of the possible crystal structures.

7.4.2 Additional criteria to lattice energy calculations

Compared to the paracetamol study (chapter 6), many more hypothetical structures have been found within an energy range of about 5 kJ/mol for the investigated carboxylic acids. For formic acid 31, for benzoic acid 10 and for tetrolic acid 39 low energy minima have been found within the energy range of possible polymorphism. The energies separating these energetically feasible crystal structures are so small that changes in the intermolecular potential or in the molecular model or entropy effects would probably change the relative energies of these structures.

For polymorphic forms with similar thermodynamic stability, the mechanical stability of the growing crystallite may be a factor in determining which of these forms are going to be observed experimentally. All lattice energy minima in tables 7.1-7.3 satisfy the Born stability criteria, but three forms for tetrolic acid have at least one small diagonal shear elastic constant and therefore are unlikely to be sufficiently mechanically stable to grow in competition with forms of similar thermodynamic stability. For formic and benzoic acid, none of the hypothetical structures seems to be particularly mechanically weak.

The attachment energy calculations lead to similar values for all hypothetical structures for each carboxylic acid and no extremes in morphology. Although all experimental structures do not have the most attractive attachment energies, the differences are so small that it is difficult to draw conclusions about the relative growth rates of their corresponding structures.

In comparison to paracetamol, the molecular carboxylic acid structures are relatively small and lack a particular shape. This generally leads to an increase in the number of low energy minima found in crystal structure prediction searches (chapter 3) but these hypothetical structures are less distinct from each other. Although the catemer and dimer motif are well known and distinct motifs in crystal engineering, our searches generated several structures based on dimers and a variety of very different arrangements of the catemer. A summary of the intermolecular hydrogen bond lengths found in these structures, tables 7.6.1-7.6.3, reveals that in most of the cases both sorts of possible hydrogen bonds, $C=O\cdots(H)-O$ and $O-(H)\cdots O-H$ are present. This is a distinct difference to the paracetamol case study, where of the four different possible hydrogen bonds only a subset was present in each hypothetical structure.

7.5 Conclusions

For all molecules, the structure corresponding to minimising the experimental structure has been found successfully in the MOLPAK/DMAREL search. However, for this comparison it was essential to use the *ab initio* optimised molecular structure in both calculations, as the differences between the *ab initio* and X-ray structure, although small, have a significant effect on the relative energetics. The structural differences are unusually large for carboxylic acids in the dimer structures because of disorder in the proton positions. Although the same structure and intermolecular potential were used,

there were sometimes small differences in the two minima for most of the molecules, which seems likely to reflect the existence of very closely related minima.

Although the search method has successfully found the experimental structures, only in two cases are these the global minimum. Thus, a prediction of the crystal structure based on energetic preferences would have correctly identified benzoic acid and possibly β -tetrolic acid. However, there is no energetic reason for picking out α -tetrolic or the known structures of formic or acetic acid from the low energy minima. The absolute values of the calculated lattice energies are quite reasonable, being within the usual threshold of significance.^{52, 53}

This is in accord with the emerging experience in crystal structure prediction methods based on searching for minima in the lattice energy that for many small systems, including the CCDC blindtest compound I (chapter 5), the 2-hydronitronylnitroxide radical (HNN)⁵⁴ and uracil⁵⁵, a set of very different packing arrangements is found with a lattice energy so close to the global minimum that there is no energetic reason for selecting the experimental structures from this set. This is the case for the smaller monocarboxylic acids, like formic, tetrolic and acetic acid, as the energetic differences between the hypothetical structures forming the catemer or the dimer motif are so small, if any, that factors other than energy seem to be responsible for the hydrogen bond motif formed.

When unknown crystal structures are predicted to be at, or close to, the global minimum, they could correspond to polymorphs that might be found in the future, as for example the high pressure forms of formic acid⁵⁶ and acetic acid.⁵⁷ Indeed, the structures predicted by such search methods could aid experimental work to identify new polymorphs.

Kinetic factors involved in nucleation and crystal growth may well favour certain low energy structures. In the case of small carboxylic acids, the consideration of the mechanical stability and growth rates, of the hypothetical minima generated by methods based on lattice energy does not help to theoretically determine which polymorphs are going to be found experimentally. This is likely to be related to the size and lack of shape of these molecular structures. As this approach gave promising results for paracetamol, it is certainly important to test it on further systems.

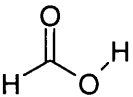
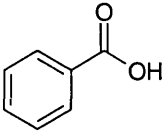
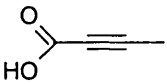
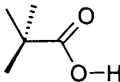
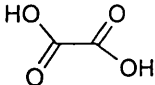
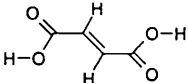
Carboxylic acid CSD REFCODE		motif	Space Group	Z'	Z	$\Delta H_{\text{sub}}^{58}$ kJ/mol
Formic acid FORMAC01		chain	$Pna2_1$	1	4	-62.1
Benzoic acid BENZAC02		dimer	$P2_1/c$	1	4	-88.7
Tetrolic acid α -form: TETROL		dimer	$P\bar{1}$	1	2	N/A
β -form: TETROL01		chain	$P2_1$	1	2	N/A
Acetic acid ACETAC01		chain	$Pna2_1$	1	4	-67.3
Oxalic acid α -form: OXALAC03		chain	$Pcab$	0.5	4	-97.9
β -form: OXALAC04		dimer	$P2_1/c$	0.5	2	-93.3
Fumaric Acid α -form: FUMAAC		chain	$P2_1/c$	1.5	6	-123.6 ‡
β -form: FUMAAC01		chain	$P\bar{1}$	0.5	1	-123.6

Table 7.1

Crystal structures used for deriving the repulsion-dispersion potential for carboxylic acids.

‡ Reference does not specify which polymorph is involved.

		cell vectors / Å (Δ / %)*			r.m.s. error (%)	cell angles / ° (Δ / %)*			cell density / g/cm ³ (Δ / %)*	U _{latt} / kJ/mol	error in U _{latt} / kJ/mol ‡	O-(H)···O / Å (Δ / %)*
		a	b	c		α	β	γ				
Formic acid	expt	10.241	3.544	5.356	0.6	90	90	90	1.57	-57.2	-4.9	2.624
	min	10.209	3.577	5.384		90	90	90	1.55	-58.2	-3.9	2.683
		(0.3)	(-0.9)	(-0.5)					(-1.3)			(2.2)
Benzoic acid	expt	5.500	5.128	21.934	2.6	90	97.4	90	1.32	-105.5	16.8	2.627
	min	5.287	5.244	21.966		90	101.2	90	1.36	-107.9	19.2	2.559
		(-3.9)	(2.3)	(0.1)			(4.0)		(3.0)			(-2.6)
Tetrolic acid α-form	expt	7.320	5.099	7.226	5.9	83.9	117.5	112.0	1.26	-77.2		2.649
	min	6.777	5.444	7.082		84.9	115.7	115.8	1.33	-79.4		2.609
		(-7.4)	(6.8)	(-2.0)		(1.2)	(-1.5)	(3.4)	(5.6)			(-1.5)
β-form	expt	7.887	7.121	3.937	1.9	90	100.2	90	1.28	-74.2		2.655
	min	7.703	6.997	3.999		90	100.5	90	1.32	-75.5		2.680
		(-2.3)	(-1.7)	(1.6)			(0.3)		(3.1)			(0.9)
Acetic acid	expt	13.310	4.090	5.769	2.6	90	90	90	1.27	-63.9	-3.4	2.627
	min	13.184	3.930	5.895		90	90	90	1.31	-64.6	-2.7	2.662
		(-0.9)	(-3.9)	(2.2)					(3.2)			(1.3)
Oxalic acid α - form	expt	6.548	7.844	6.086	1.6	90	90	90	1.91	-100.1	2.2	2.702
	min	6.709	7.910	6.036		90	90	90	1.87	-101.3	3.4	2.694
		(2.5)	(0.8)	(-0.8)					(-2.1)			(-0.3)
β-form	expt	5.330	6.015	5.436	11.4	90	115.8	90	1.91	-111.2	17.9	2.674
	min	5.309	6.692	4.558		90	112.3	90	2.00	-114.9	21.6	2.664
		(-0.4)	(11.3)	(-16.2)			(-3.0)		(4.7)			(-0.4)
Fumaric Acid α-form	expt	7.619	15.014	6.686	3.2	90	112.0	90	1.63	-128.4	4.8	2.682
	min	7.593	15.326	6.343		90	112.9	90	1.70	-131.3	7.7	2.651
		(-0.3)	(2.1)	(-5.1)			(0.8)		(4.3)			(-1.2)
β-form	expt	5.264	7.618	4.487	2.4	106.9	86.3	134.9	1.62	-128.4	4.8	2.671
	min	5.223	7.584	4.306		104.1	84.8	136.1	1.69	-134.3	10.7	2.645
		(-0.8)	(-0.5)	(-4.0)		(-2.6)	(-1.8)	(0.9)	(4.3)			(-1.0)

Table 7.2 Validation of the optimised repulsion-dispersion potential for carboxylic acids.

The experimental crystal structures (expt) are compared with the minimum in the lattice energy found starting from the experimental structure (min), using the DMA + exp-6 potential in Table 7.3.

* Quantities in parenthesis are percentage error, $\Delta / \% = 100((\text{min-expt})/\text{expt})$.

‡ This difference ($\Delta H_{\text{sub}} - U_{\text{latt}}$) is an estimate of the error in the calculated lattice energy ignoring a variety of small contributions^{52, 53} and experimental error.

Interaction	A / kJ/mol	B / Å ⁻¹	C / kJ/mol Å ⁶
C C	369743.0	3.60	2439.8
H _C H _C	11971.0	3.74	136.4
O O	230064.1	3.96	1123.6
H _O H _O	2263.3	4.66	21.5

Table 7.3: The repulsion-dispersion potential parameters used in conjunction with a DMA electrostatic model for the carboxylic acids.

	Formic acid		Benzoic acid		Tetrolic acid		
	expt FORMAC01	opt	expt BENZAC02	opt	expt, α TETROL	expt, β TETROL01	opt
Acid Group							
distances/ Å							
C=O	1.222	1.212	1.268	1.220	1.252	1.204	1.219
C-O (-H)	1.308	1.349	1.258	1.360	1.265	1.310	1.356
(C -) O - H	1.020	0.971	1.020	0.970	1.020	1.020	0.971
H - C (OOH)	1.080	1.092	-	-	-	-	-
(R -) C - C (OOH)	-	-	1.484	1.484	1.441	1.441	1.443
angles/°							
O = C - O	124	125	123	123	125	123	124
H, R - C - O	112	109	119	119	116	113	111
torsions/°							
O = C - O - H	3	0	0	0	-1	-2	0
H, R - C - O - H	180	-180	-179	-180	180	178	-180
Relative internal	10.28	0	64.08	0	43.17	75.84	0
energy / kJ/mol ‡							
Substituent R							
distances/ Å							
C(sp ²) - C(sp ²)			1.386	1.399			
			1.383	1.393			
			1.366	1.395			
			1.380	1.396			
			1.377	1.391			
			1.384	1.398			
C(sp) - C(sp ³)					1.458	1.455	1.457
C(sp) - C(sp)					1.181	1.178	1.221
angles/°							
(O=)C-C(sp)-C(sp)					178	176	176

Table 7.4: Comparison of the experimental and the *ab initio* optimised molecular structures of formic, benzoic and tetrolic acid.

‡ Δ = E (expt) – E (opt); difference in energy evaluated using an MP2 6-31G** wavefunction.

	a / Å	b / Å	Cell parameters				U _{latt} /	ΔU _{latt} ‡	Cell density/
			c / Å	α / °	β / °	γ / °	KJ/mol	/kJ/mol	g cm ⁻³
formic acid									
min _{opt}	10.325	3.583	5.376	90.0	90.0	90.0	-52.2	0.3	1.54
search	10.070	3.655	5.345	90.0	90.0	90.0	-50.8	1.7	1.55
benzoic acid									
min _{opt}	5.344	5.346	21.950	90.0	102.0	90.0	-86.4	0	1.39
search	5.465	5.424	21.620	90.0	109.5	90.0	-87.4	1.9	1.34
tetrolic acid, α - form									
min _{opt}	7.132	5.313	7.167	81.8	117.6	116.9	-67.3	4.3	1.31
search	7.406	5.313	7.167	98.2	121.4	107.3	-67.4	4.2	1.31
min _{opt} (red) ‡‡	5.313	6.690	7.167	112.4	98.2	108.1	-67.3	4.3	1.31
search (red)	5.313	6.691	7.167	112.4	98.2	108.1	-67.4	4.2	1.31
tetrolic acid, β – form									
min _{opt}	7.680	6.945	3.922	90.0	101.9	90.0	-72.0	0	1.34
search	7.508	7.073	4.159	90.0	75.9	90.0	-70.6	1.0	1.30
min _{opt} (red) ‡‡	3.922	6.945	7.680	90.0	101.9	90.0	-72.0	0	1.34
search (red)	4.159	7.073	7.507	90.0	104.1	90.0	-70.6	1.0	1.30

Table 7.5 Evaluation of the search procedure.

Comparison of closest crystal structure found (search) in the MOLPAK/DMAREL search with the corresponding minimum found (min_{opt}) starting from the experimental structure. Both minima were calculated using the same *ab initio* optimized molecular structure and model potential.

‡ ΔU_{latt} is the energy above the global minimum found in search.

‡‡ The similarity is more apparent using the Niggli reduced cells (red).⁵⁹

	Motif	Space group	Reduced cell parameters						Hydrogen bonds		$U_{\text{latt}} / \text{kJ mol}^{-1}$	$\Delta U_{\text{latt}} / \text{kJ mol}^{-1}$	Density / g cm^{-3}
			a/Å	b/Å	c/Å	$\alpha/^\circ$	$\beta/^\circ$	$\gamma/^\circ$	C=O \cdots (H)-O / Å ($^\circ$)	O-(H) \cdots O-H / Å ($^\circ$)			
Expt	chain	<i>Pna2₁</i>	3.544	5.356	10.241	90.0	90.0	90.0	2.624 (175) 3.451 (98) 3.538 (104)	3.333 (98) 3.544 (97)	-53.1	-	1.57
Min _{expt}	chain	<i>Pna2₁</i>	3.578	5.384	10.209	90.0	90.0	90.0	2.683 (172) 3.506 (99) 3.562 (99)	3.274 (96) 3.577 (92)	-53.8	-	1.55
Min _{opt}	chain	<i>Pna2₁</i>	3.583	5.376	10.325	90.0	90.0	90.0	2.701 (171) 3.537 (102) 3.578 (99)	3.253 (98) 3.583 (92)	-52.2	0.3	1.54
AV21	chain	<i>Pna2₁</i>	3.655	5.345	10.070	90.0	90.0	90.0	2.756 (171) 3.759 (105)	3.024 (94) 3.463 (96)	-50.8	1.7	1.55
Hypothetical structures found in the search which have not been experimentally characterised													
AI2	chains via OH	<i>P2₁/c</i>	4.552	6.380	6.689	95.1	90.0	90.0	2.832 (99) 3.165 (98)	2.695 (161) 3.756 (93)	-52.5	0	1.58
AI19	chains via OH	<i>P2₁/c</i>	4.550	6.528	7.361	116.1	90.0	90.0	2.812 (100) 3.162 (95)	2.692 (160) 3.740 (95) 3.910 (120)	-51.9	0.6	1.56
AQ23	chain	<i>P2₁2₁2₁</i>	5.548	5.557	6.591	90.0	90.0	90.0	2.691 (174)	3.059 (96) 3.889 (97)	-51.8	0.7	1.50
AQ15	chains via OH	<i>P2₁2₁2₁</i>	4.596	6.242	6.988	90.0	90.0	90.0	2.962 (110)	2.701 (158)	-50.9	1.6	1.53
AM20 c	dimer	<i>P$\bar{1}$</i>	3.812	7.225	7.957	66.3	80.9	86.7	2.707 (173) 2.778 (171) 2.929 (103) 3.191 (97) 3.271 (100) 3.811 (92)	3.145 (108)	-50.8	1.7	1.54
FA15	chains via OH	<i>P2₁/c</i>	3.730	3.763	14.513	90.0	90.0	104.2	2.790 (162) 2.794 (111) 3.149 (92) 3.153 (93)		-50.7	1.8	1.55
CC22	chains via OH	<i>Pbca</i>	4.524	6.417	13.600	90.0	90.0	90.0	2.976 (96) 3.375 (101)	2.671 (165) 3.778 (105) 3.778 (110)	-50.7	1.8	1.55

Table 7.6.1 (continued)

	Motif	Space group	Reduced cell parameters						Hydrogen bonds		$U_{\text{latt}} / \text{kJ mol}^{-1}$	$\Delta U_{\text{latt}} / \text{kJ mol}^{-1}$	Density / g cm^{-3}
			$a/\text{\AA}$	$b/\text{\AA}$	$c/\text{\AA}$	α°	β°	γ°	$\text{C}=\text{O}^{\cdots}(\text{H})-\text{O} / \text{\AA} (^\circ)$	$\text{O}-(\text{H})^{\cdots}\text{O}-\text{H} / \text{\AA} (^\circ)$			
AK4	dimer	$P2_1/c$	3.878	7.171	7.176	96.9	90.0	90.0	2.724 (178) 3.236 (101)	2.960 (95) 3.870 (133) 3.879 (93)	-50.5	2.0	1.54
CC5	chain	$Pbca$	4.816	6.577	12.081	90.0	90.0	90.0	2.794 (131) 2.915 (139)	3.593 (102)	-50.4	2.1	1.60
AK22	dimer	$P2_1/c$	4.647	6.640	7.024	114.1	90.0	90.0	2.815 (168) 2.864 (113)	3.615 (110)	-50.3	2.2	1.55
AQ7	chain	$P2_12_12_1$	4.587	6.071	7.098	90.0	90.0	90.0	2.774 (111) 2.794 (161)		-50.3	2.2	1.55
AF17	chain	$P2_1$	3.713	4.575	5.851	90.0	98.7	90.0	2.781 (112) 2.813 (160)	3.713 (112)	-50.2	2.3	1.56
AI4	chain	$P2_1/c$	4.596	6.317	7.373	114.4	90.0	90.0	2.791 (111) 2.809 (161) 3.216 (96)	3.166 (93)	-50.2	2.3	1.57
CC7 d	chain	$P2_1/c$	4.591	6.319	13.604	97.3	90.0	90.0	2.790 (111) 2.789 (111) 2.807 (161) 2.810 (161) 3.202 (96) 3.244 (96)	3.163 (94)	-50.1	2.4	1.56
AK23	chain	$P2_1/c$	4.865	6.465	6.672	116.0	90.0	90.0	2.785 (143) 2.855 (126)	3.491 (102) 3.951 (108) 3.951 (132)	-50.0	2.5	1.62
AK5	dimer	$P2_1/c$	5.722	5.822	6.065	95.8	90.0	90.0	2.715 (175)	3.045 (103) 3.632 (99) 3.782 (134)	-50.0	2.5	1.52
FC3	dimer	$P2_1/c$	3.727	4.935	11.300	96.7	90.0	90.0	2.722 (173) 3.235 (94) 3.989 (105)	2.791 (92) 3.853 (133)	-49.8	2.7	1.48
AB5	dimer	$P\bar{1}$	3.846	4.677	5.668	99.8	96.5	93.9	2.765 (167) 2.926 (110)	3.845 (114)	-49.8	2.7	1.54
FA14	chains via OH	$P2_1/c$	3.524	4.567	13.000	90.0	90.0	107.7	3.170 (102) 3.267 (93)	2.659 (170) 3.818 (105)	-49.8	2.7	1.53

Table 7.6.1 (continued)

	Motif	Space group	Reduced cell parameters						Hydrogen bonds		$U_{\text{latt}} / \text{kJ mol}^{-1}$	$\Delta U_{\text{latt}} / \text{kJ mol}^{-1}$	Density / g cm^{-3}
			$a/\text{\AA}$	$b/\text{\AA}$	$c/\text{\AA}$	α°	β°	γ°	$\text{C}=\text{O}^{\cdots}(\text{H})-\text{O} / \text{\AA} (^\circ)$	$\text{O}-(\text{H})^{\cdots}\text{O}-\text{H} / \text{\AA} (^\circ)$			
CA8	dimer	$P\bar{1}$	3.705	4.925	5.649	82.7	87.2	89.1	2.722 (173) 3.286 (97) 3.873 (103)	2.777 (93) 3.853 (133)	-49.8	2.7	1.50
AM17	chain	$P2_1/c$	4.864	6.409	6.729	90.0	94.9	90.0	2.692 (172) 3.155 (102)	3.913 (98)	-49.6	2.9	1.46
AZ9	chain	$P2_12_12_1$	4.422	6.565	7.374	90.0	90.0	90.0	2.729 (141)	2.745 (125)	-49.6	2.9	1.43
CB3	dimer	$Pbca$	5.579	6.128	11.777	90.0	90.0	90.0	2.710 (175) 3.550 (101) 3.764 (135)	3.069 (102)	-49.5	3.0	1.52
FA3	dimer	$P2_1/c$	3.746	4.766	11.481	90.0	90.0	92.7	2.719 (173) 3.244 (95) 3.956 (104)	2.801 (95) 3.849 (133)	-49.0	3.5	1.49
AM7	dimer	$P2_1/c$	4.660	6.297	7.089	90.0	90.0	97.5	2.780 (169) 2.852 (111)		-49.0	3.5	1.48
CB19	chain	$Pbca$	4.998	7.678	10.528	90.0	90.0	90.0	2.751 (151) 3.043 (103)	3.429 (95)	-48.4	4.1	1.51
FC9	chain	$P2_1/c$	3.793	4.565	11.385	90.7	90.0	90.0	2.766 (175) 3.030 (93)	2.972 (98)	-48.3	4.2	1.55
CB8	dimer	$Pbca$	5.390	7.747	10.204	90.0	90.0	90.0	3.545 (102)	3.897 (112)	-48.1	4.4	1.43
AZ7	chain	$P2_12_12_1$	3.890	4.496	11.463	90.0	90.0	90.0	2.765 (176) 3.016 (95)	2.985 (98)	-47.8	4.7	1.53
BD6	chain	$Pna2_1$	3.813	4.517	11.631	90.0	90.0	90.0	2.765 (175) 3.000 (91)	2.967 (97)	-47.5	5.0	1.53

Table 7.6.1 (continued) Low energy crystal structures for formic acid

The experimental crystal structure for formic acid is contrasted with lattice energy minima calculated using the same intermolecular potential. All minimisations used the *ab initio* optimised molecular structure, except Min_{expt} , where the experimental molecular structure was used (X-ray bondlengths to hydrogen atoms were standardised (ref)). The lowest energy crystal structures found in the MOLPAK/DMAREL search are designed by the MOLPAK coordination group and number, and the space group is that of the conventional setting of each structure. The Niggli reduced cell parameters (ref) are tabulated here to aid comparison.

- a The hydrogen bonds were characterised using PLUTO (ref), tabulating the $\text{O}^{\cdots}\text{O}$ distance ($< 4 \text{\AA}$) and $\text{O}-\text{H}^{\cdots}\text{O}$ angle ($> 90^\circ$).
- b ΔU_{latt} is the energy above the global minimum found in the search.
- c This $P\bar{1}$, $Z' = 2$ structure was found as a true minimum from a $P2_1/c$ starting structure.
- d This $P2_1/c$, $Z' = 2$ structure was found as a true minimum from a $Pbca$ starting structure.

	Motif	Space group	Reduced cell parameters						Hydrogen bonds ^a		$U_{\text{latt}}^b / \text{kJ mol}^{-1}$	$\Delta U_{\text{latt}} / \text{kJ mol}^{-1}$	Density / g cm^{-3}
			a/Å	b/Å	c/Å	α°	β°	γ°	C=O \cdots (H)-O / Å ($^\circ$)	O-(H) \cdots O-H / Å ($^\circ$)			
Expt	dimer	$P2_1/c$	5.128	5.500	21.934	97.0	90.0	90.0	2.627 (172)	3.431 (123)	-105.5	-	1.32
Min _{expt}	dimer	$P2_1/c$	5.244	5.287	21.570	92.7	90.0	90.0	2.559 (177)	3.276 (121)	-107.9	-	1.36
Min _{opt}	dimer	$P2_1/c$	5.344	5.345	21.229	90.0	92.2	90.0	2.641 (171)	3.658 (138)	-86.4	1.0	1.39
FC22	dimer	$P2_1/c$	5.424	5.465	20.450	95.0	90.0	90.0	2.644 (171)	3.655 (138)	-87.4	0	1.34
Hypothetical structures found in the search which have not been experimentally characterised													
AI25	dimer	$P2_1/c$	3.726	11.763	14.801	108.9	90.0	90.0	2.670 (176)	3.790 (136)	-84.5	2.9	1.32
FA16	dimer	$P2_1/c$	6.104	6.885	15.953	90.0	90.0	111.0	2.662 (167)	3.696 (137)	-84.3	3.1	1.30
AM17	dimer	$P2_1/c$	6.101	9.170	11.068	90.0	96.5	90.0	2.659 (175)	3.759 (136)	-83.7	3.7	1.32
AM24	dimer	$P2_1/c$	5.270	10.837	10.844	90.0	90.5	90.0	2.653 (171)	3.691 (137)	-83.2	4.2	1.31
AM21	dimer	$P2_1/c$	3.865	10.958	14.898	90.0	90.0	90.1	2.671 (178)	3.815 (135) 3.865 (91)	-83.2	4.2	1.29
AB9	dimer	$P\bar{1}$	5.675	6.764	8.870	91.1	102.5	108.6	2.654 (171)	3.696 (137)	-83.1	4.3	1.29
AI17	dimer	$P2_1/c$	3.731	11.929	14.781	108.4	90.0	90.0	2.656 (172)	3.688 (137)	-81.8	5.6	1.30
AB3	dimer	$P\bar{1}$	3.931	7.801	10.967	73.8	87.8	79.5	2.647 (169) 3.633 (96)	3.642 (138)	-81.1	6.3	1.28
AF5	chain	$P2_1$	3.839	7.224	11.478	106.0	90.0	90.0	2.749 (151) 3.944 (119)	-	-79.4	8.0	1.33

Table 7.6.2 Low energy crystal structures for benzoic acid

The experimental crystal structure for benzoic acid is contrasted with lattice energy minima calculated using the same intermolecular potential. All minimisations used the ab initio optimised molecular structure, except Min_{expt}, where the experimental molecular structure was used (X-ray bondlengths to hydrogen atoms were standardised). The lowest energy crystal structures found in the MOLPAK/DMAREL search are designed by the MOLPAK coordination group and number, and the space group is that of the conventional setting of each structure. The Niggli reduced cell parameters are tabulated here to aid comparison.

a The hydrogen bonds were characterised using PLUTO, tabulating the O \cdots O distance (< 4 Å) and O-H \cdots O angle (> 90°).

b ΔU_{latt} is the energy above the global minimum found in the search.

		Motif	Space group	Reduced cell parameters					Hydrogen bonds		U _{latt} / kJ mol ⁻¹	ΔU _{latt} / kJ mol ⁻¹	Density / g cm ⁻³
				a/Å	b/Å	c/Å	α°	β°	γ°	C=O ⁻ (H)-O / Å (°)			
α-form													
Expt	dimer	<i>P</i> $\bar{1}$	5.099	7.185	7.226	113.2	96.1	109.2	2.649 (168)	3.485 (136)	-77.2	-	
Min _{expt}	dimer	<i>P</i> $\bar{1}$	5.444	6.590	7.082	111.9	95.0	112.2	2.609 (172)	3.516 (154)	-79.4	-	
Min _{opt}	dimer	<i>P</i> $\bar{1}$	5.313	6.690	7.167	112.4	98.2	108.1	2.675 (172)	3.688 (137)	-67.3	4.7	1.31
CA24	dimer	<i>P</i> $\bar{1}$	5.313	6.690	7.167	112.4	98.2	108.1	2.675 (172)	3.689 (137)	-67.4	4.6	1.31
β-form													
Expt	chain	<i>P</i> 2 ₁	3.937	7.121	7.887	90.0	100.2	90.0	2.655 (164) 3.861 (93)	3.937 (100)	-74.2	-	1.28
Min _{expt}	chain	<i>P</i> 2 ₁	3.999	6.997	7.703	90.0	100.5	90.0	2.680 (169) 3.941 (91)	3.999 (99)	-75.5	-	1.32
Mint _{opt}	chain	<i>P</i> 2 ₁	3.922	6.945	7.680	90.0	101.9	90.0	2.642 (169) 3.904 (92)		-72.0	0	1.34
AF7	chain	<i>P</i> 2 ₁	4.159	7.073	7.507	90.0	104.1	90.0	2.690 (159)	-	-70.6	1.4	1.30
Hypothetical structures which have not been experimentally characterised													
AK5	chain	<i>P</i> 2 ₁ / <i>c</i>	7.143	7.824	8.098	108.4	90.0	90.0	2.682 (158)	3.002 (98)	-71.6	0.4	1.30
CA3	dimer	<i>P</i> $\bar{1}$	4.116	7.420	7.680	71.1	78.2	79.5	2.679 (169)	3.698 (137)	-71.2	0.8	1.30
CA22	dimer	<i>P</i> $\bar{1}$	3.867	7.584	7.738	76.0	82.5	82.1	2.660 (170) 3.215 (95) 3.993 (104)	3.667 (137)	-71.0	1.0	1.29
CA17	dimer	<i>P</i> $\bar{1}$	4.157	7.302	7.648	72.5	81.1	80.8	2.688 (168)	2.832 (98) 3.716 (136)	-70.2	1.8	1.29
FA6	dimer	<i>P</i> 2 ₁ / <i>c</i>	3.843	7.670	15.295	90.0	90.0	103.1	2.673 (169) 3.614 (98) 3.676 (102)	3.689 (137) 3.843 (94)	-70.0	2.0	1.27
CA10	dimer	<i>P</i> $\bar{1}$	3.854	7.802	7.819	73.0	78.3	84.6	2.672 (168) 3.623 (100) 3.701 (100)	3.692 (136) 3.853 (92)	-69.8	2.2	1.27
AQ3	chain	<i>P</i> 2 ₁ 2 ₁ 2 ₁	6.334	6.943	9.656	90.0	90.0	90.0	2.736 (166)		-69.5	2.5	1.31
AK7	chain	<i>P</i> 2 ₁ / <i>c</i>	6.362	6.949	9.630	90.0	93.1	90.0	2.735 (166)		-69.4	2.6	1.31
FC7	chain	<i>P</i> 2 ₁ / <i>c</i>	4.144	7.090	14.628	90.0	92.2	90.0	2.705 (160)		-69.4	2.6	1.30
CC15 ^c	chain	<i>P</i> 2 ₁ / <i>c</i>	7.142	7.754	16.035	93.2	90.0	90.0	2.675 (157) 2.679 (157) 3.873 (90)	2.983 (98) 3.013 (99)	-69.2	2.8	1.26
FC9	chain	<i>P</i> 2 ₁ / <i>c</i>	4.111	7.053	15.078	90.0	93.9	90.0	2.695 (162)		-69.0	3.0	1.28
AF24	chain	<i>P</i> 2 ₁	4.129	7.033	7.746	90.0	101.1	90.0	2.683 (162)		-68.4	3.6	1.27

Table 7.6.3 (continued)

	Motif	Space group	Reduced cell parameters						Hydrogen bonds		$U_{\text{latt}} / \text{kJ mol}^{-1}$	$\Delta U_{\text{latt}} / \text{kJ mol}^{-1}$	Density / g cm^{-3}
			$a/\text{\AA}$	$b/\text{\AA}$	$c/\text{\AA}$	$\alpha/^\circ$	$\beta/^\circ$	$\gamma/^\circ$	$\text{C}=\text{O}^{\cdots}(\text{H})-\text{O} / \text{\AA} (^\circ)$	$\text{O}-(\text{H})^{\cdots}\text{O}-\text{H} / \text{\AA} (^\circ)$			
AZ23	chain	$P2_12_12_1$	3.905	7.016	15.687	90.0	90.0	90.0	2.701 (169)		-68.4	3.6	1.30
FC3	dimer	$P2_1/c$	3.867	8.132	14.189	90.0	97.3	90.0	2.653 (169) 3.243 (96) 3.969 (103)	3.627 (138)	-68.4	3.6	1.26
FC22	dimer	$P2_1/c$	4.020	7.301	15.173	90.0	93.2	90.0	2.652 (170)	3.353 (106) 3.645 (137)	-68.2	3.8	1.26
FC8	chain	$P2_1/c$	3.900	7.071	15.569	93.3	90.0	90.0	2.702 (169)		-68.1	3.9	1.30
FA20	chain	$P2_1/c$	3.904	5.930	18.460	90.0	90.0	90.2	2.704 (160) 3.540 (94) 3.888 (105)		-68.0	4.0	1.31
BD6	chain	$Pna2_1$	3.895	5.959	18.583	90.0	90.0	90.0	2.702 (160) 3.494 (92) 3.927 (107)		-67.7	4.3	1.29
AM25	dimer	$P2_1/c$	4.314	7.386	13.973	90.0	90.0	99.3	2.668 (171)	3.683 (137)	-67.7	4.3	1.27
AZ8	chain	$P2_12_12_1$	4.073	6.991	15.620	90.0	90.0	90.0	2.675 (162)		-67.6	4.4	1.26
AM4	dimer	$P2_1/c$	5.896	6.961	10.523	90.0	90.0	91.1	2.684 (172)	3.694 (137)	-67.6	4.4	1.29
AI4	chain	$P2_1/c$	7.013	8.019	8.816	113.4	90.0	90.0	2.673 (160)	3.001 (97)	-67.6	4.4	1.23
AZ24	chain	$P2_12_12_1$	4.045	6.557	16.354	90.0	90.0	90.0	2.702 (171) 2.997 (91)		-67.5	4.5	1.29
BD18	chain	$Pna2_1$	4.005	6.775	16.416	90.0	90.0	90.0	2.681 (167)		-67.2	4.8	1.25
AK14	chain	$P2_1/c$	6.790	7.033	9.406	90.0	91.82	90.0	2.702 (162)		-67.1	4.9	1.24
BD12	chain	$Pna2_1$	4.089	6.993	5.859	90.0	90.0	90.0	2.671 (160)		-67.1	4.9	1.23
CC4	chain	$Pbca$	6.981	8.009	15.661	90.0	90.0	90.0	2.697 (155)		-67.1	4.9	1.28
BD25	chain	$Pna2_1$	3.886	6.901	16.373	90.0	90.0	90.0	2.696 (169)		-67.0	5.0	1.27
CA2	dimer	$P\bar{1}$	3.867	7.584	7.738	76.0	82.5	82.1	2.680 (173)	3.705 (137)	-66.9	5.1	1.27
FC16	dimer	$P2_1/c$	4.034	7.311	15.249	94.2	90.0	90.0	2.653 (169)	3.293 (108) 3.650 (137)	-66.9	5.1	1.25
BD9	chain	$Pna2_1$	3.917	5.796	19.251	90.0	90.0	90.0	2.681 (162) 3.497 (93) 3.931 (105)		-66.9	5.1	1.28
AV9	chain	$Pna2_1$	6.213	6.906	10.285	90.0	90.0	90.0	2.749 (170)		-66.8	5.2	1.27
FC17	chain	$P2_1/c$	4.060	6.663	16.117	92.3	90.0	90.0	2.704 (172) 2.986 (91)		-66.7	5.3	1.28

Table 7.6.3 (continued)

	Motif	Space group	Reduced cell parameters						Hydrogen bonds		$U_{\text{latt}} / \text{kJ mol}^{-1}$	$\Delta U_{\text{latt}} / \text{kJ mol}^{-1}$	Density / g cm^{-3}
			a/Å	b/Å	c/Å	α°	β°	γ°	C=O \cdots (H)-O / Å ($^\circ$)	O-(H) \cdots O-H / Å ($^\circ$)			
AK10	dimer	$P2_1/c$	7.429	7.618	7.824	90.0	90.0	91.8	2.667 (168)	3.638 (138)	-66.7	5.3	1.26
AK4	dimer	$P2_1/c$	7.477	7.664	7.710	90.0	90.0	93.4	2.656 (169)	3.630 (138)	-66.7	5.3	1.27
AK11	dimer	$P2_1/c$	7.468	7.338	7.796	97.4	90.0	90.0	2.659 (168)	3.625 (138)	-66.6	5.4	1.28
FC13	dimer	$P2_1/c$	4.011	7.203	15.355	95.5	90.0	90.0	2.647 (169)	3.502 (106)	-66.5	5.5	1.26

Table 7.6.3 Low energy crystal structures for tetrolic acid

The experimental crystal structures for tetrolic acid are contrasted with lattice energy minima calculated using the same intermolecular potential. All minimisations used the *ab initio* optimised molecular structure, except Min_{expt} , where the experimental molecular structures were used (X-ray bondlengths to hydrogen atoms were standardised). The lowest energy crystal structures found in the MOLPAK/DMAREL search are designed by the MOLPAK coordination group and number, and the space group is that of the conventional setting of each structure. The Niggli reduced cell parameters are tabulated here to aid comparison.

- a The hydrogen bonds were characterised using PLUTO (ref), tabulating the O \cdots O distance ($< 4 \text{ \AA}$) and O-H \cdots O angle ($> 90^\circ$).
- b ΔU_{latt} is the energy above the global minimum found in the search.
- c This $P2_1/c$, $Z' = 2$ structure was found as a true minimum from a *Pbca* starting structure.

	C ₁₁	C ₂₂	C ₃₃	C ₄₄	C ₅₅	C ₆₆	C ₁₂	C ₁₃	C ₂₃	C ₁₅	C ₂₅	C ₃₅	C ₄₆	C ₁₄	C ₁₆	C ₂₄	C ₂₆	C ₃₄	C ₃₆	C ₄₅	C ₅₆
Min _{expt}	18.6	14.9	18.2	8.8	4.5	3.2	4.8	6.0	9.5												
Min _{opt}	19.3	14.9	17.4	8.5	4.4	3.1	5.0	5.9	8.9												
AV21	21.6	11.9	12.8	9.0	3.8	3.5	7.9	5.3	9.9												
AI2	20.1	31.7	12.5	4.9	7.2	6.5	3.4	11.8	6.4	-5.9	1.8	-1.2	-2.2								
AI19	21.0	27.2	10.9	5.4	6.3	5.5	5.9	9.2	8.5	6.0	-0.9	0.2	3.4								
AQ23	28.7	14.7	16.0	2.2	5.1	2.2	3.5	6.9	4.3												
AQ15	11.3	12.2	25.9	6.8	5.7	4.8	10.8	7.4	2.0												
FA15	18.0	22.1	25.4	10.4	2.2	0.9	2.2	4.8	5.4	-0.2	1.1	0.2	0.2								
CC22	27.5	13.5	17.5	2.9	7.2	3.9	7.2	6.0	8.7												
AK4	15.4	12.4	13.6	4.3	9.8	6.1	8.7	9.3	6.0	2.5	-3.8	4.2	-2.3								
CC5	16.1	17.2	17.9	8.0	10.2	7.6	7.6	10.4	5.3												
AK22	20.2	20.5	12.7	2.1	6.9	9.5	6.7	10.6	4.2	-5.0	-1.5	0.4	-3.5								
AQ7	14.9	18.3	24.7	11.7	1.7	1.4	6.0	4.8	5.1												
AF17	13.1	27.0	19.3	11.8	8.8	3.3	4.6	11.0	8.6	-0.6	-3.6	-6.1	-5.6								
AI4	21.8	25.7	14.1	4.9	6.8	8.8	8.4	8.3	5.5	-7.1	-4.1	0.3	-5.2								
AM20	13.6	26.1	11.1	4.3	4.6	3.8	5.3	6.7	7.2	2.1	-2.6	-0.9	-0.9	0.1	0.2	-2.0	3.7	1.3	-0.3	0.8	-1.1
AK23	18.0	20.9	19.2	7.0	5.5	10.1	11.6	4.9	9.4	-0.7	0.9	-0.8	1.5								
AK5	19.7	10.7	14.3	2.9	6.2	6.3	8.9	7.9	5.6	3.7	1.0	2.2	1.6								
CC7	14.0	25.2	17.8	2.9	2.2	7.7	4.8	6.2	7.5	2.3	-0.3	1.5	0.5								
FC3	9.9	10.8	23.1	2.9	9.2	4.4	6.6	8.8	5.5	-2.0	-1.6	5.7	0.3								
AB5	18.9	11.6	19.5	7.7	8.7	4.6	5.1	7.7	10.6	-2.6	-0.7	4.0	-1.1	-2.6	-5.0	0.3	-0.2	4.5	1.0	0.2	3.2
FA14	12.5	13.8	26.0	4.2	4.3	3.9	9.2	5.9	7.3	2.7	1.0	-2.9	0.3								
CA8	15.9	11.6	21.7	4.7	7.5	3.5	5.6	7.3	7.3	-6.5	1.2	1.1	0.9	0.2	0.8	-0.7	0.1	-4.8	-0.8	0.9	0.1
AM17	8.2	24.0	10.2	1.3	3.9	2.5	5.6	6.4	2.9	0.7	1.3	-0.4	-1.1								
AZ9	11.5	10.5	22.8	5.3	6.2	6.0	6.9	8.4	5.5												
CB3	9.3	12.2	17.8	6.0	5.1	2.6	5.4	7.5	6.2												
FA3	14.2	16.3	10.3	6.5	1.8	5.8	7.6	4.7	6.6												
AM7	18.5	19.9	10.6	1.4	2.5	6.0	4.3	5.1	8.6												
CB19	14.1	32.3	9.9	3.8	2.9	21.3	16.7	6.3	9.2												
FC9	11.8	12.5	20.8	4.6	6.1	5.2	9.0	7.9	6.1												
CB8	11.8	6.1	11.6	6.0	5.0	0.8	5.4	7.8	5.4												
AZ7	13.3	11.2	10.1	5.3	5.5	8.4	7.5	9.0	8.2												
BD6	14.0	11.2	11.3	4.4	5.1	8.4	8.1	8.0	8.3												

Table 7.7.1 Calculated elastic properties (in GPa) for the minimised experimental structure (Min_{expt} and Min_{opt}) and for the energetically feasible, hypothetical structures for formic acid.

In all calculations, but for Min_{expt}, the *ab initio* optimised molecular model is used. The optical axes are defined with x parallel to a, y to b and z to c for the orthorhombic crystals (Min_{expt}, Min_{opt}, AV21, AQ23, AQ15, CC22, CC5, AQ7, CC7, AZ9, CB3, CB19, CB8, AZ7, BD6) and z parallel to c, x parallel to a* and y perpendicular to xz for the monoclinic crystals (AI2, AI19, FA15, AK4, AK22, AI4, AM20, AK23, AK5, FC3, FA14, AM17, FA3, AM7, FC9) and the triclinic systems (AF17, AB5, CA8).

	C ₁₁	C ₂₂	C ₃₃	C ₄₄	C ₅₅	C ₆₆	C ₁₂	C ₁₃	C ₂₃	C ₁₅	C ₂₅	C ₃₅	C ₄₆	C ₁₄	C ₁₆	C ₂₄	C ₂₆	C ₃₄	C ₃₆	C ₄₅	C ₅₆
Min _{expt}	11.0	12.2	16.4	10.4	5.1	3.3	9.2	10.4	9.9	-1.3	1.2	-2.7	2.6								
Min _{opt}	10.5	11.9	16.2	10.3	5.1	3.3	8.4	9.8	8.9	-0.8	1.6	-3.2	3.1								
FC22	11.9	12.1	27.7	9.0	6.2	4.5	8.4	12.2	4.7	-1.5	3.2	-11.2	3.1								
AI25	12.3	13.1	17.2	1.7	7.7	2.6	6.3	11.7	6.3	2.3	-1.9	4.3	-2.1								
FA16	11.2	10.8	11.1	8.3	1.6	1.7	6.9	6.2	9.3	0.3	-1.2	-1.1	-2.2								
AM17	10.6	11.0	12.8	4.1	4.2	3.3	10.1	9.1	8.8	-1.3	-1.6	-3.1	0.8								
AM24	11.3	11.9	13.1	6.7	4.0	3.0	10.0	8.3	7.5	0.9	1.4	2.3	0.9								
AM21	12.6	22.8	13.9	1.7	2.8	1.0	2.6	6.4	8.7	-2.1	2.7	2.4	0.8								
AB9	11.1	16.7	11.2	5.2	6.0	4.2	7.1	10.2	8.1	1.2	3.6	2.2	4.2								
DD21	18.0	14.5	11.7	10.8	2.7	1.3	0.7	2.8	12.2	-3.6	2.4	1.4	0.7								
AI17	14.2	12.1	11.7	1.6	8.8	2.6	6.6	10.7	5.8	-3.4	1.9	-3.6	2.0								
AB3	16.0	10.5	9.8	4.6	3.5	1.3	4.5	8.7	7.7	2.5	0.4	0.5	1.3	-4.4	-1.7	1.8	0.4	-1.7	-0.6	-0.6	-1.7
AF5	14.5	12.3	17.4	2.8	7.5	4.9	8.3	11.2	5.6	2.3	-2.6	7.1	-0.7								

Table 7.7.2: Calculated elastic properties (in GPa) for the minimised experimental structure (Min_{expt} and Min_{opt}) and for the energetically feasible, hypothetical structures for benzoic acid.

In all calculations, but for Min_{expt}, the *ab initio* optimised molecular model is used. The optical axes are defined with z parallel to c, x parallel to a* and y perpendicular to xz for the monoclinic crystals (Min_{expt}, Min_{opt}, FC22, AI25, FA16, AM17, AM24, AM21, DD21, AI17, AF5) and the triclinic systems (AB9, AB3).

	C ₁₁	C ₂₂	C ₃₃	C ₄₄	C ₅₅	C ₆₆	C ₁₂	C ₁₃	C ₂₃	C ₁₅	C ₂₅	C ₃₅	C ₄₆	C ₁₄	C ₁₆	C ₂₄	C ₂₆	C ₃₄	C ₃₆	C ₄₅	C ₅₆
α-form																					
Min _{expt}	19.4	20.2	13.6	2.5	2.6	20.6	17.9	3.4	5.2	-1.0	-1.2	-0.2	-0.2	-0.1	-12.1	1.0	-12.8	0.002	1.2	1.7	0.5
Min _{opt}	16.4	18.7	13.1	2.4	2.3	17.0	15.6	3.5	4.8	-1.4	-1.3	-0.2	-0.5	0.3	-9.1	0.3	-10.3	-0.2	0.7	1.3	0.8
CA24	16.4	18.7	13.1	2.4	2.3	17.0	15.6	3.5	4.8	-1.4	1.3	0.2	0.5	-0.3	-9.1	-0.3	-10.3	0.2	0.7	1.3	-0.8
β-form																					
Min _{expt}	13.3	22.2	13.0	7.1	5.7	12.7	7.5	6.7	4.9	-0.5	-3.6	0.2	-7.6								
Min _{opt}	13.6	21.8	12.8	7.2	5.9	12.9	8.3	7.6	5.2	-1.0	-3.9	0.1	-7.8								
AF7	13.4	23.3	14.8	16.6	4.8	5.8	4.0	5.6	10.7	-1.2	3.0	2.0	7.1								
AK5	12.8	27.7	12.0	8.5	5.9	13.3	9.7	6.8	7.3	1.5	5.5	-1.1	8.4								
CA3	18.1	16.1	15.6	1.9	4.6	12.7	10.8	6.2	3.2	-0.1	1.1	-1.8	0.5	-1.2	-6.5	-0.6	-5.4	0.5	1.1	1.5	-1.1
CA22	20.8	15.2	14.6	3.0	2.5	2.8	4.1	7.7	5.1	4.1	0.3	-0.1	-0.4	0.9	-2.6	1.1	0.3	-0.6	-0.1	1.3	-1.2
CA17	33.3	13.0	14.9	5.5	1.3	8.0	5.8	2.1	6.7	2.5	-0.2	-0.9	0.007	0.7	10.3	-0.5	-0.6	-3.0	-0.8	-0.6	1.2
FA6	13.4	15.9	14.2	7.2	2.8	1.1	4.1	3.4	11.4	-1.0	-1.9	-1.5	0.2								
CA10	13.1	18.6	16.3	1.4	3.8	2.1	5.0	4.7	6.4	0.4	2.5	0.1	-0.1	-0.2	1.1	-1.9	0.1	0.9	0.8	-0.7	1.0
AQ3	12.9	17.3	30.4	24.4	3.3	0.6	3.4	1.8	16.9												
AK7	11.1	29.8	11.6	18.1	4.4	9.0	5.5	7.3	12.4	-1.2	6.1	3.2	9.4								
FC7	12.1	23.6	11.4	10.5	5.5	9.4	5.8	5.6	8.0	0.1	-4.4	0.05	-8.3								
CC15	11.1	21.6	12.5	0.6	3.3	10.9	7.3	8.3	5.8	-1.1	-0.04	0.5	0.009								
FC9	11.6	20.7	9.2	9.8	3.8	8.5	5.9	6.5	9.6	0.6	-2.5	0.4	-8.0								
AF24	12.0	18.7	12.1	14.3	3.2	4.5	4.1	5.7	11.5	-1.1	1.5	2.4	6.7								
AZ23	13.2	11.7	14.1	6.9	4.2	8.8	9.1	5.8	6.1												
FC3	16.3	7.2	14.3	3.7	5.6	3.2	3.8	5.7	7.8	-0.6	-1.7	-1.4	-2.2								
FC22	12.5	11.1	14.8	6.6	4.5	2.4	4.7	5.6	6.0	0.4	-2.7	2.9	-3.4								
FC8	15.9	14.3	13.4	2.9	4.0	6.1	4.2	9.3	5.1	3.9	-1.1	0.7	-0.7								
FA20	11.4	14.1	25.4	7.1	4.6	1.4	6.1	7.0	10.4	-0.7	2.3	-0.2	-1.0								
BD6	12.9	13.6	19.2	5.7	10.3	3.2	5.6	11.4	9.1												
AM25	10.2	12.3	14.4	1.2	5.3	6.3	8.2	5.6	2.4	1.1	-1.2	0.9	1.3								
AZ8	11.1	17.4	10.7	2.0	0.4	7.0	10.8	5.7	5.0												
AM4	15.6	14.4	12.1	1.9	2.0	16.2	13.6	4.7	5.2	1.5	0.5	-1.2	2.9								
AI4	13.2	15.1	12.2	5.2	1.0	9.2	4.5	4.3	6.4	-1.6	-1.6	0.4	4.5								
AZ24	12.8	12.8	9.3	5.6	9.0	4.3	6.9	9.0	6.6												
BD18	13.6	11.3	13.8	2.0	8.0	1.7	5.7	9.0	3.4												
AK14	8.1	16.5	9.1	13.7	6.0	7.0	5.3	7.0	7.3	-1.3	3.2	1.5	8.6								
BD12	15.4	10.9	18.1	3.0	0.2	0.03	3.8	7.3	4.2												
CC4	17.2	12.0	12.6	5.2	11.2	2.2	7.9	5.9	6.8												
BD25	14.1	10.3	13.9	6.6	0.5	4.8	9.9	4.7	4.7												
CA2	13.6	22.5	10.8	6.6	5.3	15.8	14.9	6.0	7.8	3.4	7.0	-1.5	5.5	-4.4	-6.9	-3.5	-13.0	1.1	-1.3	-2.8	-7.4

Table 7.7.3 (continued)

	C ₁₁	C ₂₂	C ₃₃	C ₄₄	C ₅₅	C ₆₆	C ₁₂	C ₁₃	C ₂₃	C ₁₅	C ₂₅	C ₃₅	C ₄₆	C ₁₄	C ₁₆	C ₂₄	C ₂₆	C ₃₄	C ₃₆	C ₄₅	C ₅₆
FC16	9.5	12.0	12.2	2.6	2.9	0.5	3.9	7.0	5.1	-1.7	0.5	-1.5	0.6								
BD9	13.8	17.3	14.5	4.8	1.9	4.8	10.8	4.7	7.5												
AV9	14.1	11.2	10.0	17.3	0.4	1.4	4.8	1.3	7.3												
FC17	16.9	12.7	13.4	3.6	6.9	5.9	5.7	11.0	6.4	7.2	-0.7	2.7	0.7								
AK10	13.0	15.0	10.7	3.7	5.6	7.5	8.6	5.0	3.7	-3.1	-5.3	2.0	-4.1								
AK4	15.0	12.6	12.2	3.5	5.7	6.7	7.9	5.8	5.1	-2.6	-5.7	1.3	-4.1								
AK11	14.1	12.8	11.3	5.3	7.4	8.7	8.1	5.7	5.0	2.3	7.8	-0.5	6.0								
FC13	10.1	13.1	12.5	1.0	3.2	0.2	2.9	7.7	4.5	-2.8	0.6	-1.0	0.4								

Table 7.7.3

Calculated elastic properties (in GPa) for the minimised experimental structure

(Min_{expt} and Min_{opt}) and for the energetically feasible, hypothetical structures for tetrolic acid.

In all calculations, but for Min_{expt}, the *ab initio* optimised molecular model is used. The optical axes are defined with x parallel to a, y to b and z to c for the orthorhombic crystals (AQ3, CC15, AZ23, BD6, AV1, AZ8, AZ24, BD18, BD12, CC4, BD25, BD9, AV9,) and z parallel to c, x parallel to a* and y perpendicular to xz for the monoclinic crystals (AK5, FA6, AK7, FC7, FC9, FC3, FC22, FC8, FA20, AM25, AM4, AI4, AK14, FC16, FC17, AK10, AK4, AK11, FC13) and the triclinic systems (Min_{expt}(α), Min_{opt}(α), CA24, Min_{expt}(β), Min_{opt}(β), AF7, CA3, CA22, CA17, CA10, AF24, CA2).

formic acid	aspect ratio		Slowest growing faces for AE model (frequency of occurrence)	Attachment Energy / kJ/mol	% ^a Area of face
	BFDH	AE			
expt	2.1	2.6	200 (2)	-3.7	39.9
min _{expt}	2.1	2.9	200 (2)	-3.8	43.2
min _{opt}	2.1	2.8	200 (2)	-3.7	42.9
AV21	2.0	2.4	200 (2)	-3.8	40.5
AI2	2.2	2.3	100 (2)	-3.5	44.7
AI19	2.3	2.7	100 (2)	-3.1	44.7
AQ23	1.3	2.2	011 (4)	-4.6	57.9
AQ15	1.5	1.7	110 (4)	-5.1	52.8
FA15	2.7	4.4	020 (2)	-2.5	56.0
CC22	2.3	2.6	002 (2)	-3.3	43.5
AK4	1.8	1.7	100 (2)	-4.0	29.8
CC5	2.1	2.3	002 (2)	-2.4	42.0
AK22	2.0	2.1	100 (2)	-3.2	38.6
AQ7	1.4	2.3	110 (4)	-4.2	69.2
AF17	2.3	3.7	001 (2)	-3.3	48.5
AI4	2.4	3.6	100 (2)	-2.8	50.9
AM20	2.2	2.2	001 (2)	-2.8	32.1
AK23	2.1	2.4	100 (2)	-3.4	40.8
AK5	2.1	1.8	100 (2)	-4.0	28.9
			<i>011 (4)</i>	-4.5	52.3
CC7	4.5	6.1	001 (2)	-2.8	69.2
FC3	2.0	1.8	101 (2)	-3.5	32.8
AB5	2.0	2.2	001 (2)	-3.5	33.0
FA14	2.3	2.5	020 (2)	-3.2	43.9
CA8	2.0	2.0	100 (2)	-3.5	32.6
			*001 (2)	-4.3	
AM17	1.7	2.2	101 (2)	-5.4	38.3
			*011 (4)	-6.0	
AZ9	1.5	1.9	101 (4)	-4.9	43.8
CB3	2.1	1.7	111 (4)	-5.0	63.8
FA3	2.0	1.6	020 (2)	-4.3	22.8
			<i>111 (4)</i>	-9.3	30.5
AM7	1.5	1.6	011 (4)	-4.0	52.1
CB19	1.7	2.0	200 (2)	-4.3	35.4
FC9	2.1	2.3	002 (2)	-4.0	36.4
CB8	1.7	1.6	020 (2)	-3.5	27.6
			111 (8)	-4.4	57.9
AZ7	2.0		200 (2)	-3.9	34.0
			<i>110 (4)</i>	-4.8	41.9
BD6	2.2	2.6	200 (2)	-3.9	33.3
			110 (4)	-4.8	40.5

Table 7.8.1 Growth rate and morphology predictions for the known and hypothetical forms of formic acid.

The morphologically dominant, slowest growing face, in the attachment energy (AE) prediction of the morphology is given in normal type, along with corresponding AE. For crystals where another face provides a higher percentage of the total surface area, because of a higher frequency of occurrence, this face is given in italics. The aspect ratio for the AE predicted morphology is contrasted with that using the BFDH model, and where the BFDH model predicts a different slowest growing face, this is given after *.

^a 100×(total surface area of slowest growing face)/total surface area

benzoic acid	aspect ratio		Slowest growing face for AE model (frequency of occurrence)	Attachment Energy / kJ/mol	% Area of face (+)
	BFDH	AE			
Expt	3.0	3.9	002 (2)	-3.0	53.4
Min_{expt}	2.9	4.0	002 (2)	-3.3	53.3
Min_{opt}	2.9	3.7	002 (2)	-3.4	51.6
FC22	2.7	3.9	002 (2)	-3.3	51.1
AI25	3.3	3.5	100 (2)	-3.1	53.9
FA16	1.8	3.4	020 (2)	-4.0	47.3
AM17	1.7	2.5	011 (4)	-5.1	65.5
AM24	2.0	2.6	011 (4)	-4.8	73.2
AM21	2.6	2.3	011 (4)	-4.6	75.6
AB9	2.0	3.8	001 (2)	-3.3	50.2
AI17	3.3	3.5	100 (2)	-3.0	54.3
AB3	3.0	3.7	001 (2)	-2.9	53.4
AF5	3.4	3.8	001 (2)	-3.2	54.8

Table 7.8.2: Growth rate and morphology predictions for the known and hypothetical forms of benzoic acid.

The aspect ratio for the AE predicted morphology is contrasted with that using the BDFH model.

tetrollic acid	aspect ratio		Slowest growing faces for AE model (frequency of occurrence)	Attachment Energy / kJ/mol	% ^a Area of face
	BFDH	AE			
α-form (TETROL)					
expt	1.6	2.9	001 (2)	-4.2	34.8
min _{expt}	1.6	2.9	001 (2)	-4.3	37.0
min _{opt}	1.6	2.9	001 (2)	-4.4	34.6
CA24	1.6	2.9	10-1 (2)	-4.4	34.6
β-form (TETROL01)					
expt	2.5	3.5	100 (2)	-3.3	50.0
min _{expt}	2.5	3.9	100 (2)	-3.3	53.0
min _{opt}	2.3	3.5	100 (2)	-3.4	50.7
AF7	2.4	3.4	001 (2)	-3.5	48.6
AK5	2.1	3.6	100 (2)	-3.0	52.5
CA3	2.0	2.7	001 (2)	-3.8	32.8
			*110 (2)	-13.1	
CA22	2.2	2.9	101 (2)	-3.4	38.4
CA17	2.0	2.7	101 (2)	-3.7	34.1
			*100 (2)	-6.5	
FA6	2.3	3.8	020 (2)	-3.1	47.9
CA10	2.2	3.0	01-1 (2)	-3.3	38.6
AQ3	1.3	2.2	110 (4)	-5.2	67.2
			*011 (4)	-8.9	
AK7	1.3	2.5	10-2 (2)	-5.0	35.5
			*011 (4)	-8.9	
FC7	2.1	3.0	002 (2)	-3.6	45.0
CC15	4.2	6.4	001 (2)	-1.4	71.6
FC9	2.2	3.1	002 (2)	-3.4	46.1
AF24	2.4	3.6	001 (2)	-3.2	50.9
AZ23	2.4	3.5	200 (2)	-3.3	54.0
FC3	2.1	2.8	002 (2)	-3.7	41.1
			*011 (4)	-7.1	
FC22	2.3	3.3	002 (2)	-3.2	46.4
FC8	2.4	3.4	002 (2)	-3.3	47.4
FA20	2.9	4.7	020 (2)	-2.7	58.7
BD6	3.0	4.3	200 (2)	-2.7	59.8
AM25	2.0	2.5	011 (4)	-4.0	65.4
			*020 (2)	-6.4	
AZ8	2.3	3.3	200 (2)	-3.2	53.2
AM4	1.6	2.0	011 (4)	-4.5	66.7
AI4	2.2	4.5	100 (2)	-2.2	58.2
AZ24	2.4	3.6	200 (2)	-3.1	56.2
BD18	2.6	3.9	200 (2)	-2.9	36.7
AK14	1.3	2.4	10-2 (2)	-4.1	35.2
BD12	2.5	4.1	200 (2)	-2.8	57.7
CC4	2.1	3.3	002 (2)	-3.2	47.6
BD25	2.6	3.8	200 (2)	-3.0	56.7
CA2	1.8	3.4	101 (2)	-3.5	42.8
FC16	2.3	2.9	002 (2)	-3.1	39.5
			*100 (2)	-6.1	
BD9	3.1	4.4	200 (2)	-2.5	62.5
AV9	1.5	2.3	110 (4)	-4.5	72.8
FC17	2.4	3.6	002 (2)	-3.1	50.5

Table 7.8.3

tetrolic acid	aspect ratio		Slowest growing faces for AE model (frequency of occurrence)	Attachment Energy / kJ/mol	% ^a Area of face
	BFDH	AE			
AK10	2.2	3.3	100 (2)	-3.1	45.0
AK4	2.2	3.3	100 (2)	-3.0	47.2
AK11	2.1	3.3	100 (2)	-3.2	47.4
FC13	2.3	2.8	002 (2)	-3.3	38.7

Table 7.8.3 (continued) Growth rate and morphology predictions for the known and hypothetical forms of tetrolic acid.

The morphologically dominant, slowest growing face, in the attachment energy (AE) prediction of the morphology is given along with corresponding AE. The aspect ratio for the AE predicted morphology is contrasted with that using the BFDH model, and where the BFDH model predicts a different slowest growing face, this is given after *.

^a $100 \times (\text{total surface area of slowest growing face}) / \text{total surface area}$

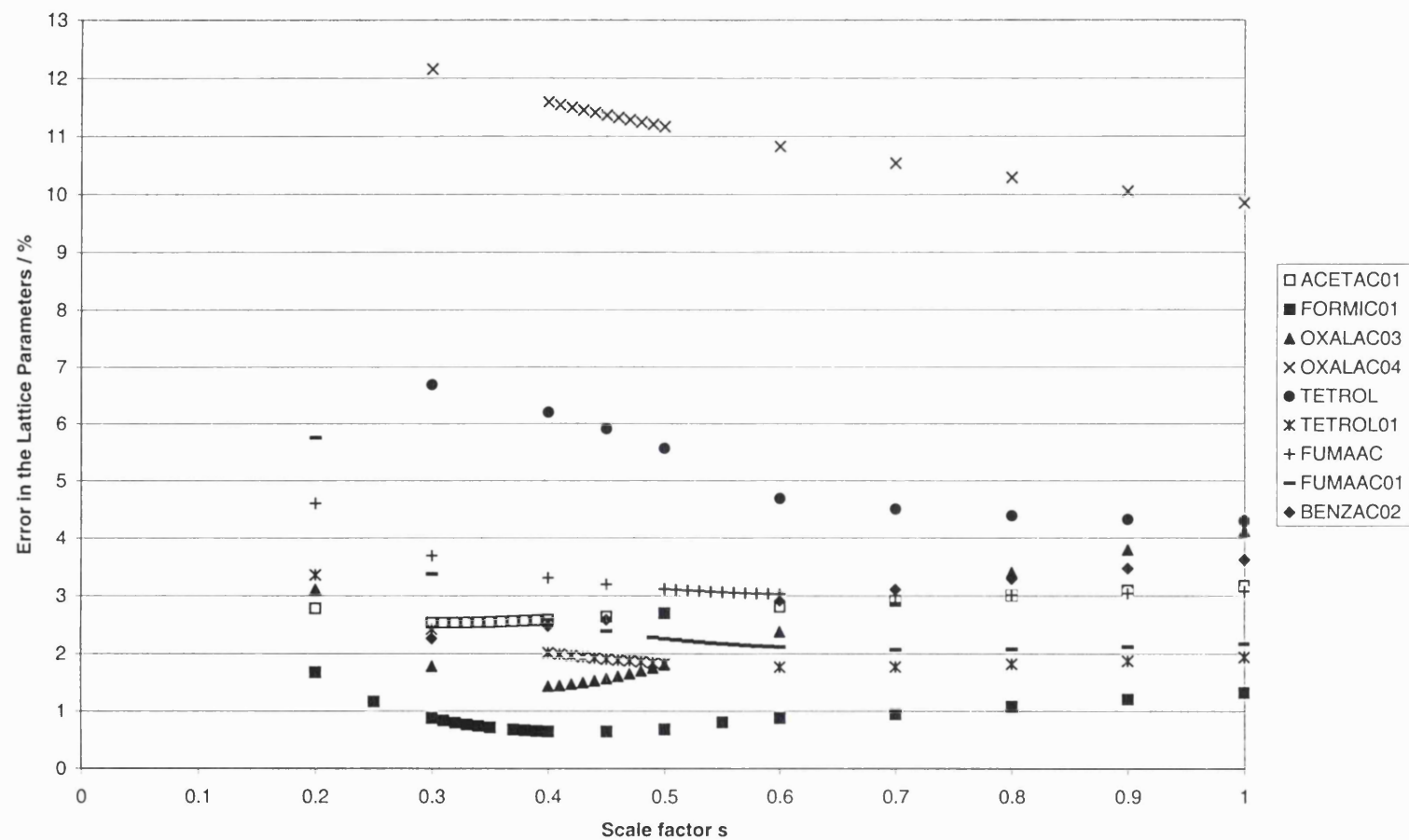


Figure 7.1a

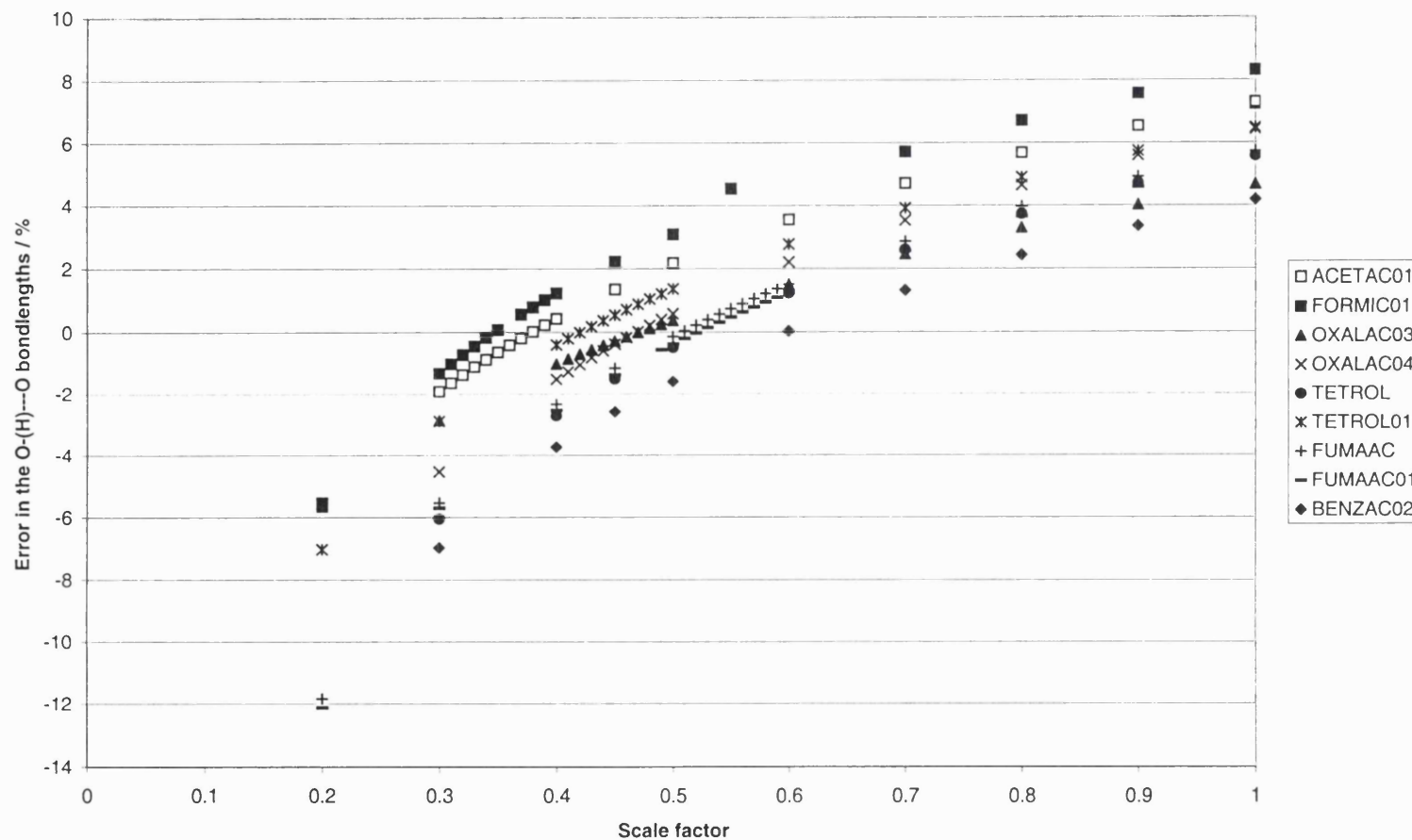


Figure 1b

Figure 1 R.m.s. percentage error in the (a) lattice parameters and in the (b) O-(H)...O bond lengths for variations in the polar hydrogen repulsion potential.

Shown as a function of the pre-exponential polar hydrogen scale factor s defined by $A_{\text{H}_0\text{H}_0} = s \times A_{\text{H}_\text{N}\text{H}_\text{N}}$.

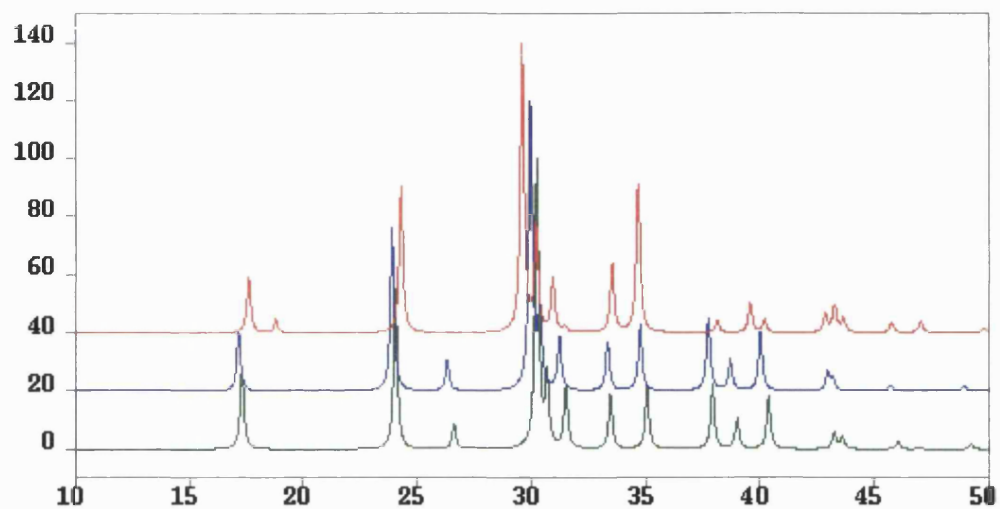


Figure 7.2a formic acid

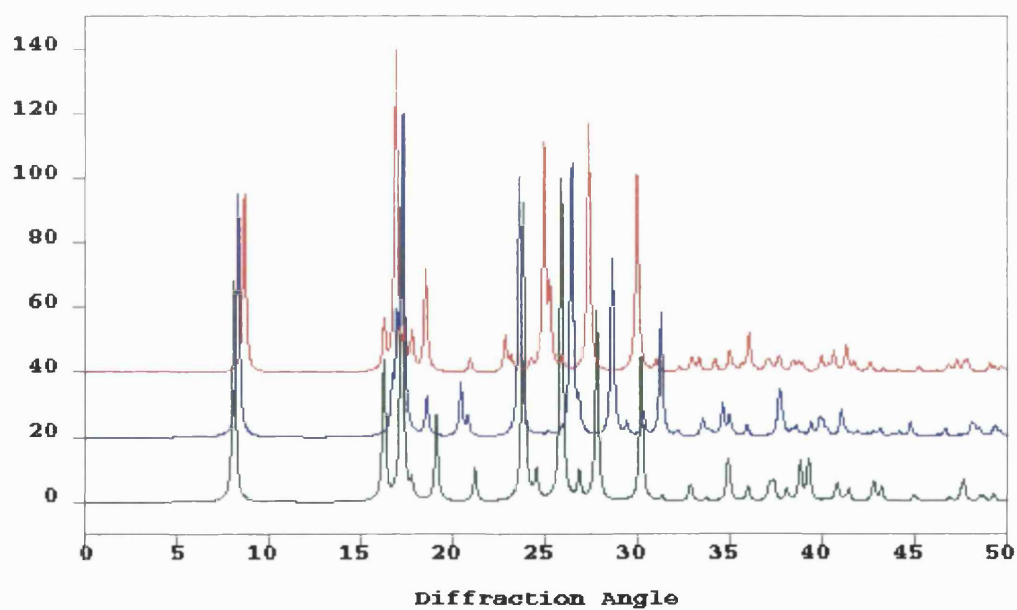


Figure 7.2b benzoic acid

(continued)

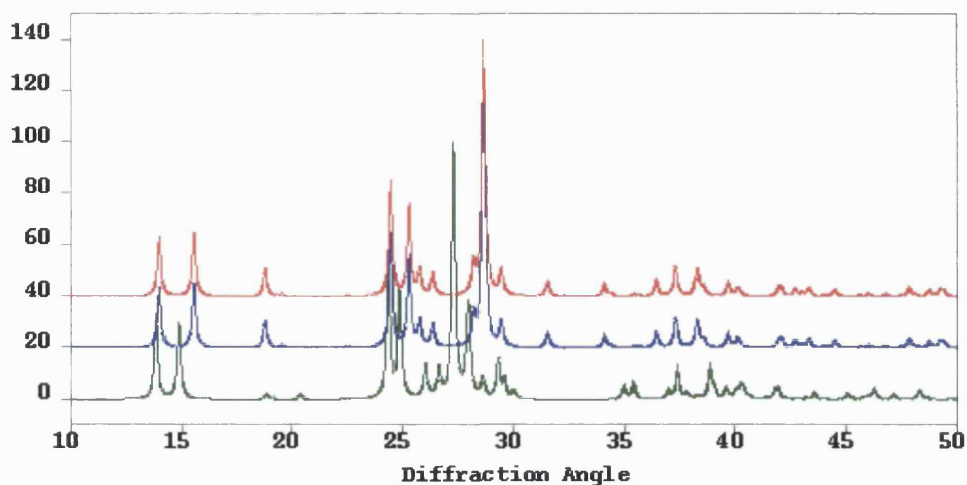
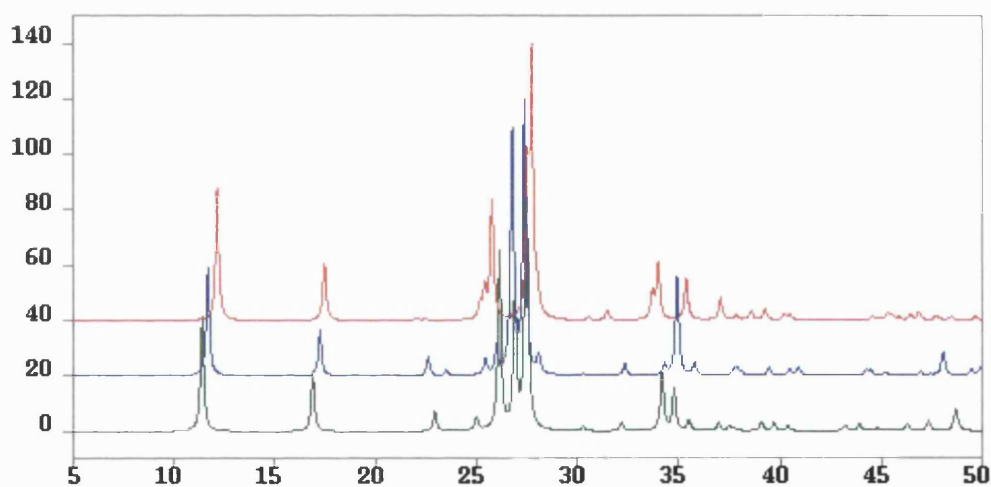
Figure 7.2c α -tetrolic acidFigure 7.2d β -tetrolic acid

Figure 7.2 Comparison of the simulated powder diffraction patterns for the experimental structure (bottom, green), the corresponding minimum in the lattice energy using the optimised molecular structure (middle, blue, offset 20), and the closest structure found in the search procedure shown in table 7.5 (top, red, offset 40) for (a) formic acid, (b) benzoic acid, (c) the α -form of tetrolic acid, and (d) the β -form of tetrolic acid.

The powder diffraction patterns were simulated for X-ray diffraction with $\lambda = 1.5418 \text{ \AA}$, using MSI Cerius2 software.⁴²

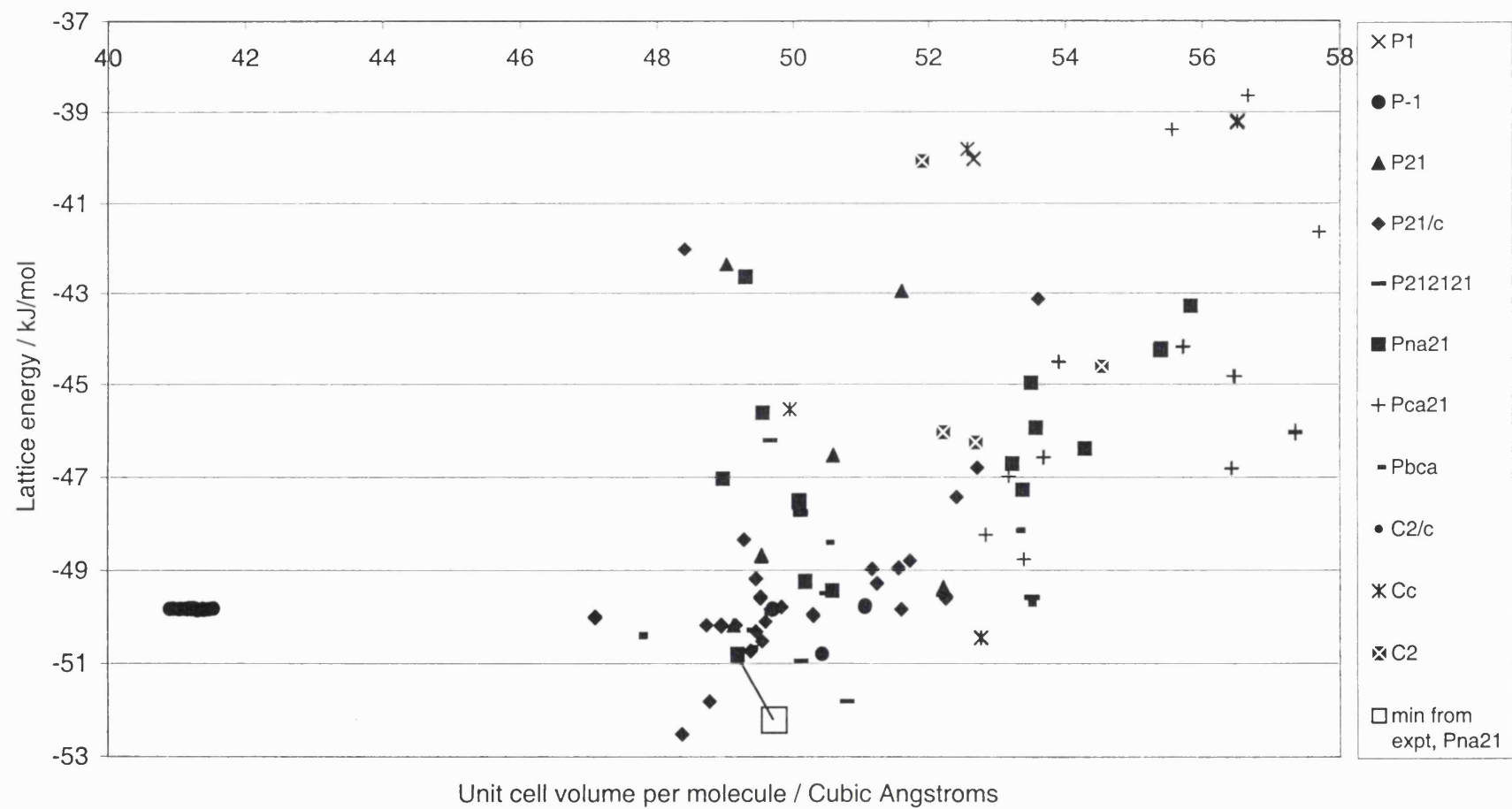


Figure 7.3a Formic acid

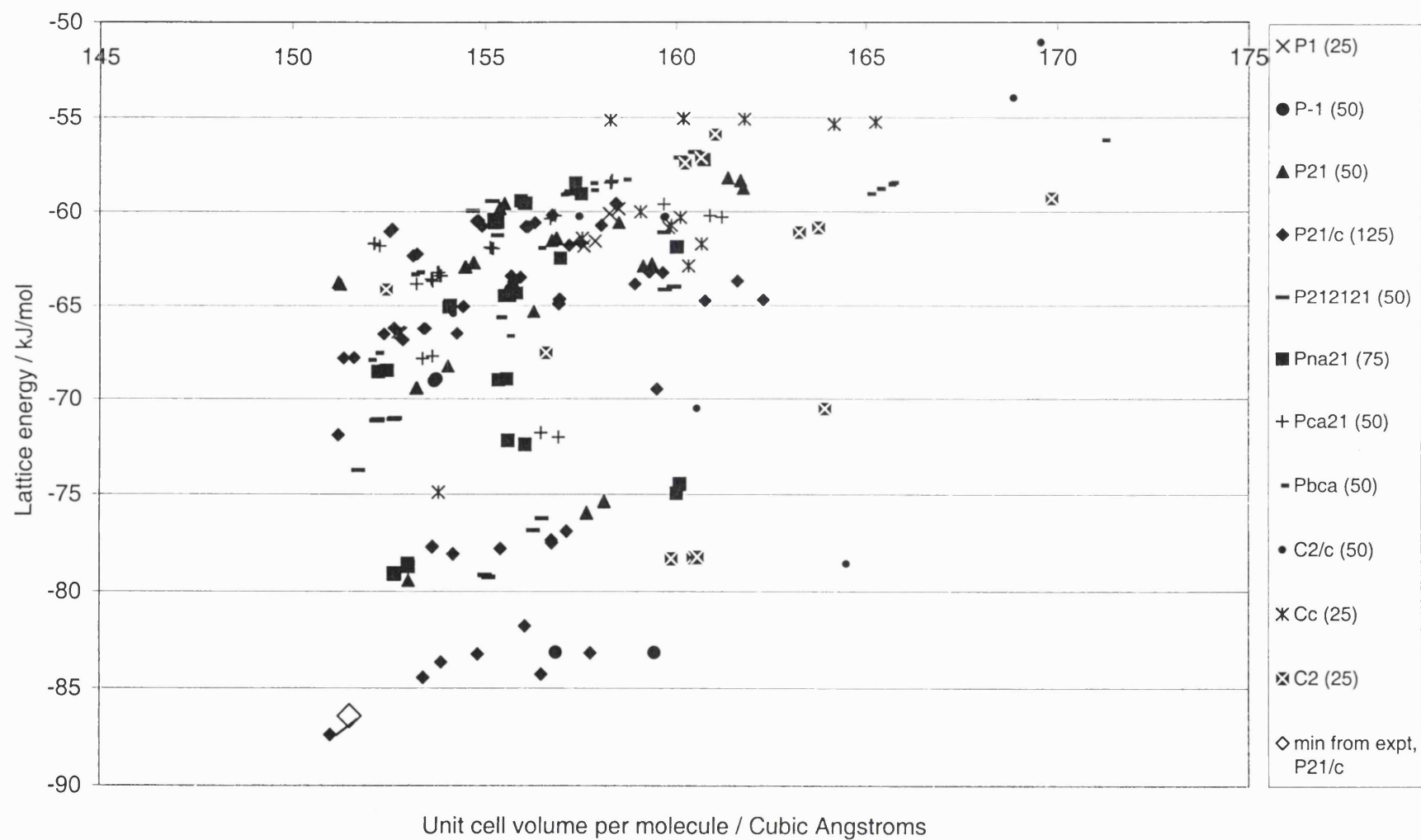


Figure 7.3b Benzoic acid

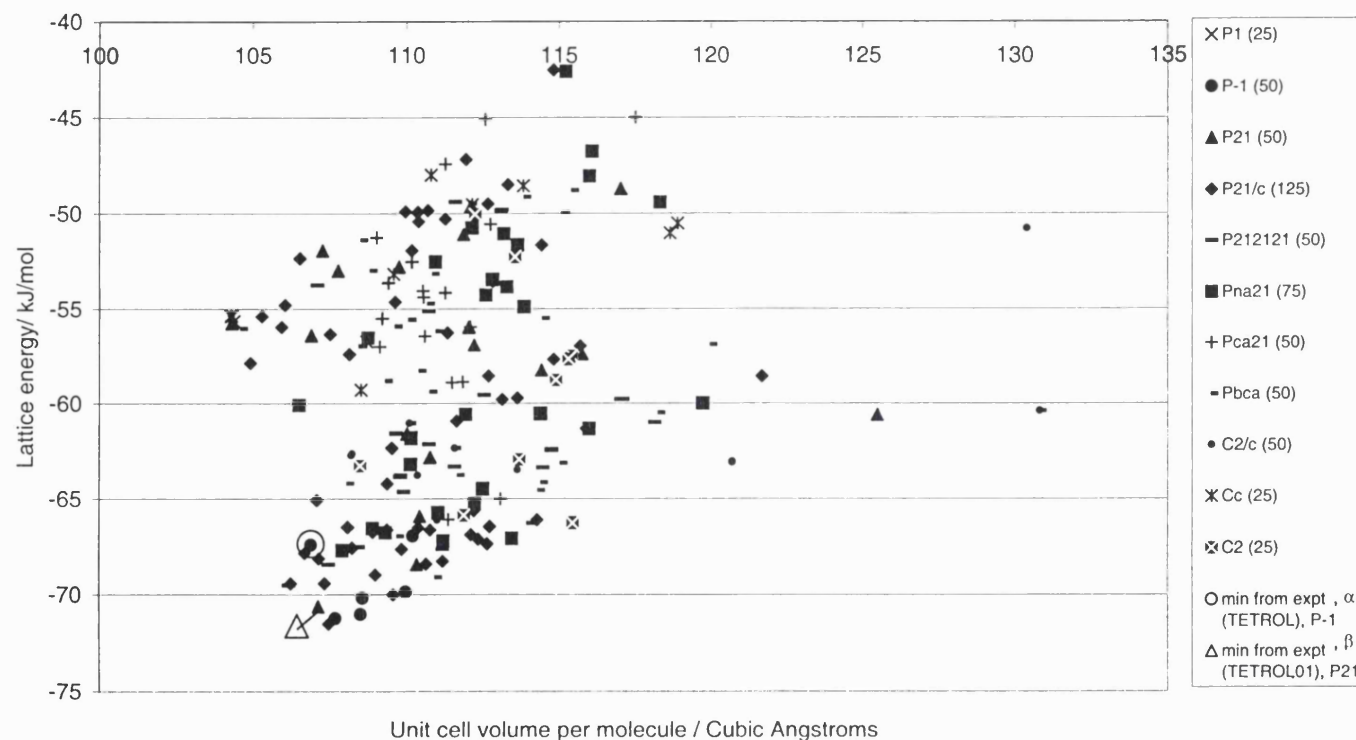


Figure 7.3c Tetrolic acid

Figure 7.3 Minima in lattice energy found by the MOLPAK/DMAREL search, using the DMA+6-exp potential and the *ab initio* molecular structure for (a) formic acid, (b) benzoic acid and (c) tetrolic acid.

The symbol denotes the space group of the MOLPAK starting structure and, in the legend, is followed by the number of minimisations performed for that space group. The minimised experimental structures using the *ab initio* optimised molecular structures are denoted by the corresponding open symbols. The line links the minimised experimental structure with the closest structure found in the search.

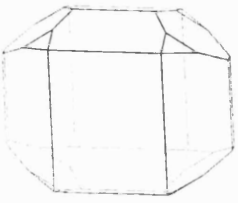
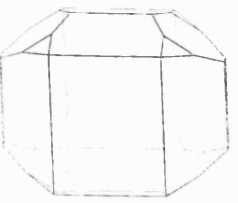
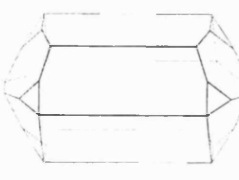
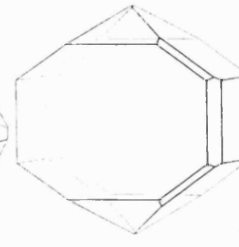
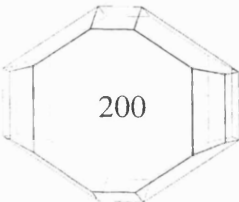
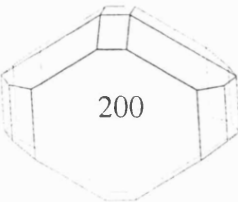
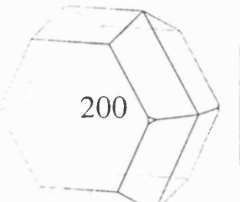
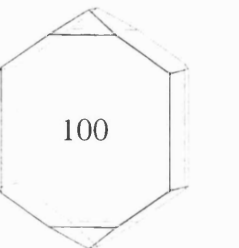
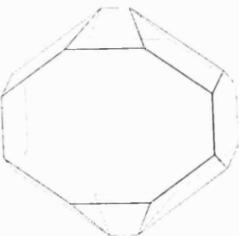
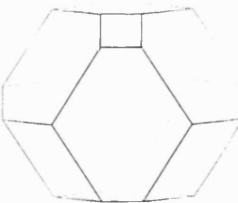
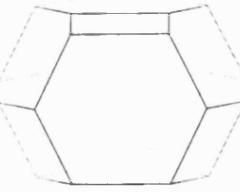
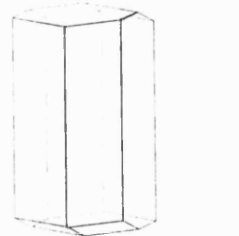
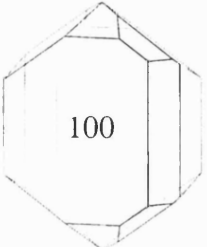
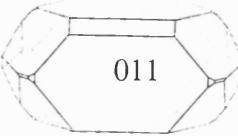
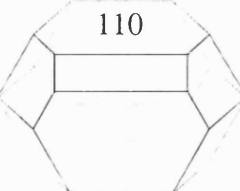
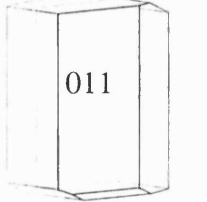
	formic (expt)	formic min _{opt}	AV21	AI2
BFDH				
AE				
	AI19	AQ23	AQ15	AM20
BFDH				
AE				

Figure 7.4.1 (continued)

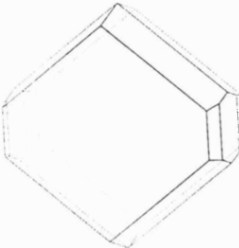
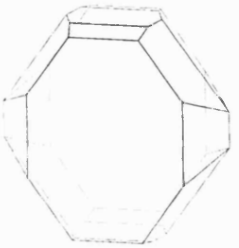
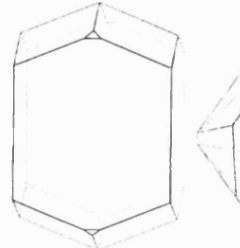
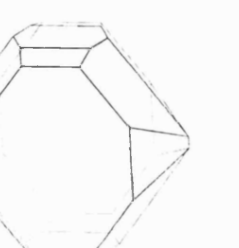

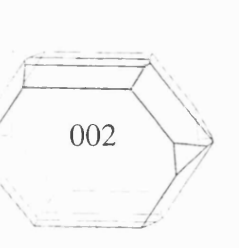
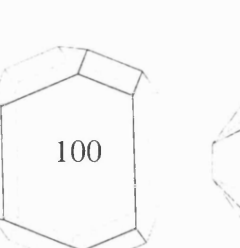

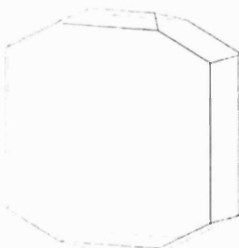
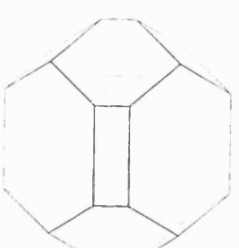
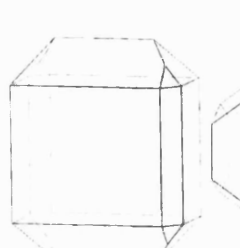
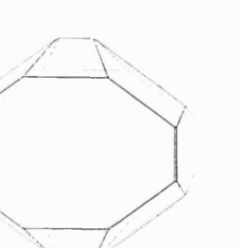
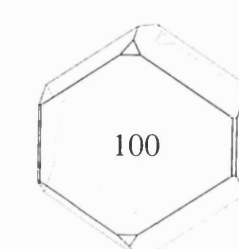
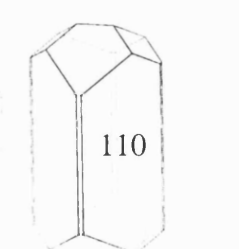
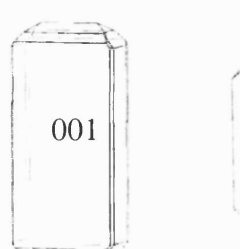
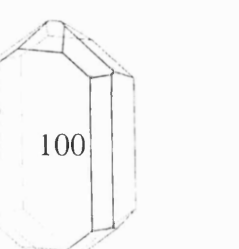
	FA15	CC22	AK4	CC5
BFDH				
AE	 020	 002	 100	 002
	AK22	AQ7	AF17	AI4
BFDH				
AE	 100	 110	 001	 100

Figure 7.4.1 (continued)

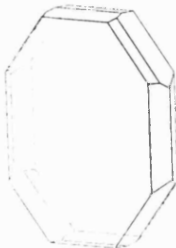
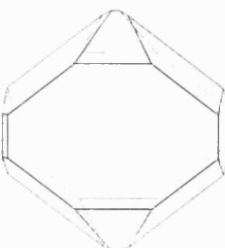
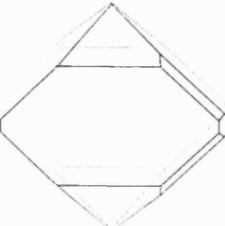
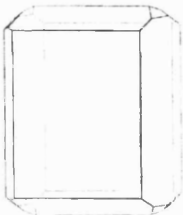
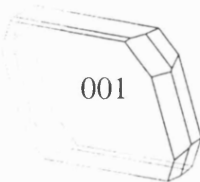
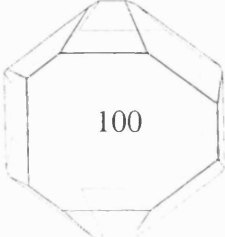
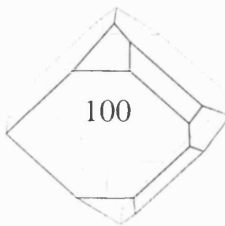
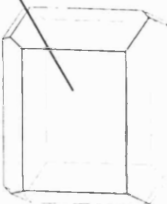
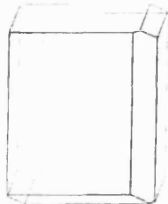
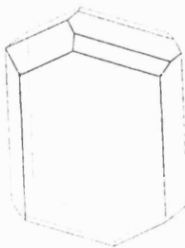
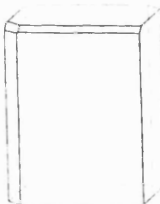
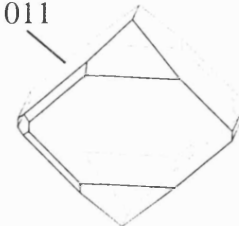
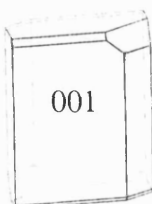
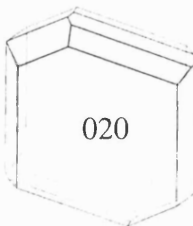
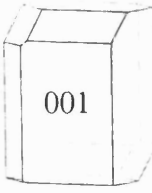
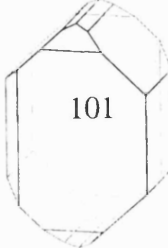
	CC7	AK23	AK5	FC3
BFDH				
AE				
	AB5	FA14	CA8	AM17
BFDH				
AE				

Figure 7.4.1 (continued)

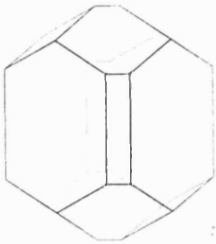
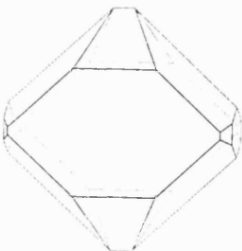
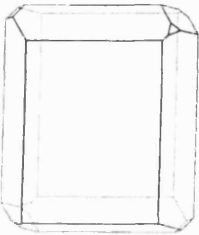
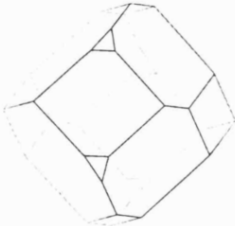
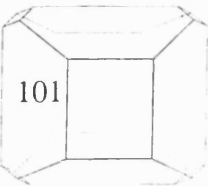
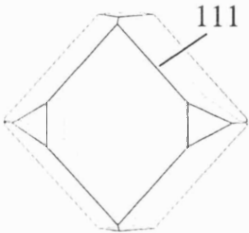
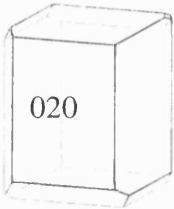
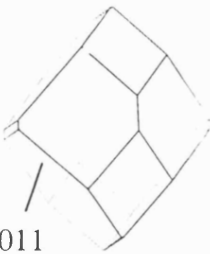
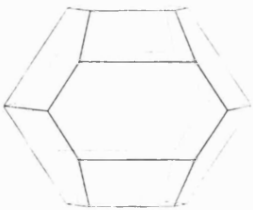
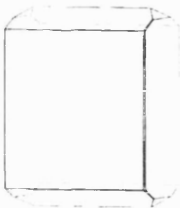
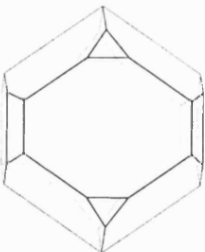
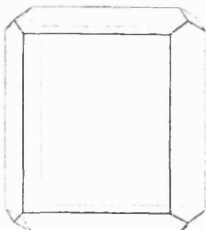
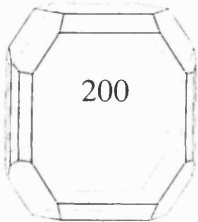
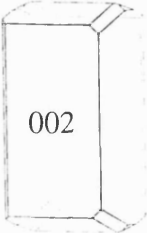
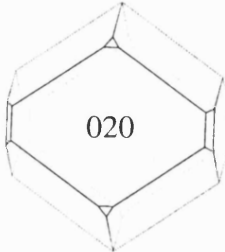

	AZ9	CB3	FA3	AM7
BFDH				
AE				
	CB19	FC9	CB8	AZ7
BFDH				
AE				

Figure 7.4.1 (continued)

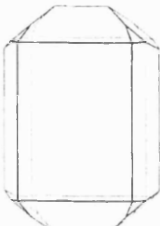
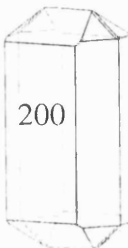
	BD6
BFDH	
AE	

Figure 7.4.1 (continued)

BFDH and attachment energy (AE) predicted morphologies for the known and hypothetical crystal structures of formic acid.

The slowest growing faces are labelled.

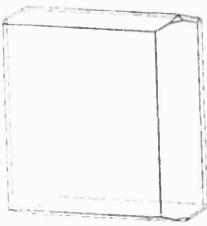
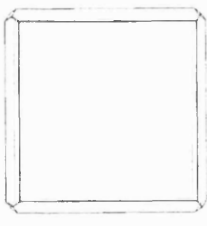
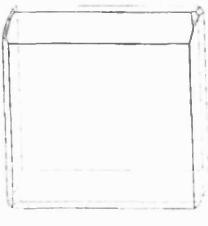


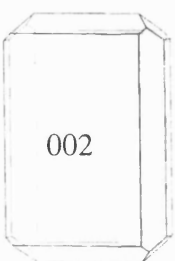

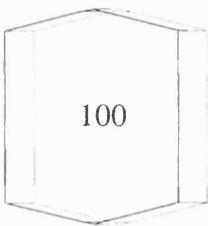
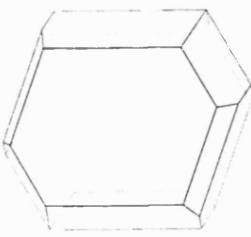
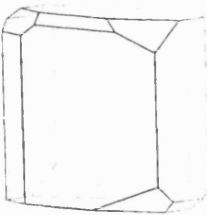
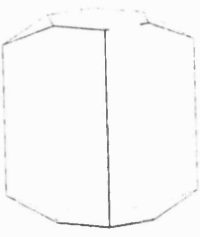


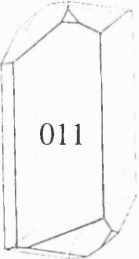
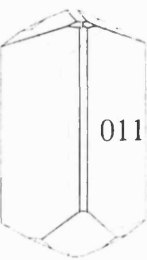
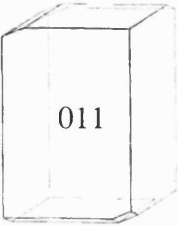
	benzoic (expt)	benzoic min _{opt}	FC22	AI25
BFDH				
AE				
	FA16	AM17	AM24	AM21
BFDH				
AE				

Figure 7.4.2 (continued)

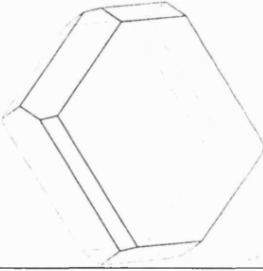
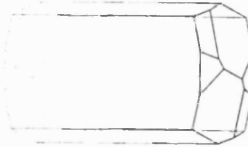
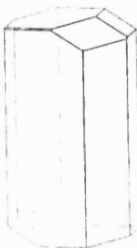
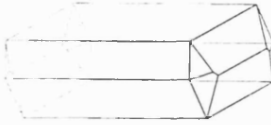
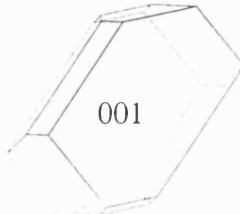
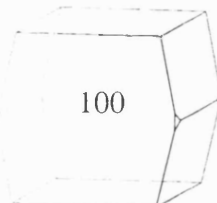
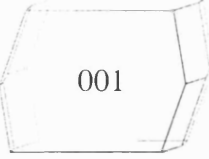
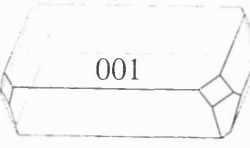
	AB9	AI17	AB3	AF5
BFDH				
AE				

Figure 7.4.2

BFDH and attachment energy (AE) predicted morphologies for the known and hypothetical crystal structures of benzoic acid.
The slowest growing faces are labelled.

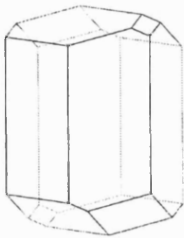
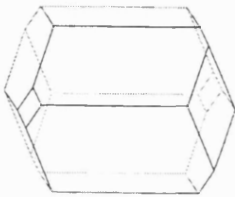
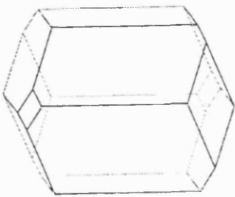
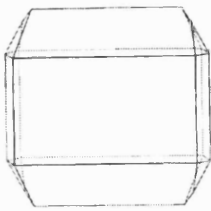
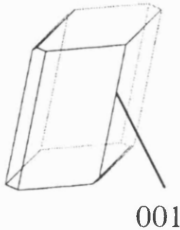
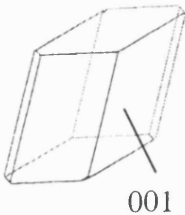
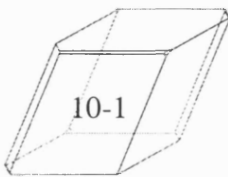
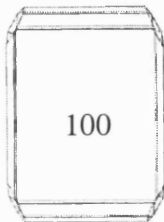
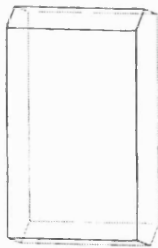
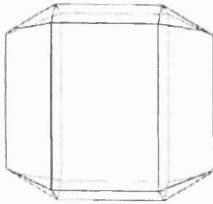
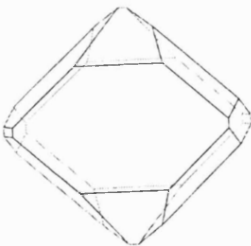

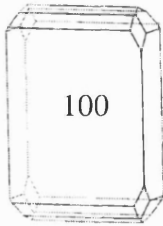
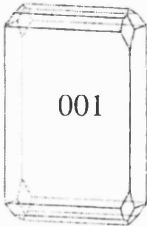
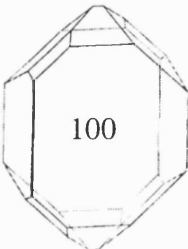
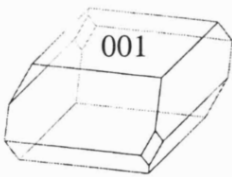
	form α (expt)	form α (\min_{opt})	CA24	form β (expt)
BFDH				
AE	 001	 001	 10-1	 100
	form β (\min_{opt})	AF7	AK5	CA3
BFDH				
AE	 100	 001	 100	 001

Figure 7.4.3



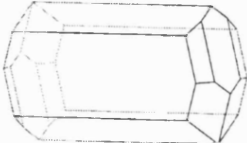

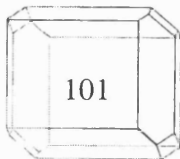
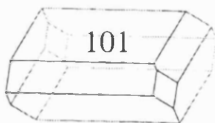
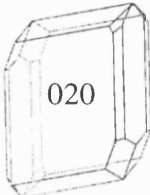
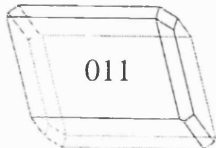
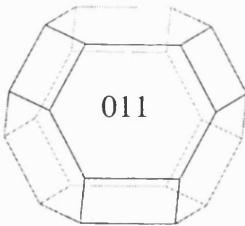
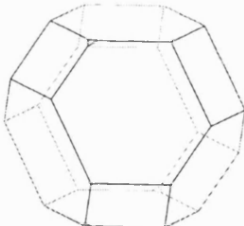
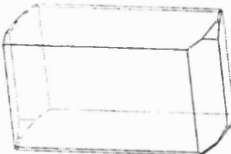
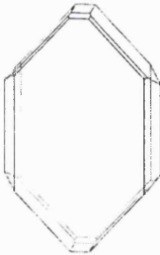
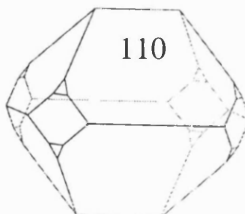
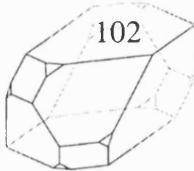
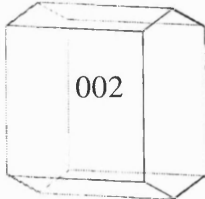
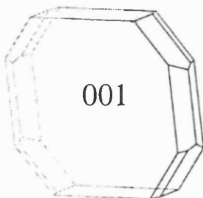
	CA22	CA17	FA6	CA10
BFDH				
AE				
	AQ3	AK7	FC7	CC15
BFDH				
AE				

Figure 7.4.3 (continued)

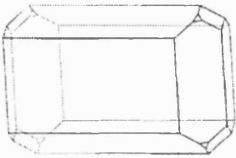
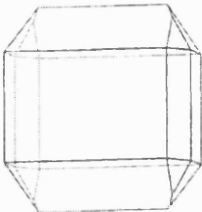


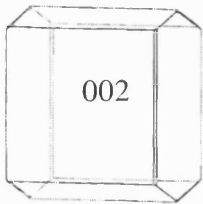
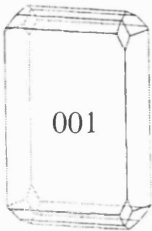
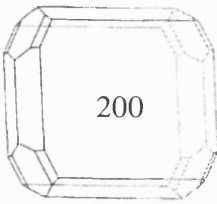
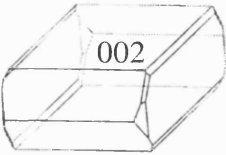

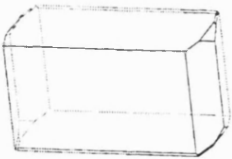
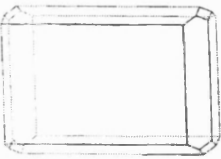
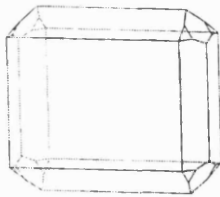
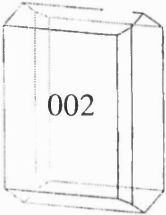
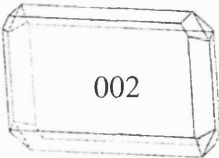
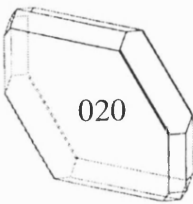
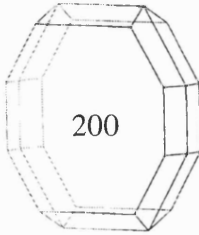
	FC9	AF24	AZ23	FC3
BFDH				
AE				
	FC22	FC8	FA20	BD6
BFDH				
AE				

Figure 7.4.3 (continued)

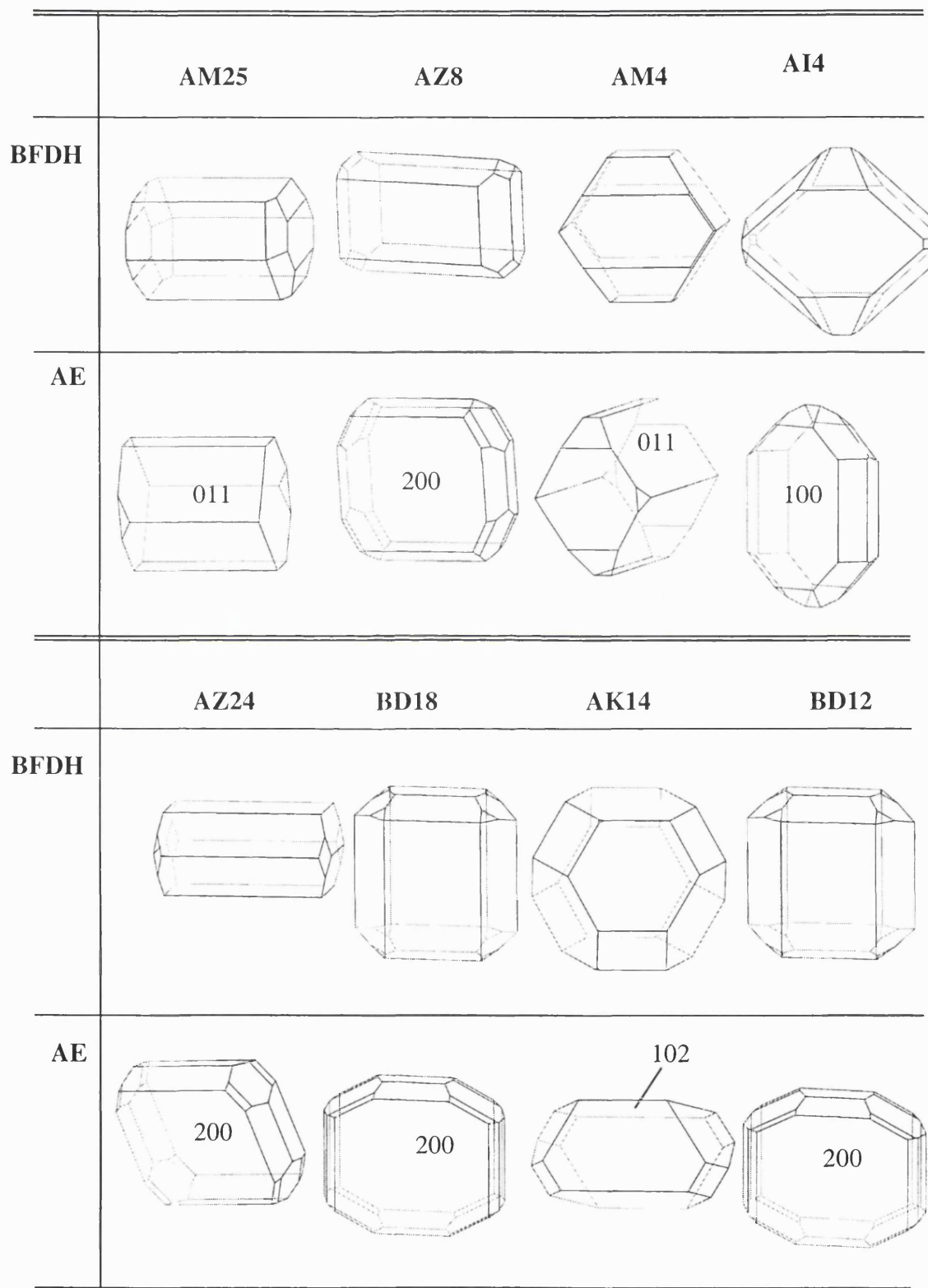


Figure 7.4.3 (continued)

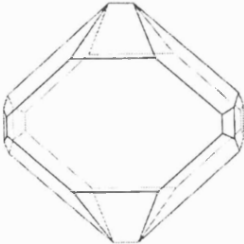

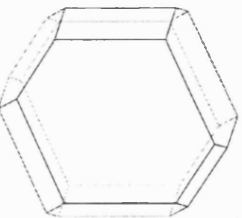
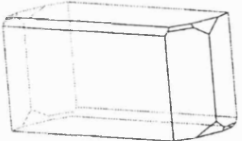
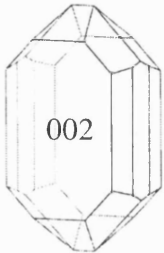
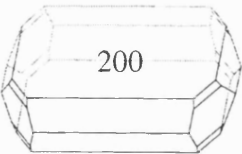
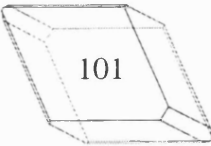
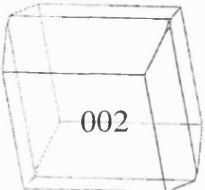
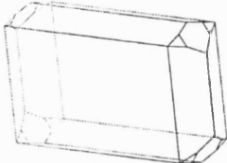
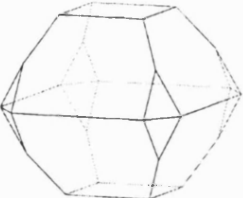

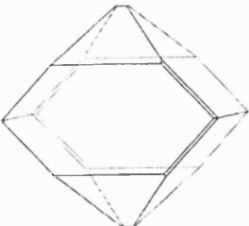
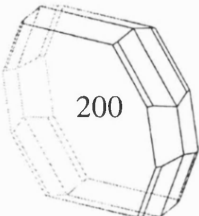
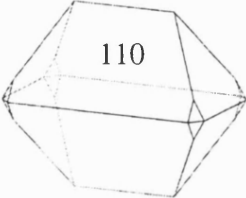
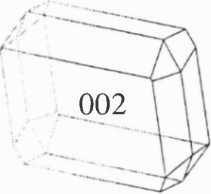
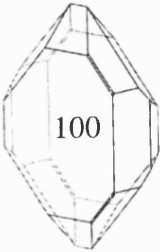
	CC4	BD25	CA2	FC16
BFDH				
AE				
	BD9	AV9	FC17	AK10
BFDH				
AE				

Figure 7.4.3 (continued)

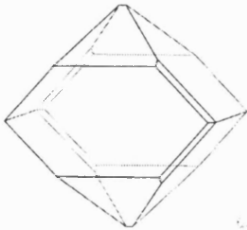
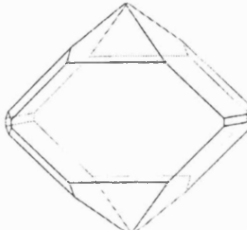

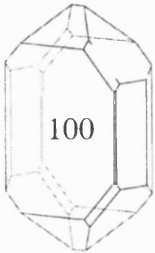
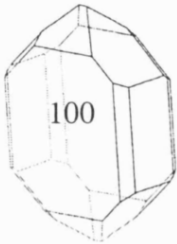
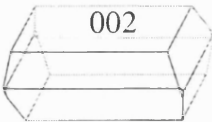
	AK4	AK11	FC13
BFDH			
AE			

Figure 7.4.3 (continued)

BFDH and attachment energy (AE) predicted morphologies for the known and hypothetical crystal structures of tetrolic acid.

The slowest growing faces are labelled.

References for chapter 7

- (1) Beyer, T.; Price, S.L., *J. Phys. Chem. B*, **2000**, *104*, 2647-2655. Dimer or catemer? Low-energy crystal packings for small carboxylic acids.
- (2) Nangia, A.; Desiraju, G.R., *Topics in Current Chemistry*, **1998**, *198*, 57-95. Supramolecular synthons and pattern recognition.
- (3) Allen, F.H.; Motherwell, W.D.S.; Raithby, P.R.; Shields, G.P.; Taylor, R., *New J. Chem.*, **1999**, 25-34. Systematic analysis of the probabilities of formation of bimolecular hydrogen-bonded ring motifs in organic crystal structures.
- (4) Gavezzotti, A.; Filippini, G., *J. Phys. Chem.*, **1994**, *98*, 4831-4837. Geometry of the intermolecular X-H...Y (X, Y = N, O) hydrogen bond and the calibration of empirical hydrogen-bond potentials.
- (5) Shimoni, L.; Glusker, J.P.; Bock, C.W., *J. Phys. Chem.*, **1996**, *100*, 2957-2967. Energies and geometries of isographic hydrogen-bonded networks. 1. The R22(8) Graph Set.
- (6) Aakeroy, C.B., *Acta Cryst.*, **1997**, *B53*, 569. Crystal engineering: Strategies and architectures.
- (7) Etter, M.C., *Acc. Chem. Res.*, **1990**, *23*, 120 - 126. Encoding and decoding hydrogen bond patterns of organic compounds.
- (8) Leiserowitz, L., *Acta Cryst.*, **1976**, *B32*, 775. Molecular packing modes. Carboxylic acids.
- (9) Derissen, J.L.; Smit, P.H., *Acta Cryst.*, **1977**, *A33*, 230-232. Could acetic acid crystallize as dimers?
- (10) Berkovitch-Yellin, Z.; Leiserowitz, L., *J. Am. Chem. Soc.*, **1982**, *104*, 4052-4064. Atom-atom potential analysis of carboxylic acids.
- (11) Hagler, A.T.; Dauber, P.; Lifson, S., *J. Am. Chem. Soc.*, **1979**, *101*, 5131-5141. Consistent force field studies of intermolecular forces in hydrogen-bonded crystals. 3. The C=O...H-O hydrogen bond and the analysis of the energetics and packing of carboxylic acids.
- (12) Mooij, W.T.M.; van Eijck, B.P.; Price, S.L.; Verwer, P.; Kroon, J., *J. Comp. Chem.*, **1998**, *19*, 459 - 474. Crystal structure predictions for acetic acid.
- (13) Payne, R.S.; Roberts, R.J.; Rowe, R.C.; Docherty, R., *J. Comp. Chem.*, **1998**, *19*, 1-20. Generation of crystal structures of acetic acid and its halogenated analogs.

- (14) McGuire, R.F.; Momany, F.A.; Scheraga, H.A., *J. Phys. Chem.*, **1972**, 76, 375-393. Energy parameters in polypeptides. V. An empirical hydrogen bond potential function based on molecular orbital calculations.
- (15) Momany, F.A.; Carruthers, L.M.; McGuire, F.; Scheraga, H.A., *J. Phys. Chem.*, **1974**, 78, 1595-1620. Intermolecular potentials from crystal data III. Determination of empirical potentials and application to the packing configurations and lattice energies in crystals of hydrocarbons, carboxylic acids, amines, and amides.
- (16) Momany, F.A., *Environmental effects on molecular structure and properties*, **1976**, Dordrecht: Reidel.
- (17) Minicozzi, W.P.; Stroot, M.T., *J. Comp. Phys.*, **1970**, 6, 95-104. On the determination of interaction energy functions. II. Crystalline formic acid.
- (18) Cox, S.R.; Hsu, L.-Y.; Williams, D.E., *Acta Cryst.*, **1981**, A 37, 293-301. Nonbonded potential function models for crystalline oxohydrocarbons.
- (19) Derissen, J.L.; Smit, P.H., *Acta Cryst.*, **1978**, A34, 842-853. Intermolecular interactions in crystals of carboxylic acids. IV. Empirical interatomic potential functions.
- (20) Smit, P.H.; Derissen, J.L.; Van Duijneveldt, F.B., *Mol. Phys.*, **1979**, 37, 521-539. Intermolecular interactions in crystals of carboxylic acids III. Non-empirical interatomic potential functions.
- (21) Lifson, S.; Hagler, A.T.; Dauber, P., *J. Am. Chem. Soc.*, **1979**, 101, 5111-5121. Consistent force field studies of intermolecular forces in hydrogen-bonded crystals. 1. Carboxylic acids, amides, and the C=O...H-hydrogen bonds.
- (22) Hagler, A.T.; Lifson, S.; Dauber, P., *J. Am. Chem. Soc.*, **1979**, 101, 5122-5130.
- (23) Hagler, A.T.; Huler, E.; Lifson, S., *J. Am. Chem. Soc.*, **1974**, 96, 5319. Energy functions for peptides and proteins. I. Derivation of a consistent force-field including the hydrogen bond from amide crystals.
- (24) Hagler, A.T.; Huler, E.; Lifson, S., *J. Am. Chem. Soc.*, **1974**, 96, 5327.
- (25) Hirshfeld, F.L., *Theor. Chim. Acta*, **1977**, 44, 129.
- (26) Lifson, S.; Hagler, A.T.; Dauber, P., *J. Am. Chem. Soc.*, **1979**, 101, 5131-5141. Consistent force field studies of intermolecular forces in hydrogen-bonded crystals. 2. A benchmark for the objective comparison of alternative force fields.
- (27) Nahrngbauer, I., *Acta Cryst.*, **1978**, B34, 315-318. A reinvestigation of the structure of formic acid (at 98 K).

- (28) Davis, R.E.; Bernstein, J., *Trans. ACA*, **1998**, *33*, 7-21. Graph set analysis of hydrogen-bond patterns in molecular crystals.
- (29) Feld, R.; Lehmann, M.S.; Muir, K.W.; Speakman, J.C., *Zeitschrift fuer Kristallographie*, **1981**, *157*, 215-231. The crystal structure of benzoic acid.
- (30) Benghiat, V.; Leiserowitz, L., *J. Chem. Soc.*, **1972**, 1763. Molecular packing modes. Part VI. Crystal and molecular structures of two modifications of tetrolic acid.
- (31) Amos, R.D.; Alberts, I.L.; Andrews, J.S.; Colwell, S.M.; Handy, N.C.; Jayatilaka, D.; Knowles, P.J.; Kobayashi, R.; Koga, N.; Laidig, K.E.; Maslen, P.E.; Murray, C.W.; Rice, J.E.; Sanz, J.; Simandiras, E.D.; Stone, A.J.; Sul, M.D., *CADPAC6: The Cambridge Analytic Derivatives Package*, Issue 6: Cambridge University, 1995.
- (32) Nahrinbauer, I., *Acta Chem. Scan.*, **1970**, 453-462. Reinvestigation of the crystal structure of acetic acid.
- (33) Brown, C.J., *Acta Cryst.*, **1966**, *21*, 1-5. The crystal structure of fumaric acid.
- (34) Bednowitz, A.L.; Post, B., *Acta. Cryst.*, **1966**, *21*, 566 - 571. Direct determination of the crystal structure of beta - fumaric acid.
- (35) Derissen, J.L.; Smit, P.H., *Acta Cryst.*, **1974**, *B30*, 2240. Refinement of crystal structures of anhydrous alpha- and beta-oxalic acids.
- (36) Allen, F.H.; Kennard, O.; Watson, D.G.; Brammer, L.; Orpen, A.G.R., *J. Chem. Soc., Perkin Trans.*, **1987**, *2*, S1 - S9. Tables of bond lengths determined by X-ray and neutron diffraction. Part 1. Bond lengths in organic compounds.
- (37) *PLUTO program*, CCDC, 12 Union Road, Cambridge CB1 2EZ (download as RPLUTO available from www.ccdc.cam.ac.uk), Cambridge, 1998, (*Download as RPLUTO available from www.ccdc.cam.ac.uk*).
- (38) Willock, D.J.; Price, S.L.; Leslie, M.; Catlow, C.R., *J. Comp. Chem.*, **1995**, *16*, 628. The relaxation of molecular crystal structures using a distributed multipole electrostatic model.
- (39) Holden, J.R.; Du, Z.; Ammon, H.L., *J. Comp. Chem.*, **1993**, *14*, 422.
- (40) Donnay, J.D.H.; Harker, D., *Am. Mineral.*, **1937**, *22*, 446.
- (41) Berkovitch-Yellin, Z., *J. Am. Chem. Soc.*, **1985**, *107*, 8239. Toward an ab initio derivation of crystal morphology.
- (42) Cerius2, Version 3.5: Molecular Simulations Inc., 1997.

- (43) Nobeli, I.; Price, S.L., *J. Phys. Chem.*, **1999**, A 103, 6448 - 6457. A non-empirical intermolecular potential for oxalic acid crystal structures.
- (44) Nagaoka, S.; Terao, T.; Imashiro, F.; Saika, A.; Hirota, N.; Hayashi, S., *Chem. Phys. Lett.*, **1981**, 80, 580-584. A study on the proton transfer in the benzoic acid dimer by ^{13}C high-resolution solid-state NMR and proton T_1 measurements.
- (45) Stoeckli, A.; Meier, B.H.; Kreis, R.; Meyer, R.; Ernst, R.R., *J. Chem. Phys.*, **1990**, 93, 1502-1520. Hydrogen bond dynamics in isotopically substituted benzoic acid dimers.
- (46) Meier, B.H.; Graf, F.; Ernst, R.R., *J. Chem. Phys.*, **1982**, 76, 767-774. Structure and dynamics of intramolecular hydrogen bonds in carboxylic acid dimers: A solid state NMR study.
- (47) Zelsmann, H.R.; Mielke, Z., *Chem. Phys. Lett.*, **1991**, 186, 501-508. Far-infrared spectra of benzoic acid.
- (48) Skinner, J.L.; Trommsdorff, H.P., *J. Chem. Phys.*, **1988**, 89, 897-907. Proton transfer in benzoic acid crystals: A chemical spin-boson problem. Theoretical analysis of nuclear magnetic resonance, neutron scattering, and optical experiments.
- (49) Docherty, R.; Clydesdale, G.; Roberts, K.J.; Bennema, P., *J. Phys. D: Appl. Phys.*, **1991**, 24, 89-99. Application of Bravais-Friedel-Donnay-Harker, attachment energy and Ising models to predicting and understanding the morphology of molecular crystals.
- (50) Frankenbach, G.M.; Etter, M.C., *Chem. Mater.*, **1992**, 4, 272-278. Relationship between symmetry in hydrogen-bonded benzoic acids and the formation of acentric crystal structures.
- (51) Kaduk, J.A.; Golab, J.T.; Leusen, F.J.J., *Crystal Engineering*, **1998**, 1, 277-290. The crystal structures of trimellitic anhydride and two of its solvates.
- (52) Gavezzotti, A. in *Theoretical Aspects of Computer Modelling of the Molecular Solid State*, A. Gavezzotti, Wiley & Sons, Chichester, **1997**, 97. Energetic aspects of crystal packing.
- (53) Pertsin, A.J.; Kitaigorodskii, A.I., *The atom-atom potential method*, **1987**, Berlin: Springer-Verlag.

- (54) Filippini, G.; Gavezzotti, A.; Novoa, J.J., **1999**, Modelling the crystal structure of the 2-hydronitronylnitroxide radical (HNN): observed and computer-generated polymorphs.
- (55) Price, S.L.; Wibley, K.S., *J Phys Chem A*, **1997**, *101*, 2198-2206. Predictions of crystal packings for uracil, 6-azauracil, and allopurinol: The interplay between hydrogen bonding and close packing.
- (56) Allan, D.R.; Clark, S.J., *Phys. Rev. Lett.*, **1999**, *82*, 3464. Impeded dimer formation in the high-pressure crystal structure of formic acid.
- (57) Allan, D.; Clark, S., *Phys. Rev. B*, **1999**, in press. Comparison of the high-pressure and low-temperature structures of ethanol and acetic acid.
- (58) Reference for the experimental heats of sublimation: NIST Chemistry Webbook (<http://webbook.nist.gov/chemistry>).
- (59) Krivy, I.; Gruber, B. *Acta Cryst.* **1976**, *A32*, 297. A Unified Algorithm for Determining the Reduced (Niggli) Cell.

Chapter 8.

Conclusions and suggestions for further work

*'Prior to solution of the crystal structure prediction problem, the principles of organic solid state chemistry should be worked out by a deliberate and coordinated effort ... a systematic study of large classes of organic solids for the establishment of thermodynamic, kinetic, and structural principles of the organization of matter in solids at a molecular level.'*¹

The reliable *ab initio* prediction of molecular crystal structures is an ambitious scientific goal. A gradual improvement of our understanding of the processes in the supermolecule 'crystal' is the key towards achieving this aim and the work in this thesis will hopefully serve as a useful building block for this learning process.

The program suite MOLPAK/DMAREL, employed in this work, is based, as the majority of currently available crystal structure prediction programs, on lattice energy minimisation. With the exception of one of the polymorphic forms of the blindtest compound I (chapter 5), this systematic rigid-body search successfully found the experimental crystal structures at or close to the global lattice energy minimum (chapter 6 and 7). This achievement confirms the importance of a thorough treatment of each of the different inputs involved in the crystal structure prediction process. As exemplified for paracetamol, small variations in the molecular model can lead to variations of a few percent in cell lengths and a few kJ mol^{-1} in the lattice energy in rigid molecular structure minima. Furthermore, an accurate potential function is crucial for the successful thermodynamic treatment of the solid state, as the order of the low energy minima is heavily dependent on the accuracy of the potential function employed. The specifically optimised intermolecular potential for carboxylic acids gave an improved reproduction of the known crystal structures and has been successfully applied in the prediction study of formic, benzoic and tetrolic acid (chapter 7). Despite these encouraging successes of this method, the underlying assumption that the thermodynamically stable experimental polymorph corresponds to the global lattice energy minimum is just a necessary but not sufficient condition for a crystal structure to be experimentally observed. Nucleation is a kinetically driven process and therefore the observed crystal structure may be metastable. In addition crystal structure prediction methods based on lattice energy minimisation may find more hypothetical low lattice energy structures within a few kJ mol^{-1} of the global minimum than there are observed polymorphs. This has been found in all case studies in this thesis (chapter 5, 6 and 7).

We have proposed that in addition to lattice energy we should consider also the growth rate, estimated by the attachment energy model, and the mechanical stability, determined by the calculated elastic constants in predicting possible polymorphs. For paracetamol (chapter 6) this approach favours the experimentally known structures and reduces the number of hypothetical structures. This promising result is based on the observation that some structures are predicted to grow as thin plates so slowly that they are unlikely to be found in competition with equally thermodynamically stable structures, and some low energy minima are only just mechanically stable according to the calculated elastic constants. In contrast, in the case of small carboxylic acids (chapter 7), the consideration of the mechanical stability and growth rates of the hypothetical minima does not help to eliminate many structures and so is less useful in

determining which polymorphs are most likely to be found experimentally. This is likely to be related to the size and lack of shape of these small molecular structures. Nevertheless, the encouraging results of this novel approach for the case of paracetamol show the importance of further applications of this method to other molecular crystals to test its general validity. Work in progress (personal communication) suggests that these criteria are useful in eliminating several hypothetical structures for pyridine, a few for chlorothalonil, but it did not help in choosing the submissions for the CCDC prediction blindtest 2001.

Overall, the work in this thesis shows that there is room for future improvements of parts of established crystal structure prediction methods. The sensitivity we have found to the rigid molecular model suggests that introducing a good model for intramolecular flexibility would improve the accuracy and reliability of the process. Although we have used state-of-the art intermolecular potentials, the implementation of theoretically derived repulsion/dispersion potentials and the explicit treatment of polarisation would be desirable (chapter 2). The consideration of thermal motion by molecular dynamics simulations might show that some hypothetical structures are unstable at higher temperatures and therefore could transform into stable related structures. The accurate evaluation of the zero point motion and the entropic contributions to the free energy might well alter the relative stability order of the remaining low energy structures. This work also emphasises the significance of other factors, as kinetics of crystallisation, which need to be taken into account before genuine prediction of crystal structures will be a feasible and robust process. A combination of a range of theoretical methods, as the pioneering molecular dynamics simulation work (chapter 3.5) on precursors in solution to reveal some insight into the first stages of nucleation, is a very promising starting point. Novel prediction methods based on the distribution of experimental data retrieved from the CSD (chapter 3.3) will reflect the effects of crystallisation conditions. Ideally, we would wish to explicitly model the effects of solvent and crystallisation method on determining the observed crystal structures. A major problem in crystal structure prediction is therefore the lack of experimental information and even if the structures are known the crystallisation conditions are not always available. The explicit experimental search for more polymorphic forms might reveal the true extent of polymorphism among molecular crystal structures. Unfortunately, this important work is often conducted on industrial grounds and the results are therefore hidden in private databases.

In general, a thorough combination of experimental experience and theoretical knowledge may provide a valuable link of already available, but so far unfortunately often not connected findings. A concerted effort of experimental and theoretical studies would not only lead to more consistency of the published literature, but it might be the only way to obtain a more global understanding of the causal relationships which govern the phenomenon 'solid state'.

Comparing the performance of different approaches to crystal structure prediction is not a simple task, as the literature survey in chapter 3 revealed. The CCDC crystal structure prediction workshop provides therefore an excellent – if not the only – way of a fair and objective tool for such a comparative study. Although the organisational effort for such an event is enormous, it is a much needed and unique opportunity to test and develop prediction methods. I expect that the results of the second prediction workshop will be in accord with the overall conclusion of this thesis.

Considerable progress in crystal structure prediction has been made, but a further understanding of the principles of the organic solid state is required before the crystal structure prediction problem will be successfully solved.

References for chapter 8

- (1) Gavezzotti, A., *Acc. Chem. Res.*, **1994**, 27, 309-314. Are crystal structures predictable?