UNIVERSITY OF SOUTHAMPTON

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

Chemistry

The Acetanilide Crystal Structures: Packing and Conformational Similarities

by

L. Susanne Coles (née Huth)

Thesis for the degree of Doctor of Philosophy

March 2011

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES
CHEMISTRY

Doctor of Philosophy

THE ACETANILIDE CRYSTAL STRUCTURES: PACKING AND CONFORMATIONAL SIMILARITIES

by
Lotte Susanne Coles (née Huth)

The dependence of crystal structure assembly on molecular changes has been studied for mono-substituted acylanilides. A library of 255 crystal structures has been investigated. 217 structures were determined during this project and 38 originated from the CSD. A subset of 64 mono-substituted acetanilides was studied in detail for crystal structure similarities using the XPac algorithm and conformational preferences were assessed. Chapter 4 presents the crystal structures and their structural similarities, whilst Chapter 5 focuses on the theoretical conformational analysis.

Three amide-amide hydrogen bond geometries and a molecular stack were frequently observed in the acetanilides. The assembly of these one-dimensional packing features governs in particular the crystal structures of the para-substituted acetanilides. The amide-amide hydrogen bonded chains were very persistent in all the acylanilides. The amide substituent had a larger effect on the chain geometry than the phenyl substituent.

In the crystal the molecular structures deviate from the expected planar conformation. Rotation about the amide to phenyl bond of up to 55˚ in ortho-*tert*butyl acetanilide is observed. Para-substituents affect the rotational barriers, whilst such correlation was not discernible for meta-substituents. Ortho substitution frequently produces intramolecular hydrogen bonding outweighing any other substituent effects. Although the experimental conformations deviate from the planar geometry they are still within the minimum energy well and energy penalties are small. These can be easily compensated for by intermolecular interactions in the crystal.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACCOMPANYING MATERIAL

APPENDIX 1-1: List of space group distribution in organic crystal structures

APPENDIX 2-1: eCrystals deposition numbers and URLs for undesired products

APPENDIX 2-2: CSD reference codes and references for structures harvested from the CSD

APPENDIX 2-3: eCrystals deposition numbers and URLs for acetanilide structures determined during this project

APPENDIX 3-1: Perl scripts

APPENDIX 4-1: Crystal data for acetanilide crystal structures determined during this project

APPENDIX 4-2: Base vectors for SCs identified by XPac

APPENDIX 4-3: Lattice energies of the acetanilide structures as calculated with the Pixel algorithm

APPENDIX 4-4: Intermolecular interaction energies of acetanilides as estimated with the Pixel method

APPENDIX 4-5: Evolution of structural similarity plots of para acetanilides

APPENDIX 5-1: PES values

all appendices are stored on the attached CD-ROM

QUICK-GUIDE: detachable guide containing fundamental diagrams and compound identifier aid for quick reference

DECLARATION OF AUTHORSHIP

I, ……*L. Susanne Coles* ……,

declare that the thesis entitled

"The Acteanilide Crystal Structures: Packing and ……

Conformational Similarities" ……..

and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- none of this work has been published before submission.

Signed: ……..

Date:…….. 31/03/2011 ……..

ACKNOWLEDGEMENTS

This work would never have been possible without the great support and enthusiasm of Terry Threlfall. My sincerest thanks go to him not only for producing all the samples for this project but also for meticulously proof reading and engaging in the most stimulating discussion of results. I would like to acknowledge my supervisor Mike Hursthouse for securing the funding for this project and for giving me the opportunity to carry out this work in his group. The financial support was provided by the Cambridge Crystallographic Data Centre and Pharmorphix Ltd, and I would like to thank both companies for their generosity. The past and present members of the National Crystallography Service were always there for a good chat and a cup of tea – thanks guys for making my time here very enjoyable! Speaking of fun, I certainly need to acknowledge Wendy for being there and everything else… Last but not least I would like to thank my family for supporting me morally in this quest. Well, and of course I am grateful to Simon for his patience and belief in me, but that almost goes without saying.

ABBREVIATIONS

Initials

CSF	Christopher S. Frampton (Pharmorphix Ltd)
TLT	Terence L. Threlfall (University of Southampton)

Abbreviations

CSD	Cambridge Structural Database
ASU	asymmetric unit
Z	number of formula units in the unit cell
Z'	number of crystallographically independent components in the ASU
uc	unit cell
0D	zero dimensional
1D	one dimensional
2D	two dimensional
3D	three dimensional
OSP	ordered set of points
COSP	corresponding ordered set of points
SC	supramolecular construct
LCAO	linerar combination of atomic orbitals
HF	Hartree Fock
DFT	Density Functional Theory
cif	crystallographic information file
HB	hydrogen bond / hydrogen bonded / hydrogen bonding
PES	potential energy surface

PREAMBLE

This thesis is concerned with the understanding of the crystalline solid state of small organic compounds. Recently the crystal structure assembly of functionalised organic molecules has attracted considerable interest due to its importance in the controlled production of pharmaceuticals and other industrially important materials as well as in the search for solids with utilisable properties. The study of the solid state behaviour of small organic molecules is also essential to the areas of intermolecular interactions, crystal engineering and crystal structure prediction.

The molecular aggregation in a crystal is governed by intermolecular forces of varying strength and the crystal structure is the result of an overall balance of these forces. Some crystal structure assemblies can be predicted or designed, especially when medium-to-strong supramolecular interactions control the particular form of assembly, overwhelming the competition from the numerous, but weaker van der Waals forces. However, crystal structure prediction and design becomes very difficult in cases where medium-to-weak intermolecular interactions dominate the assembly or where the compound contains several functional groups competing for intermolecular interaction. In addition polymorphism, the ability of a compound to crystallise in different solid state forms, challenges all predictive methods. Unfortunately, most molecules of pharmaceutical importance fall into this category, and so the possible occurrence of polymorphic forms and other similarly important systems, are generally explored via experimental screening methods.

In this work these challenges are addressed with a knowledge-based approach, whereby real structural data is used to gain a better understanding of the factors directing the crystal assembly. Ideal systems for this strategy are closely related compounds with varying small substituents such as F, Cl, Br, I, methyl, ethyl. The biggest resource for crystal structure data is the Cambridge Structural Database (CSD), but the data currently available and usable for this purpose are not entirely satisfactory as most of the database entries originate from individual compound studies and the CSD thus lacks compound libraries with significant populations.

Hence a large array of closely related compounds has been synthesised and their crystal structures determined. The chemical system investigated here is that of the mono-substituted acylanilides with particular focus on a subset of 64 crystal structures of mono-substituted acetanilides. This subset was selected from the much larger compound library of approximately 250 acylanilide crystal structures, the detailed discussion of which certainly lies beyond the scope of this thesis. For the study of the supramolecular assemblies of these compounds, crystallographic and computational methods have been combined thus offering the assessment of intermolecular interactions and of the influence of molecular structure change on crystal packing.

This thesis is organised in six major chapters. Chapter 1 covers the relevant and underlying theory for this work. The mono-substituted acylanilide compound library is introduced in Chapter 2 together with the nomenclature used to identify the individual compounds in this work. Chemical applications are also provided in this chapter. Methods and details of the experimental procedures are summarised in Chapter 3. Chapter 4 presents the crystal structures of the 64 mono-substituted acetanilides together with the results from the search for structural similarities between the individual crystal assemblies. Common packing features are discussed in detail in this chapter. The molecular structure of the acetanilides forms the focus of Chapter 5. In this chapter the findings from a computational conformational analysis are presented and discussed. The structural similarity of the acetanilide family is put into context of the bigger acylanilide library in Chapter 6 and some aspects of dealing with such a large array of compounds are mentioned. Suggestions for future work are also made in this chapter. Finally the major findings from this work are summarised in the concluding Chapter 7.

CHAPTER 1: BACKGROUND

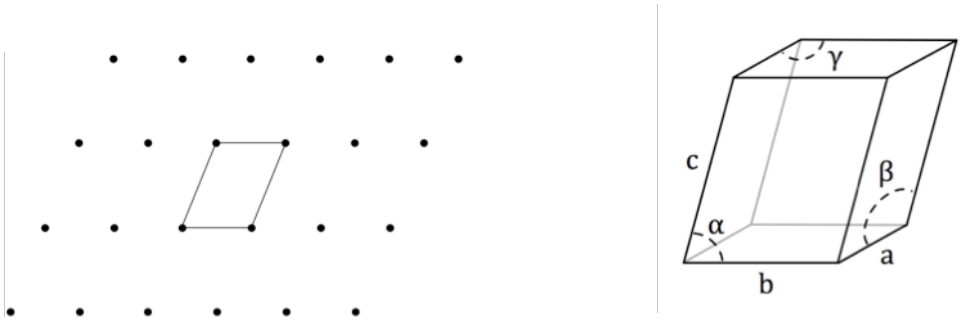
1.A. Single Crystal X-ray Crystallography

1.A.1. Crystalline solid state – molecular aggregation

A chemical substance, whether of atomic, ionic or molecular nature, can generally assume three classical states of matter dependent on ambient conditions (i.e. pressure and temperature): the vapour, liquid and solid state. In the gaseous phase molecules are considered as independent mobile entities with large and variable intermolecular distances occupying an unrestricted volume and there is no or only little intermolecular interaction. In the liquid state the occupied volume and molecular mobility is restricted, however there is sufficient movement to permit fluidity. As the spacing between molecules becomes less variable than in the vapour state, intermolecular forces become active leading to some longer range ordering with a degree of orientation and symmetry, but the molecules are not in a fixed spatial relationship. In the solid phase cohesive forces dominate and the molecules assume a rigid frame of highly-ordered and closely-packed moieties[1]; this state is also known as the crystalline state. There are a number of other non-classical states of matter, to which plasma, amorphous or glassy solids, liquid crystals and superfluids belong, but only the crystalline phase of molecular substances shall be considered further.

A molecular crystal consists of a periodic array of molecular clusters with long-range order and symmetry. Molecules within the cluster can take different orientations but are generally all related by symmetry. For convenience of describing this periodic assembly in mathematical terms, lattice theory is used. In lattice theory this array of molecular clusters is referred to a lattice and the clusters of molecules are represented by identical lattice points. Hence the crystal structure can be described as a periodic three-dimensional arrangement of similar lattice points. The lattice can now simply be represented by a single repeat unit, termed the unit cell. This is shown for a two dimensional lattice in Figure1.1.a.

This repeat unit is constructed of such a number of lattice points that the whole crystal lattice can be reproduced by infinite translation of the unit cell in all three dimensions. The shape of a unit cell is that of a parallelepiped defined by the lattice parameters, i.e. unit cell axes a, b and c, and the unit cell angles α, β and γ (Figure 1.1.b). The unit cell axes should be chosen so that the unit cell represents the smallest repeat unit for which the axes are parallel to, or coincide with, important symmetry directions in the lattice[2].



(a) **(b)**

Figure 1.1: (a) Two dimensional periodic lattice with repeat unit. **(b)** A typical unit cell with unit cell parameters a, b, c, α, β and γ.

Dependent on the number and location of the lattice points in the unit cell, there are different types of lattice centring. The smallest possible unit cell will contain one lattice point in total, i.e. one lattice point at each of the eight corners of the parallelepiped, each point being shared by eight neighbouring unit cells. Such a unit cell is called primitive (P). If the crystal structure possesses reflection and/or rotational symmetry then it might be more convenient to choose a unit cell with more than one lattice point, so that the cell axes are parallel to or coincide with the symmetry elements. In these cases the unit cell is referred to as centred. Centring can occur in different ways. An additional lattice point at the body of the cell yields body-centred unit cells (I), lattice points at the centres of all faces produce a face-centered cell (F) and side-centred unit cells contain lattice points at the centres of opposite pairs of faces (A, B, or C depending on which faces are centred).

Crystal structures are grouped according to the axial system used to describe the lattice, also called lattice system. The combination of lattice centring and lattice system then yields 14 unique lattice types, the Bravais lattices displayed in Figure 1.2.

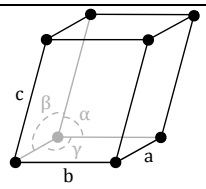
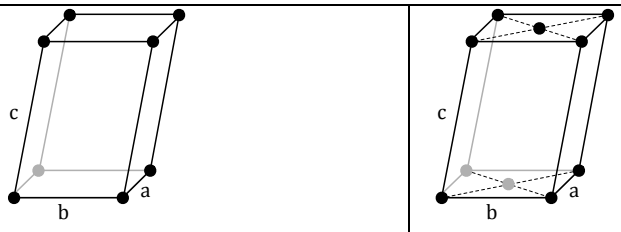
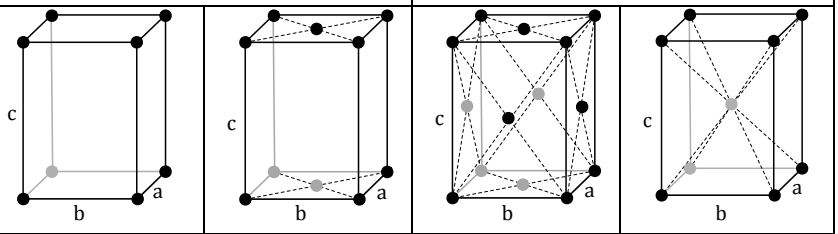
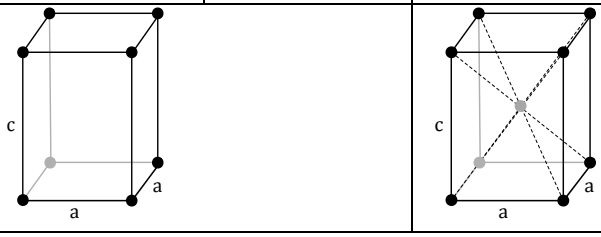
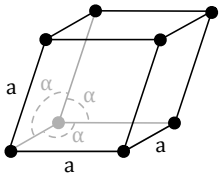
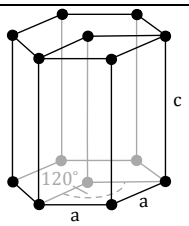
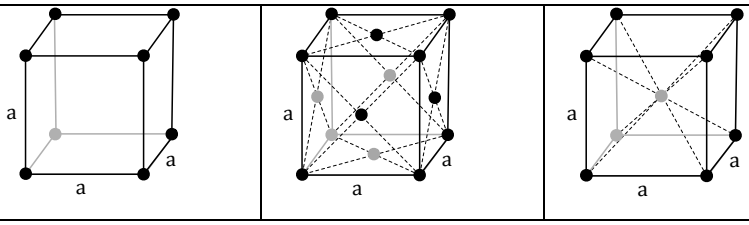
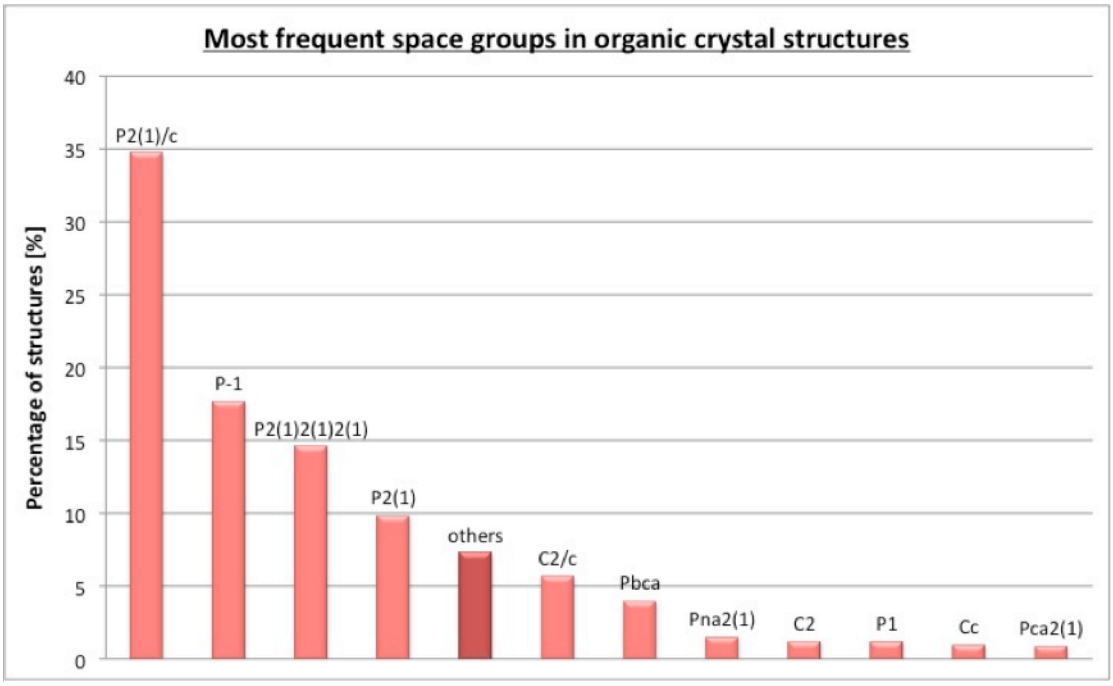
Crystal Systems	***Bravais Lattices***
triclinic P $a \neq b \neq c$ $\alpha \neq \beta \neq \gamma$	
monoclinic P, C $a \neq b \neq c$ $\alpha = \gamma = 90^\circ$ $\beta \neq 90^\circ$	
orthorhombic P, C, F, I $a \neq b \neq c$ $\alpha = \beta = \gamma = 90^\circ$	
tetragonal P, I $a = b \neq c$ $\alpha = \beta = \gamma = 90^\circ$	
trigonal P or R $a = b = c$ $\alpha = \beta = \gamma \neq 90^\circ$	
hexagonal P $a = b \neq c$ $\alpha = \beta = 90^\circ$ $\gamma = 120^\circ$	
cubic P, F, I $a = b = c$ $\alpha = \beta = \gamma = 90^\circ$	

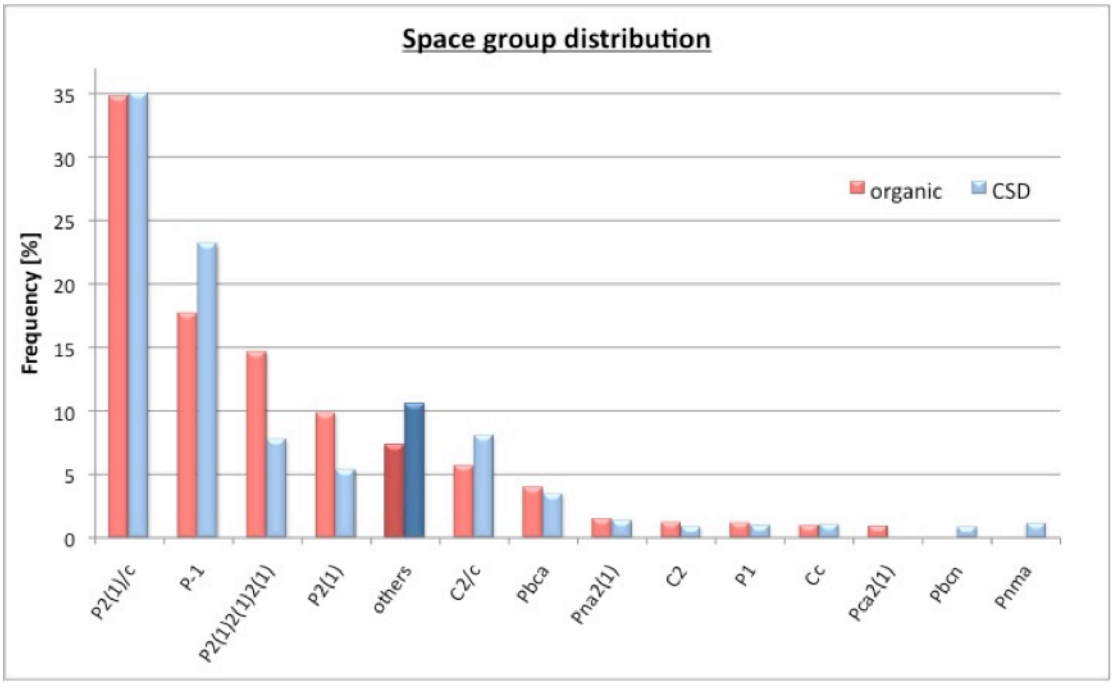
Figure 1.2: The 14 Bravais lattices. Permitted cell centring and axial relationships are included in the crystal system column.

The total symmetry of a molecule is described by its molecular point group, which is defined as the collection of all possible symmetry elements passing through a common point at the centre of the molecule. In a crystal structure the total symmetry is described by it space group. Space group symmetry takes into account not only crystallographic point group symmetry but also translational symmetry. The following symmetry operations can occur in a three-dimensional periodic lattice: rotation (both proper and improper), reflection at a mirror plane, inversion, pure translation as well as the combination of translation and rotation or reflection, which yields screw axes or glide planes, respectively. In contrast to a single molecule the symmetry elements in a crystal do not all pass through a common point, instead they lie either parallel or perpendicular or at some other special angle[3] to each other and occur at regular spacing within the unit cell. They are identically arranged in every unit cell and they can only be combined in certain ways, as is the case for an individual molecule. Due to the translational symmetry of a lattice not all molecular point group symmetry operations are possible, e.g. a 5-fold rotation axis would be incompatible with the periodic nature of a lattice. This means that for crystal structures there are exactly 230 possible space groups, but only a few of those are frequently observed in organic molecular crystal structures. The distribution of these most common space groups, i.e. with a frequency of occurrence of ≥1%, is shown overleaf in Figure 1.3.a whilst a list of all space groups observed in organic crystal structures[4] is available in Appendix 1-1. Figure 1.3.b on the other hand displays the space group distribution for the all structures contained in the Cambridge Crystallographic Database (CSD)[5] compared with the distribution shown in Figure 1.3.a.

The most notable differences between the two distributions are found in space groups $P2_12_12_1$ and $P2_1$, where the frequency of occurrence in organic structures is almost doubled, and in space groups $P\bar{1}$ and C2/c, which are found more frequently in non-organic structures. As early as in the 1950s Kitaigorodskii noticed this phenomenon of crystallisation in preferred space groups and developed a theory of close packing for molecular crystals. With this theory Kitaigorodskii was able to compile a list of most probable space groups for organic crystals taking into account how inherent molecular symmetry is most likely retained or lost in the crystal[6].



(a)



(b)

Figure 1.3: (a) Space group distribution in the set of organic structures. **(b)** Comparison of space group distribution in the set of organic structures and all structures in the CSD.

In this context it is useful to introduce a few more crystallographic terms with respect to the unit cell contents.

Considering a crystal lattice can be composed by pure translation of the unit cell or a fraction of it for centred cells, this can be further expanded. If a structure exhibits any additional symmetry elements, other than pure translation, then the atoms and molecules within a unit cell as well as between unit cells are related by these symmetry elements. Hence the unique symmetry-independent portion of a structure is only a fraction of the unit cell dependent on the symmetry elements present. This unique part of the structure is called the asymmetric unit (ASU). For a single component crystal structure the number of molecules in the unit cell is given by its multiplicity, Z. Related to this is the number of molecules in the ASU, which is expressed by the symbol Z' and usually equals 1. If a molecule in the unit cell resides on a point on a symmetry element without any translational component it is said to occupy a special position as opposed to a general position, which is any other point in the unit cell. The molecule situated on a special position must possess the appropriate symmetry. Hence a compound cannot crystallize with a molecule occupying a special position with higher symmetry than the inherent molecular symmetry, but a highly symmetrical molecule can either retain or lose some of its symmetry in the crystal. If a molecule coincides with a special position the content of the ASU is reduced with Z'<1 reflecting the symmetry-independent fraction of the molecule, e.g. if a molecule lies on a 2-fold rotation axis Z' will be 0.5 as only half of the molecule is unique whilst the other half is related by symmetry. Z' can also be greater than 1 if the ASU contains two or more molecules, that are chemically identical but cannot be related by crystallographic symmetry. The overall multiplicity, Z, of the unit cell is then lowered or increased by this fraction expressed by Z', so that $Z = Z_0 * Z'$ where Z_0 stands for the standard multiplicity of a unit cell with only general positions occupied.

Returning to the theory of close packing, Kitaigorodskii deduced that for molecules occupying special positions there is only a limited number of space groups allowing close packing. Of the space groups mentioned earlier, $P2_1$ and $P2_12_12_1$ are only compatible with molecules that have no symmetry or have lost their inherent symmetry in the crystal. $P\bar{1}$ can furthermore accommodate molecules residing on inversion centres in the crystal whilst molecules with inversion or 2-fold symmetry are close-packed in space group C2/c, i.e. molecules can occupy special positions in $P\bar{1}$ and C2/c, but not so in $P2_1$ and $P2_12_12_1$. The deductions made by

Kitaigorodskii can also be applied to understand why the space group distributions differ between organic and all crystal structures as noted above. Organic molecules are generally of irregular shape with little or no molecular symmetry, whereas ionic or metal-organic compounds are often of higher symmetry. This is reflected in the number of structures that crystallise with a non-integer number of Z', which amounts to approximately 26% of all structures in the CSD. Of these only 10% fall into the organic structure category as defined previously[4] amounting to 11% of all organic structures with molecules residing on special positions (*c.f.* 31% of all other structures). This means that a space group such as $P2_1$ or $P2_12_12_1$ is likely to be more populated for organic crystal structures compared to non-organic structures and vice versa for C2/c and $P\bar{1}$. Table 1.1 summarises the distribution of Z' in the set of organic structures to illustrate this point.

Space Group	Z' (Z)	Percentage [%]
$P2_1$	1 (2)	76.93
	2 (4)	20.79
	>2 (>4)	2.28
$P2_12_12_1$	1 (4)	92.83
	2 (8)	6.42
	>2 (>8)	0.75
$P\bar{1}$	0.5 (1)	9.43
	1 (2)	72.51
	1.5 (3)	0.19
	2 (4)	15.94
	2.5 (5)	0.02
	≥3 (≥6)	1.91
C2/c	0.5 (4)	33.56
	1 (8)	61.30
	1.5 (12)	0.76
	2 (16)	3.78
	2.5 (20)	0.03
	≥3 (≥24)	0.57

Table 1.1: Distribution of Z' in space groups $P\bar{1}$, C2/c, $P2_1$ and $P2_12_12_1$ as found in organic crystal structures.

From the table it is apparent that special positions are only occupied in $P\bar{1}$ and C2/c, especially in C2/c where over a third of structures contain a molecule that lies on a symmetry element. However, it should be kept in mind that the number of structures in C2/c is small compared to $P2_1$, $P2_12_12_1$, $P\bar{1}$ and $P2_1/c$ (*c.f.* Figure 1.3), demonstrating that the number of structures with symmetrical molecules is a small proportion of all organic structures.

It is the symmetry and long-range periodic order in a crystal that enables unambiguous structure determination from X-rays, which will be dealt with subsequently.

1.A.2. Crystal Structure Determination

The foundations of X-ray crystallography were laid with von Laue's discovery in 1912 that crystals could diffract X-rays, which not only proved the wave character of X-rays but also confirmed that crystal structures are periodic three-dimensional arrays of chemical species. The periodic crystal lattice in a single crystal acts as a diffraction grating for X-rays since the inter-atomic distances are comparable to the wavelength of X-rays. Hence exposing a single crystal to X-rays yields an interference pattern of the diffracted beam. X-rays are scattered by the electrons of each atom in the crystal and scattered waves will add constructively or destructively to an overall scattering pattern, which has continuously varying amplitudes and phases. Varying the angle of incidence of the beam, θ, over a certain range, such a pattern, i.e. set of reflections, is collected for each crystal. The geometrical condition for observing a diffraction pattern is given by Bragg's Law:

(1.1) $$n\lambda = 2d\sin\theta$$

where θ is the Bragg angle, d is the spacing between neighbouring lattice planes in the crystal and λ is the wavelength of the X-ray beam (Figure 1.4). The parameter n is an integer, but in conventional X-ray diffraction only n = 1 is considered as higher orders can be converted by dividing d by n to obtain the value of d corresponding to n = 1.

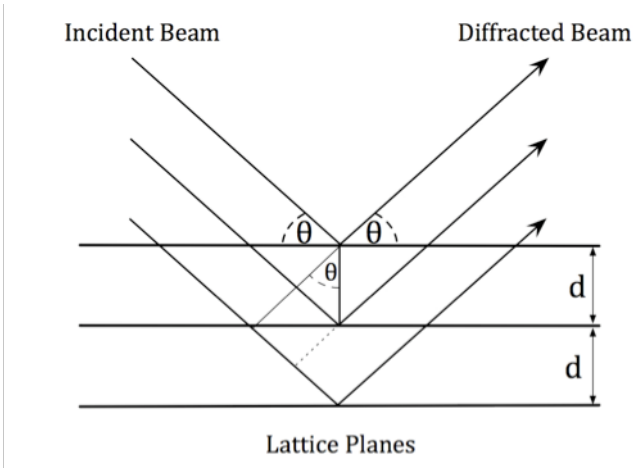


Figure 1.4: The principle underlying Bragg's Law.

The diffraction pattern is also a lattice and is related to the crystal structure via a reciprocal relationship. If the d-spacings in a crystal (real space) are large then small spacings will be observed in the reciprocal lattice (reciprocal space) and vice versa. Mathematically the relationship between the real space lattice vectors **a**, **b**, **c** and the reciprocal lattice vectors **a***, **b***, **c*** is expressed by the following series of equations:

(1.2) $$\mathbf{a}^* = \frac{\mathbf{b}\times\mathbf{c}}{V} \qquad \mathbf{b}^* = \frac{\mathbf{c}\times\mathbf{a}}{V} \qquad \mathbf{c}^* = \frac{\mathbf{a}\times\mathbf{b}}{V}$$

(1.3) $$V = \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \mathbf{b}\cdot(\mathbf{c}\times\mathbf{a}) = \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b})$$

Combining equations 1.2 and 1.3 yields equation 1.4:

(1.4) $$\mathbf{a}\cdot\mathbf{a}^* = \mathbf{b}\cdot\mathbf{b}^* = \mathbf{c}\cdot\mathbf{c}^* = 1$$

The diffraction pattern will furthermore have symmetry that is closely related to the symmetry or the crystal structure producing it. In a diffraction pattern the positions of the reciprocal lattice points (i.e. observed reflections) are determined by the pure crystal lattice whilst the intensity of the reciprocal lattice points is dependent on the type and location of the atoms in the unit cell. This means that the diffraction pattern contains information about the average[7] contents of the unit cell and by converting reciprocal space into real space the crystal structure becomes available.

The measurable quantities from a diffraction experiment are the positions and intensities of the reflections, both will be examined more closely in the following paragraphs.

The coordinates of each reflection can be uniquely identified with the three Miller indices h, k and l, by counting from the centre (position of the undiffracted beam) along the a*, b* and c* axes so that the length of the vector $(h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*)$ yields the distance of the reflection from the origin via Bragg's Law:

(1.5) $$(h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*) = \frac{1}{d_{hkl}} = \frac{2\sin\theta}{\lambda}$$

Once the diffraction pattern has been recorded the reflections are harvested and integrated to yield the reflection intensities, but before these can be used to obtain a crystal structure the intensities have to be corrected for a number of factors according to equation 1.6.

(1.6) $$I_{hkl} = kL(\theta)p(\theta)A(\theta)E(\theta)d(t)|F_{hkl}|^2$$

Some of these factors are entirely geometric corrections, such as the correction for Lorentz effect L(θ), and polarisation effects p(θ). Absorption corrections A(θ), and extinction corrections E(θ) on the other hand, depend on the chemical content, crystal size (and quality in the case of extinction) as well as morphology. Corrections are applied for sample decay with time, d(t), although this is less frequently a problem with CCD area detectors. The constant k is a scale factor.

In order to "reconstruct" an image of the real crystal structure from the reciprocal lattice the so-called structure factor, F, is used. F^2 is proportional to the intensity of a reflection as shown in equation 1.6. However, the information about the phases of the diffracted beam is lost and must be obtained indirectly in different ways, this is known as the phase-problem. Approximations employed in this respect include Patterson synthesis, direct methods, charge flipping, anomalous dispersion, isomorphous and molecular replacement[8]. In chemical crystallography only the first two techniques are widely used although charge flipping has been rapidly gaining in popularity over the last few years.

The structure factor, a reciprocal space quantity, and the real-space presentation of the unit cell contents in form of an electron density distribution (since it is the electrons that scatter the X-rays) are related via Fourier transforms. There are two important mathematical relationship between the structure factor and the contents of a single unit cell, the structure factor equation (equation 1.7) and the electron density equation (equation 1.8) listed below.

(1.7) $$F(hkl) = \sum_{j=1}^{N} f_j \exp[2\pi i(hx_j + ky_j + lz_j)]$$

There are *N* atoms in the unit cell and the fractional coordinates of the *j*th atom are x_j, y_j and z_j. The atom has a scattering factor f_j, which is related to the number of electrons associated with the atom. *F(hkl)* is a complex number consisting of amplitude and phase and it represents a diffracted beam. This Fourier transform can be used to calculate the expected diffraction pattern for any known structure.
In order to calculate the electron density distribution ρ at every point (x, y, z) in a single unit cell from an X-ray diffraction pattern a reverse Fourier transform or Fourier synthesis according to equation 1.8 is used:

$$\rho(xyz) = \frac{1}{V}\sum_{hkl} F(hkl)\exp[-2\pi i(hx + ky + lz)] \tag{1.8}$$

where *V* is the unit cell volume and *F(hkl)* is defined as for equation 1.7. Whilst equation 1.8 shows how the electron density can in principal be computed from a diffraction pattern, in practise this is not directly achievable due to the phase problem. One method commonly used in chemical crystallography to overcome the phase problem is to solve crystal structures by direct methods. The name originates from the principle that the structure factor phases are derived from a single set of X-ray intensities, which is possible since the amplitude and phase of a structure factor are not independent of each other, but are linked via the electron density.
The following summarises the fundamental steps of a structure determination with direct methods:

(1) The observed amplitudes $|F_{hkl}|$, obtained from equation 1.6, are normalised and intensity statistics are derived to aid space group determination.
(2) Phase relationships are generated for sets of three structure factors. These phase relationships arise from the knowledge that the electron density is non-negative and consists of a random distribution of discrete atoms.
(3) The structure factors with the strongest normalised amplitudes are identified for the subsequent phase determination.
(4) Starting phases are assigned, which may be generated randomly or calculated from an approximate electron density map. This is repeated several times similar

to a Monte Carlo procedure to increase the probability of identifying a suitable starting phase.

(5) Phase determination and refinement: New phases are iteratively computed from the starting phases via the tangent formula until the phases converge to stable values. Again this step is repeated many times as correct phase values are unlikely to arise from randomly assigned starting phases.

(6) A figure of merit is calculated for each set of phases obtained in the previous step. This acts as a quality measure of the calculated phase values.

(7) The best phase set identified in step 6 is then used to calculate an initial electron density map. This map is examined in terms of the expected molecular content and peaks are assigned atomic types. Sometimes not all atoms can be found on this first electron density map. In this case phases are calculated from the known atoms and are used to calculate the next map. This is repeated until the molecular structure is fully identified.

Once some content of the ASU has been assigned the next phase of a structure determination commences, the refinement of the structure. Gradually parameters are introduced to describe the crystal structure, such as atomic positions and thermal parameters to model the movement the atoms in a molecule undergo with temperature. These parameters are incorporated in a least-squares fitting procedure during which the function 1.9 is minimised:

$$\sum w(|Y_o| - |Y_c|)^2 \tag{1.9}$$

where Y is usually either |F| or F^2, but can also be the intensity I. The subscripts o and c refer to observed and calculated data, respectively. Whilst the model is progressing, structure factor amplitudes (F_{hkl}) are calculated for this "known" structure and then compared with the measured structure factor amplitudes. In order to assess the quality of the fit, residual factors or R-indices are inspected. The most commonly used are the weighted R-factor and unweighted R-factor shown in equations 1.10 and 1.11, respectively.

(1.10) $$wR = \left[\frac{\sum w(Y_o - Y_c)^2}{\sum wY_o^2} \right]^{1/2}$$

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

Another aid to assessing the fit of the structural model to the measured data is to compute a so-called difference map:

(1.12) $$\Delta\rho(x,y,z) = \frac{2}{V_c} \sum_{hkl} (|F_o| - |F_c|) \cos[2\pi(hx + ky + lz) - \phi_c]$$

Such a difference electron density, Δρ, is useful in identifying any atoms after least-squares refinement, incorrectly placed atoms or wrongly assigned atom types. Furthermore the difference density map will be relatively featureless once the model of the crystal structure is finally complete.

1.A.3. Intermolecular Interactions

Molecules in a crystal are kept in a periodic 3-dimensional arrangement by non-covalent forces arising from intermolecular interactions. As two molecules are placed in close proximity to each other both attractive and repulsive forces act on the molecules. Hence a crystal structure is a state at which these intermolecular forces are at equilibrium. The separation of the molecules corresponds to an equilibrium distance at which attraction and repulsion are balanced so that overall the assembly represents an energetic minimum (otherwise the crystal would not be stable at ambient conditions). Compared with covalent bonding the energetic gain due to intermolecular interactions is at least one order of magnitude smaller. However, the total energy of a crystal is a summation of the intermolecular forces over all molecules in the crystal and hence even weak forces can be significant contributors to the crystal assembly. A variety of intermolecular interactions have been recognized to directly affect how molecules pack in a crystal and are classified into strong and medium to weak interactions[9,10]. Strong interactions arise through ionic forces and classical hydrogen bonds, although only the latter is of relevance to this work. Medium to weak interactions include weakly directional hydrogen bonding, π-π interactions and halogen bonding. In particular the strong intermolecular interactions lead to stable and reproducible structural arrangements such as hydrogen-bonded chains and can even become determining for the whole periodic assembly. This was recognised by the early pioneers of supramolecular chemistry[11] and later structural motifs arising from intermolecular interactions were classified as structural synthons[12]. The fields of crystal engineering and crystal structure prediction make active use of known structural synthons, e.g. to design framework structures and to inform prediction procedures, respectively.

Historically intermolecular forces have been ascribed to specific atom-atom interactions, and although there are disadvantages to this approach[13] it is still the most commonly used. Of all the intermolecular interactions the classical hydrogen bond has been studied and discussed most extensively in the literature. It is an interaction of the type X-H···Y, where the donor atom X is highly electronegative (mainly N and O) and removes electron density from the hydrogen atom. This

generates a positive charge at the hydrogen atom leading to Coulomb attraction to the electronegative acceptor atom Y. Strong hydrogen bonds are directionally specific and are the most energetic amongst neutral intermolecular interactions[14] as shown in Table 1.2.

Particularly the amide hydrogen bond is of great importance – amide groups are found in biomolecules, such as peptides and proteins – and its characteristics have been discussed extensively in the literature[15]. Many functional groups contain donor and/or acceptor atoms capable of forming strong hydrogen bonds. For molecules with multiple hydrogen bond donors and/or acceptors there are hence several options for pairing the donors with the acceptors in an intermolecular interaction.

Interaction	**Distance range X···Y [Å]**	**Angle range X-H···Y [°]**	**Energy range [kJ·mol^{-1}]**
strong HB	2.5 – 3.0	165-180	10-65[16]
weak HB	3.0 – 4.0	90-180	0.3-25[17]
π-π (S)[a]	3.5 – 4.0	n/a	0-20[18]
π-π (T)[b]	ca 5.0	n/a	
π-π (PSS)[c]	*3.4 – 3.6 / 1.6 – 1.8[19]	n/a	
halogen bond	#2.7 – 3.3	160 - 180	2-33[20] / 1-15[36]

Table 1.2: Geometry and energetic value of intermolecular interactions. The possible π-π dimers are abbreviated as S for sandwich[a], T for T-shape[b] and PSS for plane slipped stack[c]. The distance range for π-π dimers refers to the separation between the centres of mass of the ring systems. The literature is not entirely conclusive about the geometry and energy range of π-π interactions and halogen bonds and references are provided for the values listed in the table. * The two value ranges given correspond to the vertical and horizontal displacement of the benzene ring centroids. #The range applies to Cl, Br and I.

The question arises whether there is a preference for forming hydrogen bonds between specific functional group, or in other words, which type of hydrogen bond is the strongest and most reliable in terms of reproducibility. This issue has fascinated many researchers, with first attempts at establishing hydrogen bonding rules being published by Etter[21] in 1990. More recently Gilli *et al*. have studied hydrogen bonds extensively and have recently introduced the concept of the pK_a slide rule, which uses differences in pK_a for hydrogen bond acceptors and donors to assess the stability of a hydrogen bond[22]. Furthermore computer based models

have been developed ascribing a statistical likelihood[23] or propensity[24] to hydrogen bond formation for competing functional groups.

The existence of so-called weak hydrogen bonds is now generally accepted, but their importance in the formation of a crystal structure has not been fully elucidated thus far. Weak hydrogen bonds are those of the type C-H···(O, N, π) or (O, N)-H···π and their geometry is more easily deformed compared with strong hydrogen bonds (Table 1.2)[25,26]. Energetically weak hydrogen bonds are less strong than the classical hydrogen bonds, but many organic molecules are C-H "rich" facilitating several of these contacts, the sum of which may then be comparable per molecule to the energy gain due to a strong hydrogen bond. There have been many studies aimed at calculating the energetic gain of weak hydrogen bonds, but the energies obtained can vary rather dramatically. Scheiner et al. investigated this problem and concluded that many calculated energies are unreliable due to poor treatment of weak interactions[27].

Organic compounds frequently contain aromatic rings and the interaction between neighbouring rings, also called π-π stacking interaction, has been the focus of many studies[28,29]. This interaction is less directional than hydrogen bonds but it is believed to influence the three-dimensional structure of biological systems such as DNA, RNA or proteins and the packing of non-linear optical materials. The geometry of a π-π assembly is somewhat less well defined than for hydrogen bonds. The rings can be stacked perfectly parallel (sandwich stacks) or with different degrees of offset within the stacks (parallel slipped stacks). The T-shaped arrangement of two phenyl rings is also often considered as π-π interaction, although this case lies at the borderline and could alternatively be classified as a C-H···π interaction. Despite the origin of the π-π interaction still being a matter of debate[30], it has been shown that long-range interactions, particularly dispersion, are a major source of attraction between benzene rings[31]. In order to gain an energy measure for this kind of interaction benzene has often been used as a model compound and the energetic contribution has been estimated to lie in the range of 2 to 20kJ·mol^{-1}, whilst energies for the benzene dimer range from 6.7 to 10.0kJ·mol^{-1} as derived experimentally[32]. However, the diversity of stacked π-systems is much larger given the various possibilities of substitution and the size of the aromatic system. These issues have been investigated in some studies[33,34],

nevertheless the compounds chosen constitute but a small set of aromatic systems. Hence the above energy should only be considered as an indication that π-π interactions are of similar strength or slightly less energetically attractive than weak hydrogen bonds.

A further weak intermolecular interaction is the so-called halogen bond. This interaction was defined in analogy to a weak hydrogen bond of the type C-X···Y, where X is chlorine, bromine or iodine and Y is typically a oxygen or nitrogen. The role of this interaction is not yet completely understood. On one hand there are many studies confirming the directionality and predictability of this interaction[35], whilst on the other hand the energy gain has been computed as being small compared to other interactions[36].

1.A.4. Polymorphism

The discovery of polymorphism goes back to Mitscherlich (1822-1823)[37], but extensive interest in this area only started rising in the second half of the last century. In 1965 McCrone defined the term 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"[38], hence polymorphism can be understood as the phenomenon of a chemical compound occurring in at least two crystalline phases. Although McCrone's definition of a polymorph appears clear at first glance, the literature contains many accounts discussing shortcomings of this definition and offering alternatives or more specialised definitions. In the context of the work discussed in this thesis polymorphic systems are those where more than one crystal structure has been identified for one chemical substance of the same molecular connectivity.

McCrone also argued that any substance can exhibit polymorphism and that it is a matter of time and money spent that determines how many polymorphs can be identified for a given compound. Nowadays it is believed that polymorphism is an inherent feature of the solid state and that the lack of observed polymorphism may be due to the transition point being higher than the melting point or influenced by external conditions not yet explored[39] or understood. Despite these conclusions exhaustive polymorph screens have not been routinely performed except within the past decade in the pharmaceutical industry and polymorphs are still found by serendipity in most cases.

Polymorphism is of great importance to the pharmaceutical industry since the physicochemical properties of polymorphs can be dramatically different. Some of the main properties affected are processability, solubilities, dissolution rate, hygroscopicity and stability[40]. Polymorphism has been shown to have effects on bioavailability and toxicity of active pharmaceutical ingredients[41,42] and hence it is imperative to gain an understanding of polymorphism and how to control polymorphic systems. Interestingly the properties of polymorphic forms can be very different even if the polymorphs share the same synthon, i.e. parts of the intermolecular interactions are equivalent, which demonstrates that there are more complex forces acting further complicating structure property relationships.

Polymorphic systems are nevertheless very useful for comparison studies since the only variable is packing of identical molecules whereas different compounds with equivalent crystal packing contain information about the influence of substituent exchange.

For this work a distinction between polymorphism is made as follows: one type is concerned with changes in intermolecular assembly, the other with variation in intramolecular conformation. The former applies if the constituent molecules of two polymorphic forms remain unchanged geometrically and the polymorphism is ascribed to a variation of the crystal packing. The change in crystal structure is then affected by different intermolecular connectivity, i.e. different intermolecular interactions are established. This could result in a completely revised intermolecular assembly or partial reorganisation of "building blocks" (e.g. hydrogen bonded synthons) common to both polymorphs. One example of the latter is polytypism, which refers to structures with differently packed identical layers – effectively this is one-dimensional polymorphism.

The other category of polymorphism considered here is that of so-called conformational polymorphism. In this phenomenon the polymorphic structures contain different conformers of the same molecule, i.e. the molecule has undergone alteration of its shape. X-ray crystallographic measurements have demonstrated over the years that bond lengths and bond angles undergo little variation in crystal structures enabling the tabulation of standard bond lengths[43] and van der Waals radii[44]. Torsional angles on the other hand have been found to vary widely, which lies at the basis of conformational polymorphism. Rotation about at least one covalent bond in the molecule alters its geometric shape and hence altered close packing in the crystal is observed. However, it is a matter of debate what defines truly different conformations: Does a rotation of e.g. 0.1° about one bond constitute a different conformer or could such a change arise from thermally induced motion? On the other hand if such small changes occur at multiple rotatable bonds in the molecule the overall geometry will be significantly changed. The distinction between individual conformers should hence be based on individual cases.

In terms of conformational polymorphism and its causes the views in the literature present a chicken and egg situation. It is not clear what affects what: does the

conformational change occur first and is hence the cause of different packing in the crystal or do changes in packing cause the molecular geometrical change?

This consideration invites the contemplation of the energies involved in crystal packing. The measure of the overall cohesion, i.e. intermolecular forces, in a crystal is the lattice energy, the calculation of which is discussed in more detail at the end of this chapter. Differences in lattice energies of two polymorphs are generally small[45,46] meaning that the potential energy surface exhibits either a broad and shallow minimum or that there are several accessible minima of similar energy. It also implies that the total of all intermolecular forces can be very similar for different sets of interaction and close contacts.

In conformational polymorphs an additional aspect is the difference in conformational energy between (a) the observed conformers and (b) the lowest energy conformation for the molecule. If the latter is non-negligible then it contributes to the individual lattice energies. Comparing two conformational polymorphs, the total energy difference is composed of the lattice energy difference and the conformational energy difference. Again the total energy differences across conformational polymorphs are found to be small and the possibility of intramolecular and intermolecular energy compensation has been investigated[47,48]. These small energy differences are one of the complications encountered in crystal structure prediction procedures, where frequently many low energy crystal structures are calculated making it difficult to predict any particular polymorphic form[49].

As already mentioned above the energetic value of intermolecular interactions lies in the region of 4 to 15 kcal·mol^{-1} for strong hydrogen bonds and in the range of 0.5 to 4 kcal·mol^{-1} for weak and dispersive interactions, respectively. On the other hand the energy required to affect rotation about single bonds varies between 1 and 3 kcal·mol^{-1} for unrestricted systems but can be as high as 8 kcal·mol^{-1} for hindered systems[48]. This means that the inter- and intramolecular changes associated with conformational polymorphs are on a similar energy scale. Hence the energy penalty of accommodating an energetically less favourable conformer in a crystal lattice can be offset by energetically more stable intermolecular interactions. This energy compensation can thus lead to overall similar lattice energies between polymorphs. However, not all polymorphs show the same

energetic behaviour and polymorphs that vary significantly in total energy have been reported[50].

Besides X-ray crystallography, thermal analysis is a valuable tool in studying polymorphism and can provide an overview of the thermodynamic relationships between the different crystalline phases.

The relative stability of polymorphs is sometimes judged based on the crystal density[51]. As a rule of thumb, the denser the crystal the lower its internal energy, hence the energetically most stable polymorph is the one were efficient packing of the molecules in the crystal maximizes the density. But there are many exceptions to this rule; for example the stability of strongly hydrogen bonded systems cannot be assessed with this simple model.

1.A.5. Structural Similarity and XPac

The analysis of structural patterns in a systematic manner requires a tool that enables the comparison of crystal structures and identifies common structural features. A number of programs have been developed using various procedures[52,] amongst which the program XPac is unique in identifying similarity not only for isostructural compounds, but also for compounds that share 0-, 1-, or 2-dimensonal structural features such as dimers, rows or layers of molecules, respectively. The method as implemented in XPac is thoroughly documented in the literature[53] and hence only the most important steps will be summarised here.

In XPac the description of a crystal structure is independent of its space group and unit cell information. A representative cluster of molecules is chosen, which consists of a central kernel molecule and a shell of close neighbours, which, in most cases are symmetry-generated from the kernel molecule but, where relevant may include symmetry unrelated molecules if the structure has Z' >1. The cut-off for the shell is determined by the sum of van der Waals radii plus t=1.5Å using intermolecular interatomic distances. Each molecule is represented as an ordered set of points (OSP), for which a number of atoms of the molecule are selected. When comparing two crystal structures the OSPs for the two different molecules should reflect a similar arrangement and are then regarded as a corresponding ordered set of points (COSP). Internal coordinates, comprising distances, angles and torsion angles, are generated between all the molecules in the individual clusters of the COSP and by comparing lists of these coordinates, similarity is established. The criteria for the existence of a similar geometrical arrangement, called supramolecular construct (SC), are defined by cut-off values applied to the mean absolute difference, δ_{ang}, δ_{dhd} or δ_{tor}, for each pair of lists. In this way XPac offers a method, which performs the systematic comparison of multiple crystal structures in a fraction of the time a purely visual inspection would take. In fact a number of publications are available that report the analysis of whole compound libraries.

The results from the XPac analysis are visualised with a structural similarity diagram with construction principles borrowed from set theory and an example diagram is provided in Figure 1.5. It should be noted that XPac performs pairwise

comparisons of crystal structures and provides a brief characterisation of any possible SCs, such as its dimensionality and base vector(s). The existence of such a construct is then verified via visual inspection. Finally this information is manually 'translated' into a similarity diagram by taking into account all pairwise relationships found with XPac. Gelbrich et al have outlined the detailed rules of constructing such a graph diagram in order to describe the relationships in a series of closely related carbamazapines[54] and hence only the very basics of interpreting such plots will be further summarised here. There are four levels of similarity and these are plotted in ascending order, so that zero-dimensionality is found in the lowest layer at the bottom of the diagram, three-dimensional SCs are situated at the top and two- and one-dimensional constructs in the middle section with one-dimensional SCs placed lower than two-dimensional features. In set theory commonality is expressed as nodes, which are drawn as circles and boxes in Figure 1.5. Each box at the top level represents a unique crystal assembly, symbolised by the letters A to E in Figure 1.5, so that structures with 3D similarity are grouped together in a single box as is shown for structures C and D in Figure 1.5. Circles at lower levels indicate SCs found in at least two crystal structures. Nodal dependencies are represented by arrows[55] in Figure 1.5 and express the relationship between the crystal structures and the constructs. To identify all structures with a certain SC the branches starting at the respective node are followed in an upward direction and similarly to find all features present in one or more crystal structures, the connecting lines are traced in a downward direction.

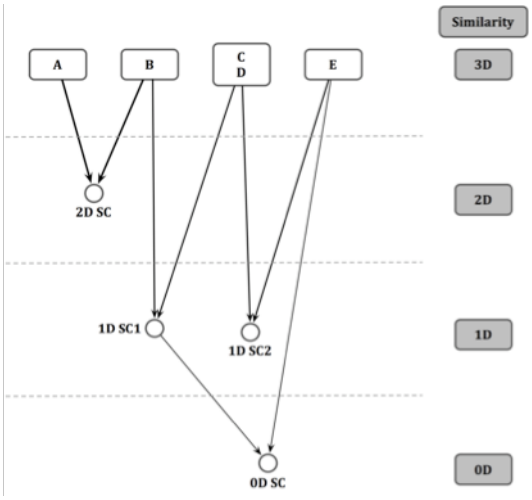


Figure 1.5: An example of a structural similarity diagram.

A comparison involving many crystal structures can potentially yield a considerable number of SCs. The similarity diagram should then be constructed so that the crystal structures and SCs are ordered with the least crossing connecting lines.

It should also be noted that the orientation of the similarity diagram presented above is arbitrary and the plot can easily be turned on its side or upside down if a different style is preferred, as demonstrated in a recent publication[56].

At this stage it is worthwhile to explain some of the terminology applied to describe the different degrees of similarity in this work. The distinct advantage of the XPac method over other related approaches is the possibility of identifying reoccurring crystal packing motifs automatically in any dimension and importantly, without bias towards any kind of intermolecular interactions, hence the term supramolecular construct is not to be confused with supramolecular synthons mentioned earlier.

In order to describe these reoccurring features appropriate terms have to be used that allow easy distinction between SCs with strong intermolecular adhesion and those without any notable interaction between neighbouring molecules in the SC. Isostructurality (at various dimensions) and homeostructurality as defined by Kálmán[57] have been used extensively in the literature for describing similar crystal structure assemblies. Especially the former term has been widely adopted, although some dispute might be caused when declaring two crystal structures of different compounds as isostructural, e.g. in cases with an unusually large variation of cell parameters or molecular dimensions, and sometimes isostructurality is confused with isomorphism. Hence in this work any of these terms were avoided and the following descriptions for the varying degrees of similarity were used instead. A 0D construct is considered as such only if there is any notable interaction present between two molecules and is then called a dimer[58]. In principle 0D constructs are not only restricted to dimeric arrangements, e.g. a reoccurring catemer would also present a 0D SC, but 0D constructs of this kind were not encountered during this work. Compounds with 3D similarity are described as crystal structure analogues or as having analogous packing. More diversity is present at the 1D and 2D level, where on one hand features may be caused by strongly interacting moieties, and on the other hand there are those SCs where aggregation cannot easily be attributed to intermolecular interactions, and it remains ambiguous whether steric factors or medium-to-weak interactions or a mixture of both are responsible for a certain packing mode. The literature contains various different approaches to

classification of such spatial arrangements. Kálmán *et al.* prefer to define similar packing in 1 or 2 dimensions as 1D and 2D isostructurality, respectively[59]. Pidcock *et al.* break down crystal assemblies into molecular boxes and introduce a numerical notation based on the exposed surface area of aggregates of these boxes[60]. This box model is actually very convenient for describing a molecular shape. Accordingly the principal dimensions of a molecule are expressed as long, medium and short or synonymously as length, width and depth, respectively. This provides an easy reference for the directionality of any SC with respect to the molecule.

Another approach to classifying spatial arrangements of molecules is that of Etter, who assigned graph sets to the topology of hydrogen-bonded molecules[61]. This latter method is somewhat more practical than Pidcock's box aggregate approach since the graph sets contain information about the nature of the feature, e.g. whether a chain is present. The box aggregate model on the other hand, is less intuitive as it requires a cross-reference to the code of a box aggregate or visual inspection of the feature to gain knowledge of the spatial arrangement of the molecules within the feature. However, the graph set method is only applicable to hydrogen-bonded systems, whilst the box aggregate model assesses the spatial assembly of the molecules irrespective of intermolecular interactions.

For this work a slightly different approach was used since the compounds studied here can facilitate hydrogen bonding, but XPac assesses similarity purely based on geometrical arrangements of molecules. In order to easily distinguish between the reoccurring SCs with and those without strong intermolecular interactions, the terminology for 1D and 2D constructs was derived as follows. The molecular arrangement in a construct is described based on the spatial object it represents. Hence a 1D arrangement of non-interacting molecules is termed a row or stack, whereas 1D SCs of hydrogen-bonded molecules are described as chains, helices, tapes or ladders. Table 1.3 summarises the terminology used in this work for describing common 1D structural motifs. It should be stressed here that this terminology is solely based on the constructs observed in this study and it may not be appropriate for systems with a greater variety of structural features or different intermolecular connectivity.

From Table 1.3 it can be seen that no distinction is made between SCs propagating along the length or width of the molecules, but different terms are assigned to the SCs with the direction of propagation being parallel to the depth of the molecules. This is because the compounds studied here contain an aromatic ring and the ring plane vectors coincide with the length and width of the molecule. Hence a SC along the depth of the molecules could potentially facilitate π-π interactions and such an arrangement is thus assigned an independent term.

Direction of Propagation	***Number of HBs***	***Name of 1D SC***
Non-interacting molecules		
length	0	row
width	0	row
depth	0	stack
Hydrogen-bonded systems		
length	1	chain
width	1	chain
depth	1	helix
length	2	tape
width	2	tape
depth	2	ladder

Table 1.3: Terminology used to describe common 1D packing features.

The descriptions in Table 1.3 are by no means novel: chains and rows or ribbons and tapes are almost interchangeably used in the literature. However, it was felt that a stricter categorization is required for this work since different connectivity is implied when considering the geometrical shape of the actual objects.

The same concept was used to derive the terminology for 2D constructs, although fewer terms are used. A 2D feature with at least two non-parallel hydrogen bonds is called a net. Those 2D SCs with no or one hydrogen bond are termed layers. 2D constructs containing only parallel hydrogen bonds are also called layers. In contrast to 1D similarity, the terms for 2D SCs were assigned irrespective of the directions of propagation with respect to the molecular dimension. Although it may at times be more convenient to describe a layer as a stack of chains or a row of helices for example. It should be pointed out that this terminology was derived purely for the purpose of this work. As such it should not be considered as a

complete terminology applicable to the general comparison or description of crystal structures.

1.B. Theoretical Calculations

The field of theoretical chemistry deals with problems such as predicting chemical properties and is relatively young compared with the more traditional subject areas of inorganic, physical and organic chemistry. One cornerstone of theoretical chemistry is undoubtedly the derivation of the Schrödinger equation in 1926[62]. However, the invention of computers and the constant development of hardware and software were equally essential to the success and accuracy of theoretical calculations, as we know it today. The vast increase in computing power over the last 20 years together with the ever decreasing cost for computational resources, have enabled theoretical chemists to extend quantum mechanics calculations from atomic or highly symmetric systems to more advanced problems and in doing so making theoretical methods available to the wider chemistry community.

In the following some of the methods employed for this project will be discussed. There are two areas of this project that rely on or involve theoretical calculations.

Firstly, the analysis of the conformational freedom of the molecules under investigation can be achieved through gas phase *ab initio* calculations. The theory and some case studies of these calculations are presented in the two subsequent sections.

Secondly, estimating lattice energies requires some energy calculation in the solid state. For this a number of 3D periodic algorithms are available, e.g. the *ab initio* software package CRYSTAL06[63], the plane wave pseudopotential package CASTEP[64] or the semi-classical density sums method available in the OPiX suite of programs[65]. The latter was found to be most suitable for this project and the underlying principles will be introduced in the final section of this chapter.

1.B.1. Gas phase calculations – theory of ab initio computations

In its most general form the Schrödinger's equation specifies how a quantum state of a physical system changes with time.

(1.13) $$i\hbar\frac{\partial\Psi}{\partial t} = \hat{H}\Psi$$

Equation 1.13 is the time-dependent version of the Schrödinger equation, but if the potential energy of a system is independent of time it can be written in its more commonly known form:

(1.14) $$\hat{H}\Psi = E\Psi$$

where $\hat{H}$ is the Hamiltonian operator, Ψ is the wavefunction (eigenfunction) and E the energy of the system (eigenvalue). Rearranging equation 1.14 enables the calculation of the energy of a system as:

(1.15) $$E = \frac{\int \Psi \hat{H} \Psi^* d\tau}{\int \Psi \Psi^* d\tau}$$

with integration over all space. The Hamiltonian consists of kinetic and potential energy terms as shown in equation 1.16:

(1.16) $$\hat{H} = -\frac{\hbar^2}{2}\sum_A \frac{1}{M_A}\nabla_A^2 - \frac{\hbar}{2m}\sum_a \nabla_a^2 - e^2\sum_A\sum_a \frac{Z_A}{r_{Aa}} + e^2\sum_A\sum_B \frac{Z_A Z_B}{r_{AB}} + e^2\sum_a\sum_b \frac{1}{r_{ab}}$$

where $$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

The first two terms in equation 1.16 represent the kinetic energies of the nuclei and the electrons, respectively (capital letters indicate nuclei and lower case electrons). The last three terms stand for the attractive potential energy of the individual nucleus-electron interactions, and the repulsive potential energy of the nucleus-nucleus and electron-electron interactions, respectively.

Equation 1.14 is the description of a system in a stationary state and chemical moieties, such as molecules or atoms, can be described by the solution to this equation. Many approaches to solving the Schrödinger equation have been developed amongst which *ab initio* methods, meaning from first principles in quantum mechanics, have been widely used. At this stage it should be noted that the Schrödinger equation cannot be solved exactly except for the hydrogen atom;

for any other chemical system approximations have to be made. The simplifications made are the Born-Oppenheimer approximation, the LCAO approximation and the Hartree-Fock independent electron approximation; furthermore the variation theorem is applied.

The Born-Oppenheimer approximation states that the electron motion is decoupled from the motion of the nucleus and as a consequence the nuclear and electronic contributions are separated so that the first term in equation 1.16 is neglected and the fourth term is considered as a constant contribution. This treatment ignores relativistic effects and leads to an electronic Hamiltonian and the electronic Schrödinger equation.

The second approximation in *ab initio* methods is the linear combination of atomic orbital (LCAO) approximation. The wavefunction of a polyelectronic system can be expressed as the product of all one electron molecular wavefunctions, ϕ_I, as follows:

(1.17) $$\Psi = \prod_{i=1}^{n} \phi_i$$

The LCAO simplification then defines the molecular wavefunctions (or molecular orbitals) as a linear combination of the known atomic orbital funtions, χ_k:

(1.18) $$\phi_i = \sum_k c_{ik}\chi_k$$

where c_{ik} are the expansion coefficients. The atomic orbitals are well described by the more accurate Slater type functions or by the more commonly used Gaussian type functions. Equation 1.19 displays the linear combination of Gaussian functions used to represent atomic orbital functions:

(1.19) $$\chi_k = \sum_{l=1}^{X} d_{lk}\varphi_l(\alpha_{lk})$$

Equation 1.19 contains d_{lk}, the coefficient of the primitive Gaussian function, φ_l, the exponent α_{lk} of the Gaussian function and X, the number of functions used in

the expansion. The expansion coefficients from equation 1.18 are estimated via the secular equations:

(1.20) $$\sum_k c_{ik}(H_{lk} - \varepsilon_i S_{lk}) = 0$$

where H is a one-electron Hamiltonian, ε_I is the energy of one particular electron in orbital I and S_{lk} is the so called overlap integral. The secular equations are usually formulated in a determinantal form. Non-trivial solutions can only be obtained if the condition in equation 1.21 holds true.

(1.21) $$\det\left|H_{lk} - \varepsilon_i S_{lk}\right| = 0$$

The quality of a wavefunction described in this way can be assessed according to the variation theorem, which states that the energy calculated from an approximated wavefunction will always be higher than the true energy. Hence the better the wavefunction the lower will be the resulting energy. This condition is applied to the expressions for the energy.

The third approximation is that of the Hartree-Fock independent electron principle. In this simplification the Fock operator is introduced, fulfilling the eigenvalue equations:

(1.22) $$\hat{F}_i \phi_i = \varepsilon_i \phi_i$$

The Fock operator is an effective one-electron Hamiltonian and expresses the sum of three individual contributions to the energy of the system. Firstly, the core term is the kinetic and potential energy contribution of each electron moving in the field of the nuclei:

(1.23) $$\hat{H}^{core}(1) = -\frac{1}{2}\nabla_1^2 - \sum_{A=1}^{M}\frac{Z_A}{r_{1A}}$$

The second term is the Coulomb operator, which corresponds to the average potential due to an electron in spin orbital ϕ_j:

(1.24) $$\hat{J}_j(1) = \int d\tau_2 \phi_j(2) \frac{1}{r_{12}} \phi_j(2)$$

And finally the term for the exchange operator, which is required to account for the Pauli principle and is defined in terms of its effect when acting on the spin orbital ϕ_i:

(1.25) $$\hat{K}_j(1)\phi_i(1) = \left[\int d\tau_2 \phi_j(2) \frac{1}{r_{12}} \phi_i(2) \right] \phi_j(1)$$

The numbers 1 and 2 indicate where an integral involves the coordinates of a single or of two electrons, respectively. Combining equations 1.23 to 1.25 then yields the Fock operator:

(1.26) $$\hat{F}_i(1) = \hat{H}^{core}(1) + \sum_{j=1}^{N} \left\{ \hat{J}_j(1) - \hat{K}_j(1) \right\}$$

The Roothan-Hall equations translate the Hartree-Fock equations 1.22, which if fully written out are in integro-differential form, into a matrix form:

(1.27) $$\mathbf{FC} = \mathbf{SCE}$$

where F is the Fock matrix, C the coefficient matrix, E the energy matrix containing the orbital energies and S is the overlap matrix. The process of solving the Roothan-Hall (and Hartree-Fock) equations is an iterative process (self-consistent field approach), since values that have to be calculated occur on both sides of the equations. The above calculations produce a set of one-electron orbital energies ε_i, which are individual sums of the core interaction of one electron and the Coulomb and exchange interactions of this electron with other electrons in the system. The

total energy of a system in the ground state with N electrons in N/2 orbitals can then be estimated from these one-electron energies as shown in equation 1.28:

(1.28) $$E = \sum_{i=1}^{N} \varepsilon_i - \sum_{i=1}^{N/2} \sum_{j=1}^{N/2} (2J_{ij} - K_{ij})$$

Whilst the energy of the ground state so obtained will be as minimal as achievable with the calculations according to the variation theorem, the Hartree-Fock method uses one-electron contributions, neglecting electron correlation. More advanced *ab initio* methods, such as Møller-Plesset perturbation theory and coupled cluster theory, include approximations of electron correlation and are often referred to as post-Hartree-Fock methods. Another problem of *ab initio* calculations is cost of computation, with even small organic molecules requiring very time-consuming calculations. This has led to the development of approximate methods over the last century, also known as semi-empirical methods. Some of the most popular and widely applied approaches nowadays are the MNDO and AM1 methods. They differ from *ab initio* methods in the way the electrons of the system are described and make use of some experimental parameters. The most time-demanding step in *ab initio* calculations is the evaluation of the integrals. Semi-empirical methods reduce the computational cost by neglecting or approximating some of these integrals. Only the valence electrons are taken into account explicitly whilst the core electrons become part of the nuclear description. The implementation of experimental parameters enables semi-empirical methods to yield accurate results for certain properties as long as the source of parameterisation and the experiment are related[66]. Despite all this *ab initio* techniques remain the closest approximation to the real system since they use the full Hartree-Fock equations and the complete Hamiltonian. With *ab initio* quantum mechanics it is possible to calculate thermodynamic and structural quantities as well as properties that depend on the electronic distribution. Even chemical reactions and excited states can be investigated.

Another approach to the electronic structure of atoms and molecules is Density Functional Theory (DFT), which has become a popular alternative to *ab initio* Hartree-Fock computational methods over the last couple of decades[67]. The theory

is based on the principle that the ground-state energy can be described by using the electron density ρ only, which means that the electron density and the energy of a system are directly connected, making DFT a wave-function-less method.

This concept is much older than one would expect; the basic notion is almost as old as quantum mechanics and was explored by Thomas, Fermi, Dirac and Wigner[68,69,70] in the late 1920s. In fact methods based on DTF have been considered useful models in solid-state physics ever since the 1950s. Finally in 1964 the proof for the one-to-one correlation of ground-state energy of an electronic system and its electron density was delivered by Hohenberg and Kohn[71] and in 1998 Walther Kohn was awarded the Nobel Prize in Chemistry[72] (shared with John A. Pople) for his remarkable work in the field.

What makes DFT so attractive is its sole dependence on the electron density, and in principle DFT could provide the exact total energy including electron correlation whereas HF energies are uncorrelated. Unfortunately the exact form of the functional connecting energy with electron density is not known and hence the aim of DFT methods is to build functionals that give satisfying approximate results. According to Hohenberg and Kohn the energy functional (expressed as square brackets) is calculated as the sum of two terms:

$$E[\rho(\mathbf{r})] = \int V_{ext}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + F[\rho(\mathbf{r})] \tag{1.29}$$

The first term describes the interaction of the electrons with an external potential $V_{ext}(\mathbf{r})$, such as the Coulomb interaction with the nuclei. The second term in equation 1.29 contains the contributions from the kinetic energy of the electrons and any interelectronic interactions. However, $F[\rho(\mathbf{r})]$ is not known, so that in 1965 Kohn and Sham[73] developed a formalism where $F[\rho(\mathbf{r})]$ is split into three functional terms:

$$F[\rho(\mathbf{r})] = E_{KE}[\rho(\mathbf{r})] + E_{H}[\rho(\mathbf{r})] + E_{XC}[\rho(\mathbf{r})] \tag{1.30}$$

The $E_{KE}[\rho(\mathbf{r})]$ is defined as the kinetic energy of a system of non-interacting electrons with the same density $\rho(\mathbf{r})$ as the real system. The second term in equation 1.30, $E_H[\rho(\mathbf{r})]$, refers to the Hartree electrostatic energy calculated from the classical interaction between two charge densities. The term $E_{XC}[\rho(\mathbf{r})]$ contains

contributions from exchange and correlation. Using the Kohn-Sham approach the ground-state energy of a system with N electrons can be written as:

$$E[\rho(\mathbf{r})] = \sum_{i=1}^{N} \psi_i(\mathbf{r}) \left(-\frac{\nabla^2}{2} \right) \psi_i(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \int \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}|_2} d\mathbf{r}_1 d\mathbf{r}_2 + E_{XC}[\rho(\mathbf{r})] - \sum_{A=1}^{M} \int \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} \rho(\mathbf{r}) d\mathbf{r} \quad (1.31)$$

including the electron-nucleus interaction with the last term. Equation 1.31 defines the exchange-correlation energy functional $E_{XC}[\rho(\mathbf{r})]$, which is related to the exchange-correlation energy V_{XC} by

$$V_{XC}[\mathbf{r}] = \left(\frac{\delta E_{XC}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} \right) \quad (1.32)$$

In the Kohn-Sham approach the density $\rho(\mathbf{r})$ is taken as the sum of the square moduli of a set of one-electron orthonormal orbitals as follows:

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} |\psi_i(\mathbf{r})|^2 \quad (1.33)$$

Hence when combined with the variational condition the one-electron Kohn-Sham equations are obtained:

$$\left\{ -\frac{\nabla_1^2}{2} - \left(\sum_{A=1}^{M} \frac{Z_A}{r_{1A}} \right) + \int \frac{\rho(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2 + V_{XC}[\mathbf{r}_1] \right\} \psi_i(\mathbf{r}_1) = \varepsilon_i \psi_i(\mathbf{r}_1) \quad (1.34)$$

The Kohn-Sham equations are solved self-consistently. This means the initial density is guessed and used in equation 1.34 yielding a set of orbitals that are then used to calculate an improved density. This process is iterated until convergence is reached. Finally the total energy is calculated in accordance with equation 1.31.
Basically DFT methods differ from each other in the form of the functional for the exchange-correlation energy and there are plenty of methods to choose from. They

range from Local Density Approximations (LDA) via Gradient Corrected Methods (or GGA for Generalized Gradient Approximation) to Hybrid Methods, such as the Adiabatic Connection Model (ACM) and the Becke 3 parameter functional (B3). The last functional is incorporated in the popular B3LYP method, which is an acronym indicating that the B3 hybrid[74] has been used in combination with the LYP correlation[75].

The development of new functionals is aimed at improving the quality of the results of a DFT calculation, but the pathway is not yet clear. So far very different functionals have delivered results[76] of comparable significance. The quality of a DFT given result should thus be assessed using experimental or high-quality wave mechanics for similar systems. Nevertheless DFT methods provide an alternative to more time consuming *ab initio* HF methods due to reduced dimensionality (computing time at the present grows with N_{at}^2 or N_{at}^3 using DFT, whereas for HF the factor is about $e^{\alpha N_{at}}$ ($\alpha \approx 1$)[77]). Hence DFT also becomes applicable to larger systems, up to sizes of $N_{at} \approx 100 - 200$.

1.B.2. Conformational Analysis

The methods described in the previous section can be used to calculate properties of the molecular energy. For small organic molecules, particularly if they are biologically active, the property of most interest is that of conformation. The degrees of freedom of a molecule increase with the number of rotatable bonds and knowledge of the energetically favourable conformations a molecule can adopt is useful for prediction procedures and the establishment of structure-property relationships.

A conformational analysis samples the conformational potential energy surface of a molecule and seeks to relate changes in conformation to energy and differences in properties. The former is achieved with a conformational search, which is aimed at identifying the preferred, i.e. lowest energy, conformations of a molecule. Different conformations are calculated for a molecule and their energy subsequently minimized, so that the relative energies can be compared against each other. The starting conformations can be achieved in various ways to which

systematic search algorithms and random approaches belong. Systematic searches produce potential energy surface maps by varying the torsion angles under investigation in a stepwise manner. Initially the rotatable bonds are identified and the respective torsion angles defined. Then the torsion angle is changed in a fixed increment and the energy minimized for this conformation. In this way a grid search of the conformational space is performed and minima on the potential energy surface become accessible. The smaller the increments the finer the sampling of the grid, but this dramatically increases the scale of the computations necessary as does an increase of the number of rotatable bonds, since:

$$\text{Number of conformations} = \prod_{i=1}^{N} \frac{360}{\theta_i} \tag{1.35}$$

where θ_i is the increment chosen for the rotatable bond i. Hence conformational analysis of highly flexible molecules using *ab initio* methods is not feasible.

Random approaches on the other hand, generate different conformations as the name suggests in a random way. In this method conformational space is explored by either changing the Cartesian coordinates of the constituent atoms or the relevant torsion angles. The minimised conformation generated in each iteration is then tested for novelty and only stored if not previously calculated. Whilst a systematic search is finite, random searches do not have a natural endpoint and hence the sampling of the potential energy surface may be fragmentary. Only systematic searches were considered in this work.

The second objective of a conformational analysis is to relate the information obtained from a conformational search with properties of the molecule. Often a series of chemical species are examined to assess the influence of a particular substituent on conformations and rotational energy barriers of the molecule. The rotational energy can be extracted from the computed potential energy surface and tested for correlation with substituent descriptors. For example the amide group has been tested for rotational stability with respect to cis/trans isomerism[78,79] fueled by its biological relevance. Furthermore conformational analysis has been employed in the study of polymorphism[80] and is of great relevance to medicinal chemistry and drug design[81].

1.B.3. Lattice Energy Calculations

Molecules assemble in a 3-dimensional array due to favourable energetics. One variable associated in particular with the assembly of molecules in a crystal lattice is the lattice energy, which is generally defined as the energy released when gaseous molecules form a solid. Since interactions between molecules are lost as a solid sample converts to the gaseous phase the lattice energy is also a measure of the total of the intermolecular forces in a crystal. The calculation of the lattice energy of a crystal structure can be achieved in a number of ways ranging from relatively easy methods using empirical force fields to more complex methods utilising *ab initio* calculations. For this work the Pixel method by Gavezzotti was used throughout and hence only this approach will be discussed here further.

The Pixel method or semi-classical density sums (SCDS) method belongs to the so-called full density models because it incorporates the numerical integration over the whole molecular electron density (as opposed to more historical atom-centred approaches). For a full description of the Pixel formulation the reader is referred to the literature[82] . Here only the key points of this method will be summarised.

The all-underlying principle is utilising the electron density of a molecule in a crystal. The charge density is readily available through the calculation of a molecular orbital wavefunction and can then be partitioned into pixels. A relatively large number of pixels are required to fully describe the electron density (overall the total electron number has to be reproducible). However, the number of pixels has to be reduced to minimise the computational cost and this is achieved by creating super-pixels, i.e. cubic cluster of (n × n × n) pixels. The parameter n is termed the condensation level. Furthermore a cut-off is applied to exclude super-pixels of negligible charge.

The total lattice energy, of an organic molecular crystal can be split into four parts according to intermolecular association: the Coulombic, polarisation, dispersion and repulsion terms as outlined in equation 1.36:

(1.36) $$E_{tot} = E_{Coul} + E_{pol} + E_{disp} + E_{rep}$$

The Coulomb interaction is calculated as the sum over all pixel-pixel, pixel-nucleus and nucleus-nucleus pairs. For this each pixel is assigned a charge q_x, which is defined as the electron density ρ_x in a volume V_x (centered at point x):

(1.37) $$q_x = \rho_x V_x$$

These pixels are associated with atoms in the molecule dependent on pixel-nucleus distances.

Next the Coulombic interaction between two molecules A and B is considered. In molecule A the atomic nuclei are located at positions j (=[x_j, y_j, z_j]) and carry a charge Z_j. Analogously for molecule B the nuclei are positioned at points m (=[x_m, y_m, z_m]) with charge Z_m. The associated pixels are found at points k (=[x_k, y_k, z_k]) with a charge of q_k in A and at points i (=[x_i, y_i, z_i]) with a charge of q_i in B, respectively. Hence the total electrostatic potential between A and B is:

(1.38) $$E_{elst,AB} = \frac{1}{4\pi\varepsilon_0}\left\{\sum_i q_i\left[\sum_k \frac{q_k}{R_{ik}} + \sum_j \frac{Z_j}{R_{ij}}\right] + \sum_m Z_m\left[\sum_k \frac{q_k}{R_{km}} + \sum_j \frac{Z_j}{R_{im}}\right]\right\}$$

where R is the separation between the atomic nuclei, the distance between the pixels and the respective distance of each pixel from the nuclei of A and B.

The electrostatic field ε exerted by molecule A on molecule B causes polarisation, α, of the pixels (at point i) in B and vice versa. The linear polarisation energy takes the form:

(1.39) $$E_{pol,i} = -\frac{1}{2}\alpha_i \varepsilon_i^2$$

The definition of the pixel polarisability makes use of atomic charges and polarisabilities:

(1.40) $$\alpha_i = \frac{q_i}{Z_{atom}}\alpha_{atom}$$

In order to account for contributions at very short distances a damping function, d_i, is applied:

(1.41) $$E_{pol,i} = -\frac{1}{2}\alpha_i(\varepsilon_i d_i)^2 \quad \text{with} \quad d_i = \exp-\left[\frac{\varepsilon_i}{(\varepsilon_{\max} - \varepsilon_i)}\right]$$

Summation over all points i yields the total polarisation energy for the perturbation of molecule B by molecule A.

The combination of these polarisabilities with an oscillator strength, E_{OS}, enables the modelling of the dispersion energy. Again, a damping function is introduced to avoid singularities due to very short inter-pixel distances.

The dispersion energy term is hence formulated as:

(1.42) $$E_{disp,AB} = -\frac{3}{4}\sum_{k,A}\sum_{i,B} E_{OS} f(R)\frac{\alpha_k\alpha_i}{(4\pi\varepsilon^\circ)^2(R_{ik})^6} \quad \text{with} \quad f(R) = \exp\left[-\left(\frac{D}{R_{ik}} - 1\right)^2\right]$$

where D is the so called damping threshold distance.

Finally the repulsion energy between two molecules A and B is derived from the overlap of their electron densities, S_{AB}, as:

(1.43) $$E_{rep,AB} = K(S_{AB})^\gamma \quad \text{with} \quad S_{AB} = \sum_{k,A}\sum_{i,B}\left[\rho_k(A)\rho_i(B)\right]V$$

At this point it is convenient to examine the repulsive component of the lattice energy more closely. Electrostatic attractive forces between molecules bring them into close contact whereas the repulsive forces between the outer electrons of the molecules counteract the approach. Both attraction and repulsion energies increase sharply at short intermolecular distances and converge to zero at large distances. Hence the intermolecular separation observed in a crystal is the distance at which attractive and repulsive forces balance so that a further approach of the molecules would result in steeply increased repulsion and further separation would be accompanied by a significant decrease in attraction. In either case the total energy will be less favourable than at the equilibrium distance in the

crystal. In the Pixel algorithm the repulsion and Coulombic energy both depend on overlap or penetration effects of the electron density envelope[83]. Hence as the Coulombic energy increases the repulsion energy will become more destabilising. Furthermore the calculation of the repulsion energy with the Pixel method is not ideal for systems in which the interacting electron densities overlap extensively, such as hydrogen bonded systems. However, this effect is only of particular concern if second row elements are involved[84]. Hence in this work the repulsion energy will not be extensively discussed.

It should be pointed out that the energy terms above refer to the interaction of two molecules and that the determination of the total energy of a crystal involves the summation of these terms over all molecules in the ensemble.

There are four empirical parameters that can be adjusted by the user for Pixel calculations, namely ε_{max} (E_{pol}), D (E_{disp}), K and γ (both E_{rep}). Typical values used by the program developer can be found in the literature[85].

The total lattice energy computed in this way can then be partitioned into molecule-molecule contributions making this a useful tool for identifying significant intermolecular interactions. Particularly in polymorphs where the total energy difference is small, this method allows the easy recognition of energetic differences in the intermolecular framework. The suitability of this approach to calculating the lattice energies of molecular crystals has been thoroughly tested in the literature[86].

1.C. References

[1] S. Dunstan, ***"Principles of Chemistry"***, Camelot Press Ltd., London and Southampton (1968).

[2] M. Ladd, R. Palmer, ***"Structure Determination by X-ray Crystallography"***, Kluwer Academic/Plenum Publishers, New York (2003).

[3] Angles different to 90˚ apply to the trigonal/hexagonal and cubic crystal systems.

[4] This data was obtained by searching the Cambridge Crystallographic Database V5.31 (including updates November 2009, February 2010 and May 2010) and yielded 124095 crystal structures (approximately 24% of 512881 total entries). The following filters were applied to the search: organic structures only; excluding: metal atoms, B, Si, As, Te, Se, At; R≤0.75; no errors, no ions; no structures from powder diffraction; 3D coordinates determined.

[5] These are the space group statistics as obtained from the CCDC website and refers to the 498,123 CSD entries present on 1st January 2010, for which the space group is fully defined.

[6] A. I. Kitaigorodskii, ***"Organic Chemical Crystallography"***, Consultants Bureau Enterprises, Inc., New York (1961) – *translated from the Russian original published by the Press of the Academy of Sciences, Moscow (1955)*.

[7] As the diffraction pattern is measured over a finite time and is obtained from a whole crystal the resulting contents of a single unit cell are averaged over time and space, i.e. all unit cells in the crystal structure.

[8] W. Clegg, A. J. Blake, R. O. Gould, P. Main, ***"Crystal Structure Analysis Principles and Practice***", Oxford University Press, Oxford (2001).

[9] C. A. Hunter, *Angewandte Chemie International Edition*, **43**, 5310-5324 (2004).

[10] H.-J. Schneider, *Angewandte Chemie International Edition*, **48**, 3924-3977 (2009).

[11] In recognition of their pioneering work in supramolecular chemistry Donald J. Cram, Jean-Marie Lehn and Charles J. Pedersen received the Nobel Prize in Chemistry in 1987:
http://nobelprize.org/nobel_prizes/chemistry/laureates/1987

[12] G. R. Desiraju, *Angewandte Chemie International Edition in English*, **34**(21), 2311-2327 (1995).

[13] J. D. Dunitz, A. Gavezzotti, *Angewandte Chemie International Edition*, **44**, 1766-1787 (2005).

[14] ed. G. R. Desiraju, ***"The Crystal as a Supramolecular Entity"***, in *Perspectives in Supramolecular Chemistry Volume 2*, John Wiley & Sons Ltd., Chichester, England, (1995).

[15] A series of articles published by R. Taylor and O. Kennard probes the amide-amide hydrogen bond in detail based on existing structures in the CSD: (a) R. Taylor, O. Kennard, W, Versichel, *Journal of the American Chemical Society*, **105**, 5761-5766 (1983); (b) R. Taylor, O. Kennard, W, Versichel, *Journal of the American Chemical Society*, **106**, 244-248 (1984); (c) R. Taylor, O. Kennard, W, Versichel, *Acta Crystallographica*, **B40**, 280-288 (1984); (d) R. Taylor, O. Kennard, W, Versichel, *Accounts of Chemical Research*, **17**, 320-326 (1984).

[16] C. B. Aakeröy, K. R. Seddon, *Chemical Society Reviews*, **22**(6), 397-407 (1993).

[17] G. R. Desiraju, T. Steiner, ***"The Weak Hydrogen Bond"***, Oxford University Press, Oxford, UK (1999).

[18] M. O. Sinnokrot, C. D. Sherrill, *Journal of Physical Chemistry A*, **110**, 10656-10668 (2006).

[19] E. A. Meyer, R. K. Castellano, F. Diederich, *Angewandte Chemie International Edition*, **42**(11), 1210-1250 (2003).

[20] P. Politzer, P. Lane, M. C. Concha, Y. Ma, J. S. Murray, *Journal of Molecular Modelling,* **13**, 305-311 (2007).

[21] M. C. Etter, *Accounts of Chemical Research*, **23**(4), 120-126 (1990).

[22] P. Gilli, L. Pretto, V. Bertolasi, G. Gilli, *Accounts of Chemical Research*, **42**(1), 33-44 (2009).

[23] H. Y. Ando, L. Dehaspe, W. Luyten, E. Van Craenenbroeck, H. Vandecasteele, L. Van Meervelt, *Molecular Pharmaceutics*, **3**(6), 665-674 (2006).

[24] P. T. A. Galek, L. Fábián, W. D. S. Motherwell, F. H. Allen, N. Feeder, *Acta Crystallographica*, **B63**, 768-782 (2007).

[25] G. R. Desiraju, *Accounts of Chemical Research*, **24**, 290-296 (1991).

[26] T. Steiner, *Angewandte Chemie International Edition*, **41**, 48-76 (2002).

[27] S. Scheiner, ***"Hydrogen bonding. A theoretical perspective."***, Oxford University Press, Oxford (1997).

[28] C. A. Hunter, J. K. M. Sanders, *Journal of the American Chemical Society*, **112**, 5525-5534 (1990).

[29] C. G. Claessens, J. F. Stoddart, *Journal of Physical Organic Chemistry*, **10**, 254-272 (1997).

[30] A slightly different opinion on π-π stacking: S. Grimme, *Angewandte Chemie International Edition*, **47**, 3430-3434 (2008).

[31] S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, K. Tanabe, *Journal of the American Chemical Society*, **124**(1), 104-112 (2002).

[32] S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, K. Tanabe, *Journal of the American Chemical Society*, **124**(1), 104-112 (2002).

[33] M. O. Sinnokrot, C. D. Sherrill, *Journal of Physical Chemistry A*, **107**(41), 8377-8379 (2003).

[34] A. Gavezzotti, *Journal of Chemical Theory and Computation*, **1**, 834-840 (2005).

[35] A few comprehensive reviews and articles on this matter are listed here: (a) P. Auffinger, F. A. Hays, E. Westhoff, P. Shing Ho, *Proceedings of the National Academy Sciences of the United States of America*, **101**(48), 16789-16794 (2004); (b) P. Metrangolo, H. Neukirch, T. Pilati, G. Resnati, *Accounts of Chemical Research*, **38**, 386-395 (2005); (c) P. Metrangolo, T. Pilati, G. Resnati, *CrystEngComm*, **8**, 946-947 (2006); (d) P. Metrangolo, F. Meyer, T. Pilati, G. Resnati, G. Terraneo, *Angewandte Chemie International Edition*, **47**, 6114-6127 (2008); (e) P. Metrangolo, G. Resnati, *Science*, **321**, 918-919 (2008).

[36] A. Gavezzotti, *Molecular Physics*, **106**(12-13), 1473-1485 (2008).

[37] E. Mitscherlich, *Abhandlungen der Königlichen Preußischen Akademie der Wissenschaften zu Berlin*, 43-48 (1822-1823).

[38] W. C. McCrone, ***"Polymorphism"***, in *Physics and chemistry of the organic solid state*, Vol. 2, 725-767, Wiley Interscience, New York (1965).

[39] N. N. Sirota, *Crystal Research and Technology*, **17**, 661-691, (1982).

[40] D. Giron, *Journal of Thermal Analysis and Calorimetry*, **64**, 37-60 (2001).

[41] J. Haleblian, W. McCrone, *Journal of Pharmaceutical Sciences*, **58**(8), 911-929 (1969).

[42] D. Giron, *STP Pharma Pratiques*, **4**, 330-340 (1988).

[43] F. H. Allen, O. Kennard, D. G. Watson, L. Brammer, A. G. Orpen, R. Taylor, *Journal of the Chemical Society, Perkin Transactions 2*, 12, S1-S19 (1987).

[44] http://www.ccdc.cam.ac.uk/products/csd/radii/table.php4#name

[45] A. Gavezzotti, G. Filippini, *Journal of the American Chemical Society*, **117**, 12299-12305 (1995).

[46] S. L. Price, *International Reviews in Physical Chemistry*, **27**(3), 541-568 (2008).

[47] A. Gavezzotti, *Journal of the American Chemical Society*, **113**, 4622-4629 (1991).

[48] A. Nangia, *Accounts of Chemical Research*, **41**(5), 595-604 (2008).

[49] S. L. Price, *Accounts of Chemical Research*, **42**(1), 117-126 (2009).

[50] J. Starbuck, R. Docherty, M. H. Charlton, D. Buttar, *Journal of the Chemical Society, Perkin Transactions 2*, 2, 677-691 (1999).

[51] J. D. Dunitz, *Pure and Applied Chemistry*, **63**(2), 117-185 (1991).

[52] (a) A. Kálmán, L. Fábián, G. Argay, G. Bernáth, Z. Gyarmati, *Journal of the American Chemical Society*, **125**, 34-35 (2003); (b) A. V. Dzyabchenko, *Acta Crystallographica*, **B50**, 414-425 (1994); (c) L. Fábián, A. Kálmán, *Acta Crystallographica*, **B55**, 1099-1108 (1999); (d) P. Verwer, F. J. J. Leusen, *Computer simulations to predict possible crystal polymorphs*, Wiley-VCH, New York, 1988. (e) B. P. van Eijck, J. Kroon, *Journal of Computational Chemistry*, **18**(8), 1036-1042 (1997); (f) E. L. Willighagen, R. Wehrens, P. Verwer, R. de Gelder, L. M. C. Buydens, *Acta Crystallographica*, **B61**, 29-36 (2005); (g) H. R. Karfunkel, J. H. Noordik, F. J. J. Leusen, R. J. Gdanitz, G. Rihs, *Journal of Computational Chemistry*, **14**(10), 1125-1135 (1993); (h) R. de Gelder, R. Wehrens, J. A. Hageman, *Journal of Computational Chemistry*, **22**(3), 273-289 (2001); (i) J. A. Chisholm, S. Motherwell, *Journal of Applied Crystallography*, **38**, 228-231 (2005).

[53] T. Gelbrich, M. B. Hursthouse, *CrystEngComm*, **7**(53), 324-336 (2005).

[54] T. Gelbrich, M. B. Hursthouse, *CrystEngComm*, **8**(6), 448-460 (2006).

[55] Simple connecting lines equally express this information.

[56] S. L. Childs, P. A. Wood, N. Rodriguez-Hornedo, L. S. Reddy, K. I. Hardcastle, *Crystal Growth & Design*, 2009, **9(4)**, 1869-1888.

[57] A. Kálmán, L. Párkányi, G. Argay, *Acta Crystallographica*, **B49**, 1039-1049 (1993).

[58] Each coordination sphere consists of approximately 14 molecules, which are closely packed. The likelihood of finding a pair of molecules with similar spatial arrangements in two related crystal structures is quite high compared with higher dimensionality SCs since less internal coordinates are considered in the comparison algorithm.

[59] L. Fábián, A. Kálmán, *Acta Crystallographica*, **B60**, 547-558 (2004).

[60] E. Pidcock, W. D. S. Motherwell, *Crystal Growth and Design*, **4**(3), 611-620 (2004).

[61] M. C. Etter, J. C. MacDonald, J. Bernstein, *Acta Crystallographica*, **B46**, 256-262 (1990).

[62] http://nobelprize.org/nobel_prizes/physics/laureates/1933/schrodinger.html

[63] R. Dovesi, V. R. Saunders, R. Roetti, R. Orlando, C. M. Zicovich-Wilson, F. Pascale, B. Civalleri, K. Doll, N. M. Harrison, I. J. Bush, P. D'Arco, and M. Llunell, **CRYSTAL06** (CRYSTAL06 User's Manual. University of Torino, Torino, 2006)

[64] S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson, M. C. Payne, *Zeitschrift für Kristallographie*, **220**(5-6), 567-570 (2005).

[65] A.Gavezzotti, ***"OPiX, A computer program package for the calculation of intermolecular interactions and crystal energies"***, University of Milano (2007).

[66] G. H. Grant, W. G. Richards, ***"Computational Chemistry"***, in *Oxford Chemistry Primers*, Oxford University Press, Oxford (1995).

[67] T. Ziegler, *Chemical Reviews*, **91**, 651-667 (1991).

[68] L. H. Thomas, *Mathematical Proceedings of the Cambridge Philosophical Sociecty*, **23**(5), 542-548 (1927).

[69] E. Z. Fermi, *Zeitschrift für Physik,* **48**(1-2), 73-79 (1928).

[70] P. A. M. Dirac, *Mathematical Proceedings of the Cambridge Philosophical Sociecty*, **26**(3), 376-385 (1930).

[71] P. Hohenberg, W. Kohn, *Physical Reviews A*, **136**(3), 864-871 (1964).

[72] http://nobelprize.org/chemistry/laureates/1998/kohn-autobio.html, From *Les Prix Nobel. The Nobel Prizes 1998,* Editor Tore Grängsmyr, [Nobel Foundation], Stockholm, 1999.

[73] W. Kohn, L. J. Sham, *Physical Review A*, **140**(4), 1133-1138 (1965).

[74] A. D. Becke, *Journal of Chemical Physics*, **98**(7), 5648-5652 (1993).

[75] C. Lee, W. Yang, R. G. Parr, *Physical Review B*, **37**(2), 785-789 (1988).

[76] F. Jensen, ***"Introduction to Computational Chemistry"***, John Wiley & Sons Ltd, Chichester, UK (1999).

[77] W. Kohn, A. D. Becke, R. G. Parr, *Journal of Physical Chemistry*, **100**, 12974-12980 (1996).

[78] A. T. Hagler, L. Leiserowitz, M. Tuval, *Journal of the Amercian Chemical Society*, **98**(15), 4600-4611 (1976).

[79] S. Ilieva, B. Hadjieva, B. Galabov, *Journal of Organic Chemistry*, **67**, 6210-6215 (2002).

[80] J. Bernstein, A. T. Hagler, *Journal of the American Chemical Society*, **100**(3), 673-681 (1978).

[81] K. A. Brameld, B. Kuhn, D. C. Reuter, M. Stahl, *Journal of Chemical Information and Modeling*, **48**(1), 1-24 (2008).

[82] A series of papers explains the philosophy and the method of the Pixel approach: (a) A. Gavezzotti, *CrystEngComm*, **5**(76), 429-438 (2003); (b) A. Gavezzotti, *CrystEngComm*, **5**(77), 439-446 (2003); (c) A. Gavezzotti, *Zeitschrift für Kristallographie*, **220**, 499-510 (2005).

[83] J. D. Dunitz, A, Gavezzotti, *Angewandte Chemie International Edition*, **44**, 1766-1787 (2005).

[84] A. Gavezzotti, ***"Molecular Aggregation"***, Oxford University Press, Oxford, UK (2007).

[85] Table 2 in A. Gavezzotti, *Zeitschrift für Kristallographie*, **220**, 499-510 (2005).

[86] Gavezzotti has shown in many publications how the Pixel method is applicable to organic crystals: (a) A. Gavezzotti, *Structural Chemistry*, **16**(3), 177-185 (2004); (b) F. Cozzi, S. Bacchi, G. Filippini, T. Pilati, A. Gavezzotti, *Chemistry – A European Journal*, **13**(25), 7177-7184 (2007); (c) C. J. Eckhardt, A. Gavezzotti, *Journal of Physical Chemistry B*, **111**, 3430-3437 (2007); (d) A. Gavezzotti, *CrystEngComm*, **10**, 389-398 (2008); (e) A. Gavezzotti, *Acta Crystallorgaphica*, **D64**, 905-908 (2008); (f) F. Cozzi, S. Bacchi, G. Filippini, T. Pilati, A. Gavezzotti, *CrystEngComm*, **11**(6), 1122-1127 (2009); (g) A. Gavezzotti, *Acta Crystallographica*, **B66**, 396-406 (2010).

CHAPTER 2: THE COMPOUND LIBRARY

2.A. Introduction

This project explores how differences in molecular structure affect the solid state assembly of small organic compounds. Gaining a better understanding of intermolecular interactions and the balance between electrostatic and steric forces should offer insight into crystal structure formation and may assist in associating solid state properties with specific structural features. In order to establish widely applicable structure property relationships a large proportion of chemical space has to be considered. Chemical space is vast and complex hence this field is still in its early stages and variation in solid state structure is not well understood. Recent studies in this area have been aimed at identifying the pivotal factors in the relationship between molecular and crystal structure[1].

To be able to understand and possibly even predict the variation in the crystal structures of two different chemical compounds one has to be able to associate the changes made on the molecule with the changes observed in the crystal packing. Any deductions formed may only be specific to this one comparison and thus need to be probed and refined systematically so as to achieve generalization and applicability to other systems.

A suitable system for such a systematic study comprises a group of closely related chemical compounds that share a common core structure and only differ in a small number of substituents. The core should be easy to functionalise to yield as many similar compounds as possible so that any findings may be quantified. The exchange of substituents should not drastically vary the overall shape of the molecule and the introduction of functional groups should be kept to a minimum. Close packing, which is affected by molecular shape and intermolecular interaction, should hence be achievable in a similar manner for these similar molecules whereas differences may be directly linked to the contribution of a specific substituent.

The class of compounds forming the basis for this project is the family of mono-substituted acylanilides in general and mono-substituted acetanilides in particular. The following sections focus on the chemical structure of the compound library, nomenclature used in this work and application of some of the library members.

2.B. The chemical structure and composition of the compound library

The molecular core of the acylanilides shown in Figure 2.1 effectively consists of two functional parts, providing the following potential interaction sites: the amide group can facilitate both intra- and intermolecular hydrogen bonding and the aromatic ring provides a further site for intermolecular interactions, such as π-π stacking. Also shown in Figure 2.1 are the points of substitution **X** and **Y**, which can occur at either end of the core molecule and in the case of **X** the substitution can either be at the ortho, meta or para position on the phenyl ring.

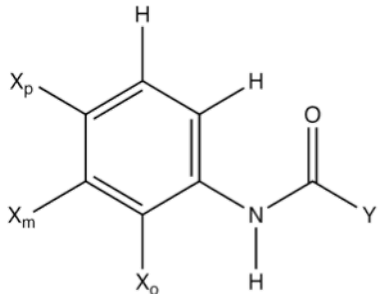


Figure 2.1: The acylanilide core with the positions of substitution, X and Y. Only one position X is substituted at any one time while the other two positions are occupied by hydrogen atoms.

Variation of the substituent at the amide carbon atom (position **Y** in Figure 2.1) introduces different parent compounds, the acylanilides, whereas mono-substitution at the phenyl ring[2] forms a subset of compounds for each parent.

Substitution at the amide nitrogen atom was not considered for this project since this would have eradicated the hydrogen bonding potential of the amide group also.

The general acylanilide shape can be considered as that of a two-dimensional object, especially since the aromatic ring and the amide group form a conjugated system (resonance structures shown in Figure 2.2) and according to hybridization theory the molecule should be near co-planar.

Such acylanilides are easy to obtain via acylation of the respective substituted aniline and over the course of twelve months Terry Threlfall kindly produced more than 400 compounds following the substitution matrix shown in Table 2.1.

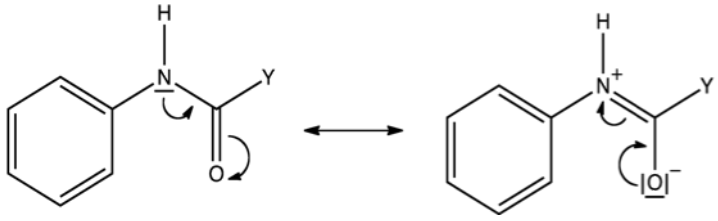


Figure 2.2: Resonance structures for the amide group as present in acylanilides. Hydrogen atoms are shown explicitly for amide nitrogen atoms only.

Table 2.1 also shows which synthesis yielded crystalline material and where structure determination from single crystal X-ray diffraction was possible.

X \ Y	H	CH_3	C_2H_5	C_3H_7	$C(Me)_3$	CF_3	OMe	OEt	NH_2	Cl
H	❄	👍	👍	◇?	👍	◇?	👎	👎	👍	
p-CF_3O	⌛	👍	👍	⌛	👍	👍	👍	👍	👍	
p-EtO	❄	👍	👍	❄	👍	👍	👍	⌛	◇?	
p-MeO	👍	👍	⌛	👍	👍	👍	👍	👍	👍	
p-Et	⌛	👍	👍	👍	👍	👎	👎		◇?	
p-^{i}Pr	⌛	👍	⌛		👍	❄	👍		👍	
p-^{t}Bu	⌛	👍	👍	⌛	👍	👍	👍		⌛	
p-CH_3	👍	👍	👍	👍	👍	👍	👍	👎	👍	
p-F	❄	👍	👍	❄	👍	◇?	👍	❄	👍	❄
p-Cl	👍	👍	👍	👍	👍	👍	👎	👎	👍	👎
p-Br	👍	👍	👍	👍	👍	👍	👍	👍	👍	👎
p-I	👍	👍	👍	◇?	👍	◇?	👍	👍	👎	
p-CF_3	👍	👍	👍	👍	👍	👍	👍	👍	👍	⌛
p-CN	❄	👍	👍	👍	👍	👍	👍	👍	✕	
p-NO_2	❄	👍	👍	👍	👍	👍	❄	👍	✕	
p-OH	⌛	👍	⌛		⌛					
p-COOH		👍	👎		⌛	👍	⌛			
p-COOMe		👍			👍	👍	❄			
p-COOEt	⌛	👍			👍	👍	⌛			
p-COMe	❄	👍			◇?	⌛	⌛			
p-NH_2		👍								
p-NHCOMe	⌛	◇?			◇?	◇?	⌛			
p-OCOMe		👍								
p-SO_2OMe		👎								
p-SO_2OEt		👎 / ◇?								
p-SO_2NH_2		👍								

X \ Y	H	CH_3	C_2H_5	C_3H_7	$C(Me)_3$	CF_3	OMe	OEt	NH_2	Cl
m-MeO	👍	⌛	⌛	⌛	👍	⌛	👎	⌛	👍	
m-Et	⌛	⌛	⌛	⌛	👍	⌛	⌛	⌛	👍	
m-CH_3	⌛	👍	👍	❄	👍	👎	👍	👎	👍	
m-F	👍	👍	👍	👍	👍	👍	👍	⌛	👍	❄
m-Cl	👍	👍	👍	👍	👍	👍	👎	❄	👍	⌛
m-Br	👍	👍	👍	👎	👍	👍	👎	👎	👍	❄
m-I	👍	👍	👍	⌛	👍	◇?	👎	⌛	◇?	
m-CF_3	❄	👍	❄	⌛	👍	👍	👎	❄	👍	❄
m-CN	❄	👍	👍	⌛	👍	👍	👎	👍	⌛	
m-NO_2	❄	👍	⌛	⌛	👍	👍	👍	👎	✕	
m-OH	⌛	👍	👍		👍	⌛	⌛			
m-COOH		❄	⌛		⌛					
m-COOEt	❄	👍	👍	⌛	👍	👍	👍		⌛	
m-COMe	❄	👍	👍	👍	⌛	👍	⌛		👍	
m-NH_2		👎[1]			⌛					
m-NHCOMe		◇?	✕		❄					
m-OCOMe		❄	❄							

X \ Y	H	CH_3	C_2H_5	C_3H_7	$C(Me)_3$	CF_3	OMe	OEt	NH_2	Cl
o-CH_3	❄	👍	👍	👍	👍	👍	👎	❄	❄	
o-F	❄	👍	⌛	👍	👍	👍	⌛	⌛	👍	👎
o-Cl	✕	👍	👍	👍	👍	👎	👎	❄	👍	❄
o-Br	👎	👍	👍	👍	👍	👍	👎	⌛	👍	👎
o-I	👍	👍	⌛	⌛	👍	?	👍	⌛	👍	
o-CF_3	👍	👍	👍	👍	👍	👍	👍	❄	👍	⌛
o-CN	✕	👍	👎	👎	👍	⌛	👍	👍	⌛	
o-NO_2	❄	👍	⌛	❄	👍	👍	❄	⌛	⌛	
o-^{t}Bu	⌛	👍	❄		❄	👍	❄	⌛	⌛	

X \ Y	H	CH_3	C_2H_5	C_3H_7	$C(Me)_3$	CF_3	OMe	OEt	NH_2	Cl
2-pyridyl	👍	👎			❄	❄	❄		👎	
3-pyridyl	👍	👎			❄	👎	❄		❄	
4-pyridyl	❄	👍	❄		❄	❄	❄		❄	
2-pyrimidyl	❄	👍			❄		❄		❄	

👍 = crystal structure determined

? = no solution obtained yet

❄ = crystallised, awaiting data collection

👎 = not desired reaction product

✕ = crystal structure determination failed[2]

⌛ = are crystallising

*[1] This compound reacted further to yield the doubly acylated product **Me:m-NHCOMe**.*
[2] Failure due to one of the following: micro-crystalline material, no diffraction observable or undesired product.

Table 2.1: Matrix used for synthesis of mono-substituted acylanilides. The symbols in each cell indicate whether synthesis yielded crystalline material and whether structure determination resulted in the desired product.

The major limiting factor for structure determination was crystallization time, in this context meaning the time from the initial reaction to occurrence of crystalline material (For a further discussion of this matter and detailed synthesis *c.f.* Threlfall et al.[3]). In summary, out of 403 reactions, crystalline material was obtained for 333 samples from which 269 crystal structure determinations originated contributing 212 structures to the acylanilides library. So far only six of 333 crystalline samples have failed to yield a structure due to poor diffraction. 58 samples are still awaiting data collection. Furthermore 37 crystal structures of undesired products were

determined and their crystallographic data have been deposited on the eCrystals archive[4]. A full list of deposition numbers can be found in Appendix 2-1.

Originally only acetanilides (**Y** = CH_3) were considered, which is reflected in the near-complete respective column of the compound matrix. The systematic study of the crystal structures of this sub-family forms the major part of this thesis with Chapters 4 and 5 discussing the results obtained. However, the investigation was soon extended to other acylanilides to probe the influence of substitution at the amide carbon atom on intermolecular hydrogen bonding. Conformational flexibility was increased by longer alkyl residues (e.g. C_3H_7), steric hindrance created by bulkier substituents (e.g. ^{t}Bu), the electronic character altered (e.g. OMe) and competing interaction sites introduced (e.g. NH_2). Some of the findings from the systematic analysis of the crystal structures of this extended compound library are presented in Chapter 6. The selection of substituents **X** can be divided into three categories: (1) Rotationally symmetric substituents, (2) rotationally asymmetric substituents and (3) substituents with hydrogen bonding functionality. However, a limiting factor for the choice of **X** was commercial availability of the mono-substituted aniline. Following the principle of a systematic series in the more traditional chemical sense, the substituents included alkyl and alkoxy residues of variable length and branching as well as the halogen atoms (categories 1 and 2). Also functional groups (e.g. hydroxy, amino, cyano) were chosen to investigate the competition between hydrogen bonding sites of varying proximity (category 3). Furthermore the compound matrix for synthesis contains pyridine and sulfonyl derivatives, which change the functionality of the aromatic ring significantly. These are part of a separate study not discussed in this work and are only shown for completeness.

In addition to synthesis, the CSD was searched for crystal structures of any of the ten acylanilides shown in Table 2.1 yielding a total of 71 structures[5]. Of these, 33 crystal structures were re-determined at low temperature (120K) during this project to obtain parameters of increased accuracy, so that 38 structures were included as found on the CSD (a full list of structures and references can be found in Appendix 2-2). Tables 2.2 to 2.4 summarise the compound library at the time of writing as follows: circles refer to newly determined crystal structures, empty stars represent structures only found on the CSD and filled stars denote re-

determined structures. Some of the acylanilides were found to be polymorphic and this is indicated by multiple symbols within one cell. It should be noted here that the objective of this project was not to screen for polymorphism in any of the compounds, but multiple crystal structures were simply observed during synthesis and/or crystallisation or were identified from the literature.

Hence by combining high throughput crystal structure determination and database mining the acylanilide library contained 255 unique crystal structures at the time of writing. Of these, 64 crystal structures belong to the acetanilide family, with 44 structures determined during this project (27 newly identified and 17 re-determined structures, which have been deposited on the eCrystals archive *c.f.* Appendix 2-3 for deposition details) and 20 other acetanilide structures only found on the CSD.

In this way, an extensive wealth of crystallographic data was obtained in a very short time enabling several structure comparisons: Within each parent series, variation in crystal packing can be assessed not only for the different types of phenyl substituents but also for the position of substitution. Furthermore, for a specific substituent **X** at position **s**, crystal structures can be compared across the set of parents **Y**. These comparisons form the basis for the final chapter. It is worth mentioning that the compound library is by no means complete and extension is possible in many ways as long as the molecular structures remain comparable (i.e. the chemical nature of the compound is maintained). This matter is discussed in more detail in subsequent chapters.

PARA	H:p-X	Me:p-X	Et:p-X	Pr:p-X	^{t}Bu:p-X	CF_3:p-X	OMe:p-X	OEt:p-X	NH_2:p-X
X \ Y	H	CH_3	C_2H_5	C_3H_7	$C(CH_3)$	CF_3	OCH_3	OC_2H_5	NH_2
H	☆	☆	●		★			☆	★
F		★	●		●		●		●
Cl	★	★	●	●	★	●			●
Br	●	● ★	★	●	●	●	●	●	●
I	●	●	●		●		●	●	
CF_3	●	●	●	●	●	●	●	●	●
OCF_3		●	●		●	●	●	●	●
CH_3	●	★ ☆	●	●	★	●	●		★
C_2H_5		●	●	●	●				
^{i}Pr		●			●		●		●
^{t}Bu		●	●		●	●	●		
CN		●	●	●	●	●	●	●	
NO_2	☆	●	★	●	●	●		●	
NH_2		★							
OH		★ ☆ ☆							
OCH_3	★	★		●	★	●	●	●	●
OC_2H_5		★	●		●	●	●		
OC_3H_7		☆							☆
$O^{i}Pr$		☆							
OBu		●							
COOH	☆	★		☆		●			
COOMe		●			●	●			
COOEt		●			●	●			
COMe		●							
OCOMe		● ☆							
^{c}Pr		☆							
COOOH		☆	☆	☆					
$CBrCH_2$		☆							
NMe_2		☆					☆		
SF_5		☆							
NHCOY					☆	☆			

Table 2.2: The library of crystal structures of para-substituted acylanilides including the unsubstituted parent acylanilides. The legend is as follows:
☆ = found on CSD; ● = new determination; ★ = redetermination

META	H:m-X	Me:m-X	Et:m-X	Pr:m-X	^{t}Bu:m-X	CF_3:m-X	OMe:m-X	OEt:m-X	NH_2:m-X
Y / X	H	CH_3	C_2H_5	C_3H_7	$C(CH_3)$	CF_3	OCH_3	OC_2H_5	NH_2
F	●	●	●	●	●	●	●		●
Cl	●	★	●	●	★	●			● ●
Br	●	● ☆	●		●	●			●
I	●	●	●		●				
CF_3		●			●	●			●
CH_3		★	●		★		●		● ☆
C_2H_5					●				●
CN		●	●		●	●		●	
NO_2		★			●	●	●		
NH_2									
OH		★	●		●				
OCH_3	●				●				●
COOH		☆							
COOMe									
COOEt		●	●		●	●	●		
COMe		●	●	●		●			● ●
OCOMe									
NHCOY		●							
SF_5		☆							

Table 2.3: The library of crystal structures of meta-substituted acylanilides. The legend is as follows:
☆ = found on CSD; ● = new determination; ★ = redetermination

ORTHO	H:o-X	Me:o-X	Et:o-X	Pr:o-X	^{t}Bu:o-X	CF_3:o-X	OMe:o-X	OEt:o-X	NH_2:o-X
Y / X	H	CH_3	C_2H_5	C_3H_7	$C(CH_3)$	CF_3	OCH_3	OC_2H_5	NH_2
F		●		●	●	●			●
Cl		★ ☆	●	●	★				● ●
Br		★	★	●	●	●			●
I	●	●			●		●		● ●
CH_3	☆ ☆	● ☆	●	●	★	●			☆
CF_3	●	●	●	●	●	●	●		●
CN		★			●		●	●	
NO_2		★			●	●			
COOH		☆							
COMe		☆							
NHCOY	☆	☆					☆		
OH						☆			
^{t}Bu		●				●			

Table 2.4: The library of crystal structures of ortho-substituted acylanilides. The legend is as follows:
☆ = found on CSD; ● = new determination; ★ = redetermination

2.C. Nomenclature

With such a variety of compounds the use of a compound identifier that is simple yet intuitive and fully descriptive of the substitution pattern is vital for quick unambiguous identification of the chemical in question. Such identifiers were derived based on Figure 2.1 and have the general syntax **Y:s-X**, using the chemical formula or its common abbreviation (e.g. CH_3 is commonly abbreviated as Me) as notation for the substituents **X** and **Y**, so that the acylanilide parent is specified first, followed by the position of aryl substitution and then the ring substituent. If the chemical formula is used for **X** or **Y** then the order of elements in this abbreviation should be such that the element directly bonded to the core molecule is written first and subsequent elements should be in an order that resembles the connectivity where possible. The principle is demonstrated in Figure 2.3.

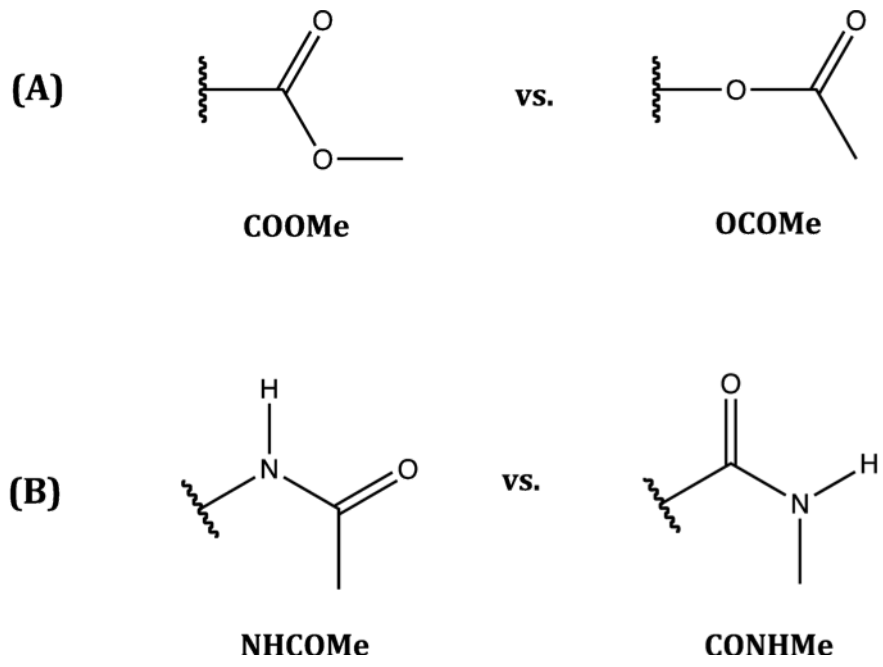


Figure 2.3: Two examples (A) and (B) of isomeric substituents and their labels as used in this study.

Finally **s** indicates the position of substitution on the phenyl ring and can be o, m or p for ortho-, meta- and para-substitution, respectively. A summary of the parent acylanilides chosen for this project can be found in Table 2.5.

Y	*IUPAC* name	Name used in study	Compound ID
H	N-phenylformamide	formanilides	**H**
CH_3	N-phenylacetamide	acetanilides	**Me**
C_2H_5	N-phenylpropanamide	propionanilides	**Et**
C_3H_7	N-phenylbutanamide	butyrylanilides	**Pr**
^{t}Bu	2,2-dimethyl-N-phenylpropanamide	trimethylacetanilides	**^{t}Bu**
CF_3	2,2,2-trifluoro-N-phenylacetamide	trifluoroacetanilides	**CF_3**
OCH_3	methyl phenylcarbamate	methyl phenylcarbamates	**OMe**
OC_2H_5	ethyl phenylcarbamate	ethyl phenylcarbamates	**OEt**
NH_2	N-phenylurea	N-phenylureas	**NH_2**
Cl	phenylcarbamic chloride	phenylcarbamic chlorides	**Cl**

Table 2.5: Summary of parent acylanilides synthesised as part of this study including IUPAC names and short abbreviations used for the compound identifiers.

According to this procedure the compound identifier for para-chloro acetanilide is **Me:p-Cl**. This can also be applied in a more general manner so that all meta-substituted trifluoroacetanilides are abbreviated as **CF_3:m-X**, a range of ortho-methyl acylanilides are denoted as **Y:o-Me** and if no distinction is made between the position of substitution on the phenyl ring, e.g. iodo formanilides, then the generalised syntax **H:s-I** is used. In cases where more than one crystal structure is known, the differentiation between polymorphs is made based on differing crystal systems, e.g. **Me:p-Me** has a monoclinic and an orthorhombic form, so that the first letter of the crystal setting is included in parenthesis, e.g. **Me:p-Me(m)** and **Me:p-Me(o)**[6]. Should the crystal setting be the same for several polymorphic structures, then a simple integer count indicates the number of varying crystal assemblies, e.g. **Me:p-OCOMe(m1)** and **Me:p-OCOMe(m2)** where the numbering follows the sequence of occurrence. The pure parent acylanilide, i.e. X=H, is expressed as **Y:H** leaving **s** undefined since all ring positions are occupied by hydrogen only, e.g. methylcarbamate is abbreviated as **OMe:H**.

2.D. Chemical Applications

The library introduced above contains such a large number of chemical compounds that compiling an exhaustive review of the uses and applications of each individual one is beyond the scope of this project. This actually raises an interesting almost philosophical issue regarding the nature of publishing research data and the way chemical studies are conducted from a crystallographer's point of view. Traditionally a chemical investigation[7] deals with very few compounds at a time and focuses on synthesis and characterization, whilst the study is seldom truly application-driven. This is not because there has been a lack of interest in the latter, but producing and analysing a 'new molecule' is time consuming and application tests often involve specialized equipment different to synthesis, so that traditionally the search for a new chemical and the search for a new material are focus of separate research areas without much overlap. Technical advances over the past few decades have introduced automation to both the synthetic and analytical processes so that nowadays chemicals can be studied rapidly[8] and the more recent literature demonstrates this trend containing frequently studies of sets of compounds[9] rather than isolated cases. On the other hand this increased output generates the problem of how to publish all this data[10]. Furthermore the challenge of an ever increasing need for new functional material, whether in environmental science, healthcare or nanotechnology to name but a few[11], requires the 'modern' chemist to employ an interdisciplinary approach. Again the recent literature reflects this development to some extend; it is not sufficient any longer to publish a crystal structure on its own, but value is added through analysis of properties, such as electric or magnetic behaviour, in the search for chemicals with enhanced performance over materials in present use.

However, it is still early days and there is little literature available concentrating on a whole class of compounds and their relevance as functional material. There are a few exceptions, e.g. the Merck Index[12], the Dictionary of Substances and their Effects (DOSE)[13] and various Encyclopedias (such as the Encyclopedia of Agrochemicals[14]), but these are specialist reference works mainly aimed at industrialists and access is seldom granted via library subscriptions.

Another aspect comes into consideration here: assuming a sizable set of compounds was available and that tests for utilizable properties have been performed, only 'positive' results are deemed publishable[15]. Unsuccessful attempts are poorly documented in the public domain, as they are not break-through findings or not what was set out to achieve, so why waste the time on a write-up? As the number of chemicals produced (and the amount of experimental data available) increases ever more rapidly, it has become feasible to mine available data and search for generally applicable rules. Access to both positive and negative results is thus paramount if a thorough understanding is to be reached. Only then can the field of chemoinformatics fully emerge as a powerful area[16]. Data repositories and archives have been designed[17] to address this point and current developments are aimed at 'stringing' all this data together to enable comprehensive searches for chemicals, e.g. ChemSpider[18]. A resource that should not go without mention in this context is Wikipedia[19]. This web-based, free-content encyclopedia is compiled by Internet users, whose articles may be reviewed and changed at any time by anybody. Over just nine years Wikipedia has grown into a source containing millions of articles in multiple languages. Although not peer-reviewed as the traditional paper-based literature, Wikipedia relies on verification of contributions from published sources and the presented material must contain relevant supporting citations. Articles evolve over time and one has to be critical of any information within, but nevertheless, Wikipedia is an excellent resource to find out about the basic chemistry and applications of many compounds.

The digression above discusses the difficulties encountered whilst reviewing the literature in search of applications of acylanilides and hence the following section constitutes a brief summary of the relevance of the phenyl amide core and of the application fields for a few individual compounds that are extensively described in the literature.

The phenyl amide core plays an important role in many different areas. Many biological moieties contain amide bonds[20], termed peptide bonds, linking amino acids together to form polypeptides and proteins. Hence phenyl amides present good model compounds for these much more complex biological systems[21,22]. Studying structural phenomena such as hydrogen bonding (HB)[23] and cis-/trans-

isomerism[24] helps in understanding the chemical behaviour of peptides and proteins[25] and stabilization of their secondary structure[26].

The ability to form chains of molecules linked together by an amide bond is also utilized in the production of rubbers and fibres, for example the synthetic fibre Kevlar is an aromatic polyphenylamide[27]. Other polymers widely used are polyurethanes, where the units are linked through carbamate bonds[28]. These polymers are produced by addition of polyisocyanates. Methyl N-phenyl carbamates are precursors for the manufacture of polyisocyanates, which are also employed in pesticide production[29].

Commercially many arylamide modifications are manufactured as pesticides, fungicides and herbicides, with the so-called chloroacetanilides being among the most commonly used herbicides in agriculture[30]. Further related herbicides widely used are such containing formanilides[31], acylanilides, carbanilates, carbamates and phenylureas[32,33]. These industrial herbicides are somewhat more complex derivatives of the acylanilides studied in this project but as for peptides the library presented here contains many suitable model compounds for these commercially important chemicals[34,35]. Acylanilides have also been used in the development of agrophore models for the discovery of new potential leads as chemical hybridizing agents (CHAs), which induce male sterility in self-pollinating crops hence allowing pollination from other sources[36].

Formanilides are another important class of chemicals and they have found wide applications including the synthesis of biologically active compounds, e.g. quinolone antibacterials and chemotherapeutic agents[37]. They can be further utilized as precursor of biphenyl aniline derivatives, which have been recognized for their anti-phlogistic, analgesic, anti-obesity and anti-tumor activity[38]. Formanilides can also be used in the synthesis of poly(imino-1,4-phenylene), a material employed in electrochemical devices and battery electrodes[39]. In addition compounds containing the N-formyl functional group have been found in extracts of fragrant flowers, hence being of relevance to the perfume industry, and they play an important role as Lewis bases[40].

Another major area that makes use of arylamides is the pharmaceutical industry. The acetamido moiety is one of the most frequently occurring organic substituents

in drug molecules[41] and acylamides are commonly used as protective groups for amines in synthesis[42].

The most prominent member of the compound library is para-hydroxy acetanilide, more commonly known as Paracetamol. It belongs to a group of analgesic and antipyretic acetanilides, which also includes acetanilide [**Me:H**] and para-ethoxy acetanilide [**Me:p-OEt**], introduced on the market as Antifebrin[43] and Phenacetin[44], respectively, and para-bromo acetanilide [**Me:p-Br**][45]. Initially Paracetamol was disregarded as commercial drug due to adverse side effects, and Antifebrin and Phenacetin were widely used as pain and fever relief from the end of the 19th century until the last half of the 20th century[46]. However, several decades after the biological activity of Paracetamol was established it was shown that Paracetamol is in fact a metabolite of Antifebrin and Phenacetin and much better tolerable by the human body than its precursors[47]. Furthermore it was proven that the initially observed side effects were spurious findings and subsequently Paracetamol entered the commercial market whilst Antifebrin and Phenacetin were withdrawn due to their toxic properties [48]. The original patent on Paracetamol has long since expired and nowadays Paracetamol and many formulations thereof are available worldwide under a multitude of brand names[49]. Another para-substituted acetanilide that is connected with Paracetamol is para-acetoxyacetanilide [**Me:p-OCOMe**], which was developed as a prodrug for Paracetamol[50]. Despite the beneficial effects Paracetamol exhibits a few undesired properties, which are mainly poor compaction behaviour and bad taste. Para-acetyloxyacetanilide showed promise and similar bioavailability to Paracetamol and posed a potential derivatisation to overcome the bad taste problem[51]. However, the compound exhibited polymorphism[52] (albeit no phase transitions were detected upon grinding) and it is not clear whether **Me:p-OCOMe** ever entered into commercial production. Incidentally it has been shown that Paracetamol has two polymorphic forms stable under ambient conditions and one metastable polymorphic form[53]. The stable polymorphs are monoclinic [**Me:p-OH(m)**] and orthorhombic [**Me:p-OH(o)**] Paracetamol, respectively, with the monoclinic form being thermodynamically favoured. The crystal structure of the third form has been elusive for many years and been focus of many discussions[54], however recent efforts have concentrated on elucidating the solid state structure[55] and

characterising this third polymorph[56]. Of the two kinetically stable polymorphs the orthorhombic form is the preferred for tableting since the molecules stack in layers in the crystal and the inherent slip planes enable plastic deformation of the solid as occurs in tablet formulation[57]. This example demonstrates how crystal structure features can directly affect physical properties of a solid. Paracetamol finds further application in the manufacture of azo dyes and photographic material[12].

Other commercially important acetanilides are the aminoacetanilides [**Me:s-NH_2**], which occur as intermediates in the production of azo dye products and in pharmaceutical manufacture[13], and ortho-cyano acetanilide [**Me:o-CN**], which is used in pharmaceutical heterocycle synthesis[58], whilst ortho-acetamidobenzoic acid [**Me:o-COOH**] has been studied for its triboluminescent and birefringent properties[59].

2.E. References

[1] *A few selected articles from groups leading in the field are listed subsequently.* L. Fabian, A. Kalman, *Acta Crystallographica* **B60**, 547-558 (2004); T. Gelbrich, M. B. Hursthouse, T. L. Threlfall, *Acta Crystallographica* **B63**, 621-632 (2007); S. M. S. V. Wardell, M. V. N. de Souza, T. R. A. Vasconcelos, M. de L. Ferreira, K. L. Wardell, J. N. Low, C. Glidewell, *Acta Crystallographica* **B64**, 84-100 (2008); K. Donnelly, J. F. Gallagher, A. J. Lough, *Acta Crystallographica* **C64**, o335-o340 (2008).

[2] In this project only single substitution was considered either at ortho, meta or para position as indicated by X_o, X_m and X_p, respectively, in Figure 2.1.

[3] M. B. Hursthouse, L. S. Huth, T. L. Threlfall, *Organic Process Research & Development*, **13**, 1231-1240 (2009).

[4] http://ecrystals.chem.soton.ac.uk/ (accessed 2010).

[5] F. H. Allen, *Acta Crystallographica* **B58**, 380-388 (2002). CSD version 5.31 (2009, including latest update May 2010).

[6] To distinguish between triclinic and trigonal setting the abbreviations t and tg are used, respectively.

[7] Note that this work is only concerned with small organic molecules and as such the views expressed are concerned with organic chemistry, in particular chemical crystallography, but may be applicable to other areas in chemistry as well.

[8] For a crystallography perspective: M. B. Hursthouse, *Crystallography Reviews,* **10**, 85-96 (2004).

[9] An exception worth mentioning is the field of spectroscopic studies, such as IR investigations, where molecular descriptors have been correlated with signal shifts for many years, typically requiring tens of compounds to establish any trends and relationships.

[10] A study based on crystallographic data: F. H. Allen, *Crystallography Reviews*, **10**, 3-15 (2004).

[11] Funding bodies such as the Engineering and Physical Sciences Research Council (http://www.epsrc.ac.uk/) or the Natural Environment Research Council (http://www.nerc.ac.uk/) have identified priority research areas and are

encouraging interdisciplinary research with schemes for funding across research councils.

[12] ed. M. J. O'Neil, P. E. Heckelman, C. B. Koch, K. J. Roman, C. M. Kenny, M. R. D'Arecca, ***"The Merck Index" (14th Edition)***, Merck & Co. Inc., Whitehouse Station, NJ, USA (2006).

[13] S. D. Gangolli, ***"Dictionary of Substances and their Effects (DOSE)"***, The Royal Society of Chemistry, UK (1999).

[14] J. R. Plimmer, D. W. Gammon, N. R. Ragsdale, ***"Encylopedia of Agrochemicals"***, Wiley & Sons, Inc., New York (© 1999-2010).

[15] An essay found with Google using the search term "The Journal of Failed Experiments" (25/09/2007): T. Goetz, http://www.wired.com/science/discoveries/magazine/15-10/st_essay .

[16] S. J. Coles, J. G. Frey, M. B. Hursthouse, M. E. Light, A. J. Milsted, L. A. Carr, D. DeRoure, C. J. Gutteridge, H. R. Mills, K. E. Meacham, M. Surridge, E. Lyon, R. Heery, M. Duke, M. Day, *Journal of Chemical Information and Modelling*, **46**(3), 1006-1016 (2006).

[17] http://www.ukoln.ac.uk/projects/ebank-uk/ (2003-2008, accessed 2010); http://www.reciprocalnet.org/ (accessed 2010).

[18] ChemSpider is provided by the Royal Society of Chemistry (RSC) http://www.chemspider.com/ (accessed 2010).

[19] Wikipedia – The Free Encyclopedia, http://www.wikipedia.org/ (accessed 2010).

[20] G. A. Jeffrey, W. Saenger, ***"Hydrogen Bonding in Biological Structures"***, Springer-Verlag, Berlin, Germany (1991).

[21] A. Stash, V. Tsirelson, *Crystallography Reports,* **50**(2), 177-184 (2005).

[22] R. Pollack, T. Dumsha, *Journal of the American Chemical Society,* **97**(2), 377-380 (1974).

[23] S. Jalili, M. Akhavan, *Journal Of Theoretical & Computational Chemistry,* **3**(4), 527-542 (2004).

[24] S. Ilieva, B. Hadjieva, B. Galabov, *Journal of Molecular Structure*, **508**, 73-80 (1999).

[25] B. Galabov, D. Cheshmedzhieva, S. Ilieva, B. Hadjieva, *Journal Of Physical Chemistry A,* **108**(51), 11457-11462 (2004).

[26] G. Careri, U. Buontempo, F. Galluzi, A. Scott, E. Gratton, E. Shyamsunder, *Physical Review B,* **30**, 4689-4702 (1984).

[27] http://en.wikipedia.org/wiki/Kevlar (accessed 2010).

[28] http://en.wikipedia.org/wiki/Polyurethane (accessed 2010).

[29] X. Wan, Z. Ma, B. Li, K. Zhang, S. Cao, S. Zhang, Z. Shi, *Journal of the American Society:Communications,* **128**, 7416-7417 (2006).

[30] B. Oosterhuis, K. Vukman, E. Vagi, H. Glavinas, I. Jablonkai, P. Krajcsi, *Toxicology,* **248**, 45-51 (2008).

[31] B.-C. Chen, M. S. Bednarz, R. Zhao, J. E. Sundeen, P. Chen, A. P. Skoumbourdis, J. C. Barnish, *Tetrahedron Letters*, **41**, 5453-5456 (2000).

[32] A. Wood, http://www.alanwood.net/pesticides/index.html (1995-2010, accessed 2010).

[33] J. L. Hilton, L. L. Jansen, H. M. Hull, *Annual Review of Plant Physiology*, **14**, 353-384 (1963).

[34] D. Villarreal, R. Turco, A. Konopka, *Applied And Environmental Microbiology,* **60**(11), 3939-3944 (1994).

[35] J. Sorci, D. Macalady, *Journal of Agricultural and Food Chemistry,* **41**, 1760-1766 (1993).

[36] K. Chakraborty, C. Devakumar, *Journal of Agricultural and Food Chemistry,* **54**, 6800-6808 (2006).

[37] K. Kobayashi, S. Nagato, M. Kawakitu, O. Morikawa, H. Konishi, *Chemistry Letters*, **24**(7), 575-577 (1995).

[38] L. Zhu, M. Patel, M. Zhang, *Tetrahedron Letters*, **49**, 2734-2737 (2008) and references therein.

[39] R. Koppang, *Journal of Fluorine Chemistry*, **127**, 1505-1509 (2006) and references therein.

[40] T. Pratap, S. Baskaran, *Tetrahedron Letters,* **42**, 1983-1985 (2001) and references therein.

[41] P. Ertl, *Journal of Chemical Information and Computer Sciences,* **43**(2), 374-380 (2003).

[42] P. G. M. Wuts, T. W. Greene, ***"Greene's protective groups in organic synthesis"***, Wiley & Sons, Chichester, England (2007).

[43] A. Cahn, P. Hepp, *Zentralblatt für Klinische Medizin*, **7**, 561-564 (1886).

[44] J. von Mering, *Therapeutische Monatsschrift*, **7**, 577-587 (1893).

[45] M. Caira, D. Dodds, *Journal of Inclusion Phenomena and Macrocyclic Chemistry,* **34**, 19-29 (1999).

[46] A. Bertolini, A. Ferrari, A. Ottani, S. Guerzoni, R. Tacchi, S. Leone, *CNS Drug Reviews*, **12**(3-4), 250-275 (2006)

[47] B. B. Brodie, J. Axelrod, *Journal of Pharmacology and Experimental Therapeutics*, **94**(1), 22-28 (1948); B. B. Brodie, J. Axelrod, *Journal of Pharmacology and Experimental Therapeutics*, **94**(1), 29-38 (1948); F. B. Flinn, B. B. Brodie, *Journal of Pharmacology and Experimental Therapeutics*, **94**(1), 76-77 (1948); B. B. Brodie, J. Axelrod, *Journal of Pharmacology and Experimental Therapeutics*, **97**(1), 58-67 (1949).

[48] W. Sneader, ***"Drug Discovery: A History"***, Wiley & Sons Ltd., Chichester, England (2005).

[49] http://en.wikipedia.org/wiki/List_of_paracetamol_brand_names (accessed 2010).

[50] S. Finnie, A. Kennedy, K. Prasad, R. Ristic, D. Sheen, J. Sherwood, J., *Acta Crystallographica* **C55***, 234-236* (1999).

[51] M. Caira, F. Wet, J. Gerber, *Journal of Chemical Crystallography,* **29**(2), 175-178 (1999).

[52] D. Bhalla, J. Lalla, *Drug Development and Industrial Pharmacy,* **16**(1), 115-135 (1990).

[53] G. Perlovich, T. Volkova, A. Bauer-Brandl, *Journal of Thermal Analysis and Calorimetry,* **89**(3), 767-774 (2007).

[54] T. Beyer, G. Day, S. Price, *Journal of the American Chemical Society,* **123**(21), 5086-5094 (2001); M. Peterson, S. Morissette, C. McNulty, A. Goldsweig, P. Shaw, M. LeQuesne, J. Monagle, N. Encina, J. Marchiona, A. Johnson, J. Gonzalez-Zugasti, A.

V. Lemmo, S. J. Ellis, M. J. Cima, O. Almarsson, *Journal of the American Chemical Society: Communications,* **124**, 10958-10959 (2002) and references therein.

[55] M.-A. Perrin, M. Neumann, H. Elmalch, L. Zaske, *ChemCommun*, 3181-3183 (2009).

[56] J. Burler, M. Duer, R. Stein, R. Vrcelj, *European Journal of Pharmaceutical Sciences,* **31**, 271-276 (2007); S. Gaisford, A. Buanz, N. Jethwa, *Journal of Pharmaceutical and Biomedical Analysis,* **53**(3), 366-370 (2010).

[57] G. Nichols, C. S. Frampton, *Journal of Pharmaceutical Sciences,* **87**(6), 684-693 (1998).

[58] B. Arslan, C. Kazak, C. Kirilmis, M. Koca, O. Büyükgüngör, *Acta Crystallographica* **E61**, 1652-1653 (2005).

[59] Y. Mascarenhas, V. de Almeida, J. Lechat, N. Barelli, *Acta Crystallographica* **B36**, 502-504 (1980).

CHAPTER 3: EXPERIMENTAL

This chapter summarises all routine procedures used in the various crystallographic and theoretical experiments in order to avoid unnecessary repetition in latter chapters. Deviations from any of these routine procedures are pointed out where applicable.

3.A. Crystal Structure Determination from Single Crystal X-ray Diffraction

X-ray diffraction data were collected by means of combined phi and omega scans on a Bruker-Nonius KappaCCD area detector situated at the window of a rotating anode (Mo-K_α radiation, λ = 0.71073Å; graphite monochromated radiation or 10cm confocal mirrors radiation). The structures were solved by direct methods with SHELXS-97[1] and refined on F^2 using SHELXL-97[1] as implemented in the WinGX[2] suite of programs. Anisotropic displacement parameters were assigned to all non-hydrogen atoms. Hydrogen atoms bonded to carbon atoms were included at calculated positions and the thermal parameters were constrained to ride on the atom to which they are bonded. Any other hydrogen atoms (mostly amide hydrogens) were located on the difference map and the bond distances were restrained to idealised values. Thermal parameters for those hydrogen atoms were refined isotropically in the majority of structures, or, if badly behaved, were refined as riding on the atom to which they are bonded. The data were corrected for crystal anomalies and absorption effects using SADABS V2.10[3].

3.B. Crystal Structure Comparison with XPac

In order to perform an XPac analysis the molecular shape of the various compounds has to be described adequately and thus an ordered set of corresponding points (COSP) has to be defined for the library. Given that the acetanilide compounds discussed in this thesis only vary in one substituent position, one COSP choice could comprise the whole core, i.e. the acetamido group together with the phenyl ring giving a total of 10 points. However, there is an apparent spread of torsion angles between the amide group and the phenyl ring ranging from near planarity to approximately 80° rotational offset especially in the ortho acetanilide family of compounds. Such a 10-point COSP would yield a high consistency index δ_{con} indicating that the described arrangement of points lacks similarity due to the different torsion angles. At this stage the objectives of the structure comparison should be considered and incidentally the question asked whether conformational changes of the backbone are relevant to assessing common crystal packing features. For this work this is not the case. In order to accommodate rotational freedom of the phenyl ring, the COSP was constructed so that the ortho and meta carbon atoms of the phenyl rings were excluded, hence giving a general COSP of 6 atomic positions for all acetanilides as shown in Figure 3.1. In crystal structures with more than one molecule in the asymmetric unit (ASU) an ordered set of points (OSP) was defined for each independent molecule. During the actual XPac analysis the three different default cut-off levels, summarised in Table 3.1, were applied to allow for angular and longitudinal variations across the series of mono-substituted acetanilides and acylanilides. The optimum setting was then used in a final crystal structure comparison cycle.

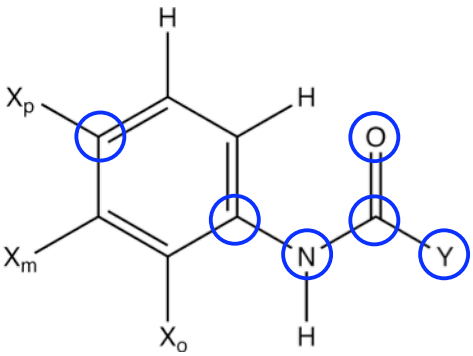


Figure 3.1: Standard COSP as used in the crystal structure comparison of mono-substituted acetanilides and acylanilides.

cut-off level	δ_{ang} [°]	δ_{dhd} [°]	δ_{tor} [°]
low	7	13	13
medium	10	18	18
high	12	25	25

Table 3.1: The default cut-off levels as available in XPac.

3.C. Theoretical Calculations

All theoretical calculations were performed using the Magellan Cluster provided by the National Service for Computational Chemistry Software (NSCCS)[4]. This is a 224-core Silicon Graphics Altix 4700 with 1.6GHz Montecito Itanium2 processors, 896GB memory and 15TB of disk space.

3.C.1. Conformational Analysis

During the course of the study it was found that the dihedral angle between the aromatic ring and the amide group varied significantly, hence a conformational analysis was commenced. In order to evaluate the potential energy surface with regards to the rotation about α (Figure 3.2) theoretical calculations in the gas phase were performed as subsequently detailed. The molecular geometry as determined in the crystal structure was used as starting configuration in all calculations, which were performed with the *Gaussian 03* suite of programs[5]. The experimental parameters were allowed to relax in a geometry optimisation except for the torsion angle α, which was restrained using the *AddRedundant* option implemented in *g03*. The torsion angle α was sampled in 36 steps with an incremental step size of 10˚. Smaller increments would have increased the computational cost beyond the scope of this project. Geometry optimisations were set up as multi-stage optimisations, starting with the semi-empirical method AM1, allowing fast principal adjustment of α to the desired conformation. In the next step the AM1 optimised geometric parameters were further optimised switching to *ab initio* Hartree Fock (HF) methods in order to increase the accuracy of the calculated parameters. The basis set was gradually increased from 3/21G to 6/311G and lastly to the highest level of theory used in

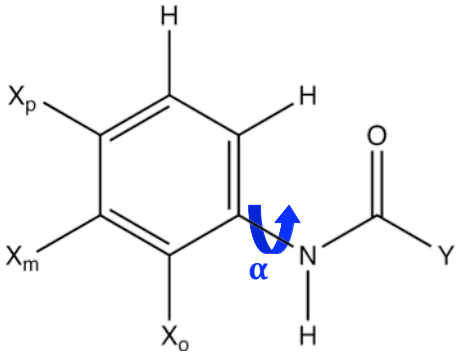


Figure 3.2: The torsion angle α is defined as the angle of rotation about the C-N bond between the aromatic ring system and amide group.

this study 6/311G**. Finally single point energies were computed employing Density Functional Theory (DFT) at the 6/311G** level to allow for electron correlation. The Becke three-parameter exchange functional was used in combination[6] with the gradient corrected exchange-correlation potential of Lee, Yang and Paar[7] (B3LYP).

The choice of basis sets was governed by the availability of basis set parameters for all elements found in the acylanilides. However iodine is rarely parameterised in standard basis sets and hence iodoacetanilides could not be included in the calculations.

3.C.2. Lattice Energy Calculations

Lattice energies were calculated with Gavezzotti's OPiX suite of programs[8] following the Pixel method as described in Chapter 1. The typical workflow for such a calculation is depicted in Figure 3.3. The modules RETCIF, COOR and MAT2 are essentially file format converters, whereby the initial input is in form of a crystallographic information file (cif) and the output of MAT2 yields an instruction file for *Gaussian03*. The last module, PIXEL, then performs the actual lattice energy summation.

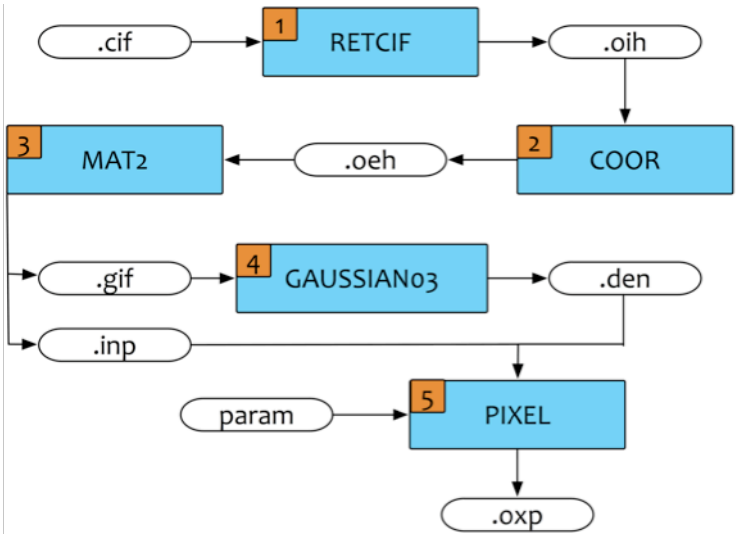


Figure 3.3: A typical workflow for calculation of lattice energies using the pixel method.

RETCIF is designed to convert cifs in CSD format, but the largest proportion of cifs used in this study were produced in the format of SHELXL-97[1], the refinement software. Hence manual editing was required before the cifs could be processed by RETCIF. Given the large number of files this step was automated with the Perl[9] script *prepcif.pl*. After conversion of the crystal structure coordinates from cif to the *Gaussian03*[5] input format (steps 1 to 3 in Figure 3.3) the molecular electron density was then calculated with *Gaussian03 (g03)*. The default level of theory, second order Møller Plesset perturbation theory (MP2) with a 6/31G** basis set, was adopted throughout to compute the valence density ('cube=frozencore' option

in *g03*). Finally intermolecular energies were calculated by the semi-classical density sums using the PIXEL module. PIXEL also produced the total lattice energy as the sum of intermolecular electrostatic, repulsion, polarisation and dispersion contributions. Default values for all variable parameters were used throughout. The workflow from steps 1 to 4 in Figure 3.3 was automated with the Perl script *runOPiX.pl* and the energy calculation (step 5) was controlled with the Perl script *runpixelc.pl*. This latter script initiated lattice energy calculations for those compounds only, for which the electron density was successfully computed with *g03* in the previous step. A log file was produced highlighting erroneous *g03* runs for further manual inspection. Once all calculations were completed the intermolecular interaction energy and total lattice energy values (in $kJmol^{-1}$) were extracted from the Pixel output file using the Perl script *getEnergies.pl*. All Perl scripts mentioned previously were designed as part of this project and can be found in Appendix 3-1.

Lattice energies could not be estimated for iodoacetanilides and **Me:p-COOOH** since iodine is not part of the standard basis set used during the molecular electron density computation and **COOOH** is not parameterized for the Pixel summation. Furthermore the algorithm was found unsuitable for structures containing two independent molecules in the ASU. In these cases problems were encountered during the summation of the intermolecular interaction energies leading to gross over-estimation of the energies. This affected the structures **Me:m-Cl**, **Me:m-Br(m)**, **Me:m-Br(o)**, **Me:p-SF_5**, **Me:p-NO_2**, **Me:p-NMe_2**, **Me:p-OH(o2)** and **Me:p-O^{i}Pr**. Lastly, the lattice energy could not be calculated for **Me:m-NO_2** as this compound crystallises with four independent molecules in the ASU hence exceeding the Z' limit of the pixel code.

3.D. References

[1] G. M. Sheldrick, "***SHELX suite of programs for crystal structure solution and refinement"***, University of Göttingen, Germany (1997).

[2] L. J. Farrugia, *Journal of Applied Crystallography,* **32**, 837-838 (1999).

[3] G. M. Sheldrick, ***SADABS V2.10*** (2003).

[4] http://www.nsccs.ac.uk/ (accessed 2010).

[5] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. W. Ochterski, G. A. Petersson, P.Y. Ayala, Q.Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, F. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, A. Gonzalez, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, J. A. Pople, ***GAUSSIAN 03*** (2003).

[6] A. D. Becke, *J. Chem. Phys.*, **98**, 5648-5652 (1993).

[7] C. Lee, W. Yang, R. G. Parr, *Physical Reviews B*, **37**, 785-789 (1988).

[8] A. Gavezzotti, ***"OPiX: A computer program package for the calculation of intermolecular interactions and crystal energies"***, University of Milano, Italy (2003).

[9] http://www.perl.org/

CHAPTER 4: STRUCTURAL SYSTEMATICS OF ACETANILIDES

During this project a compound library of 64 crystal structures of mono-substituted acetanilides was compiled, of which 44 were determined by single crystal X-ray diffraction. Subsequently all crystal structures in the library were compared with each other using the XPac algorithm in the search for packing similarity. In the first part of this chapter some details of the acetanilide crystal structures are summarised whereas the second part of this chapter focuses on the results from the crystal structure comparison and their interpretation. It should be noted that hydrogen atoms were omitted for clarity from most pictures in this chapter with the exception of hydrogen atoms on HB functional groups. The molecular connectivity found in the crystal structures is not primarily part of the discussion in this chapter but is further considered in Chapter 5.

4.A. Crystal structures of 64 mono-substituted acetanilides

The 64 crystal structures in the acetanilide compound library can be divided into 9 ortho-substituted, 12 meta-substituted and 23 para-substituted newly determined acetanilide crystal structures with the remaining 20 structures harvested from the CSD. It should be noted here that of the 23 para-substituted acetanilide structures those of **Me:p-Br(m)** and **Me:p-Br(o)** were determined by Chris Frampton (CSF)[1] and the resulting cifs kindly provided for this study. However, during this study two different crystalline samples of **Me:p-Br** were produced and their unit cells were confirmed to match the two structures obtained from CSF.

The reflection data for **Me:p-OH(m)** were collected at a lower temperature and to a higher resolution than is the standard laboratory routine. This was done for the purpose of a charge density study not further mentioned in this work. The known monoclinic crystal structure was unambiguously confirmed, but in order to ensure the crystal structures used in the similarity search were generated at comparable

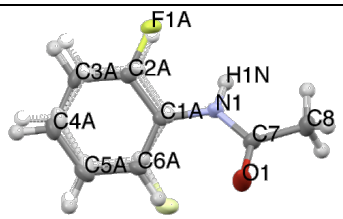
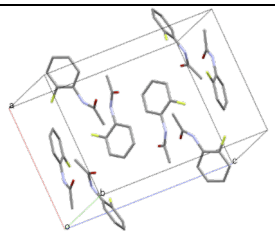
ambient conditions, one of the **Me:p-OH(m)** crystal structures available on the CSD was used here instead.

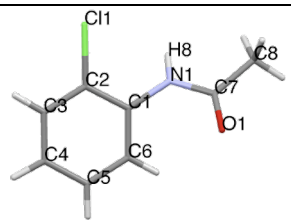
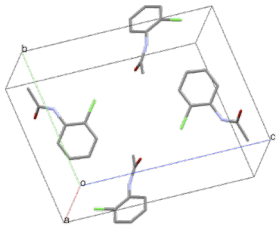
In addition, the unit cell of a crystalline sample of **Me:H** was confirmed to be the same as those already in the CSD. Since there is an abundance of identical crystal structures on the database for this compound, the crystal structure was not re-determined and a low temperature structure representative from the CSD used instead.

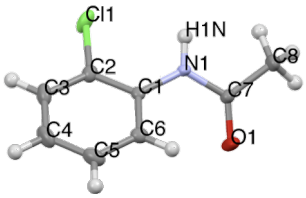
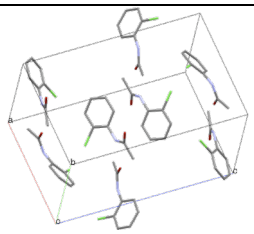
As mentioned in the previous chapter, during the crystallisation process polymorphism was observed for some compounds, which are shown in Tables 2.2 to 2.4 as multiple entries per individual cell. The acetanilides crystallising in two polymorphic forms are **Me:o-Cl**, **Me:o-Me**, **Me:m-Br**, **Me:p-Br**, **Me:p-Me**, and **Me:p-OCOMe**. Of these, four structures were previously unknown [**Me:o-Me(m)**, **Me:m-Br(m)**, **Me:p-Br(m)** and **Me:p-OCOMe(m2)**]. **Me:p-OH** is known to be at least trimorphic[2], but the meta-stable form has been elusive and basis for many speculations over the years[3]. Recently, a crystal structure of **Me:p-OH(o2)**obtained from powder diffraction data and molecular simulation has been published[4]. Although all other structures considered in this work were determined from single crystals, it was deemed worthwhile to probe the **Me:p-OH(o2)** crystal assembly for any similarity with the other polymorphs of the compound in particular and with the acetanilide structures in general. The polymorphic behaviour of acetanilides is discussed in more detail in the subsequent chapter.

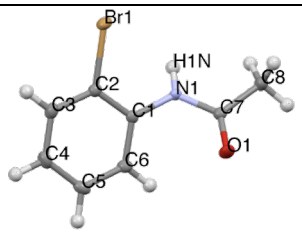
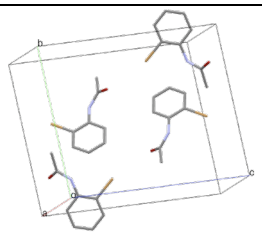
In the following the most relevant crystal structure details for all structures in the acetanilide library are presented as follows: (I) compound identifier, IUPAC name and sum formula; (II) the content and labelling scheme of the ASU is shown in a thermal ellipsoid plot for those structures determined during this project or in a capped stick representation for the structures from the CSD; (III) unit cell content in capped stick representation with any disorder and hydrogen atoms omitted for clarity; (IV) table containing unit cell parameters, unit cell volume, Z and Z' as well as the data collection temperature. The graphics were produced using Mercury Materials[5] and the IUPAC names were generated with the ADC/IUPAC Name Web Service available through the Chemical Database Service at Daresbury[6]. The full details for all determined crystal structures can be found in Appendix 4-1.

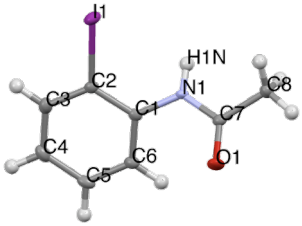
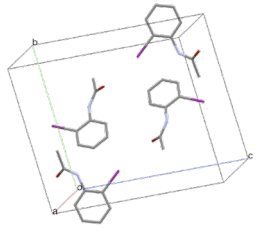
4.A.1. Ortho-substituted acetanilide crystal structures

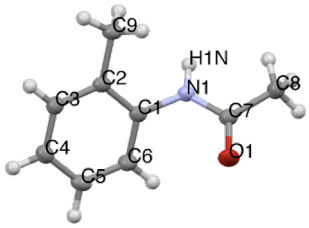
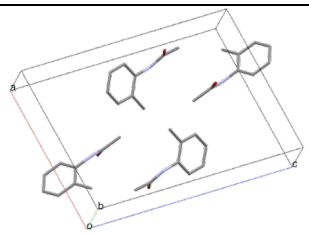
Me:o-F	**C_8H_8FNO**	***N-(2-fluorophenyl)acetamide***			
					
The substituted phenyl ring was found to be disordered over two positions and the disorder ratio refined to 82:18%. The minor component is shown in faded colours in the thermal ellipsoid plot on the left.					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
10.5132(4)	9.4393(3)	15.1635(5)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
Pbca	1504.78(9)	8	1	120(2)	

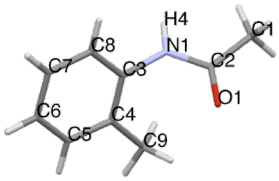
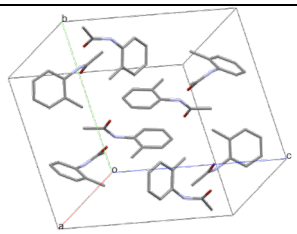
Me:o-Cl(m)	**C_8H_8ClNO**	***N-(2-chlorophenyl)acetamide***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
4.7468(4)	11.699(1)	14.640(2)	90	95.74(2)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/n$	808.924	4	1	100	

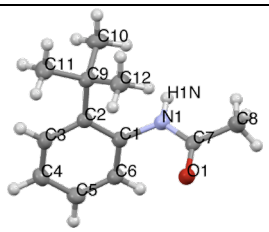
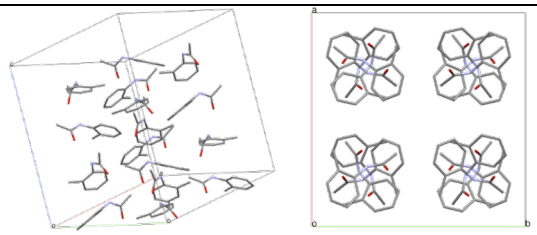
Me:o-Cl(o)	**C_8H_8ClNO**	***N-(2-chlorophenyl)acetamide***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
10.5174(2)	9.4252(4)	16.2656(6)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
Pbca	1612.39(10)	8	1	120(2)	

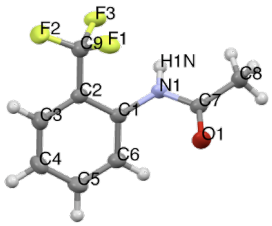
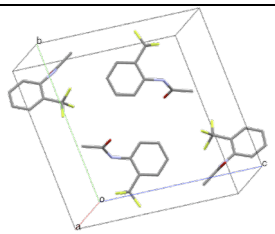
Me:o-Br	**C_8H_8BrNO**	***N-(2-bromophenyl)acetamide***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
4.7808(1)	11.9344(5)	14.6859(5)	90	96.820(2)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/n$	831.99(5)	4	1	120(2)	

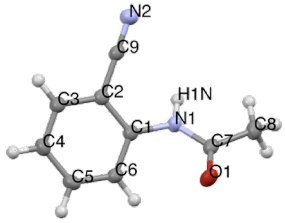
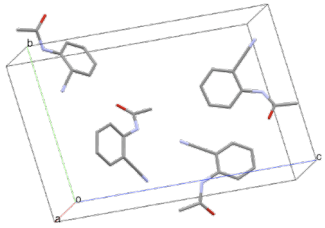
Me:o-I	**C_8H_8INO**	***N-(2-iodophenyl)acetamide***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
4.7969(1)	12.4977(3)	14.6790(3)	90	98.336(1)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/n$	870.71(3)	4	1	120(2)	

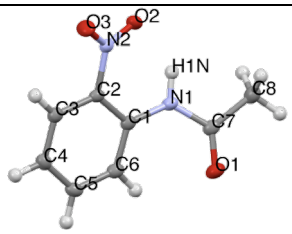
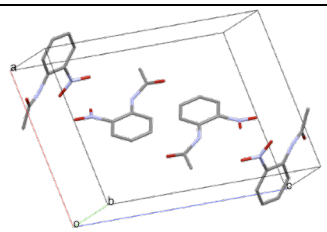
Me:o-Me(m)	**$C_9H_{11}NO$**	***N-(2-methylphenyl)acetamide***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
11.1083(6)	4.8283(3)	15.2681(7)	90	99.170(3)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/n$	808.43(8)	4	1	120(2)	

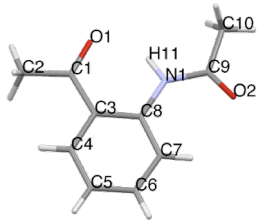
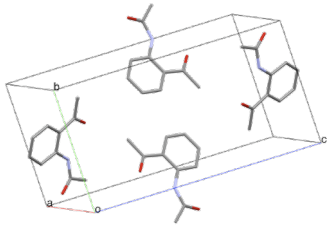
Me:o-Me(o)	**$C_9H_{11}NO$**	***N-(2-methylphenyl)acetamide***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
8.9608(10)	12.4784(17)	15.012(2)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
Pbca	1678.588	8	1	298	

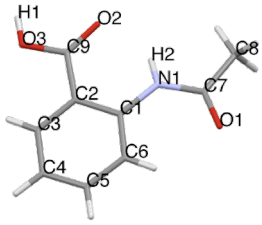
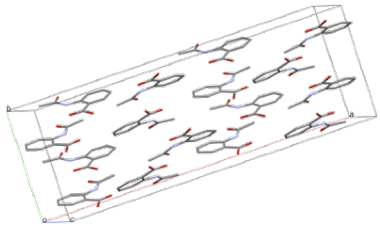
Me:o-^{t}Bu	**$C_{12}H_{17}NO$**	***N-(2-tert-butylphenyl)acetamide***			
					
The unit cell content is shown from two perspectives; a random unit cell orientation (left) and viewed along the c-axis of the unit cell (right).					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
16.2982(6)	16.2982(6)	17.0833(6)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$I4_1cd$	4537.9(3)	16	1	120(2)	

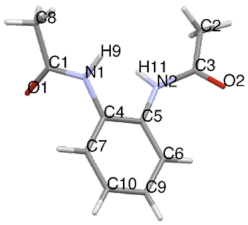
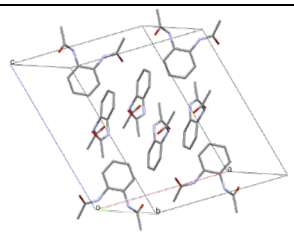
Me:o-CF_3	**$C_9H_8F_3NO$**	***N-[2-(trifluoromethyl)phenyl]acetamide***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
4.7821(4)	13.4170(16)	13.9793(15)	90	92.527(7)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/n$	896.06(16)	4	1	120(2)	

Me:o-CN	**$C_9H_8N_2O$**	***N-(2-cyanophenyl)acetamide***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
3.8114(4)	11.3360(14)	18.3264(17)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_12_12_1$	791.81(15)	4	1	120(2)	

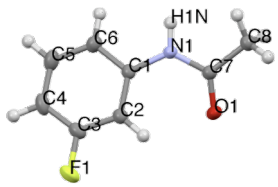
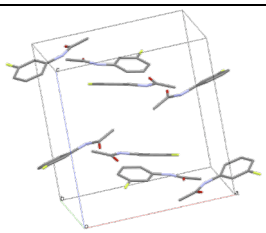
Me:o-NO_2	**$C_8H_8N_2O_3$**	***N-(2-nitrophenyl)acetamide***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
10.5564(4)	4.9758(2)	15.4117(5)	90	97.134(2)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/n$	803.26(5)	4	1	120(2)	

Me:o-COMe	**$C_{10}H_{11}NO_2$**	***N-(2-acetylphenyl)acetamide***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
7.765(7)	8.699(7)	15.805(13)	90	119.35(7)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/c$	930.558	4	1	298	

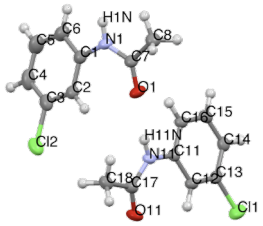
Me:o-COOH	$C_9H_9NO_3$	*2-acetamidobenzoic acid*			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
30.014(3)	10.5150(12)	10.5839(12)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
Fdd2	3340.249	16	1	100	

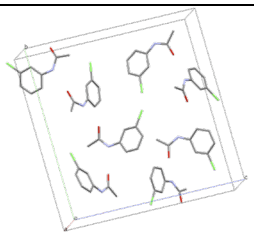
Me:o-NHCOMe	$C_{10}H_{12}N_2O_2$	*N,N'-1,2-phenylenediacetamide*			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
14.490(1)	9.246(1)	15.928(1)	90	113.051(10)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
C2/c	1963.565	8	1	173	

4.A.2. Meta-substituted acetanilide crystal structures

Me:m-F	C_8H_8FNO	*N-(3-fluorophenyl)acetamide*			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
12.1766(4)	9.5024(3)	12.9638(5)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
Pbca	1500.00(8)	8	1	120(2)	

Me:m-Cl	**C_8H_8ClNO**	***N-(3-chlorophenyl)acetamide***

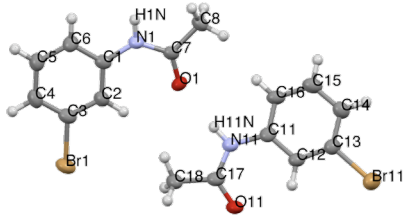


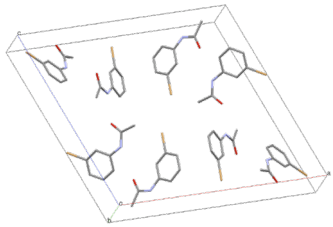


There are two crystallographically independent molecules in the ASU. The two molecules are of similar geometry, except for a difference in the dihedral angle between the plane of the aromatic ring and the amide group. Although the molecules are conformationally distinct they are related by pseudo-glide symmetry.

a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
4.7667(1)	18.4984(4)	18.6318(4)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_12_12_1$	1642.88(6)	8	2	120(2)	

Me:m-Br(m)	**C_8H_8BrNO**	***N-(3-bromophenyl)acetamide***

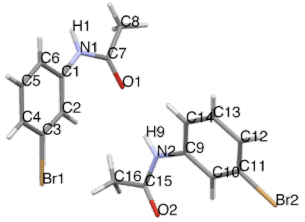


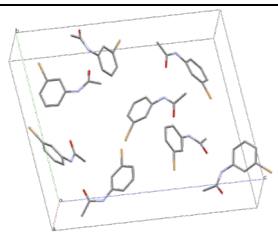


There are two crystallographically independent molecules in the ASU. The two molecules are of similar geometry, except for a difference in the dihedral angle between the plane of the aromatic ring and the amide group. Although the molecules are conformationally distinct they are related by pseudo-2_1 screw symmetry.

a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
19.855(2)	5.002(1)	18.336(2)	90	108.655(4)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/c$	1725.4(3)	8	2	120(2)	

Me:m-Br(o)	**C_8H_8BrNO**	***N-(3-bromophenyl)acetamide***

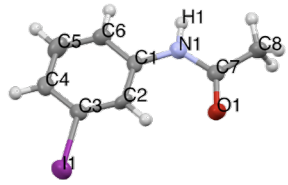


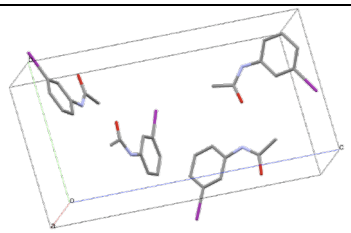


There are two crystallographically independent molecules in the ASU. The two molecules are of similar geometry, except for a difference in the dihedral angle between the plane of the aromatic ring and the amide group. Although the molecules are conformationally distinct they are related by pseudo-glide symmetry.

a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
4.7836(6)	18.765(1)	19.379(2)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_12_12_1$	1739.541	8	2	299	

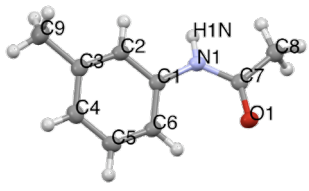
Me:m-I	**C_8H_8INO**	***N-(3-iodophenyl)acetamide***

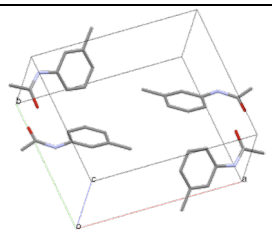




a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
4.9076(5)	9.7504(8)	18.6116(11)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_12_12_1$	890.58(13)	4	1	120(2)	

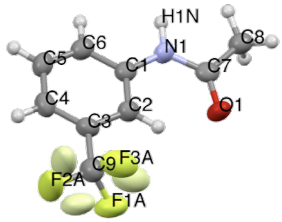
Me:m-Me	**$C_9H_{11}NO$**	***N-(3-methylphenyl)acetamide***

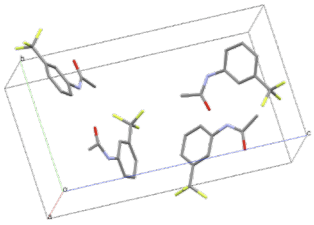




a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
12.2837(4)	9.4548(4)	7.3046(3)	90	99.932(2)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/c$	835.64(6)	4	1	120(2)	

Me:m-CF_3	**$C_9H_8F_3NO$**	***N-[3-(trifluoromethyl)phenyl]acetamide***

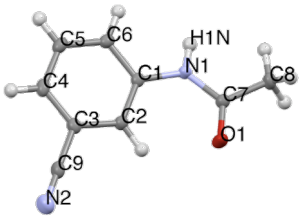


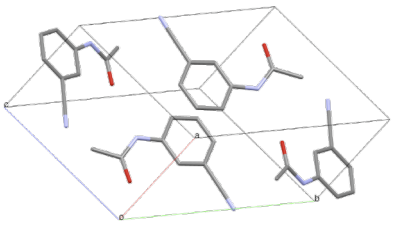


The trifluoromethyl group is disordered over two positions and the disorder ratio refined to 94:06%. The orientation of the minor component is shown in faded colours in the thermal ellipsoid plot on the left.

a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
5.1717(8)	9.7527(13)	17.811(2)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_12_12_1$	898.4(2)	4	1	120(2)	

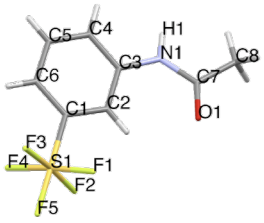
Me:m-CN	**$C_9H_8N_2O$**	***N-(3-cyanophenyl)acetamide***

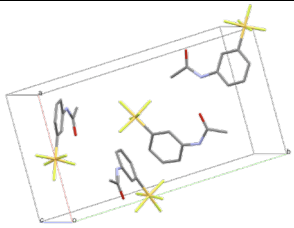




a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
7.0527(3)	12.7252(6)	9.6648(4)	90	111.315(3)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/n$	808.05(6)	4	1	120(2)	

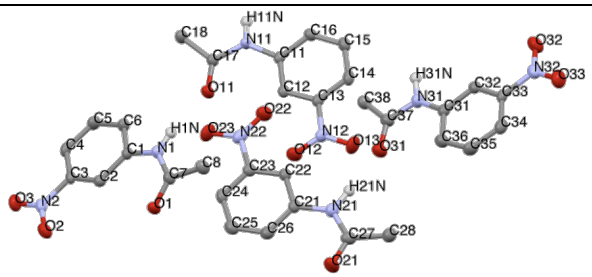
Me:m-SF_5	**$C_8H_8F_5NOS$**	***N-[3-(pentafluoro-l⁶-sulfanyl)phenyl]acetamide***

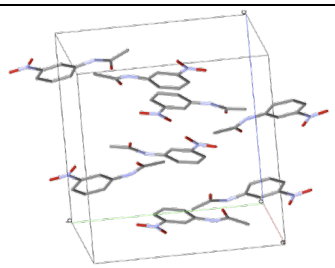




a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
9.8003(11)	17.894(2)	5.9141(7)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_12_12_1$	1037.135	4	1	293	

Me:m-NO_2	**$C_8H_8N_2O_3$**	***N-(3-nitrophenyl)acetamide***

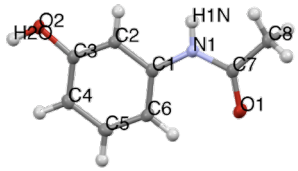


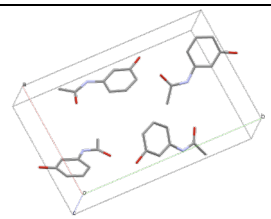


There are 4 crystallographically independent molecules in the ASU, which are of similar geometry, except for differences in the dihedral angle between the plane of the phenyl and amide group. In the molecules with N1 and N11 the nitro groups are *syn* to the amide oxygen whilst in the molecules containing N21 and N31 the nitro groups are *anti* to the amide oxygen. Although conformationally distinct the molecules within the *syn* and *anti* pair, respectively, are related via pseudo-glide symmetry, whereas a pseudo-inversion relates the *syn* with the *anti* molecules. Hydrogen atoms were omitted from the thermal ellipsoid plot for clarity.

a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
9.7285(5)	13.3392(7)	12.9849(5)	90	103.791(3)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1$	1636.48(14)	8	2	120(2)	

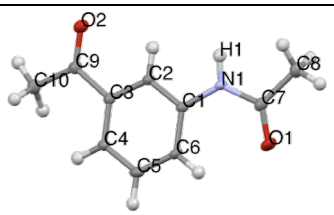
Me:m-OH	**$C_8H_9NO_2$**	***N-(3-hydroxyphenyl)acetamide***

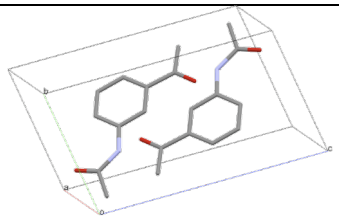




a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
10.5129(4)	17.0435(4)	4.0675(1)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$Pna2_1$	728.80(4)	4	1	120(2)	

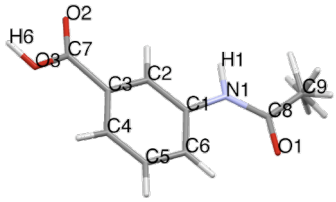
Me:m-COMe $C_{10}H_{11}NO_2$ ***N-(3-acetylphenyl)acetamide***

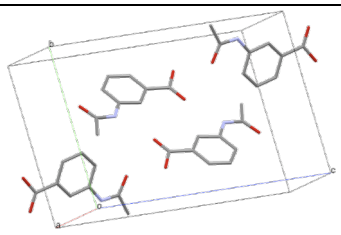




a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
5.5061(2)	7.2607(3)	11.4567(5)	103.709(2)	93.064(3)	100.528(3)
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P\bar{1}$	435.23(3)	2	1	120(2)	

Me:m-COOH $C_9H_9NO_3$ ***3-acetamidobenzoic acid***

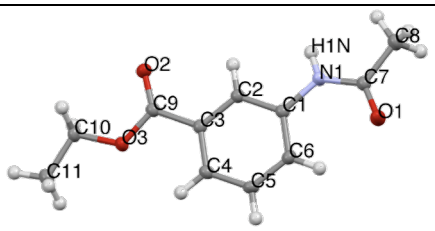


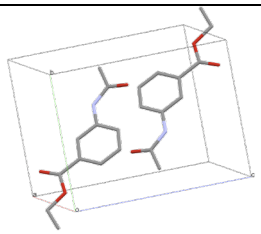


The methyl substituent attached to the amide group was modelled in two orientations (ratio unknown).

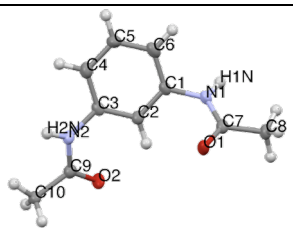
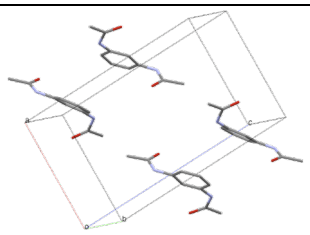
a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
3.9522(15)	10.699(4)	19.831(7)	90	93.393(8)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/n$	837.076	4	1	293	

Me:m-COOEt $C_{11}H_{13}NO_3$ ***ethyl 3-acetamidobenzoate***

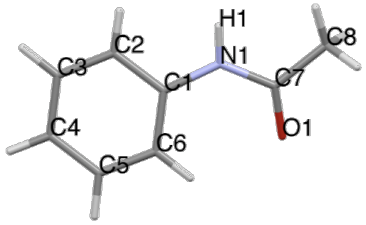
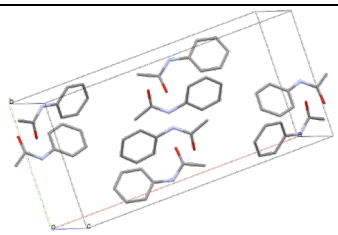


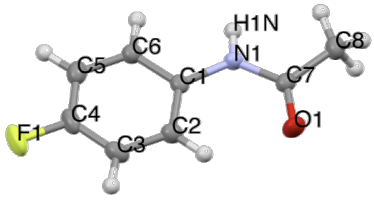
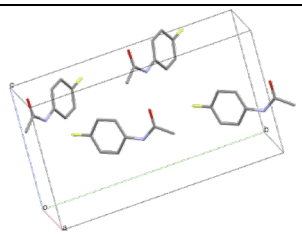


a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
5.7366(2)	8.5492(5)	10.6373(6)	93.974(2)	99.351(3)	95.314(3)
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P\bar{1}$	510.65(5)	2	1	120(2)	

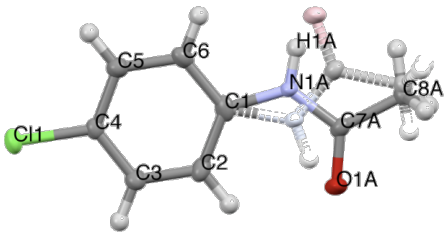
Me:m-NHCOMe	$C_{10}H_{12}N_2O_2$	***N,N'-1,3-phenylenediacetamide***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
8.9722(6)	6.8630(5)	15.3273(8)	90	90.092(5)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/c$	943.80(11)	4	1	120(2)	

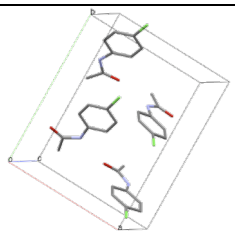
4.A.3. Para-substituted acetanilide crystal structures

Me:H	C_8H_9NO	***N-phenylacetamide***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
19.509(11)	9.364(8)	7.778(10)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
Pbca	1420.903	8	1	113	

Me:p-F	C_8H_8FNO	***N-(4-fluorophenyl)acetamide***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
4.6142(2)	17.0417(9)	9.5621(6)	90	93.378(3)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
Cc	750.60(7)	4	1	120(2)	

Me:p-Cl	**C_8H_8ClNO**	***N-(4-chlorophenyl)acetamide***

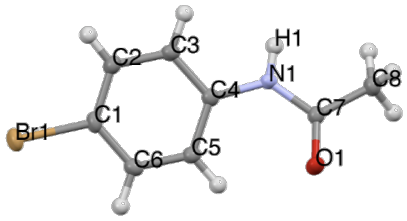


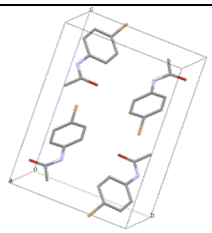


The acetamido group was found to be disordered over two positions and the disorder ratio refined to 79:21%. The minor component is shown in faded colours in the thermal ellipsoid plot on the left.

a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
9.6978(4)	12.3874(4)	6.5757(2)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$Pna2_1$	789.94(5)	4	1	120(2)	

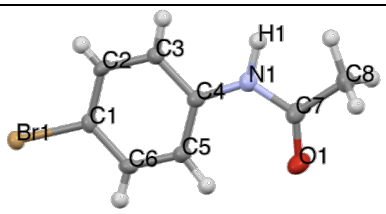
Me:p-Br(m)	**C_8H_8BrNO**	***N-(4-bromophenyl)acetamide***

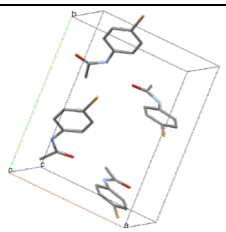




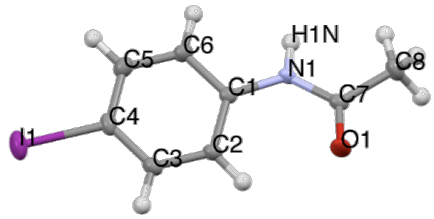
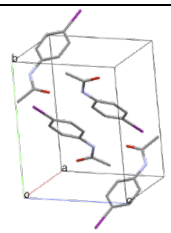
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
6.7180(4)	9.4048(5)	12.6994(7)	90	90.867(2)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/n$	802.27(8)	4	1	120(2)	

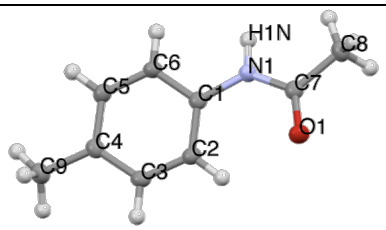
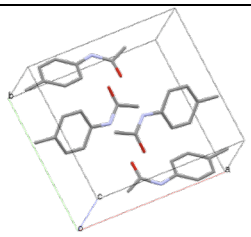
Me:p-Br(o)	**C_8H_8BrNO**	***N-(4-bromophenyl)acetamide***

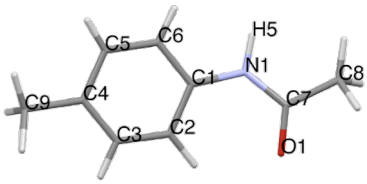
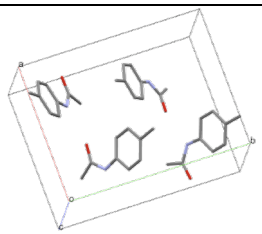




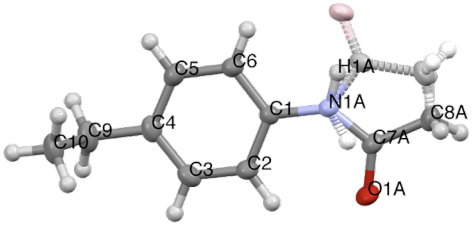
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
9.7734(6)	12.5038(7)	6.7209(4)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$Pna2_1$	821.33(8)	4	1	120(2)	

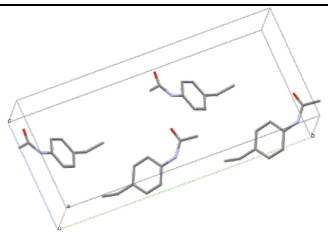
Me:p-I	**C_8H_8INO**	***N-(4-iodophenyl)acetamide***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
9.5782(3)	10.3047(3)	9.4258(2)	90	109.285(2)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/c$	878.13(4)	4	1	120(2)	

Me:p-Me(m)	**$C_9H_{11}NO$**	***N-(4-methylphenyl)acetamide***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
11.6779(9)	9.4902(7)	7.4186(6)	90	106.515(5)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/c$	788.25(11)	4	1	120(2)	

Me:p-Me(o)	**$C_9H_{11}NO$**	***N-(4-methylphenyl)acetamide***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
9.899(5)	12.956(6)	6.541(4)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$Pna2_1$	838.893	4	1	293	

Me:p-Et	**$C_{10}H_{13}NO$**	***N-(4-ethylphenyl)acetamide***

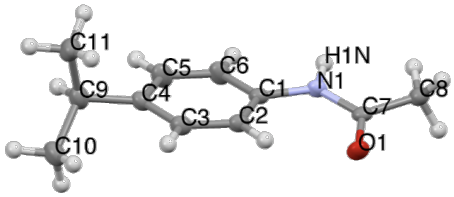


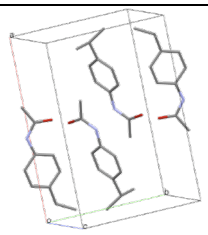


The acetamido group was found to be disordered over two positions and the disorder ratio refined to 50:50%. The second orientation is shown in faded colours in the thermal ellipsoid plot on the left.

a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
4.5852(2)	21.3953(10)	9.4310(4)	90	90.001(3)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
Cc	925.20(7)	4	1	120(2)	

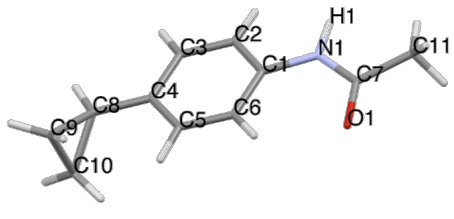
Me:p-iPr	**$C_{11}H_{14}NO$**	***N-(4-isopropylphenyl)acetamide***

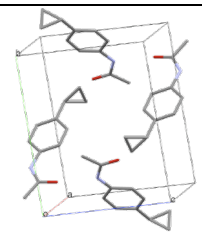




a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
14.0052(4)	9.3488(3)	7.8095(2)	90	90.711(2)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/c$	1022.43(5)	4	1	120(2)	

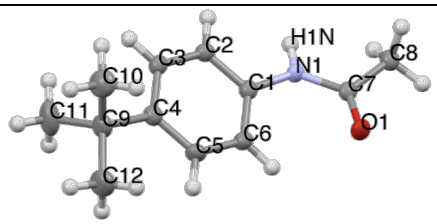
Me:p-cPr	**$C_{11}H_{13}NO$**	***N-(4-cyclopropylphenyl)acetamide***

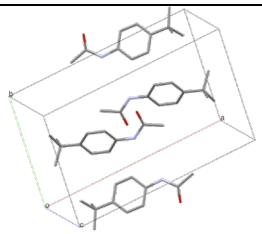




a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
9.557(4)	10.980(3)	9.641(2)	90	106.34(3)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/c$	970.824	4	1	173	

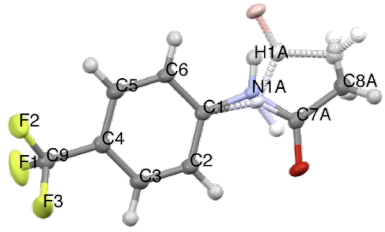
Me:p-^{t}Bu **$C_{12}H_{17}NO$** ***N-(4-tert-butylphenyl)acetamide***

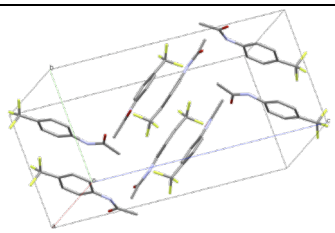




a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
16.4165(6)	9.6159(3)	7.1985(3)	90	97.315(2)	90
space group	**V [Å^3]**	**Z**	**Z'**	**T [K]**	
$P2_1/c$	1127.10(7)	4	1	120(2)	

Me:p-CF_3 **$C_9H_8F_3NO$** ***N-[4-(trifluoromethyl)phenyl]acetamide***

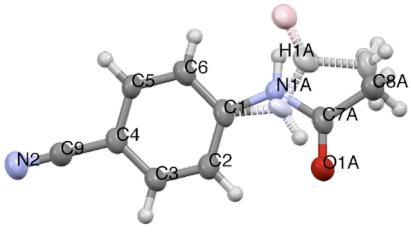


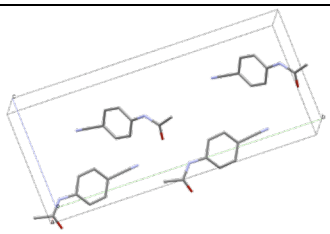


The acetamido group was found to be disordered over two positions and the disorder ratio refined to 83:17%. The minor component is shown in faded colours in the thermal ellipsoid plot.

a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
9.6630(3)	9.2888(2)	19.8221(5)	90	90	90
space group	**V [Å^3]**	**Z**	**Z'**	**T [K]**	
Pbca	1779.19(8)	8	1	120(2)	

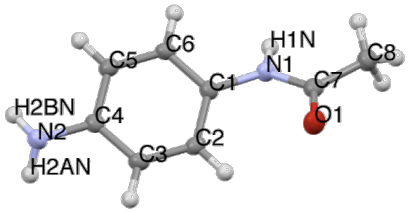
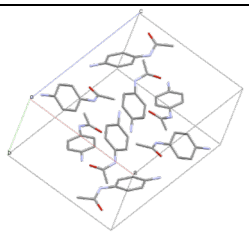
Me:p-CN **$C_9H_8N_2O$** ***N-(4-cyanophenyl)acetamide***

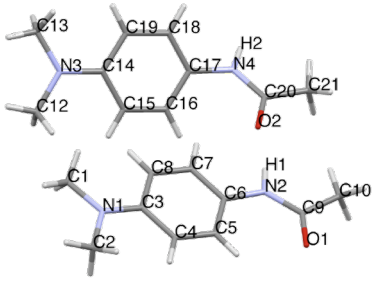
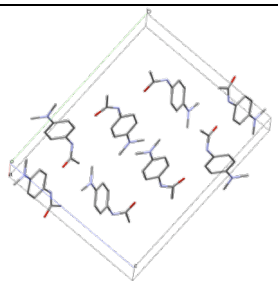




The acetamido group was found to be disordered over two positions and the disorder ratio refined to 50:50%. The second orientation is shown in faded colours in the thermal ellipsoid plot.

a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
3.8456(10)	22.190(5)	9.448(2)	90	90.048(15)	90
space group	**V [Å^3]**	**Z**	**Z'**	**T [K]**	
Cc	806.2(3)	4	1	120(2)	

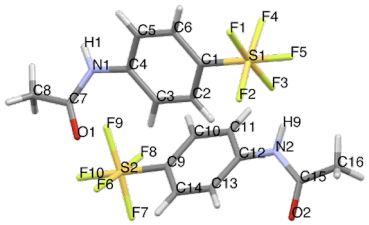
Me:p-NH_2	**$C_8H_{10}N_2O$**	***N-(4-aminophenyl)acetamide***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
11.9225(7)	7.2026(2)	17.4291(1)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
Pbca	1496.69(13)	8	1	120(2)	

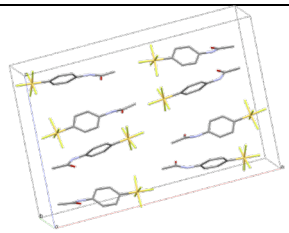
Me:p-NMe_2	**$C_{10}H_{14}N_2O$**	***N-[4-(dimethylamino)phenyl]acetamide***			
					

There are two crystallographically independent molecules in the ASU. The molecules are of similar geometry, except for a difference in the dihedral angle between the plane of the aromatic ring and the amide group. Although the molecules are conformationally distinct they are related by pseudo-translation.

a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
7.5101(3)	18.3314(8)	14.9572(7)	90	99.243(2)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/n$	2032.431	8	2	293	

Me:p-SF_5	**$C_8H_8F_5NOS$**	***N-[4-(pentafluoro-l^6-sulfanyl)phenyl]acetamide***

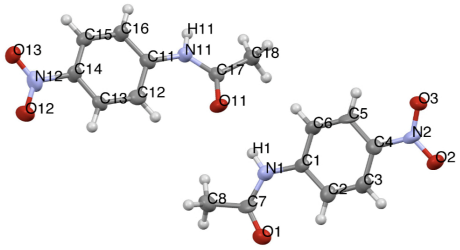


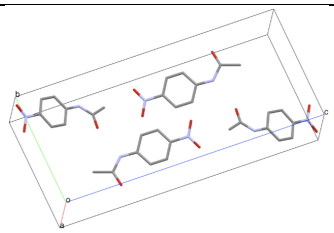


There are two crystallographically independent molecules in the ASU. The two molecules are of similar geometry, and the difference in the dihedral angle between the plane of the aromatic ring and the amide group is small compared to other structures in the library with Z'=2. Pseudo symmetry elements are less apparent in this structure although there are indications of a pseudo-glide relationship between the two independent molecules.

a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
23.706(2)	5.2534(4)	15.9163(13)	90	90	90
space group	**V [Å^3]**	**Z**	**Z'**	**T [K]**	
$Pca2_1$	1982.170	8	2	296	

Me:p-NO_2	**$C_8H_8N_2O_3$**	***N-(4-nitrophenyl)acetamide***

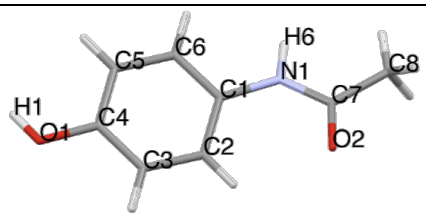


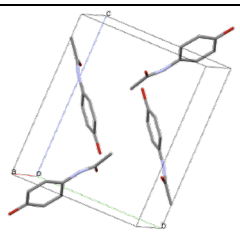


There are two crystallographically independent molecules in the ASU. The two molecules are of similar geometry, except for a difference in the dihedral angle between the plane of the aromatic ring and the amide group. Although the molecules are conformationally distinct they are related by pseudo-glide symmetry.

a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
3.821(5)	9.726(14)	22.67(3)	92.151(16)	93.359(18)	100.409(16)
space group	**V [Å^3]**	**Z**	**Z'**	**T [K]**	
$P\bar{1}$	826.2(19)	4	2	120(2)	

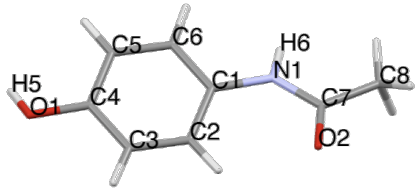
Me:p-OH(m)	**$C_8H_9NO_2$**	***N-(4-hydroxyphenyl)acetamide***

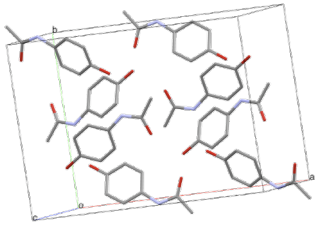




a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
7.0941(12)	9.2322(11)	11.6196(10)	90	97.821(10)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/n$	753.937	4	1	123	

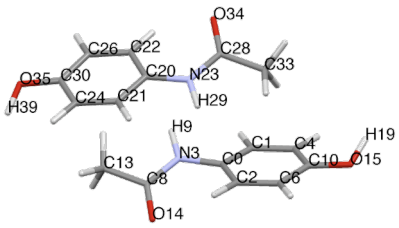
Me:p-OH(o1)	**$C_8H_9NO_2$**	***N-(4-hydroxyphenyl)acetamide***

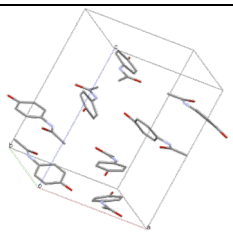




a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
17.1657(12)	11.7773(11)	7.212(2)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
Pbca	1458.018	8	1	123	

Me:p-OH(o2)	**$C_8H_9NO_2$**	***N-(4-hydroxyphenyl)acetamide***

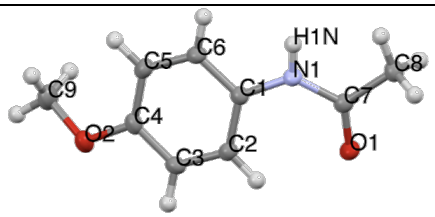


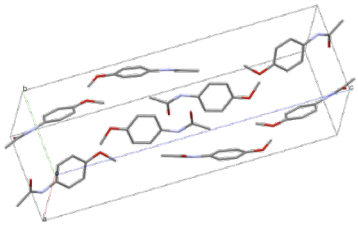


There are two crystallographically independent molecules in the ASU. The two molecules are of similar geometry, except for a difference in the dihedral angle between the plane of the aromatic ring and the amide group. Although the molecules are conformationally distinct they are related by pseudo-inversion.

a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
11.8352	8.5601	14.8199	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$Pca2_1$	1051.411	8	2	n/a	

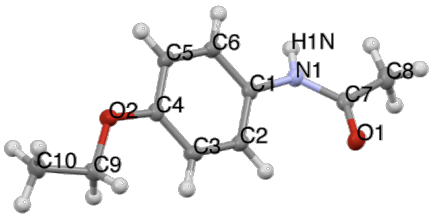
Me:p-OMe $C_9H_{11}NO_2$ ***N-(4-methoxyphenyl)acetamide***

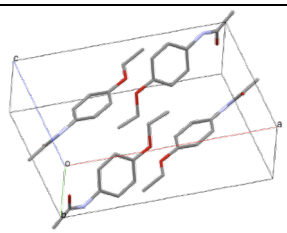




a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
9.1172(2)	7.4950(2)	24.6628(5)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
Pbca	1685.29(7)	8	1	120(2)	

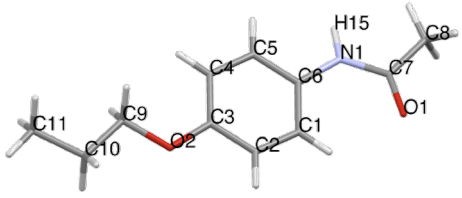
Me:p-OEt $C_{10}H_{13}NO_2$ ***N-(4-ethoxyphenyl)acetamide***

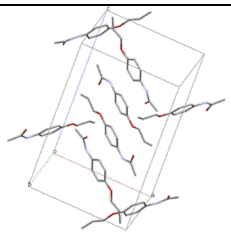




a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
13.2432(3)	9.5856(3)	7.4783(2)	90	103.345(2)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/c$	923.69(4)	4	1	120(2)	

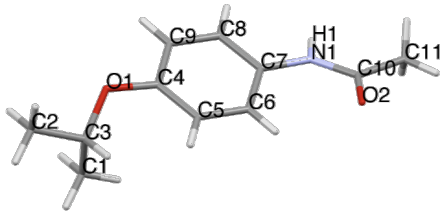
Me:p-OPr $C_{11}H_{15}NO_2$ ***N-(4-propoxyphenyl)acetamide***

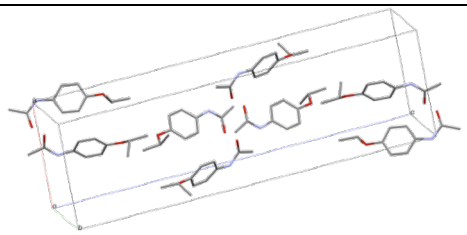




a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
11.7776(11)	9.4779(8)	18.8180(18)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
Pbcn	2100.595	8	1	93	

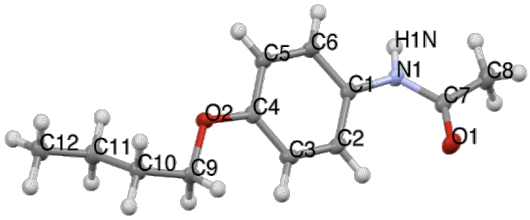
Me:p-O^{i}Pr $C_{11}H_{15}NO_2$ ***N-(4-isopropoxyphenyl)acetamide***

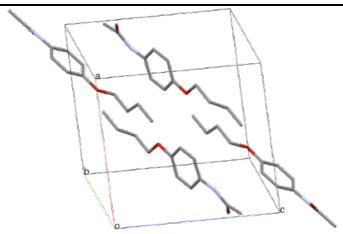




a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
9.3010(19)	7.6490(15)	31.394(6)	90	90	90
space group	**V [Å^3]**	**Z**	**Z'**	**T [K]**	
Pbca	2233.474	8	1	294	

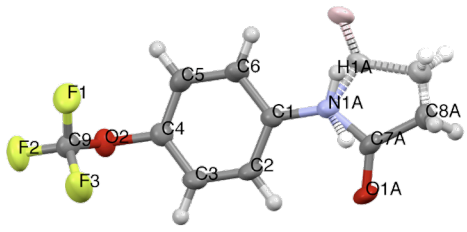
Me:p-OBu $C_{12}H_{17}NO_2$ ***N-(4-butoxyphenyl)acetamide***

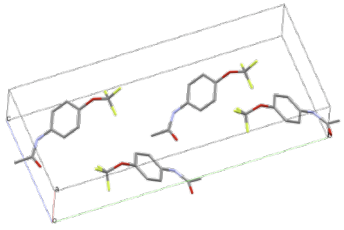




a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
10.5164(3)	9.8032(3)	10.9034(2)	90	95.691(2)	90
space group	**V [Å^3]**	**Z**	**Z'**	**T [K]**	
$P2_1/c$	1118.54(5)	4	1	120(2)	

Me:p-OCF$_3$ $C_9H_8F_3NO_2$ ***N-[4-(trifluoromethoxy)phenyl]acetamide***

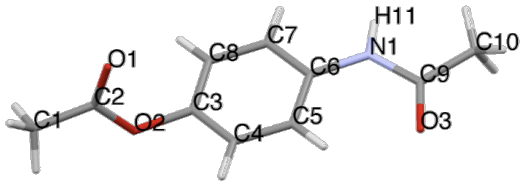


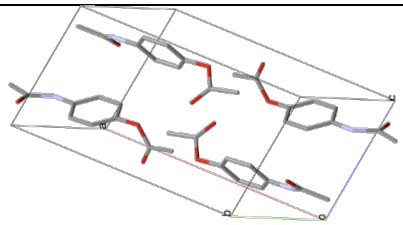


The acetamido group was found to be disordered over two positions and the disorder ratio refined to 51:49%. The minor component is shown in faded colours in the thermal ellipsoid plot on the left.

a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
4.4884(5)	22.168(3)	9.5284(13)	90	90.048(7)	90
space group	**V [Å^3]**	**Z**	**Z'**	**T [K]**	
Cc	948.1(2)	4	1	120(2)	

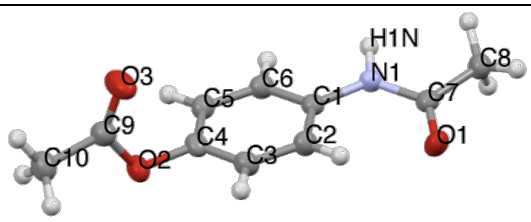
Me:p-OCOMe(m1) $C_{10}H_{11}NO_3$ ***4-acetamidophenyl acetate***

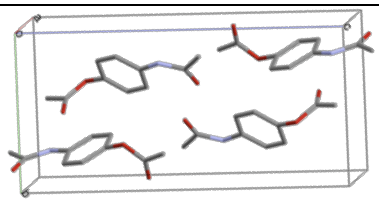




a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
7.2264(9)	8.0250(18)	16.5902(12)	90	92.052(8)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/n$	961.480	4	1	293	

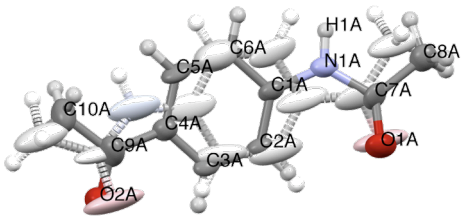
Me:p-OCOMe(m2) $C_{10}H_{11}NO_3$ ***4-acetamidophenyl acetate***

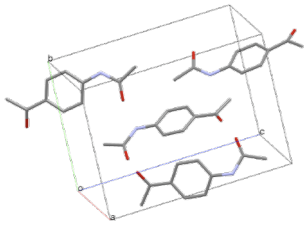




a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
13.8770(4)	9.5769(4)	7.4415(3)	90	103.022(2)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/c$	963.53(6)	4	1	120(2)	

Me:p-COMe $C_{10}H_{11}NO_2$ ***N-(4-acetylphenyl)acetamide***

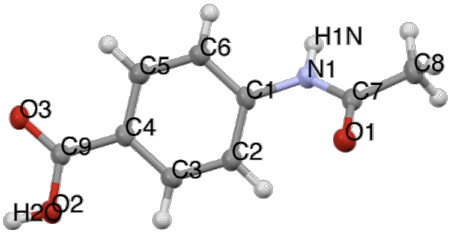


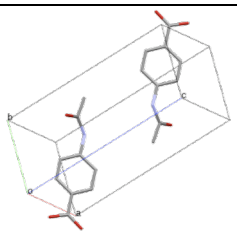


The whole molecule was found to be disordered over two positions and the disorder ratio refined to 57:43%. The minor component is shown in faded colours in the thermal ellipsoid plot.

a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
7.5138(5)	9.3146(7)	12.8999(11)	90	90	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_12_12_1$	902.84(12)	4	1	120(2)	

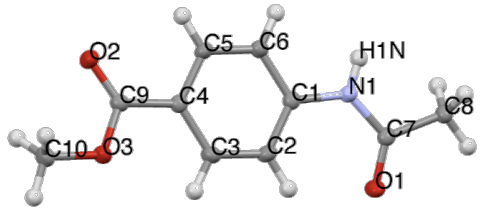
Me:p-COOH $C_9H_9NO_3$ ***4-acetamidobenzoic acid***

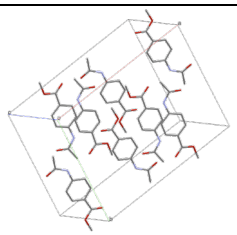




a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
5.0214(3)	6.8411(5)	12.2219(10)	89.465(4)	80.477(4)	79.234(4)
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P\bar{1}$	406.67(5)	2	1	120(2)	

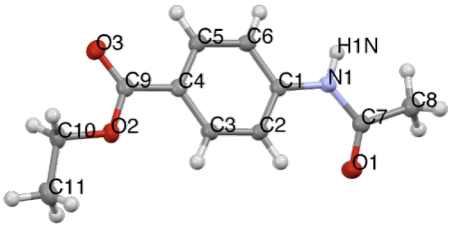
Me:p-COOMe $C_{10}H_{11}NO_3$ ***methyl 4-acetamidobenzoate***

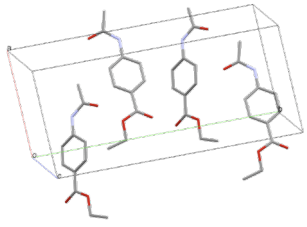




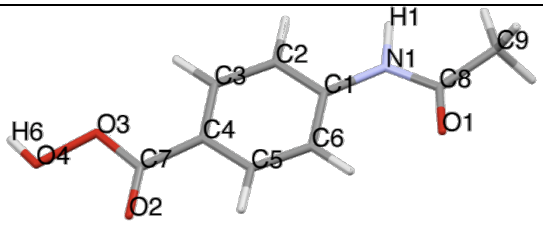
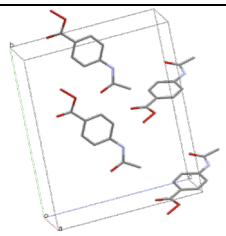
a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
17.5403(5)	13.7630(3)	7.7717(2)	90	101.876(1)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
C2/c	1835.99(8)	8	1	120(2)	

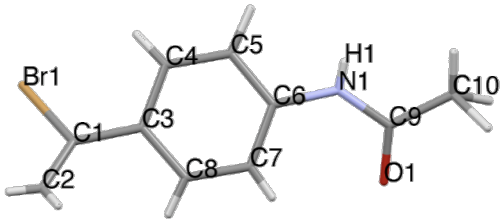
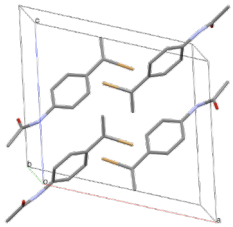
Me:p-COOEt $C_{11}H_{13}NO_3$ ***ethyl 4-acetamidobenzoate***





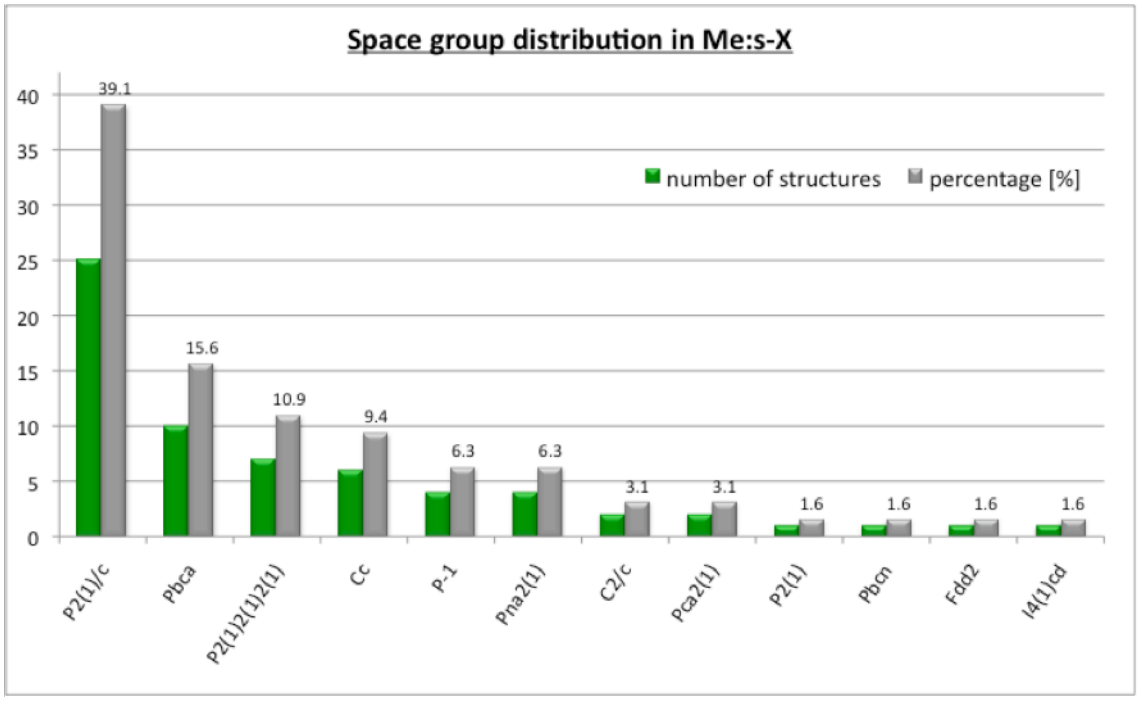
a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
8.4447(3)	19.2810(12)	7.3108(5)	90	117.051(4)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
Cc	1060.14(11)	4	1	120(2)	

Me:p-COOOH	**$C_9H_9NO_4$**	***4-acetamidobenzenecarboperoxoic acid***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
5.054(1)	14.747(5)	12.568(3)	90	92.88(2)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
Cc	935.527	4	1	293	

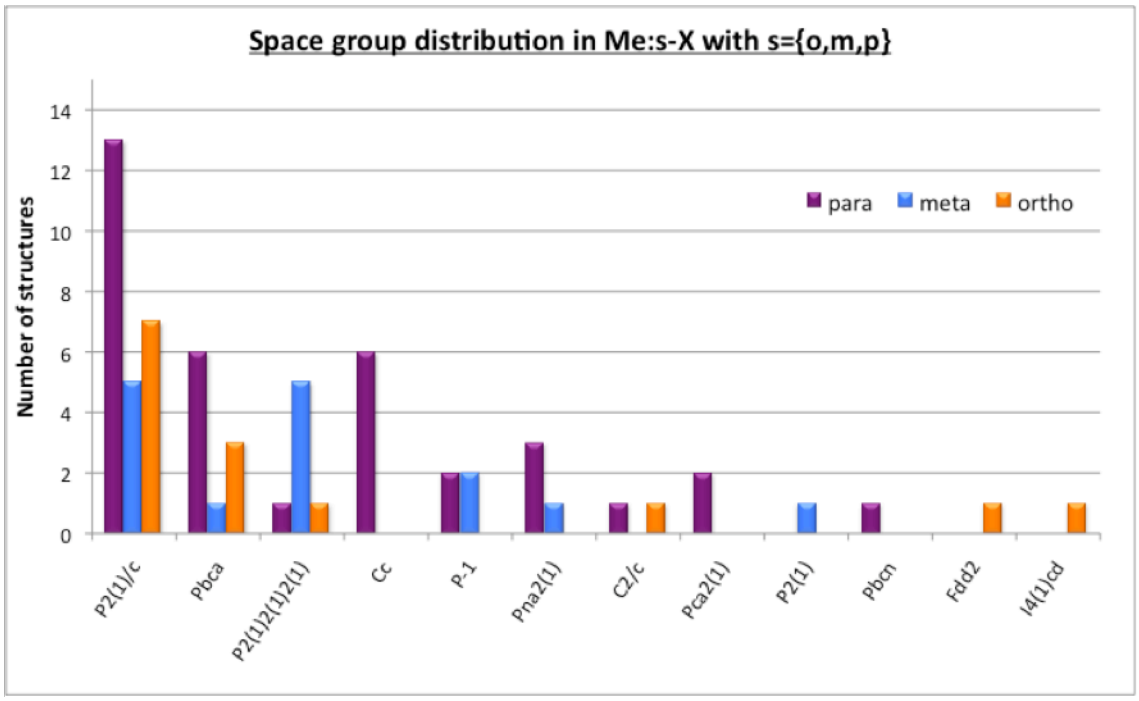
Me:p-CBrCH$_2$	**$C_{10}H_{10}BrNO$**	***N-[4-(1-bromovinyl)phenyl]acetamide***			
					
a [Å]	**b [Å]**	**c [Å]**	**α [°]**	**β [°]**	**γ [°]**
11.151(6)	9.237(5)	10.142(5)	90	93.393(8)	90
space group	**V [Å³]**	**Z**	**Z'**	**T [K]**	
$P2_1/c$	1003.216	4	1	293	

4.A.4. Discussion of space group distribution

The 64 crystal structures in the acetanilide library are distributed over 12 space groups as depicted in Figure 4.1.a. This distribution is further deconvoluted into the position of substitution on the phenyl ring in Figure 4.1.b.



(**a**)

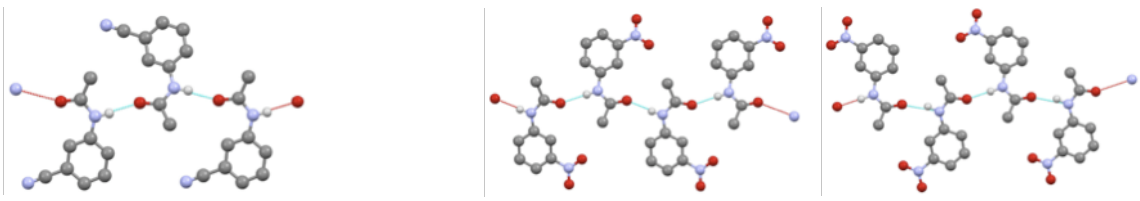


(**b**)

Figure 4.1: The distribution of space groups as found in all **Me:s-X** crystal structure (a) and broken down into the phenyl substitution pattern (b).

Figure 4.1.b shows that only 3 of the total 12 space groups in the acetanilide library occur irrespective of the position of substitution **s**. Two thirds of all acetanilide structures crystallise in $P2_1/c$, Pbca or $P2_12_12_1$ and whilst it may be argued that the library size is not sufficient to detect statistically significant trends, the spread nevertheless demonstrates that these space groups offer efficient close packing for many different acetanilides. The variety of space groups is largest in the para-substituted series, where 9 space groups are found compared with 6 in the meta as well as in the ortho series, but this may be an artefact of the number of crystal structures in the individual series.

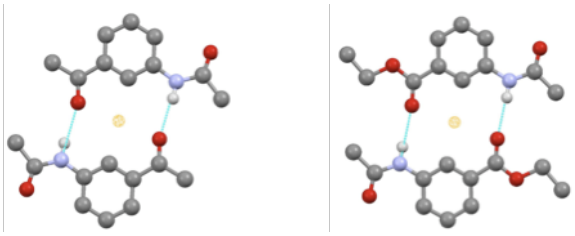
From Figure 4.1 it can further be seen that the majority of **Me:s-X** structures crystallise in space group $P2_1/c$, which is the most common space group for solid state structures as mentioned in Chapter 1. However, for the meta series the trend is somewhat different with an equal number of structures possessing space groups $P2_1/c$ and $P2_12_12_1$, clearly dominating over any other space groups. The meta acetanilides furthermore seem to preferably adopt space groups containing 2_1 screw elements compared to the ortho and para series. Only two meta structures break this trend, namely **Me:m-COMe** and **Me:m-COOEt**, which crystallise in $P\bar{1}$. A contributing factor to this space group choice may be the hydrogen bonding of these compounds compared with the other compounds in the series. Strong hydrogen bonding is possible for all acetanilides due to the amide group in the core molecule. Amongst the substituents **m-X** there are seven with additional HB functionality. Firstly the group of structures where **X** is a HB acceptor are considered. In the case of **X=CN**, **NO_2**, the substituents are not involved in strong hydrogen bonds and only the amide-amide hydrogen bond prevails as strong intermolecular interaction as displayed in Figure 4.2.



(**a**) (**b**)

Figure 4.2: Hydrogen bonded chain involving the amide group only shown for **Me:m-CN** (a) and for **Me:m-NO_2** (b). In (b) two chains formed by the individual pairs of syn and anti molecules are shown on the left and right hand side, respectively. Hydrogen atoms not involved in hydrogen bonding have been omitted for clarity.

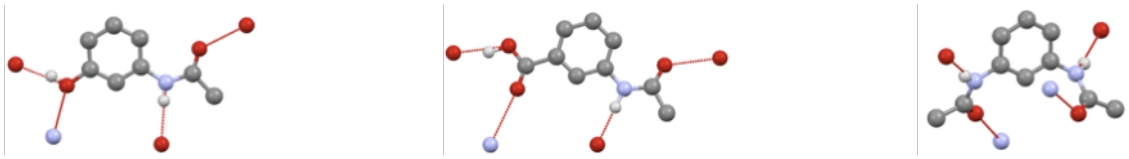
In the crystal assemblies **Me:m-COMe** and **Me:m-COOEt** hydrogen bonded dimers are formed between pairs of molecules as shown in Figure 4.3. This cyclic dimer possesses inversion symmetry and is of the type $R_2^2(14)$ in Etter notation. Of the two available HB acceptors, only O2 is involved in strong HB interactions in these structures.



(**a**) (**b**)

Figure 4.3: The cyclic hydrogen bonded dimer is shown for **Me:m-COMe** (**a**) and **Me:m-COOEt** (**b**). The inversion centre is indicated by the yellow dot in both dimers.

The other group containing HB active substituents **X**, consists of **Me:m-OH**, **Me:m-COOH** and **Me:m-NHCOMe**. In these structures all HB donors (substituent and amide) engage in strong HB interactions as depicted in Figure 4.4. Strong intermolecular hydrogen bonds are structural synthons that persist both in the liquid and solid state, where periodic ordering has to occur in the crystal.



(**a**) (**b**) (**c**)

Figure 4.4: The hydrogen bonded contacts are indicated for **Me:m-OH** (**a**), **Me:m-COOH** (**b**) and **Me:m-NHCOMe** (**c**).

Hence, if considering hydrogen bonded systems, a crystal structure may also be regarded as an assembly of individual HB synthons. Compounds containing a HB cyclic dimer may then impose the HB synthon symmetry on their crystal structure. For **Me:m-COMe** and **Me:m-COOEt** this means that the solid state structure is dominated by the inversion dimer and overall close packing can be achieved without higher symmetry than found in $P\bar{1}$.

Another point worthy of note is that 3 of 6 para-acetanilide crystal structures in Cc show disorder of the amide group. The two alternative orientations of the amide

residue are related by a 180° rotation about the amide-phenyl bond. Cc is a polar space group, i.e. there is no inversion centre and there are no other symmetry operators present that fix the unit cell origin. This means that the crystal structure is polarised or in other words that the unit cell has a significant dipole moment as the molecules and hence their electron densities are aligned in the same direction. The disorder in **Me:p-CN**, **Me:p-Et** and **Me:p-OCF$_3$** counteracts this effect since the directionality of the amide group is inverted for one half of the total structure.
Figure 4.5 compares the overall space group distribution in **Me:s-X** with all organic structures as defined in Chapter 1.

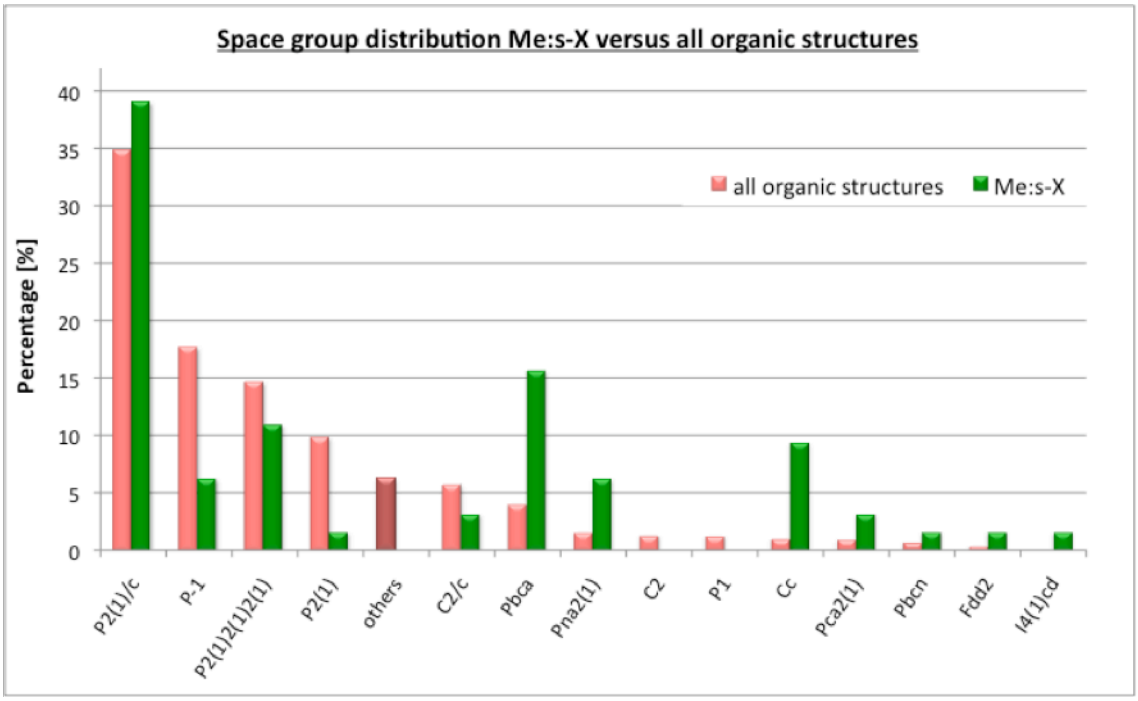


Figure 4.5: The most frequent space groups and their distribution in organic crystal structures are shown in comparison to the 12 space groups observed in **Me:s-X**.

This chart shows that except for $P2_1/c$ the compounds **Me:s-X** follow a quite different space group distribution than identified for organic structures in general. Especially the orthorhombic space groups Pbca, $Pna2_1$ and $Pca2_1$ as well as the monoclinic space group Cc occur more frequently whilst the space groups $P\bar{1}$ and $P2_1$ are less common in acetanilides than could be expected for crystal structures of organic molecules.

With the exception of $P\bar{1}$ and Cc this distribution is consistent with the acetanilide structures exhibiting a preference for 2_1 screw symmetry combined with either further 2_1 screw or glide symmetry. As mentioned above any HB synthons and their symmetry in the solid state are likely to influence the overall symmetry of the crystal structure. The hydrogen bonding in this family of compounds is further discussed in the next section of this chapter.

4.B. Structural similarity in mono-substituted acetanilides

The search for packing similarity in the family of mono-substituted acetanilides was facilitated with the program XPac as outlined in Chapters 1 and 3. Firstly the library of crystal structures was divided into three series according to the substitution on the phenyl ring. The structures within these series were then compared with each other using XPac. In a final comparison all structures, irrespective of the position of substitution, were considered. The number of comparisons, K, scales with the number of structures, N, according to:

$$K = [(N^2 - N)/2] \quad (4.1.)$$

N is related to the number of components in the ASU, so that each component of the ASU counts as one structure. Table 4.1 provides a summary of the comparisons per series and for the whole library.

Series	N	K
ortho	14	91
meta	21	210
para	38	703
all	73	2628

Table 4.1: Summary of the number of structure components N and the number of comparisons K for the different sets of structures.

In the following the results from the XPac analysis of the crystal structures are presented and discussed. For this particular study the 0D and 1D SCs are labelled according to their shape and connectivity including primary symmetry operations of the construct. For example a translational stack is described in shorthand notation as $\mathbf{S1_{[t]}}$, where S stands for stack, the numeral 1 refers to the type and the symmetry operation is given in square brackets as a subscript. The letter used to describe the shape of the similarity feature is chosen according to the terminology defined in Chapter 1. A hydrogen bonded chain of type 1 with glide symmetry would for example be denoted as $\mathbf{C1_{[g]}}$ and an inversion dimer as $\mathbf{D1_{[\bar{1}]}}$.

4.B.1. Structural Relationships in Me:o-X

The findings from the XPac analysis for the ortho acetanilide series are summarised in the similarity diagram in Figure 4.6 overleaf. The individual crystal structures are distinguished via the abbreviated identifier **s-X** and are listed at the top of Figure 4.6.

At a first glance there is only a little similarity in this group of crystal structures with six structures listed on the right hand side of Figure 4.6 having no packing feature in common with any other structures. However, the remaining 8 crystal structures contain recurring arrangements of molecules; two groups of analogous crystal structures and three one-dimensional SCs were identified.

There are two groups of analogous crystal structures: **Me:o-Cl(m)**, **Me:o-Br** and **Me:o-I** form group 1 and **Me:o-F** and **Me:o-Cl(o)** form group 2. In the following the crystal packing of group 1 and 2 is briefly discussed.

The 1D constructs identified in this series are two hydrogen bonded chains of the same type but with different symmetry, $\mathbf{C1_{[t]}}$ and $\mathbf{C1_{[g]}}$, and a single row of molecules, $\mathbf{R1_{[g]}}$. These features will be discussed in subsequent sections. The base vectors of all constructs can be found in Appendix 4-2. The intermolecular assembly will also be assessed with respect to energetic contributions. The total lattice energies are provided in Appendix 4-3 so that the pairwise interaction energies can be put into context of the overall crystal stability.

4.B.1.1 Three-dimensional similarity

The structures of **Me:o-Cl(m)**, **Me:o-Br** and **Me:o-I** all have the same 3D packing mode and crystallise in space group $P2_1/n$. **Me:o-F** and **Me:o-Cl(o)** are also analogously packed but in the higher symmetry space group Pbca. The unit cell parameters of the packing analogues are recorded in Table 4.2.

Structure	a [Å]	b [Å]	c [Å]	β [°]	V [Å³]
Me:o-Cl(m)	4.7468(4)	11.699(1)	14.640(2)	95.74(2)	808.924
Me:o-Br	4.7808(1)	11.9344(5)	14.6859(5)	96.820(2)	831.99(5)
Me:o-I	4.7969(1)	12.4977(3)	14.6790(3)	98.336(1)	870.71(3)
Me:o-F	10.5132(4)	9.4393(3)	15.1635(5)	90	1504.78(9)
Me:o-Cl(o)	10.5174(2)	9.4252(4)	16.2656(6)	90	1612.39(10)

Table 4.2: Unit cell parameters and cell volume of the ortho acetanilide packing analogues.

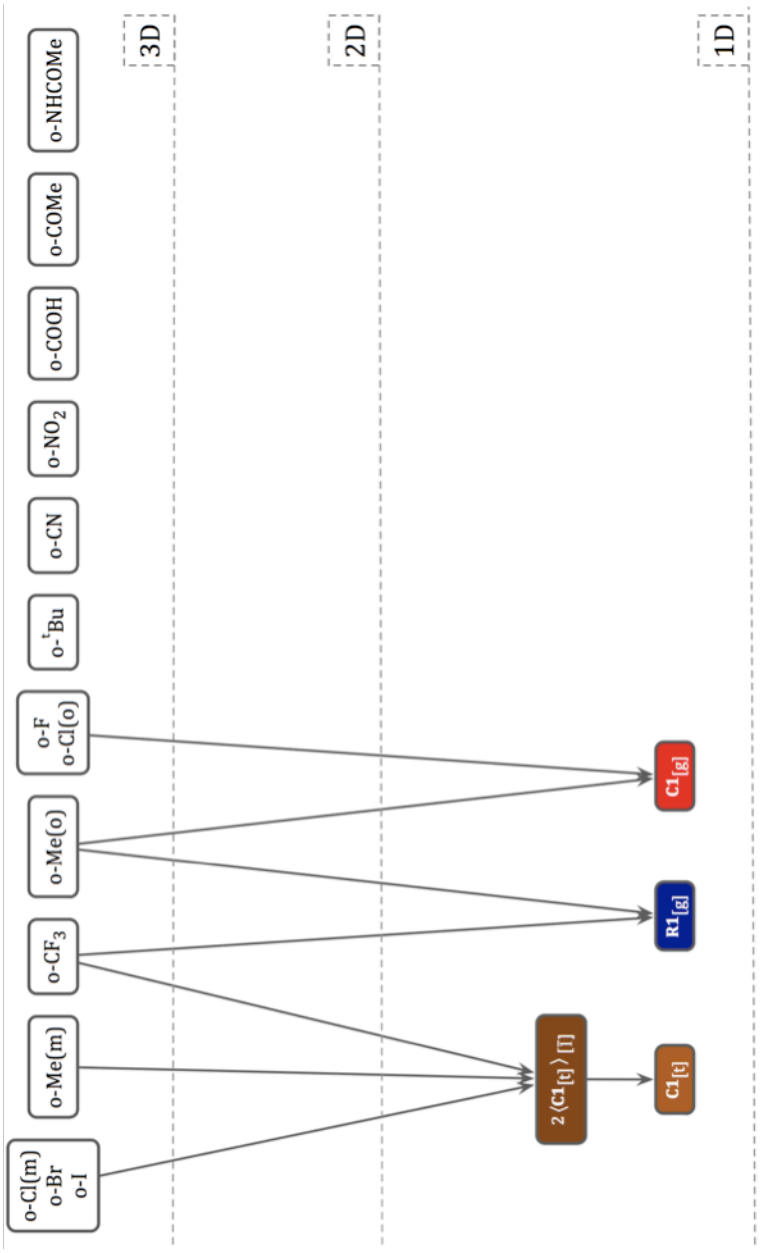


Figure 4.6: Structural similarity plot for ortho-substituted acetanilides.

As expected the parameters are very similar within the groups and the cell volume increases with size of the substituent. Interestingly **Me:o-Cl** has the ability to crystallise in both crystal lattices, whereby the unit cell of the orthorhombic polymorph of **Me:o-Cl** is twice the size of the unit cell in the monoclinic form.

The crystal packing in the two groups will be discussed further in the following using one representative for each group, namely **Me:o-Br** and **Me:o-F**, respectively.

The crystal assembly in **Me:o-Br** and **Me:o-F** is dominated by strong hydrogen bonding between neighbouring amide groups. The HB chains are displayed in Figure 4.7.

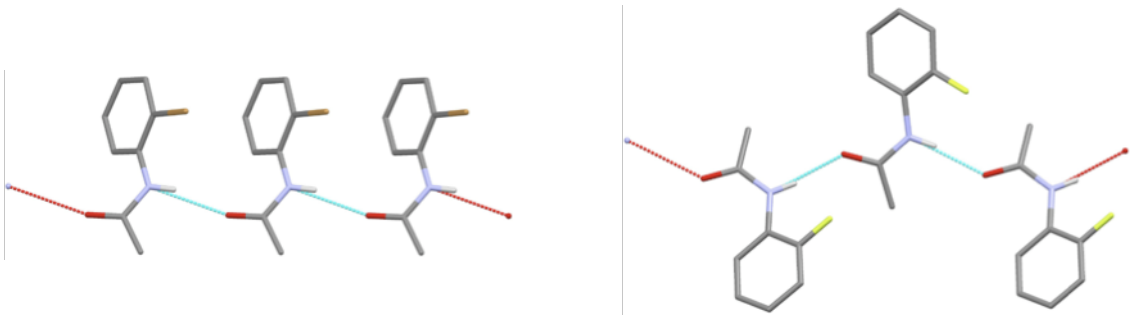


Figure 4.7: The hydrogen bonded chain is shown for **Me:o-Br** (left) and for **Me:o-F** (right). The disorder in **Me:o-F** was omitted for clarity.

The chains are of the same type: $\mathbf{C_1^1(4)}$ in Etter's graph set notation. They differ, however, in their symmetry of propagation. In **Me:o-Br** the HB chain has only translational symmetry and this SC can be described as $\mathbf{C1_{[t]}}$ whereas in **Me:o-F** the HB chain has glide symmetry and the SC is hence termed $\mathbf{C1_{[g]}}$. Both HB features are discussed in more detail in Section 4.B.1.3. In **Me:o-Br** neighbouring chains $\mathbf{C1_{[t]}}$ primarily pack with inversion and 2_1 screw symmetry. Glide symmetry is also present in the crystal structure, but the molecular assembly with this symmetry element has to be seen as less influential on the overall packing.

The close packing of chain $\mathbf{C1_{[g]}}$ in **Me:o-F** on the other hand involves predominantly glide and inversion symmetry with 2_1 screw symmetry playing a secondary role. This assessment of the 'importance' of the intermolecular symmetry relationships is based on lattice energy calculations using Gavezzotti's Pixel method as described in Chapters 1 and 3. This method not only provides an estimate of the total lattice energy, but also allows the partitioning of this total sum

into the contribution from pairwise molecular interactions. However, it should be kept in mind that the intermolecular energies do not arise due to the symmetry elements *per se*, but are of course the result of the geometrical arrangement of the molecules with respect to each other and the varying regions of charge concentration and depletion within the molecular electron densities. However the spatial intermolecular assemblies are easily identified based on their symmetry.

Table 4.3 lists the most significant pairwise interaction energies between molecules A and B for **Me:o-Br** and **Me:o-F**, whereby the numbering of the molecules was generated by the Pixel calculation routine.

A	B	Symmetry B	r_{A-B}	E_{tot}	E_{coul}	E_{pol}	E_{disp}	E_{rep}
Me:o-Br								
1	11	x-1, y, z **[t]**	4.781	-39.8	-37.8	-14.4	-33.5	46
1	37	1.5-x, y-½, ½-z **[2_1]**	7.038	-20.6	-11	-3.7	-22.6	16.7
1	57	2-x, -y, -z **[$\bar{1}$]**	6.183	-18.7	-8.9	-3.9	-14.7	8.8
1	53	1-x, -y, -z **[$\bar{1}$]**	4.287	-14.5	-2.8	-3	-30.8	22.1
1	2	½-x, y+½, ½-z **[2_1]**	7.803	-8.2	-0.8	-2.1	-15	9.7
1	55	1-x, 1-y, -z **[$\bar{1}$]**	10.59	-3.3	-1.6	-0.5	-4.9	3.8
1	73	x-½, ½-y, z-½ **[g]**	8.42	-2.8	-0.4	-0.2	-2.4	0.2
Me:o-F								
1	66	½-x, y-½, z **[g]**	6.222	-35.9	-48.1	-18.3	-19.4	49.9
1	41	-x, -y, 1-z **[$\bar{1}$]**	3.961	-15.7	-1.7	-2.2	-27.4	15.5
1	48	x-½, y, 1.5-z **[g]**	6.518	-13.2	-3.1	-2.1	-18	10
1	27	-x, y-½, 1.5-z **[2_1]**	6.21	-8.7	-2.4	-2	-8.7	4.4
1	8	-½-x, y-½, z **[g]**	7.999	-8.5	-3.6	-0.8	-8.8	4.7
1	31	x -½, -½-y, 1-z **[2_1]**	7.654	-8.2	-2.4	-2.3	-9.9	6.4
1	45	1-x, -y, 1-z **[$\bar{1}$]**	10.047	-2	-0.8	-0.8	-6.5	6.1

Table 4.3: Pairwise intermolecular interaction energies between molecules A and B are given in kJ·mol^{-1} for **Me:o-Br** and **Me:o-F**. The distance between the centres of mass, r_{A-B}, is given in Å. The entries in brackets refer to the symmetry element as follows: translation **[t]**, inversion **[$\bar{1}$]**, 2_1 screw axis **[2_1]**, glide **[g]**.

Appendix 4-4 contains the corresponding interaction energies of **Me:o-Cl(m)** and **Me:o-Cl(o)**.

Figure 4.8 displays the molecules of **Me:o-Br** and **Me:o-F** corresponding to the values in Table 4.3.

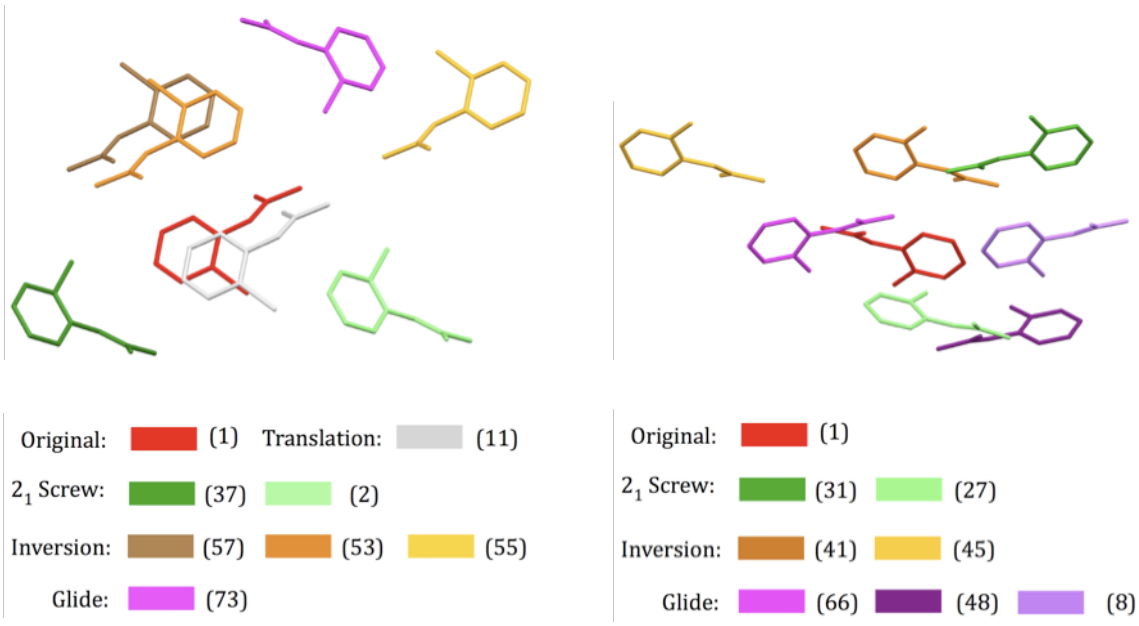


Figure 4.8: Partial coordination spheres of **Me:o-Br** (left) and **Me:o-F** (right). The molecule numbering is consistent with Table 4.3.

The red molecule is the reference positioned at x, y, z and all other molecules are symmetry-generated from this molecule. The interaction energies are estimated between this reference molecule and its symmetry equivalents.

The values in Table 4.3 reveal that the strongest intermolecular interaction in the two crystal assemblies arises from the hydrogen bonded chains, **C1$_{[t]}$** and **C1$_{[g]}$** with 39.8 kJ·mol^{-1} and 35.9 kJ·mol^{-1}, respectively. The next strongest interactions in **Me:o-Br** are found between molecules related by 2_1 screw (20.6 kJ·mol^{-1}) and inversion symmetry (18.7 kJ·mol^{-1}). In comparison the highest energy between **Me:o-Br** molecules related by glide symmetry is smaller by one order of magnitude with 2.8 kJ·mol^{-1}, hence this symmetry element is considered as secondary in the Aufbau principle of the crystal structure in **Me:o-Br**.

In **Me:o-F** molecular pairs with inversion (15.7 kJ·mol^{-1}) and glide symmetry (13.2 kJ·mol^{-1}) are the strongest contributors to the overall lattice energy after the HB interaction with glide symmetry. Although they are slightly weaker compared with the primary interactions in **Me:o-Br**. It is noticeable that there is a less clear division with respect to the symmetry of the intermolecular arrangement in **Me:o-F** than there is in **Me:o-Br** and molecular pairs related via 2_1 screw symmetry (8.7 kJ·mol^{-1}) are not very different from arrangements with glide or

inversion symmetry. Hence the assembly of chain **C1$_{[g]}$** in **Me:o-F** is not dominated by specific intermolecular interactions but instead it is achieved through utilising many energetically similar interactions simultaneously.

Table 4.3 also lists the individual components of the total intermolecular energy as obtained with the Pixel method. Both HB chains are primarily stabilised by Coloumb and polarisation forces with energies typical for a N–H···O hydrogen bond interaction[7]. Dispersive forces further enhance the intermolecular energy of **C1$_{[t]}$**, whereas the dispersive energy is only moderate in **C1$_{[g]}$** and this hydrogen bond is hence slightly less energetic than **C1$_{[t]}$** by approximately 3 kJ·mol^{-1}. The attractive energies between molecules without intermolecular hydrogen bonding on the other hand, are mainly dominated by dispersive forces as can be deduced from the values in Table 4.3. Repulsion is another noticeable factor demonstrating that the close packing of molecules is associated with an energy loss due to unfavourable proximity of the molecules and is counterbalanced by attractive forces to yield an overall stable arrangement.

The energy values calculated for **Me:o-Cl(m)** (*c.f.* Appendix 4-4) are very similar to those of **Me:o-Br** indicating that the substituent plays a minor role in the assembly of the molecular core in the crystal lattice. This is further supported by the overall lattice energies, which are very similar with 99.7 kJ·mol^{-1} and 100.6 kJ·mol^{-1} for the crystal structures of **Me:o-Br** and **Me:o-Cl(m)**, respectively. The influence of the substituent is however reflected in the somewhat shorter separation of the molecular centres of mass between the symmetry related molecules, which can be ascribed to the smaller van der Waals radius of chlorine. Hence the unit cell of **Me:o-Cl(m)** is more condensed, i.e the volume is smaller, than that of **Me:o-Br** and the separation of the centres of mass in the hydrogen bond chain in **Me:o-Cl(m)** is shorter by 0.032 Å compared with **Me:o-Br**.

The interaction energies of **Me:o-Cl(o)** (Appendix 4-4) are also very similar to those of its packing analogue **Me:o-F**, except for one interaction between the molecular pair 1-28 in **Me:o-Cl(o)** which are related by 2_1 screw symmetry. In this pair the substituent is directed towards the centre of the neighbouring phenyl ring similar to the geometry of the T-shaped benzene dimer as shown in Figure 4.9.

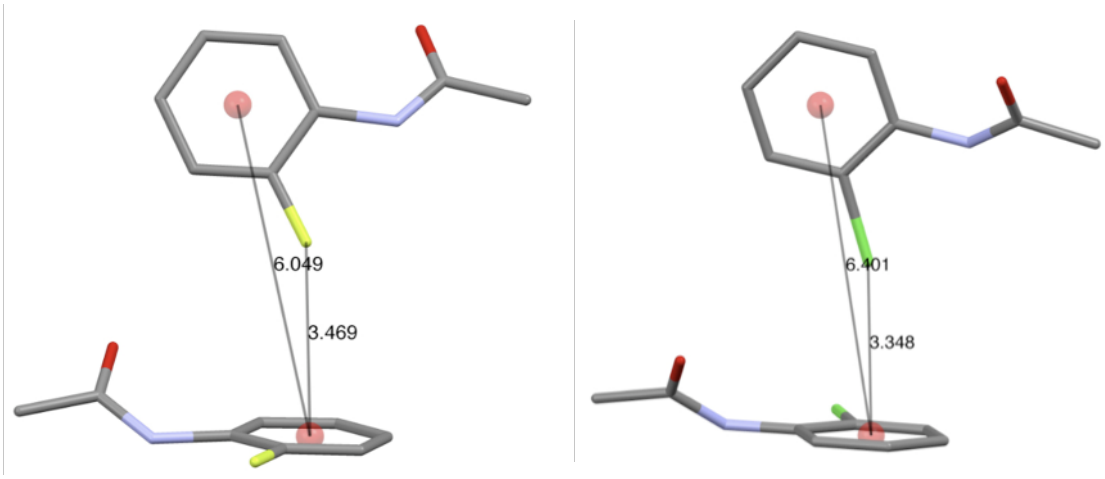


Figure 4.9: The T-shaped dimer with 2_1 screw symmetry shown for **Me:o-F** (left) and for **Me:o-Cl(o)** (right). Distances between ring centroids and substituent to centroid are also included. The disorder in **Me:o-F** was omitted for clarity.

Whilst in **Me:o-F** the separation between the ring centroids exceeds the range of π-π interaction, this distance is well within the π-π interaction range in **Me:o-Cl(o)**[8] leading to the higher dispersive energy for this particular intermolecular arrangement.

According to the overall lattice energies **Me:o-Cl(o)** is more stable than **Me:o-F** (99.5 kJ·mol^{-1} vs 91.1 kJ·mol^{-1}, respectively), but this should be treated with caution. The lattice energy of **Me:o-F** was calculated for the major component only and the disorder mentioned above, although small in amount, may have some further stabilising or destabilising effects on the crystal structure.

4.B.1.2. The polymorphs Me:o-Cl(m) and Me:o-Cl(o)

The lattice energy difference between the polymorphs **Me:o-Cl(m)** and **Me:o-Cl(o)** is too small to allow the conclusive assignment of the thermodynamic stability. Although the hydrogen bond is of the same type in both polymorphs the symmetry of the chain results in quite different geometrical arrangements. The chain **C1$_{[g]}$** is energetically marginally less favourable than **C1$_{[t]}$** (ΔE = 2.8 kJ·mol^{-1}), but this is compensated for by utilising intermolecular interactions differently, so that the total lattice energies are comparable within the accuracy of the calculation method. In fact the discussion of the intermolecular interaction energies in the

preceding section has demonstrated that different crystal assemblies of the same molecule can be achieved at no great energetic cost. For example the spatial arrangement of two molecules can be quite different yet the weak intermolecular interactions are energetically very similar. Also possible is the tuning of the intermolecular geometry so as to maximise the interaction energy. In this way many different possibilities are available for the close packing of a single molecule and hence the observation of polymorphic forms is not surprising.

At the molecular level the two polymorphs differ slightly in the dihedral angle between the phenyl ring and the amide group: in **Me:o-Cl(m)** this angle is 42.65˚, whilst in **Me:o-Cl(o)** it is 38.23˚. The role of this dihedral angle is further analysed in Chapter 5.

4.B.1.3 Description of 1D SCs

XPac identified three 1D features common to at least two crystal structures as shown in the structural similarity diagram in Figure 4.6. These are the HB chains **C1**$_{[g]}$ and **C1**$_{[t]}$, already mentioned in the previous sections, and a row of molecules related by glide symmetry, **R1**$_{[g]}$. These 1D features are further discussed in the following sections.

4.B.1.3.a The chain C1$_{[g]}$

In addition to **Me:o-F** and **Me:o-Cl(o)** the hydrogen bonded chain **C1**$_{[g]}$ is also present in **Me:o-Me(o)** as shown in Figure 4.10.

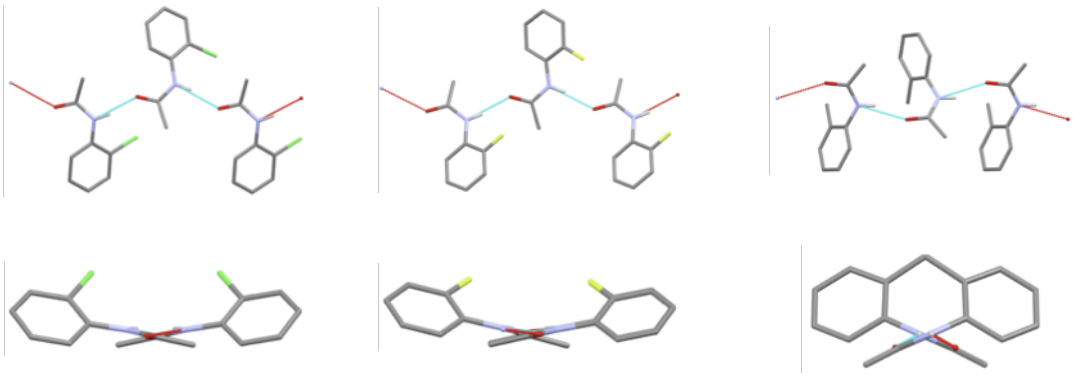


Figure 4.10: The HB chain **C1**$_{[g]}$ is shown for **Me:o-Cl(o)** (left), **Me:o-F** (middle) and **Me:o-Me(o)** (right) as viewed perpendicular to the chain (top) and along the hydrogen bond (bottom). The phenyl ring disorder in **Me:o-F** is not included.

Due to the glide symmetry the SC assumes the shape of a V. It can be seen in Figure 4.10 that the phenyl ring in **Me:o-Me(o)** takes a different orientation and the angle of the V is more acute in **Me:o-Me(o)** compared to **Me:o-Cl(o)** and **Me:o-F**. In the crystal structure of **Me:o-Me(o)** the methyl group is positioned on the same side of the molecules as the amide oxygen (*syn*), whereas the phenyl substituents are located on the opposite side to the amide oxygen (*anti*) in **Me:o-F** and **Me:o-Cl(o)** as displayed in Figure 4.11.

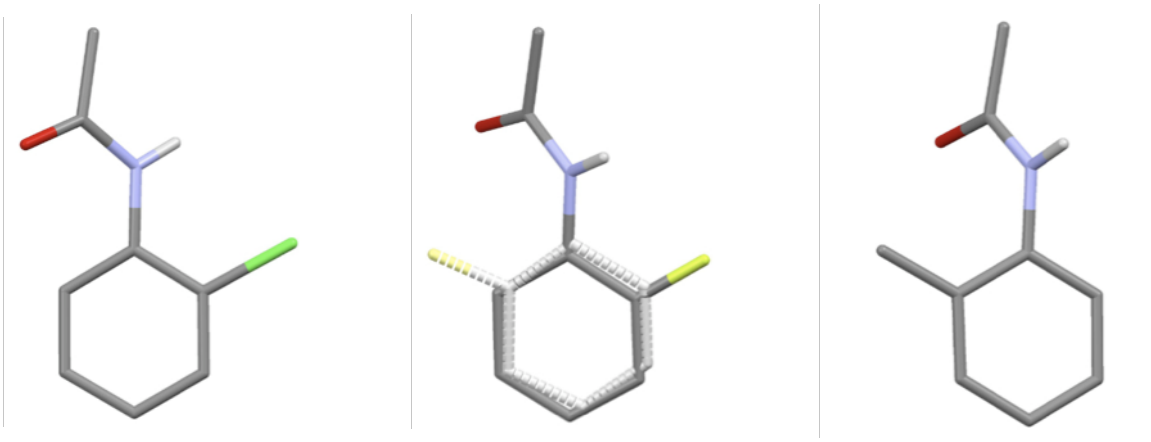


Figure 4.11: Molecular configurations in **Me:o-Cl(o)** (left), **Me:o-F** (middle) and **Me:o-Me(o)** (right). The phenyl ring disorder in **Me:o-F** is shown.

Since the methyl group is bulkier than fluorine or chlorine it could be assumed that the configurational difference is caused by steric hindrance between the methyl hydrogen atoms and the amide hydrogen atom and indeed the largest deviation from molecular planarity occurs in **Me:o-Me(o)**. However, in the polymorph **Me:o-Me(m)** the molecules adopt the *anti* configuration as found in the structures with smaller substituents **X**. Furthermore the minor component of the disordered **Me:o-F** phenyl ring also assumes the *syn* configuration. Hence other factors must be responsible for the variation in configuration.

Table 4.4 contains the hydrogen bond distances and the associated energies for **C1**$_{[g]}$ are recorded in Table 4.5.

Structure	$r_{N\cdots O}$ [Å]	$\angle_{N-H\cdots O}$ [°]	Symmetry
Me:o-F	2.854(1)	171(2)	glide
Me:o-Cl(o)	2.882(2)	168(2)	glide
Me:o-Me(o)	2.801	154	glide

Table 4.4: Hydrogen bond distances for SC **C1**$_{[g]}$.

Structure	r_{A-B} [Å]	E_{tot}	E_{coul}	E_{pol}	E_{disp}	E_{rep}
Me:o-F	6.222	-35.9	-48.1	-18.3	-19.4	49.9
Me:o-Cl(o)	6.256	-36.7	-46.1	-17.8	-23.4	50.6
Me:o-Me(o)	4.786	-40.1	-43.4	-17.4	-30.6	51.4

Table 4.5: Intermolecular interaction energies for SC **C1**$_{[g]}$. The energies are given in kJ·mol^{-1}.

The hydrogen bond distance in **Me:o-Me(o)** is the shortest and most energetic in this trio of structures. Looking at the individual interaction energies it can be seen that although the Coloumbic term is the smallest in **Me:o-Me(o)** the overall interaction energy is enhanced by the dispersive term. This increased dispersion interaction arises due to the weak interaction between the methyl group and a neighbouring phenyl ring in **C1$_{[g]}$**. As illustrated in Figure 4.12 the methyl group is positioned directly over the centre of the phenyl ring and the ring centroid to methyl carbon distance falls into the range of weak HB interactions (this type of interaction is often described as CH···π interaction). This interaction is only possible for molecules in the *syn* configuration and at acute angles of **C1$_{[g]}$**. The particular shape of **C1$_{[g]}$** in **Me:o-Me(o)** may hence be the result of maximising the intermolecular interaction energies.

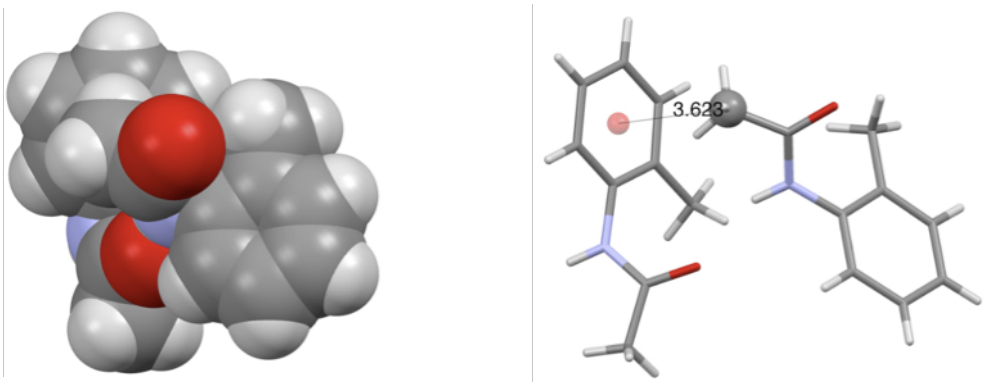


Figure 4.12: *Left:* The overlap between phenyl ring and methyl group in **C1$_{[g]}$** as found in **Me:o-Me(o)**. *Right:* The corresponding distance between methyl carbon atom and phenyl ring centroid is displayed.

The Coloumb energy of **C1$_{[g]}$** is the strongest in **Me:o-F** which indicates that the amide hydrogen is the most acidic in this compound, i.e. the charge separation between amide oxygen and amide hydrogen is the most pronounced. Fluorine is the strongest electron withdrawing of the three substituents and due to the conjugated molecular backbone this effect is most likely responsible for the charge depletion at the amide hydrogen atom.

4.B.1.3.b The chain C1[t]

The HB chain **C1[t]** occurs in the crystal structures **Me:o-Cl(m)**, **Me:o-Br**, **Me:o-I**, **Me:o-CF_3** and **Me:o-Me(m)**. The construct is shown in Figure 4.13 for each of the structures. The construct propagates with the depth of the molecules which coincides with the shortest unit cell axes (*c.f.* base vectors in Appendix 4-2 and unit cell dimensions in Section A of this chapter).

This translational feature is the most commonly observed SC in the series of ortho-substituted acetanilides and hence presents an intermolecular arrangement flexible enough to accommodate substituents of varying size and electronegativity.

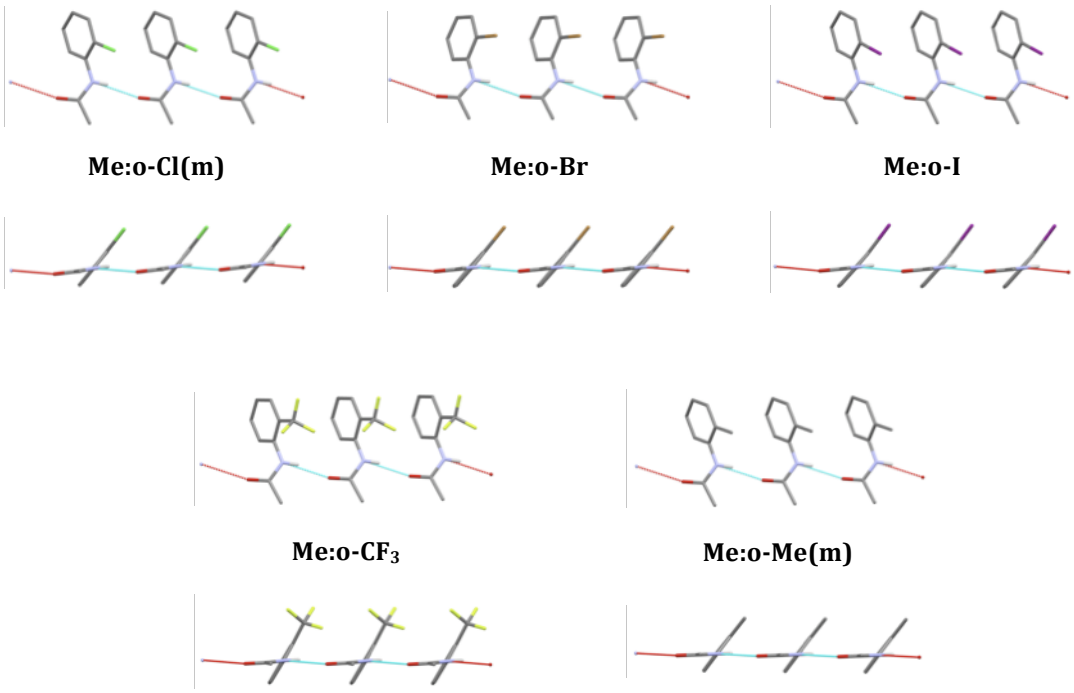


Figure 4.13: C1[t] as it occurs in **Me:o-Cl(m)** (top left), **Me:o-Br** (top middle), **Me:o-I** (top right), **Me:o-CF_3** (bottom left) and **Me:o-Me(m)** (bottom right). Two perspectives as viewed perpendicular to the hydrogen bond are shown.

The hydrogen bond parameters and the interaction energies are listed in Tables 4.6 and 4.7, respectively.

Structure	$r_{N\cdots O}$ [Å]	$\angle_{N-H\cdots O}$ [°]	Symmetry
Me:o-Cl(m)	2.863	162	translation
Me:o-Br	2.895(2)	161(2)	translation
Me:o-I	2.915(2)	159(2)	translation
Me:o-CF$_3$	2.898(2)	165(3)	translation
Me:o-Me(m)	2.941(2)	165(2)	translation

Table 4.6: Hydrogen bond distances for SC **C1$_{[t]}$**.

Structure	r_{A-B} [Å]	E_{tot}	E_{coul}	E_{pol}	E_{disp}	E_{rep}
Me:o-Cl(m)	4.747	-39.5	-44.2	-17.6	-33.1	55.3
Me:o-Br	4.781	-39.8	-37.8	-14.4	-33.5	46.0
Me:o-CF$_3$	4.782	-37.3	-40.3	-15.2	-26.5	44.7
Me:o-Me(m)	4.828	-42.3	-43.3	-15.6	-30.3	47.0

Table 4.7: Intermolecular interaction energies for SC **C1$_{[t]}$**. Energies are given in kJ·mol^{-1}.

The spatial arrangement of **C1$_{[t]}$** results in strong Coulombic and polarisation interactions with energies typical for hydrogen bonded systems. Compared with **C1$_{[g]}$** however, the dispersion energy is more favourable. This can be accredited to the parallel arrangement of the phenyl rings facilitating some interaction involving the π system. At a first glance the separations between the ring centroids of neighbouring molecules appears too large for typical organic π-π stacking (*c.f.* Chapter 1.A.3.) implying that a direct π-π interaction between the rings is unlikely. However, inspection of a space filling diagram of a pair of translationally related HB bonded molecules of **Me:o-Br** (Figure 4.14) for example, shows that the stack is off-set so that the substituent is positioned almost perfectly above the centre of a neighbouring phenyl ring. The parameters associated with the separation of the substituted π-systems in **C1$_{[t]}$** are recorded in Table 4.8. It can now be seen that the distances between the substituents and the neighbouring ring centroids are close to the sum of the van der Waals radii of the constituent atoms.

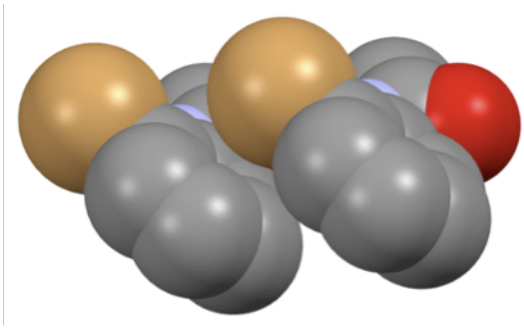


Figure 4.14: Space filling diagram of neighbouring **Me:o-Br** molecules in **C1$_{[t]}$**.

Structure	d_{ring} [Å]	d_{offset} [Å]	$d_{X\cdots ring}$ [Å]
Me:o-Cl(m)	3.476	3.233	3.547
Me:o-Br	3.500	3.257	3.606
Me:o-I	3.527	3.251	3.740
Me:o-CF$_3$	3.974	2.660	4.076/3.224*
Me:o-Me(m)	3.631	3.182	3.629

Table 4.8: d_{ring} = distance between neighbouring ring planes in **C1$_{[t]}$**; d_{offset} = offset of phenyl ring with respect to neighbouring ring; $d_{X\cdots ring}$ = distance between substituent **X** and ring centroid in neighbouring phenyl ring. * The first value refers to the distance between ring centroid and the CF_3 carbon whilst the second value is measured from the F atom closest to the ring.

This close proximity may thus be responsible for the attractive dispersion between the hydrogen bonded molecules. Interestingly when comparing the interaction energy values of **C1$_{[t]}$** and **C1$_{[g]}$** for the polymorphs of **Me:o-Cl** and **Me:o-Me** it transpires that **C1$_{[t]}$** is a slightly more stable arrangement than **C1$_{[g]}$**. Whilst the Coulomb, polarisation and dispersion energies are very similar in the case of **Me:o-Me** the repulsive energy is higher if **Me:o-Me** aggregates in **C1$_{[g]}$** – possibly a penalty for the shorter HB distance in **C1$_{[g]}$**.

For **Me:o-Cl** the most pronounced energy difference is found in the dispersive energy when comparing **C1$_{[t]}$** with **C1$_{[g]}$**. This is due to the missing overlap of substituent and π-system in the **C1$_{[g]}$** arrangement of **Me:o-Cl** as already discussed above.

The 1D feature **C1$_{[t]}$** can actually be extended to a double chain. As shown in the similarity plot (Figure 4.6) the highest order SC common to the structures **Me:o-Cl(m)**, **Me:o-Br**, **Me:o-I**, **Me:o-CF$_3$** and **Me:o-Me(m)** is abbreviated as **$2<C1_{[t]}>_{[\bar{1}]}$**. This is the shorthand notation for a double chain with inversion symmetry between neighbouring chains of the kind **C1$_{[t]}$**. Figure 4.15 shows this feature for **Me:o-Br**.

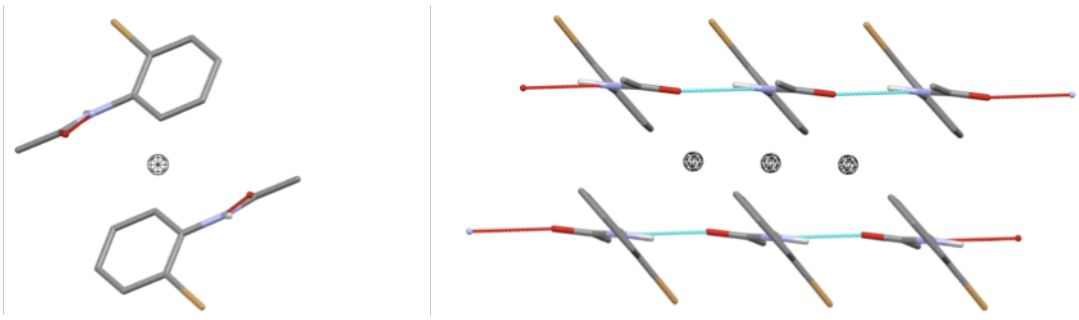


Figure 4.15: $\mathbf{2{<}C1_{[t]}{>}_{[\bar{1}]}}$ as it occurs in **Me:o-Br** shown as viewed along (left) and perpendicular to (right) the hydrogen bond. The inversion centres are included as black circles.

4.B.1.3.c The row $R1_{[g]}$

The 1D SC **$R1_{[g]}$** occurs in two structures only, namely **Me:o-Me(o)** and **Me:o-CF_3**. The construct as found in these two structures is shown in Figure 4.16. Neighbouring molecules in this construct are related by glide symmetry and there is no strong HB interaction present in this row.

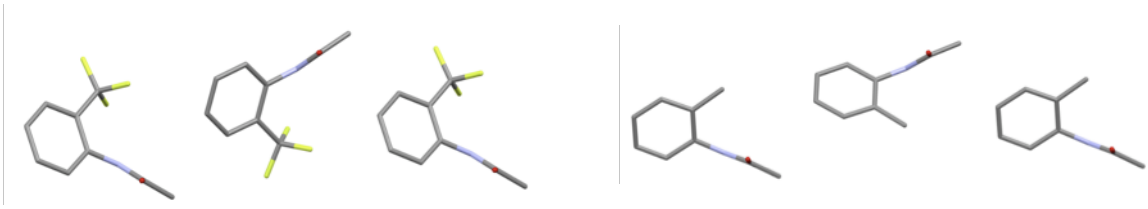


Figure 4.16: Construct **$R1_{[g]}$** as it occurs in **Me:o-CF_3** (left) and **Me:o-Me(o)** (right).

As can be seen in Figure 4.16 the direction of propagation is along the diagonal between the length and width of the molecule. The intermolecular energy of this arrangement is -12.5 kJ·mol^{-1} for **Me:o-Me(o)** and -11.9 kJ·mol^{-1} for **Me:o-CF_3** with the strongest contribution from dispersive interactions. In comparison with the HB SCs the feature **$R1_{[g]}$** results in a low energetic gain. Other intermolecular arrangements without hydrogen bonding are associated with similar energies as was analysed in Section 4.B.1.1. Hence **$R1_{[g]}$** can only be regarded as a minor packing feature, which is further supported by the low frequency of occurrence of this SC. It is remarkable however that this geometric arrangement does not appear to require the same configuration of the molecule. It was already mentioned that the molecule **Me:o-Me(o)** assumes a *syn* configuration between substituent and amide oxygen. **Me:o-CF_3** on the other hand exists in the *anti* configuration. The molecular shape is only similar in so far as the amide group is twisted out of the phenyl ring plane considerably in both compounds. The detection of **$R1_{[g]}$** hence demonstrates how the XPac analysis is based purely on geometrical criteria. With respect to the ortho acetanilides **$R1_{[g]}$** can be seen as a flexible packing feature able to accommodate different molecular configurations.

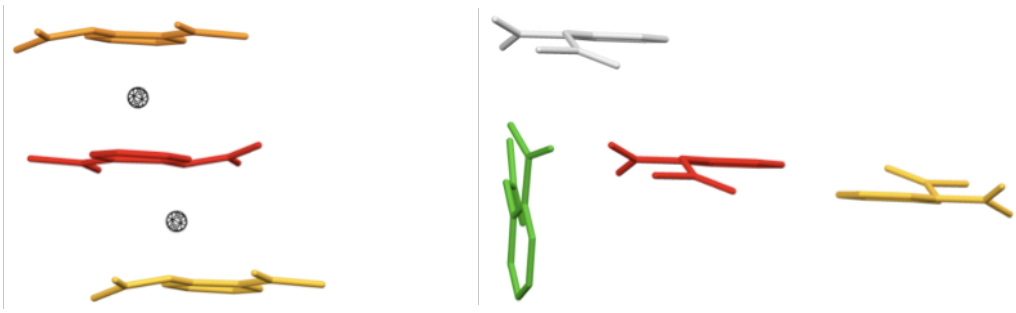
4.B.1.4 Structures without similarity

According to the XPac analysis the crystal structures **Me:o-CN**, **Me:o-NO_2**, **Me:o-COOH**, **Me:o-COMe**, **Me:o-NHCOMe** and **Me:o-^{t}Bu** are unique and do not share any packing features with other ortho structures. For all these crystal assemblies except for **Me:o-^{t}Bu**, the substituent **X** provides an additional hydrogen bond acceptor and/or donor affecting the amide hydrogen bonding observed in the structures discussed thus far. Due to the close proximity of the ortho-substituent and the amide group, intramolecular interactions are facilitated providing steric hindrance is minimal. Intramolecular hydrogen bonding is observed in **Me:o-NO_2**, **Me:o-COMe** and **Me:o-COOH**, whereas the hydrogen bond acceptor is unavailable for intramolecular interaction in **Me:o-CN** and **Me:o-NHCOMe** due to its shape and size, respectively. In all cases (except for **Me:o-^{t}Bu**) the hydrogen bonding between neighbouring core amide groups is interrupted due to the substituent **X**. **Me:o-NO_2** and **Me:o-COMe** contain no additional strong hydrogen bonding donor for the amido oxygen and weak interactions are utilised instead. This is reflected in both the total lattice energies, 89.4 kJ·mol^{-1} and 80.3 kJ·mol^{-1}, respectively, and the pairwise intermolecular interaction energies. Table 4.9 overleaf lists the interaction energies for the compounds discussed in this section.

A	B	Symmetry B	r_{A-B}	E_{tot}	E_{coul}	E_{pol}	E_{disp}	E_{rep}
Me:o-COMe								
1	29	-x, 1-y, 2-z **[$\bar{1}$]**	3.655	-43.4	-18.1	-5.5	-43.2	23.5
1	32	1-x, 1-y, 2-z **[$\bar{1}$]**	4.234	-23.6	-14.2	-6.7	-40.7	38.1
1	17	x, -1.5-y, ½+z **[2_1]**	8.260	-11.5	-4.3	-2.0	-10.2	5.0
Me:o-NO_2								
1	9	x, y-1, z **[t]**	4.976	-21.9	-12.9	-4.5	-25.7	21.2
1	32	1-x, 2-y, 1-z **[$\bar{1}$]**	6.882	-20.9	-19.1	-6.3	-16.3	20.8
1	20	1.5-x, y-½, ½-z **[2_1]**	5.429	-20.0	-12.5	-4.2	-22.8	19.4
1	30	1-x, 1-y, 1-z **[$\bar{1}$]**	5.213	-12.9	1.0	-1.5	-21.5	9.0
Me:o-CN								
1	53	x-½, ½-y, 1-z **[2_1]**	6.705	-32.8	-34.8	-13.3	-16.1	31.4
1	3	-x, y+½, ½-z **[2_1]**	7.269	-21.9	-14.5	-5.1	-16.4	14.0
1	10	x-1, y, z **[t]**	3.811	-20.6	-3.1	-5.4	-41.5	29.5
1	54	x-½, 1.5-y, 1-z **[2_1]**	8.550	-14.9	-10.4	-2.5	-8.6	6.7
Me:o-COOH								
1	31	¼-x, ¼+y, -¾+z **[g]**	8.680	-38.1	-77.3	-39.0	-14.4	92.6
1	4	¼-x, ¼+y, ¼+z **[g]**	7.460	-20.4	-3.1	-3.3	-32.3	18.3
1	6	-x, ½-y, ½+z **[2_1]**	7.563	-14.5	-7.0	-3.5	-19.7	15.6
Me:o-NHCOMe								
1	24	-x, 1-y, -z **[$\bar{1}$]**	5.170	-90.0	-105.5	-39.0	-48.3	101.9
1	5	-x, -y, -z **[$\bar{1}$]**	5.081	-87.9	-97.6	-38.2	-50.7	98.7
1	40	½-x, -½+y, ½-z **[2_1]**	7.732	-21.6	-13.5	-5.3	-17.4	14.6
1	2	½+x, ½+y, z **[C]**	8.594	-14.1	-8.1	-2.9	-14.1	11.0

Table 4.9: The strongest intermolecular interaction energies of the unique ortho structures. The energies are given in kJ·mol^{-1}. The symbols in square brackets indicate the type of symmetry operation. [C] stands for C-centring.

In **Me:o-COMe** there are only two major interactions above 20 kJ·mol^{-1} contributing 43.4 kJ·mol^{-1} and 23.6 kJ·mol^{-1} as can be seen from Table 4.9. These two interactions prevail between molecules related by inversion so that the molecular mean planes face each other as displayed in Figure 4.17.a.

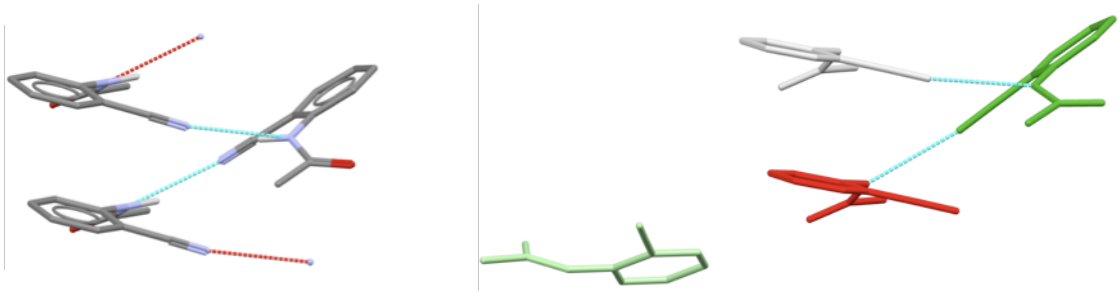


(a) **(b)**

Figure 4.17: The reference molecule is shown in red; all others are symmetry generated from the reference molecule. Intermolecular interaction energies are calculated between the reference A and its symmetry equivalent B *c.f.* Table 4.9. (a) **Me:o-COMe**: 29 (orange), 32 (yellow). Inversion centres are indicated as black circles. (b) **Me:o-NO$_2$**: 9 (white), 32 (yellow), 20 (green).

The intermolecular interaction energies in **Me:o-NO$_2$** are more evenly distributed (*c.f.* Table 4.9) and arise primarily from molecular pairs related by translation, inversion and 2_1 screw symmetry as depicted in Figure 4.17.b. The values in Table 4.9 show that these are interactions due to dispersive forces.

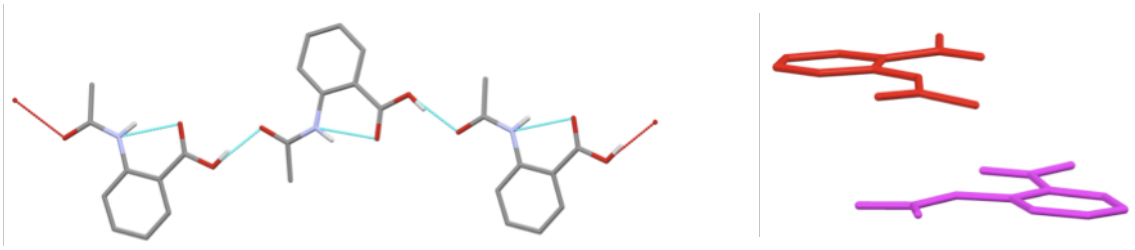
In **Me:o-CN** the substituent **CN** disrupts the HB interaction between the amide groups and instead amido-cyano hydrogen bonding occurs as illustrated in Figure 4.18.a. The intermolecular energy of this interaction amounts to 32.8 kJ·mol^{-1}, which is actually slightly less energetic than the hydrogen bonds **C1$_{[t]}$** and **C1$_{[g]}$**. Apart form this interaction there are two more above 20 kJ·mol^{-1}. The corresponding molecular pairs are shown in Figure 4.18.b.



(a) **(b)**

Figure 4.18: (a) Cyano-amide hydrogen bonding in **Me:o-CN**. (b) Molecular pairs with strongest interaction energies *c.f.* Table 4.9. Reference molecule shown in red; other colours refer to symmetry generated molecule with numbers 10 (white), 53 (green), 3 (light green).

The substituent **COOH** contains both a HB donor and acceptor group. The carboxyl acceptor engages in intramolecular hydrogen bonding with the amide nitrogen, whereas the carboxyl donor is utilised in intermolecular HB interaction with the amide oxygen atom of a neighbouring molecule in. The hydrogen bonding found in **Me:o-COOH** is displayed in Figure 4.19.a.

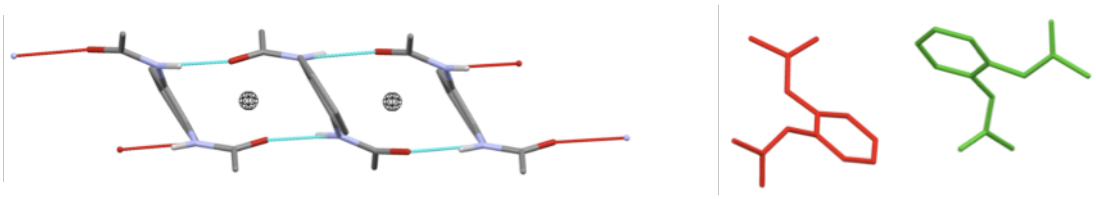


(a) **(b)**

Figure 4.19: (a) Intra- and intermolecular carboxy-amide hydrogen bonding in **Me:o-COOH**. (b) Molecular pair with strongest interaction energy *c.f.* Table 4.9 Reference molecule shown in red; molecule in magenta refers to molecule 4 in Table 4.9. Molecule 31 is found in HB in (a).

The carboxyl-amide intermolecular HB has glide symmetry and is associated with 38.1 kJ·mol^{-1}, which is similar to the HB interactions between amide groups. However, the Coulombic gain is almost doubled in comparison with the chains of type **C1**. Otherwise there is only one further intermolecular interaction exceeding 20 kJ·mol^{-1}. The molecular pair is shown in Figure 4.19.b.

In **Me:o-NHCOMe** hydrogen bonding is established between both amide groups in a ladder motif. This ladder has inversion symmetry, so that the core amide group hydrogen bonds to the substituent amide group and vice versa as shown in Figure 4.20.a.



(a) **(b)**

Figure 4.20: (a) Ladder HB motif in **Me:o-NHCOMe**. Inversion centres are included as black circles. (b) Molecular pair with strongest interaction energy *c.f.* Table 4.9. Reference molecule shown in red; molecule in green refers to molecule 40 in Table 4.9. Molecules 24 and 5 participate in HB in (a).

This ladder motif is the most energetic discussed thus far as can be seen from Table 4.9. It should be noted that within the ladder the distance between a molecule and the inversion centres on either side of the molecule is not equivalent leading to slightly different energy values (87.9 and 90.0 kJ·mol^{-1}) for the HB interaction of one molecule with the neighbouring molecules on either side of it. The next strongest intermolecular interaction with 21.6 kJ·mol^{-1} is much less significant and occurs between molecules related by 2_1 screw symmetry. The corresponding pair of molecules is displayed in Figure 4.20.b.

In the crystal structure **Me:o-^{t}Bu** the molecules pack in the tetragonal space group $I4_1cd$ – the highest symmetry in the ortho acetanilides. Neighbouring molecules interact via the amide-amide hydrogen of type $\mathbf{C}_1^1(\mathbf{4})$ in Etter notation, but forming a helix with 4 fold symmetry in this structure (Figure 4.21).

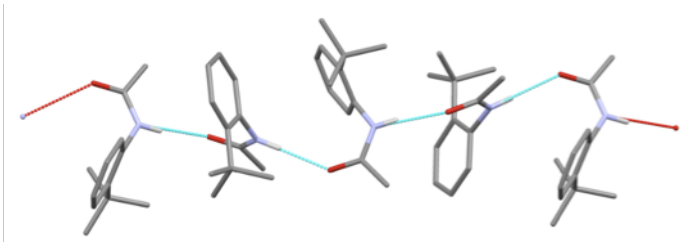


Figure 4.21: HB helix in **Me:o-^{t}Bu**.

The pairwise interaction energy between molecules in this helix amounts to 56.3 kJ·mol^{-1} (Table 4.10) and is significantly higher than **C1$_{[t]}$** and **C1$_{[g]}$**.

Structure	$r_{N\cdots O}$ [Å]		$\angle_{N-H\cdots O}$ [°]		Symmetry	
Me:o-^{t}Bu	2.931(3)		171(3)		4_1	
	r_{A-B} [Å]	E_{tot}	E_{coul}	E_{pol}	E_{disp}	E_{rep}
	4.727	-56.3	-51.1	-19.9	-48.3	63.0

Table 4.10: Structural parameters and intermolecular interaction energies for the hydrogen bond in **Me:o-^{t}Bu**. The energies are given in kJ·mol^{-1}.

Both Coulomb and dispersive contributions to this interaction are high. The substituent **^{t}Bu** is the bulkiest within the series of ortho-acetanilides and the close proximity to the amide group results in a near perpendicular positioning of the

amide group with respect to the phenyl ring. This means that the molecules are almost intertwined within the helix so that a larger area of the molecule can interact than is the case in $\mathbf{C1}_{[t]}$ and $\mathbf{C1}_{[g]}$ hence leading to the stronger interaction energy.

4.B.2. Structural Relationships in Me:m-X

The structural similarity in the meta acetanilide series is summarised in the diagram in Figure 4.22 shown overleaf. The structural relationships in the meta series are a lot more complex and the number of packing features occurring in at least two crystal structures is much higher than in the ortho acetanilides. XPac identified four groups of packing analogous, four 2D SCs, five 1D features and three 0D constructs. Only one crystal structure, **Me:m-NHCOMe**, lacks similarity with any of the other fourteen meta acetanilides. It should be noted here that the 1D packing feature $\mathbf{C2_{[g]}}$ was not identified by XPac during the comparison of the meta structures, but was added manually to aid the understanding of the 2D assembly **L1**. The structural similarity relationships in the meta acetanilide series are discussed subsequently. Corresponding base vectors are provided in Appendix 4-2. The total lattice energies can be found in Appendix 4-3 providing a reference for the importance of the pairwise interaction energies.

4.B.2.1. 3D similarity in Me:m-X

The XPac analysis revealed analogous packing in three pairs and one triplet of meta acetanilide crystal structures. The unit cell parameters are summarised in Table 4.11.

The parameters are very similar for the first two groups, but less so for the last two groups in Table 4.11. It should be noted that the crystal structures of at least one of the compounds in each of the first three sets of packing analogues are presented in the literature[9] and hence will not be discussed in detail here.

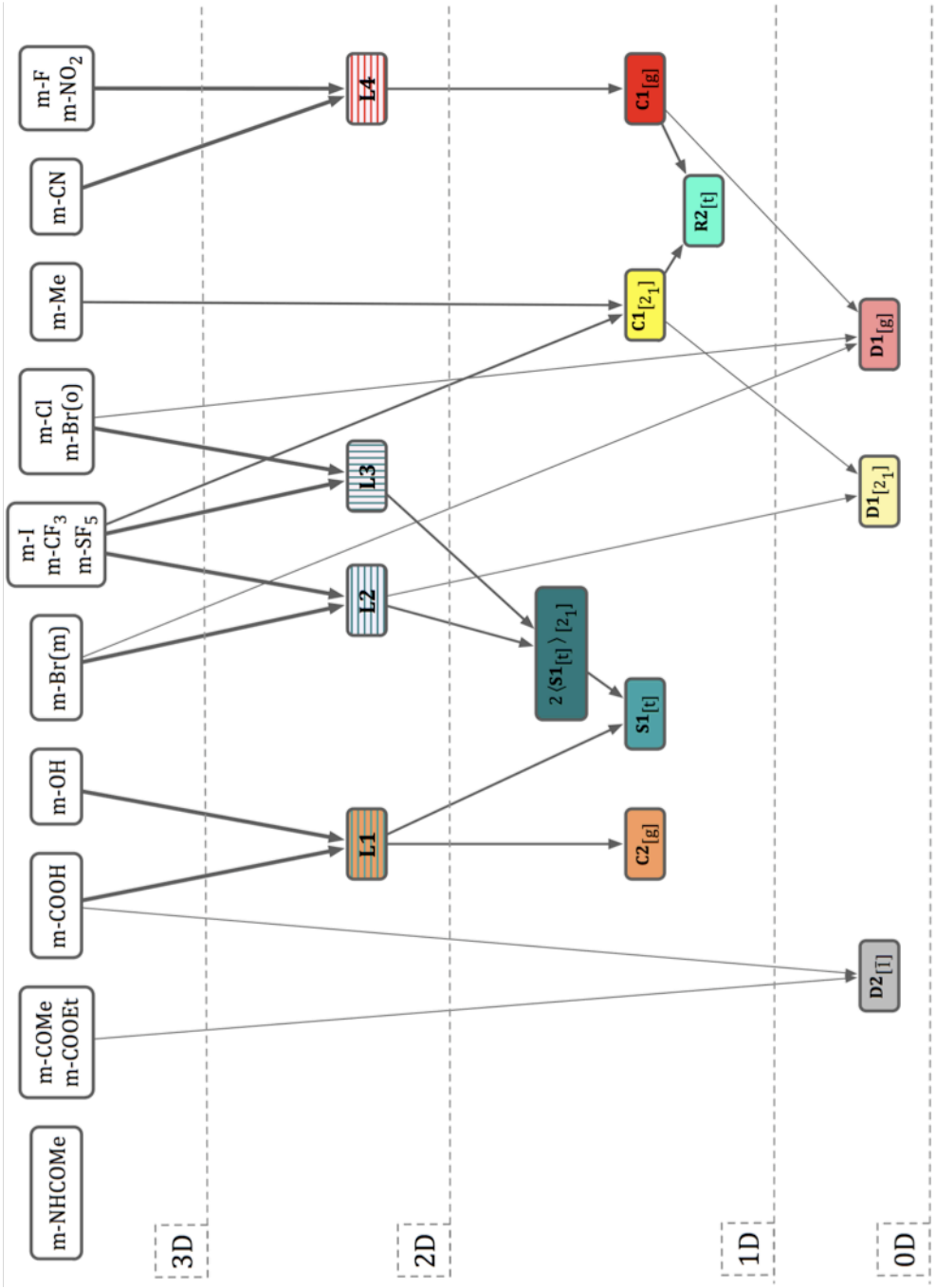


Figure 4.22: Structural similarity plot for meta-substituted acetanilides.

Structure	*Me:m-I*	*Me:m-CF_3*	*Me:m-SF_5*	*Me:m-Cl*	*Me:m-Br(o)*
a [Å]	4.9076(5)	5.1717(8)	9.8003(11)	4.7667(1)	4.7836(6)
b [Å]	9.7504(8)	9.7527(13)	17.894(2)	18.4984(4)	18.765(1)
c [Å]	18.6116(11)	17.811(2)	5.9141(7)	18.6318(4)	19.379(2)
α [°]	90	90	90	90	90
β [°]	90	90	90	90	90
γ [°]	90	90	90	90	90
Z'	1	1	1	2	2
Space Group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$

Table 4.11: Unit cell parameters for the groups of packing analogues in the meta acetanilide series.

Structure	*Me:m-F*	*Me:m-NO_2*	*Me:m-COMe*	*Me:m-COOEt*
a [Å]	12.1766(4)	9.7285(5)	5.5061(2)	5.7366(2)
b [Å]	9.5024(3)	13.3392(7)	7.2607(3)	8.5492(5)
c [Å]	12.9638(5)	12.9849(5)	11.4567(5)	10.6373(6)
α [°]	90	90	103.709(2)	93.974(2)
β [°]	90	103.791(3)	93.064(3)	99.351(3)
γ [°]	90	90	100.528(3)	95.314(3)
Z'	1	4	1	2
Space Group	Pbca	$P2_1$	$P\bar{1}$	$P\bar{1}$

Table 4.11 *continued*.

Me:m-F crystallises in the orthorhombic space group Pbca whereas **Me:m-NO_2** crystallises in the monoclinic space group $P2_1$. At a first glance this seems to exclude 3D similarity, but this discrepancy is compensated for by an increase in Z' – there are four molecules in the ASU of **Me:m-NO_2** (Z'=4) compared with Z'=1 in **Me:m-F** (*c.f.* Ortep of ASU in Chapter 4.A.). In the ASU of **Me:m-NO_2** there are two pairs of crystallographically independent molecules interacting via amide-amide hydrogen bonding. The molecules within each HB pair are related by pseudo-glide symmetry, whilst the pairs of molecules are related by pseudo-inversion. Space group $P2_1$ is furthermore a subgroup of Pbca and together with the pseudo-symmetry relationships between the molecules in the ASU, the spatial arrangement in **Me:m-NO_2** mimics completely the higher symmetry of **Me:m-F**. Figure 4.23 shows the unit cell contents of **Me:m-F** and **Me:m-NO_2** as viewed parallel to the HB direction. Shown also in the picture is the position of the 2_1 screw axis common to both structures.

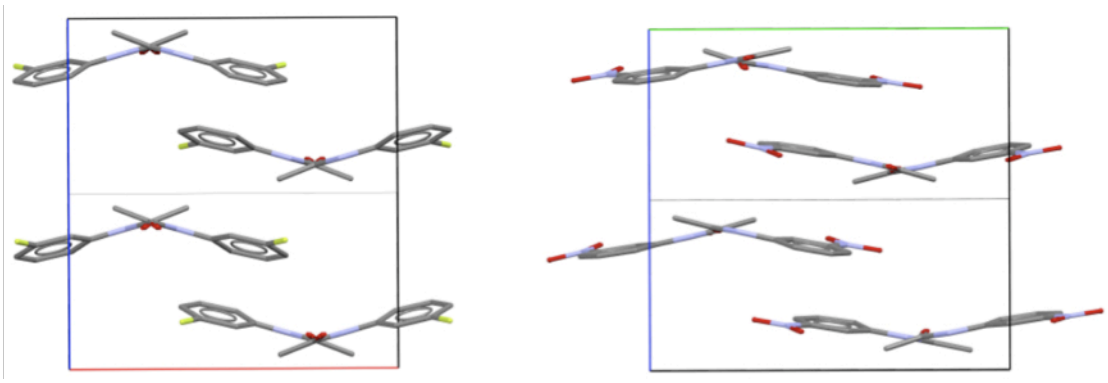


Figure 4.23: Packing of the molecules within one unit cell of **Me:m-F** (left) and **Me:m-NO_2** (right).

The packing mechanism seems very robust with respect to the phenyl ring orientation, since two molecules of the ASU in **Me:m-NO_2** adopt the *syn* whilst the other two assume the *anti* configuration, i.e. the position of substituent **X** with respect to the amide oxygen atom. In **Me:m-F** on the other hand, the molecule is present in the *syn* configuration only (*c.f.* Chapter 4.A.).

The largest differences in unit cell parameters between packing analogues are found in the last group of structures, **Me:m-COMe** and **Me:m-COOEt**. Both structures crystallise in the triclinic space group $P\bar{1}$ with one molecule in the ASU. The difference in the unit cell parameters can be explained by considering the molecular shape. The molecule **Me:m-COOEt** is much longer than **Me:m-COMe** as can be seen from the molecular overlay in Figure 4.24, the difference in length being approximately 2 Å. Hence the different unit cell dimensions are associated with accommodating such differently sized molecules in a highly similar crystal lattice. In the crystal structure neighbouring molecules of **Me:m-COMe** and **Me:m-COOEt** associate in HB dimers with inversion symmetry. This dimer $\mathbf{D2}_{[\bar{1}]}$ also occurs in **Me:m-COOH** and will be discussed further in the following section. The dimer is a closed unit in terms of hydrogen bonding and no further strong intermolecular interactions are possible in **Me:m-COMe** and **Me:m-COOEt**. Close packing of these dimers is achieved by translation along the depth of the molecule

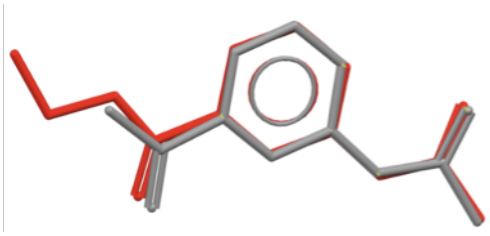


Figure 4.24: Molecular overlay of **Me:m-COOEt** (red) and **Me:m-COMe** (grey).

and inversion symmetry along the length and width of the molecules. The packing is depicted in Figure 4.25 as viewed parallel to the depth of the molecule.

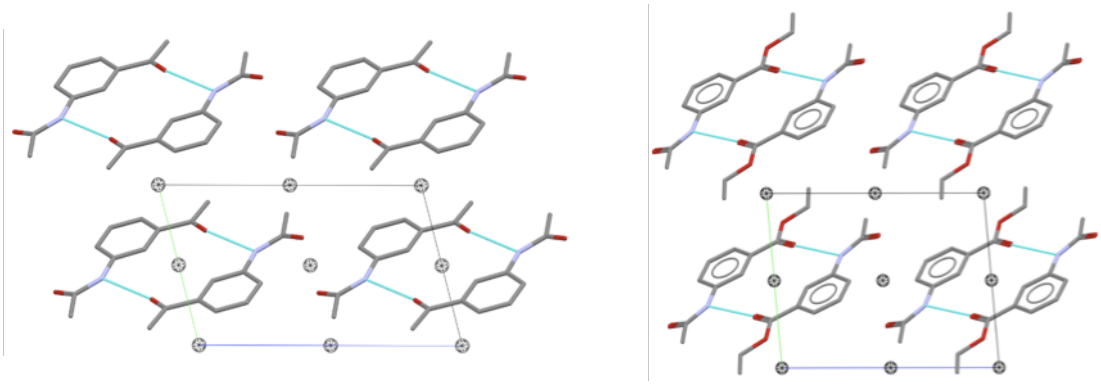


Figure 4.25: Crystal packing diagrams for **Me:m-COMe** (left) and **Me:m-COOEt** (right). The location of the inversion centres within the unit cell is indicated with the black circles.

The location of the inversion centres are also shown in the packing diagram. The difference in unit cell dimensions between **Me:m-COMe** and **Me:m-COOEt** is clearly visible in the figure.

4.B.2.2. Hydrogen bonding in Me:m-X and associated SCs

Some of the SCs identified with XPac are combinations of lower dimensionality SCs, e.g. the 2D layer **L1** is obtained by combining the 1D stack $\mathbf{S1}_{[t]}$ (discussed in the subsequent section) with the 1D chain $\mathbf{C2}_{[g]}$. In this respect the 0D SCs $\mathbf{D1}_{[2_1]}$, $\mathbf{D1}_{[g]}$ and $\mathbf{D2}_{[\bar{1}]}$ are the lowest order SCs and can be seen as fundamental building blocks that find incorporation in most meta acetanilide crystal structures. The only structures that do not contain one of these 0D SCs are **Me:m-OH** and **Me:m-NHCOMe**, the latter lacking similarity with any of the other structures altogether (see Section 4.B.2.5).

4.B.2.2.a Dimers

$\mathbf{D1}_{[2_1]}$, $\mathbf{D1}_{[g]}$ and $\mathbf{D2}_{[\bar{1}]}$ are HB dimers of varying intermolecular connectivity. $\mathbf{D2}_{[\bar{1}]}$ is an inversion dimer of planar geometry in which the amide group interacts with the available carbonyl group of the substituent, i.e. a N-H···O=C HB is formed. The inversion centre is located in the same plane as the molecular backbone and hence the molecules within the dimer are placed side-by-side along their width. Figure 4.26 displays the inversion dimer in **Me:m-COMe**, **Me:m-COOEt** and **Me:m-COOH**.

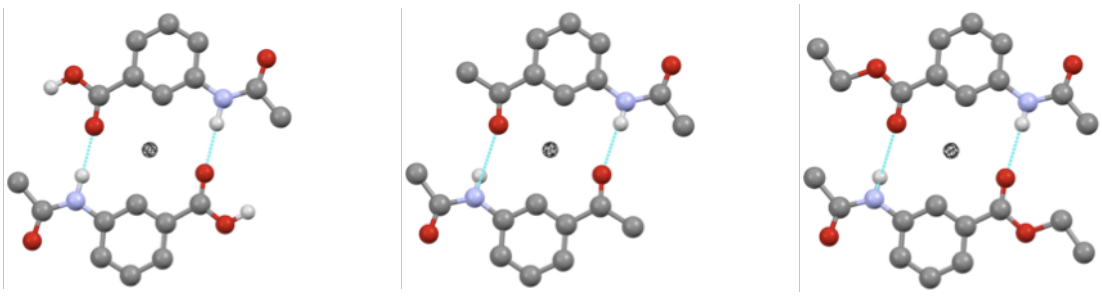


Figure 4.26: Dimer $\mathbf{D2}_{[\bar{1}]}$ in **Me:m-COOH** (left), **Me:m-COMe** (middle) and **Me:m-COOEt** (right). The inversion centre is shown as black circle.

The construct occurs if **X** contains at least one additional HB acceptor. Hence it is found in **Me:m-COMe**, **Me:m-COOEt** and **Me:m-COOH**. However, **CN**, **NO_2** and **OH** also exhibit HB acceptor functionality, but $\mathbf{D2}_{[\bar{1}]}$ is not observed in the respective crystal structures. These functional groups are smaller and more electron withdrawing than **COMe**, **COOEt** and **COOH** and it may hence be concluded that the formation of $\mathbf{D2}_{[\bar{1}]}$ is dependent on both the geometrical and electrostatic characteristics of substituent **X**.

Constructs $\mathbf{D1}_{[2_1]}$ and $\mathbf{D1}_{[g]}$ are more frequently observed. $\mathbf{D1}_{[2_1]}$ occurs in five structures, namely in **Me:m-Br(m)** and **Me:m-Me** as well as in **Me:m-I** and its two packing analogues. $\mathbf{D1}_{[g]}$ on the other hand is present in the six crystal structures **Me:m-CN**, **Me:m-F**, **Me:m-NO_2**, **Me:m-Cl**, **Me:m-Br(o)** and **Me:m-Br(m)**. The structure of the monoclinic polymorph of **Me:m-Br** hence contains both dimers $\mathbf{D1}_{[2_1]}$ and $\mathbf{D1}_{[g]}$. The two dimers are of the same connectivity type, i.e. adjacent amide groups are hydrogen bonded, but they differ in the symmetry relationship

between the two molecules in the dimer. Due to the 2_1 screw symmetry in $\mathbf{D1}_{[2_1]}$ the dimer is a flat object in which the phenyl rings are approximately co-planar and found on opposite sides of the hydrogen bond. Figure 4.27 shows construct $\mathbf{D1}_{[2_1]}$ from two perspectives as present in the five structures. It can be seen from the figure that the molecular configuration in **Me:m-Me** is *anti* with respect to the amide oxygen, whereas all other compounds assume the *syn* configuration. This shows that the geometrical arrangement in the HB dimers occurs irrespective of the orientation of the phenyl ring.

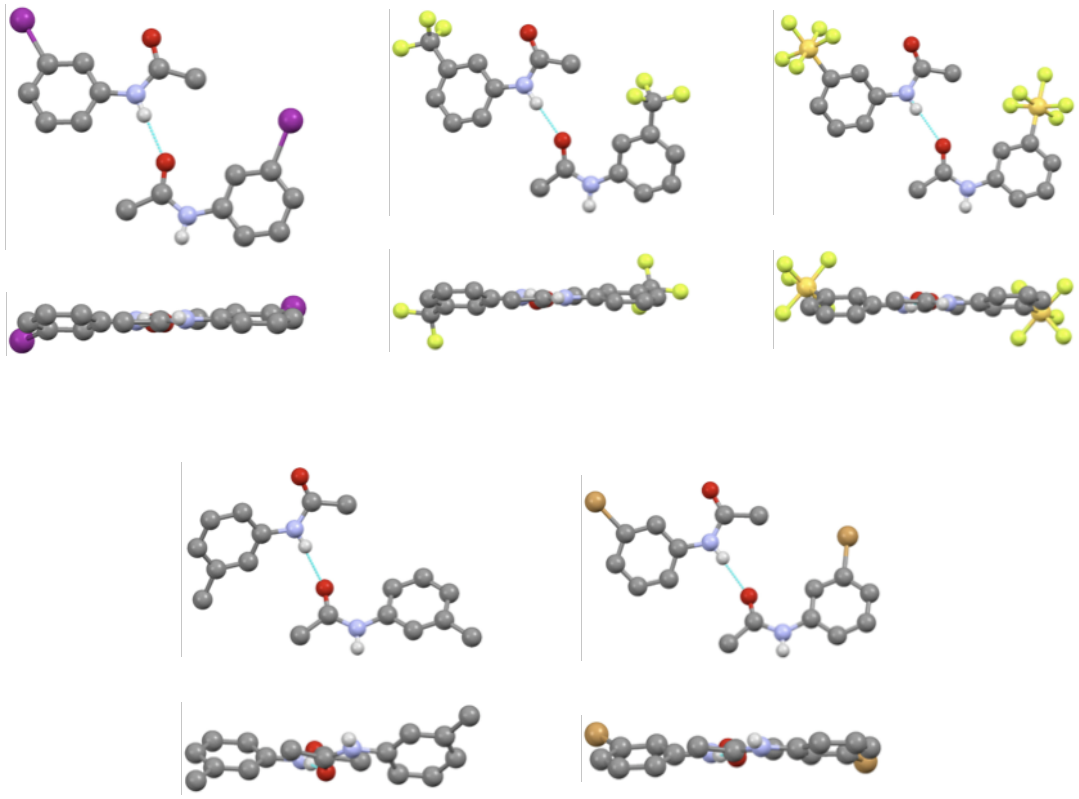


Figure 4.27: Dimer $\mathbf{D1}_{[2_1]}$ in meta acetanilides viewed perpendicular and parallel to the hydrogen bond. Top row from left to right: **Me:m-I**, **Me:m-CF_3** and **Me:m-SF_5**. Bottom: **Me:m-Me** (left) and **Me:m-Br(m)** (right).

Since the molecules in $\mathbf{D1}_{[g]}$ are related by glide symmetry (and as they reside at an angle to the mirror plane rather than coinciding with the mirror plane) the dimer $\mathbf{D1}_{[g]}$ assumes a wedge or V-shape with varying span. Figure 4.28 depicts the HB dimer $\mathbf{D1}_{[g]}$ for the six structures it occurs in.

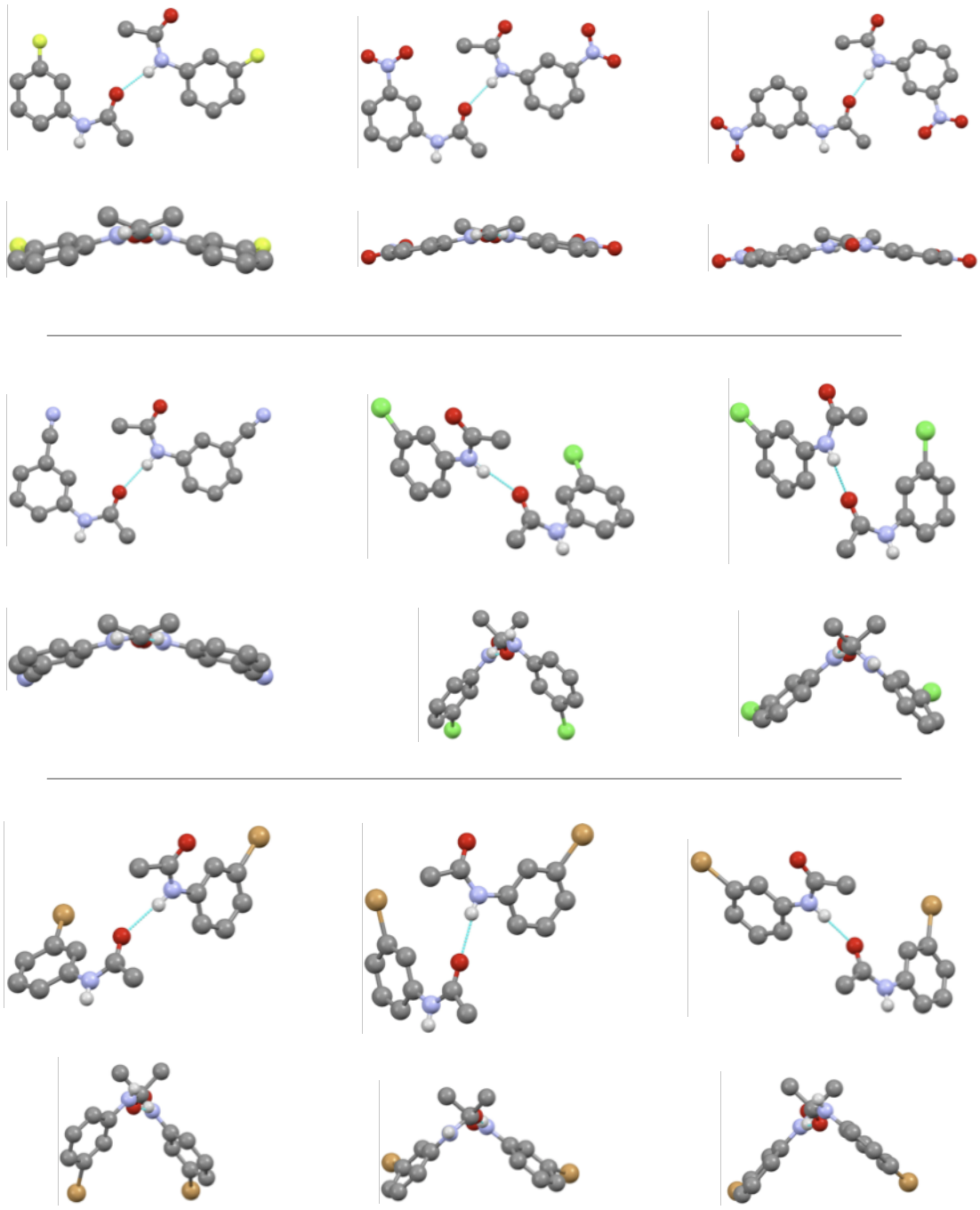


Figure 4.28: Dimer **D1$_{[g]}$** in meta acetanilides viewed perpendicular and parallel to the hydrogen bond. Top row from left to right: **Me:m-F** and the dimer between the two individual pairs of molecules in the ASU of **Me:m-NO$_2$**. Middle row from left to right: **Me:m-CN** and dimers for each independent molecule in the ASU of **Me:m-Cl**. Bottom row from left to right: dimers for each independent molecule in the ASU of **Me:m-Br(o)** (left and middle) and **Me:m-Br(m)** (right).

The difference in the acuteness of the V-shape is clearly visible. Whilst the hydrogen bonded molecules form a very wide V in **Me:m-CN**, **Me:m-F** and

Me:m-NO_2, the V is much narrower in **Me:m-Cl** and the two polymorphic forms of **Me:m-Br**.

The intermolecular energies associated with the HB dimers are specified in Table 4.12[10] and the respective HB parameters are recorded in Table 4.13.

Structure	r_{A-B} [Å]	E_{tot}	E_{coul}	E_{pol}	E_{disp}	E_{rep}
$D1_{[g]}$						
Me:m-F	6.201	-55.7	-46.4	-38.2	-20.5	49.4
Me:m-CN	6.334	-53.3	-42.0	-33.6	-22.8	45.1
$D1_{[2_1]}$						
Me:m-CF_3	6.816	-50.4	-39.9	-30.3	-23.6	43.4
Me:m-SF_5	7.298	-48.0	-36.7	-24.6	-23.9	37.3
Me:m-Me	6.545	-52.8	-43.3	-31.2	-22.9	44.6
$D2_{[\bar{1}]}$						
Me:m-COOH	6.136	-73.6	-58.1	-36.3	-20.5	45.8
Me:m-COMe	5.704	-77.0	-60.1	-40.9	-31.3	55.4
Me:m-COOEt	5.627	-78.2	-58.2	-40.1	-35.6	55.7

Table 4.12: Intermolecular interaction energies for the HB dimers found in the meta series. Energies are given in kJ·mol^{-1}.

Structure	$r_{N\cdots O}$ [Å]	$\angle_{N-H\cdots O}$ [°]	Symmetry
Me:m-F	2.831(2)	168(2)	glide
Me:m-NO_2	2.874(3)	172(4)	pseudo-glide
	2.887(3)	180(4)	pseudo-glide
	2.858(3)	175(2)	pseudo-glide
	2.875(3)	159(3)	pseudo-glide
Me:m-CN	2.868(2)	169(2)	glide
Me:m-Cl	2.813(2)	167(2)	pseudo-glide
	2.901(2)	171(2)	pseudo-glide
Me:m-Br(o)	2.886	166	pseudo-glide
	2.953	169	pseudo-glide
Me:m-Br(m)	2.823(5)	172(5)	pseudo-glide
	2.892(5)	175(6)	pseudo 2_1 screw
Me:m-I	2.902(10)	174(2)	2_1 screw
Me:m-CF_3	2.907(5)	166(4)	2_1 screw
Me:m-SF_5	2.975	170	2_1 screw
Me:m-Me	2.891(2)	174(2)	2_1 screw
Me:m-COOH	3.022	173	inversion
Me:m-COMe	2.950(2)	154(2)	inversion
Me:m-COOEt	3.031(1)	161(2)	inversion

Table 4.13: Hydrogen bond parameters for dimers **D1** and **D2** present in the meta series.

On average the HB distances in dimer $\mathbf{D2}_{[\bar{1}]}$ are the longest but energetically this construct is the most stable owing to the fact that there are two hydrogen bonds in dimer $\mathbf{D2}_{[\bar{1}]}$ in contrast to only one HB in the other two dimers. From the breakdown into individual energies it can also be seen that for $\mathbf{D2}_{[\bar{1}]}$ the dispersive term increases with the size of the aliphatic residue in substituent **X**, i.e. a larger area is available for dispersive interaction. In terms of energy dimer $\mathbf{D1}_{[g]}$ is slightly stronger than $\mathbf{D1}_{[2_1]}$, which may be related to the HB distance on average being shorter than in $\mathbf{D1}_{[2_1]}$. The shortest HB distances occur in the acetanilides with the strongest electron withdrawing substituents so that a link exists between molecular structure and intermolecular association. From Table 4.13 another trend is visible: the meta acetanilides with large "spherical" substituents adopt the $\mathbf{D1}_{[2_1]}$ interaction mode, whereas the meta acetanilides with smaller substituents arrange in the $\mathbf{D1}_{[g]}$ motif. **Me:m-Br(m)** may present the borderline case, where both dimers coexist.

4.B.2.2.b 1D HB Chains

Dimers $\mathbf{D1}_{[2_1]}$ and $\mathbf{D1}_{[g]}$ are part of extended infinite chains, which are established via amide-amide hydrogen bonding. At this stage the energies discussed in the preceding section need to be put into perspective. The inversion dimer $\mathbf{D2}_{[\bar{1}]}$ is the most energetic when it comes to pairwise HB interaction, but since $\mathbf{D1}_{[2_1]}$ and $\mathbf{D1}_{[g]}$ find incorporation in infinite chains the overall stability of the crystal is enhanced by the latter two HB dimers.

In most cases the symmetry of the dimer also prevails throughout the HB chain, hence the 1D SCs $\mathbf{C1}_{[2_1]}$ and $\mathbf{C1}_{[g]}$ are formed (*c.f.* Figure 4.22). The feature $\mathbf{C1}_{[2_1]}$ is present in **Me:m-Me** and in the group of analogous structures containing **Me:m-I** as shown in Figure 4.29.

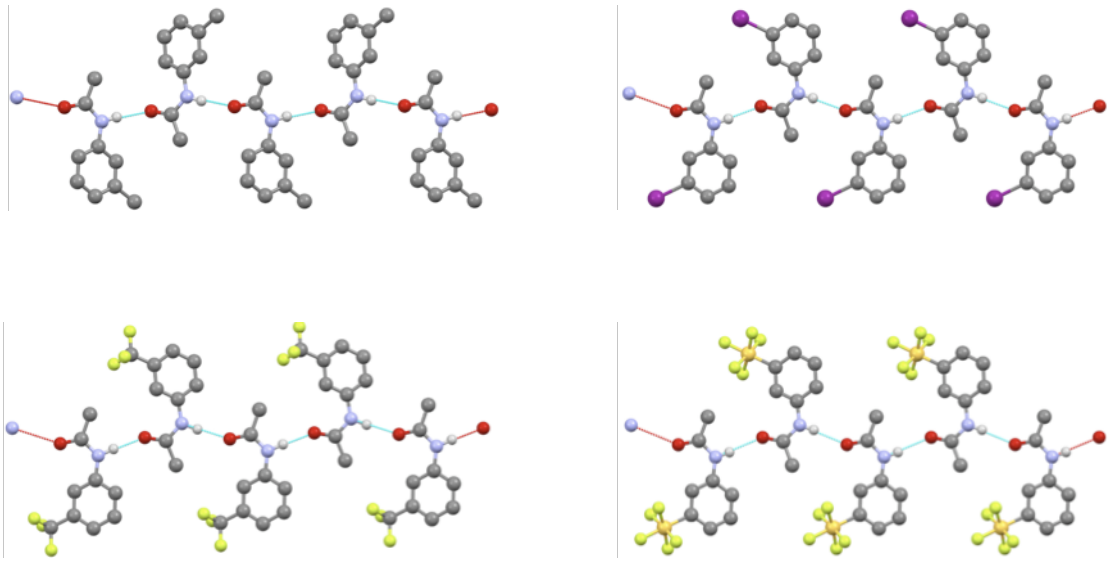


Figure 4.29: HB chain $\mathbf{C1}_{[2_1]}$ in meta acetanilides viewed perpendicular to the hydrogen bond. **Me:m-Me** (top left), **Me:m-I** (top right), **Me:m-CF_3** (bottom left) and **Me:m-SF_5** (bottom right).

Construct $\mathbf{C1}_{[g]}$ exists in the crystal structures **Me:m-CN**, **Me:m-F** and **Me:m-NO_2** and the construct is displayed in Figure 4.30.

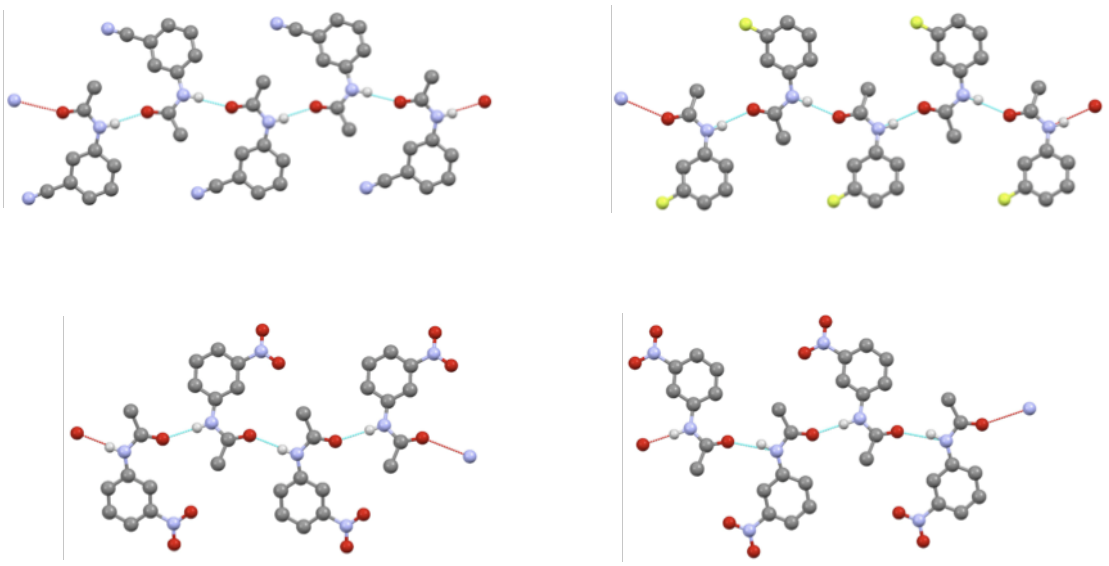


Figure 4.30: HB chain $\mathbf{C1}_{[g]}$ in meta acetanilides viewed perpendicular to the hydrogen bond. **Me:m-CN** (top left), **Me:m-F** (top right). Two chains shown for **Me:m-NO_2** as formed between pairs of independent molecules in the ASU with the same molecular configuration (bottom).

The intermolecular interaction energies and HB parameters are identical with those already discussed for the dimeric fragment so that **C1**$_{[g]}$ can be considered as slightly more energetically stable than **C1**$_{[2_1]}$.

XPac also identified 1D similarity between the structures containing constructs **C1**$_{[2_1]}$ and **C1**$_{[g]}$, namely the row **R2**$_{[t]}$. Surprisingly XPac assigned this translational row of molecules the same base vectors as for the HB chains. Upon closer inspection it transpired that the similarity is concerned with one side of the HB chains only. In the chains the phenyl rings of adjacent molecules are positioned on opposite sides of the hydrogen bond (Figure 4.31).

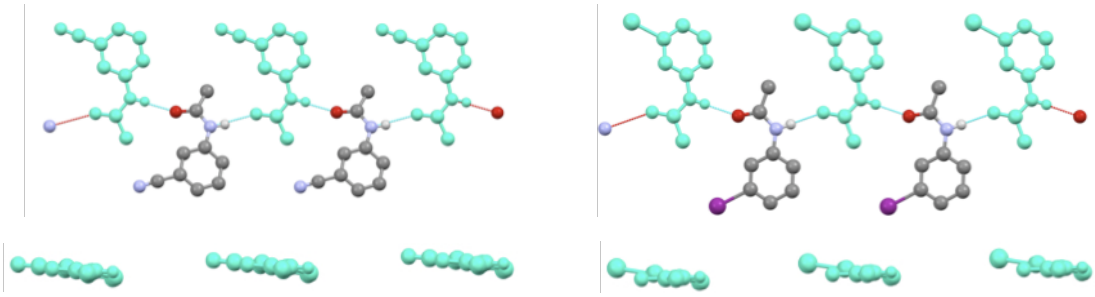
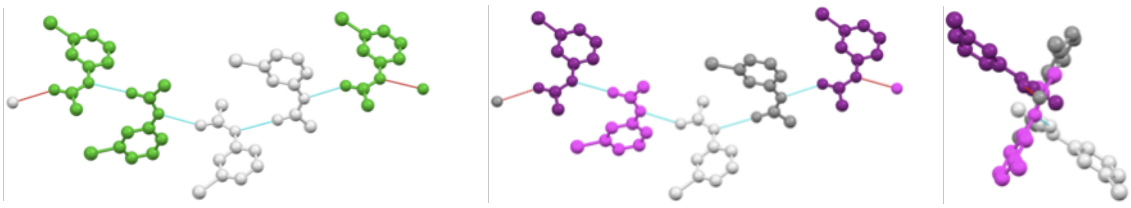


Figure 4.31: Row **R2**$_{[t]}$ shown for **Me:m-CN** (left) and **Me:m-I** (right) as viewed from the top and side on, respectively.

Given the symmetry of the chain, every other molecule is obtained by translation and hence placed on the same side of the chain. The finding of construct **R2**$_{[t]}$ thus indicates the similarity of the individual sides of the amide HB chains, but it also emphasises the different geometric arrangement of the two sides within the constructs **C1**$_{[2_1]}$ and **C1**$_{[g]}$. It can thus only be regarded as a minor similarity feature.

The structures **Me:m-Cl**, **Me:m-Br(o)** and **Me:m-Br(m)** also contain infinite amide HB chains, but the compounds crystallise with two crystallographically independent molecules in the ASU (Z'=2), which affects the symmetry of the chain. The molecules in the ASU are conformationally different and are thus related by pseudo symmetry, i.e. they do not fully satisfy crystallographic symmetry. In **Me:m-Cl** and **Me:m-Br(o)** the molecules in the ASU are related by a pseudo-glide plane and the characteristic V-shape of **D1**$_{[g]}$ is observed. Within the amide HB

chain such a pair of molecules is then related to the next pair of molecules via crystallographic 2_1 screw symmetry. The symmetry difference is visualised with different colours in Figure 4.32.a.

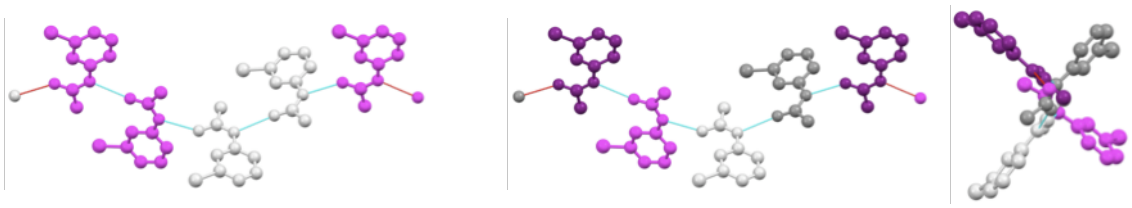


(a) **(b)**

Figure 4.32: (a) Amide-amide hydrogen bonding chain in **Me:m-Br(o)** with molecules coloured according to symmetry. (b) HB chain in **Me:m-Br(o)** as viewed from the top and along the HB, respectively. Colours indicate pseudo-symmetry relationships.

The independent molecules in the ASU are drawn in the same colour and the green and white molecules are related by 2_1 screw symmetry. Although neighbouring pairs of molecules are generated by a 2_1 screw axis, the symmetry relationship between individual adjacent molecules remains as pseudo-glide symmetry throughout the amide HB chain. The pseudo-symmetry relationships in **Me:m-Br(o)** are shown in Figure 4.32.b. The molecules shown in white and grey symbolise the two components of the ASU, which are related by pseudo-glide symmetry. As going from white to magenta and grey to purple, molecules are related by a further pseudo-glide plane.

By contrast in polymorph **Me:m-Br(m)**, the molecules in the ASU are related by a pseudo 2_1 screw axis forming the flat unit of $\mathbf{D1}_{[2_1]}$. In the HB chain neighbouring pairs of molecules are related by glide symmetry as highlighted with the white and magenta molecules in Figure 4.33.a. Hence the hydrogen bond symmetry between adjacent molecules within the HB chain alternates between glide and 2_1 screw symmetry (*c.f.* magenta/purple to white/grey and white to grey in Figure 4.33.b, respectively). This means in contrast to **Me:m-Cl** and **Me:m-Br(o)** both dimers $\mathbf{D1}_{[2_1]}$ and $\mathbf{D1}_{[g]}$ coexist in the infinite amide HB chain in **Me:m-Br(m)**.



(a) **(b)**

Figure 4.33: (a) Amide-amide hydrogen bonding chain in **Me:m-Br(m)** with molecules coloured according to symmetry. (b) HB chain in **Me:m-Br(m)** as viewed from the top and along the HB, respectively. Colours indicate pseudo-symmetry relationships.

A third HB chain was found upon visual inspection of the crystal structures of **Me:m-COOH** and **Me:m-OH** and it was added to the similarity plot as feature $\mathbf{C2_{[g]}}$ (Figure 4.22). The connectivity in this HB feature is different to the amide-amide HB chains discussed thus far since it involves the interaction between a hydroxy group and the amide carbonyl. The topology of $\mathbf{C2_{[g]}}$ expressed in Etter's notation is $\mathbf{C_1^1(8)}$ for **Me:m-OH** and $\mathbf{C_1^1(9)}$ for **Me:m-COOH**. Strictly speaking the topology is not identical, but since the interacting atomic groups are very similar and in order to allow for differences in the substituents the HB interaction was assigned the same construct descriptor. The chain $\mathbf{C2_{[g]}}$ propagates with the length of the molecule (as opposed to the width in SCs $\mathbf{C1_{[2_1]}}$ and $\mathbf{C1_{[g]}}$) and the interaction occurs in a head-to-tail arrangement of the molecules as displayed in Figure 4.34.

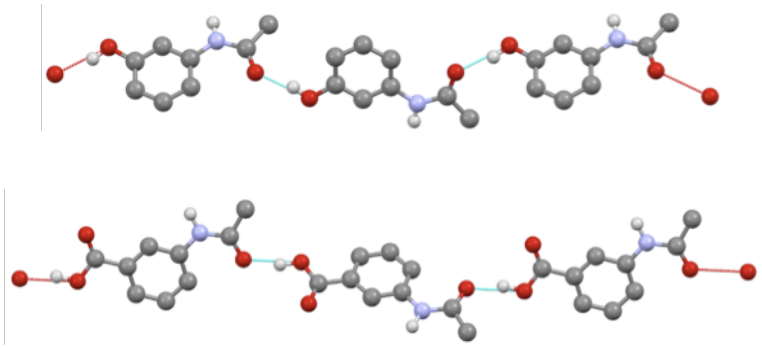


Figure 4.34: HB chain $\mathbf{C2_{[g]}}$ as found in **Me:m-OH** (top) and **Me:m-COOH** (bottom).

Tables 4.14 and 4.15 contain the HB parameters and interaction energies, respectively.

Structure	$r_{O\cdots O}$ [Å]	$\angle_{N-H\cdots O}$ [°]	Symmetry
Me:m-OH	2.6288(13)	166(2)	glide
Me:m-COOH	2.715	164	glide

Table 4.14: Hydrogen bond parameters for **C2$_{[g]}$** in **Me:m-OH** and **Me:m-COOH**.

Structure	r_{A-B} [Å]	E_{tot}	E_{coul}	E_{pol}	E_{disp}	E_{rep}
Me:m-OH	8.856	-28.3	-68.8	-33.7	-18.1	92.2
Me:m-COOH	10.324	-64.1	-53.5	-53.6	-9.9	52.9

Table 4.15: Intermolecular interaction energies for **C2$_{[g]}$** in **Me:m-OH** and **Me:m-COOH**. Energies are given in kJ·mol^{-1}.

The HB length is considerably shorter in **Me:m-OH**, but its interaction energy is only approximately 44% of the energy in **Me:m-COOH**. The largest differences are found in the repulsion and polarisation energies. The former arises due to the short HB distance and the latter indicates that the carboxyl moiety is more readily polarisable than the hydroxy group. Hence the HB arrangement **C2$_{[g]}$** is only associated with a moderate energy gain in **Me:m-OH**.

4.B.2.3. Other 1D common packing features – the stack S1[t]

In the series of meta acetanilides there is only one commonly occurring intermolecular arrangement that does not involve hydrogen bonding, namely the stack **S1[t]**. Although there appears to be a lack of strong intermolecular interaction this construct is the most frequently observed. It can be found in eight of the fifteen crystal structures if the arrows are traced from the box **S1[t]** in the reverse direction. The stack has translational symmetry and affords the packing of the molecules along their depth coinciding with the shortest unit cell axis. Figure 4.35 displays two perspectives of **S1[t]** as present in **Me:m-COOH**, **Me:m-OH**, **Me:m-Br(m)**, **Me:m-I** and **Me:m-Cl**. The latter two structures are to be considered as representative for the two groups of analogous meta structures discussed previously.

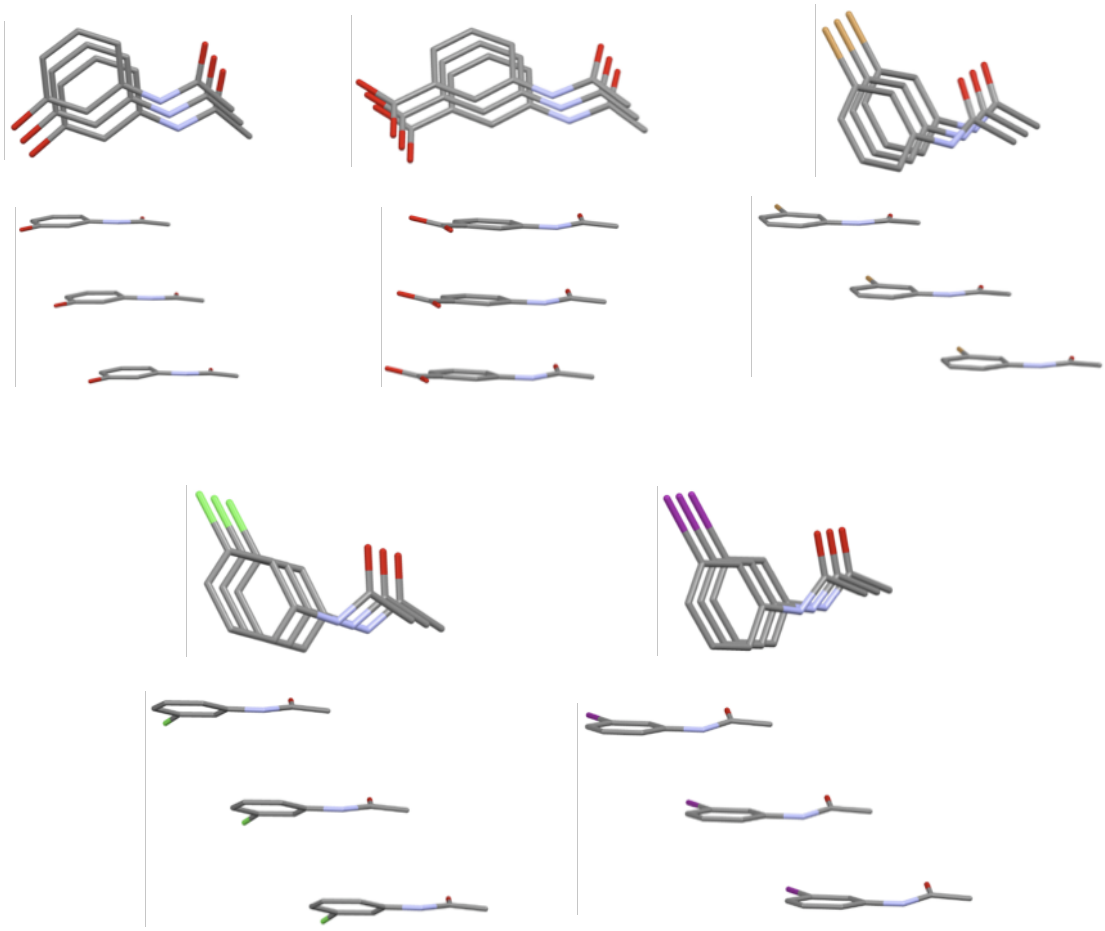


Figure 4.35: Stack **S1[t]** in **Me:m-OH**, **Me:m-COOH**, **Me:m-Br(m)**, **Me:m-I** and **Me:m-Cl** (moving clockwise starting in top left corner). Two perspectives are shown: viewed parallel and perpendicular to the stack direction, respectively.

It can be seen from Figure 4.35 that the stack is off-set so that the π-systems hardly overlap. Instead the amide group is positioned above the phenyl ring in adjacent pairs of molecules in **S1$_{[t]}$**.

There seems to be no direct correlation between the off-set in the stack and the type of substituent **X**. **Me:m-COOH** constitutes the only exception to this geometry since the off-set is much smaller in this stack so that the phenyl rings overlap. It should be mentioned that stack **S1$_{[t]}$** occurs for both molecules in the ASU for the Z'=2 structures. This packing feature is again independent from the molecular configuration so that both *syn* and *anti* arrangements of substituent and amide oxygen can be accommodated.

Table 4.16 contains the pairwise energies for this intermolecular geometry.

Structure	r_{A-B} [Å]	E_{tot}	E_{coul}	E_{pol}	E_{disp}	E_{rep}
Me:m-CF$_3$	5.172	-13.1	1.8	-6.9	-28.8	20.8
Me:m-SF$_5$	5.914	-12.1	5.2	-3.0	-23.1	8.7
Me:m-OH	4.067	-13.4	-4.0	-6.8	-34.4	31.8
Me:m-COOH	3.952	-18.2	-4.1	-8.6	-40.2	34.6

Table 4.16: Pair wise interaction energies for **S1$_{[t]}$** in **Me:m-CF$_3$**, **Me:m-SF$_5$**, **Me:m-OH** and **Me:m-COOH**. Energies are given in kJ·mol^{-1}.

Although the interaction energies could not be calculated for **Me:m-Cl**, **Me:m-Br(m)** and **Me:m-Br(o)** since they are Z' = 2 structures, it can nevertheless be seen that the stack **S1$_{[t]}$** is much less energetic than any of the HB assemblies as would be expected. Hence this packing feature can be regarded as a secondary construct compared with the HB constructs. Interestingly in **Me:m-CF$_3$** and **Me:m-SF$_5$** the contribution from the Coulomb energy is positive, i.e. repulsive. This can be attributed to the fact that stacking of the bulky substituent groups **CF$_3$** and **SF$_5$** above each other is unfavourable. However, the intermolecular interaction is dominated by dispersive forces due to the overlap of the phenyl ring and the amide group so that the stacked assembly is energetically stable overall. In fact this construct is the next strongest in energy after any hydrogen bonding feature and can in so far be seen as structure directing for these meta acetanilides.

In six of the eight structures mentioned thus far the stack **S1$_{[t]}$** is actually part of a double stack **$2 < S1_{[t]} >_{[2_1]}$**. In this arrangement two neighbouring stacks **S1$_{[t]}$** are related by 2_1 screw symmetry. The double stack is highlighted in the crystal lattice

of **Me:m-Br(m)**, **Me:m-I** and **Me:m-Cl** in Figure 4.36, whereby the latter two are taken as representatives for the respective groups of analogous structures.

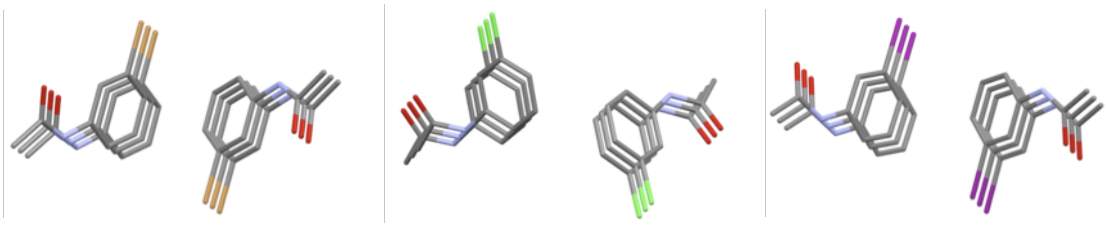


Figure 4.36: The double stack $\mathbf{2<S1_{[t]}>_{[2_1]}}$ in **Me:m-Br(m)** (left) **Me:m-Cl** (middle) and **Me:m-I** (right).

4.B.2.4. 2D similarity and the polymorphs of Me:m-Br

Four 2D SCs were revealed during the XPac analysis. All four 2D constructs are layers of molecules and are hence abbreviated with **L** and the integer is used to distinguish between unique layers.

Three of the layers, **L1**, **L2** and **L3**, contain the 1D stack $\mathbf{S1_{[t]}}$ or its double stack $\mathbf{2<S1_{[t]}>_{[2_1]}}$. Layer **L1** is a combination of $\mathbf{S1_{[t]}}$ and $\mathbf{C2_{[g]}}$ and occurs in **Me:m-COOH** and **Me:m-OH** as shown in Figure 4.37.

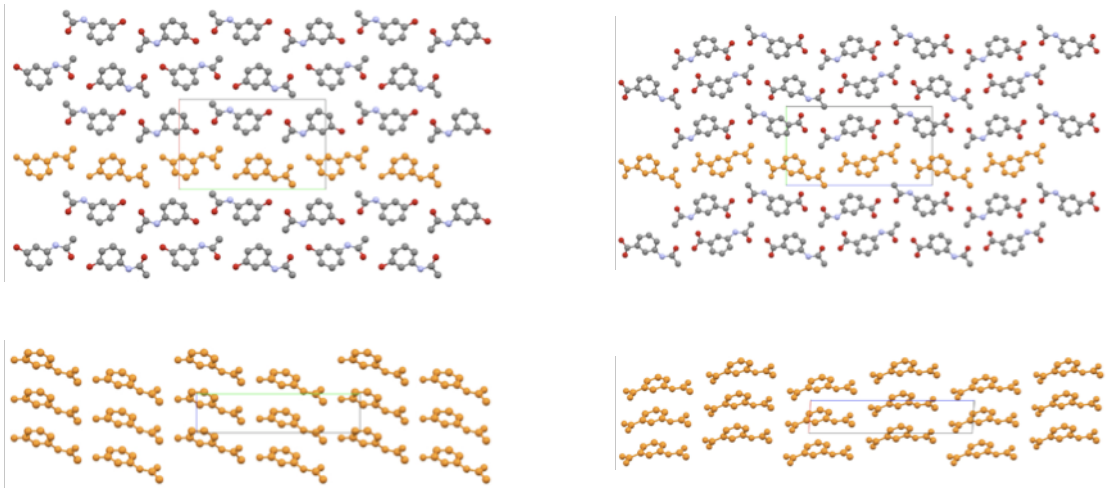


Figure 4.37: Layer **L1** as viewed along the stack direction in the crystal packing (top) and the single layer rotated by 90° (bottom). **Me:m-OH** and **Me:m-COOH** are shown on the left and right hand side, respectively.

The two structures only differ in the packing of **L1**: In **Me:m-COOH** neighbouring layers are linked via the HB dimer $\mathbf{D2}_{[\bar{1}]}$, whereas in **Me:m-OH** the layers are connected through hydrogen bonding between the amide and the hydroxyl group. In this hydrogen bond the amide nitrogen acts as the HB donor and the hydroxyl oxygen atoms as the HB acceptor. A HB chain approximately propagating with the depth of the molecule is established and adjacent molecules in this chain are related by 2_1 screw symmetry. This difference in packing of **L1** is depicted in Figure 4.38.

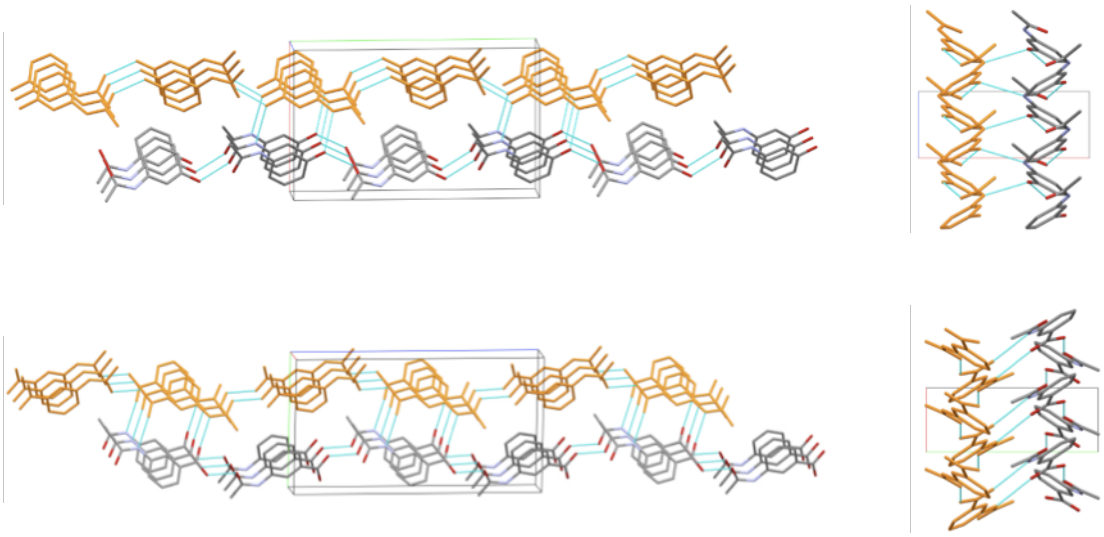


Figure 4.38: Difference in packing of **L1** in **Me:m-OH** (top) and **Me:m-COOH** (bottom) as viewed parallel to the stack (left) and along the HB chain $\mathbf{C2}_{[g]}$ (right).

The constructs **L2** and **L3** represent alternative ways of extending the double stack $\mathbf{2{<}S1}_{[t]}\mathbf{{>}}_{[2_1]}$ into two dimensions. Layer **L2** occurs in the group of packing analogues of **Me:m-I** and in **Me:m-Br(m)**. The 2D assembly is achieved by combining the double stack with the HB dimer $\mathbf{D1}_{[2_1]}$. This leads to a row of stacks wherein adjacent double stacks are linked through hydrogen bonds. Layer **L2** is shown as viewed parallel to the stack in Figure 4.39 for **Me:m-Br(m)** and **Me:m-I**. It is clear from Figure 4.39 how the layer **L2** propagates along a unit cell diagonal in **Me:m-I** rather than along a unit cell axis as in **Me:m-Br(m)**. As noted in the discussion of the hydrogen bonding in **Me:m-Br(m)** above the HB dimer is actually

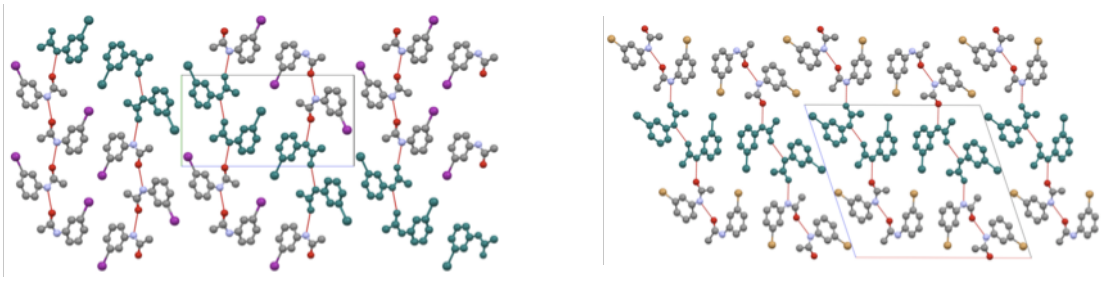


Figure 4.39: Layer **L2** as viewed parallel to stack direction in **Me:m-I** (left) and **Me:m-Br(m)** (right).

established between the two independent molecules in the ASU and has pseudo 2_1 screw symmetry.

Layer **L3** is also observed in **Me:m-I** and its analogous structures. In addition the construct is detected in **Me:m-Cl** and **Me:m-Br(o)**. This 2D feature also contains rows of the double stack $\mathbf{2 < S1}_{[t]} \mathbf{>}_{[2_1]}$, but they differ from those in **L2**. Figure 4.40 depicts the 2D construct **L3** as it occurs in **Me:m-Br(o)** and **Me:m-I** as representative for the other analogous structures.

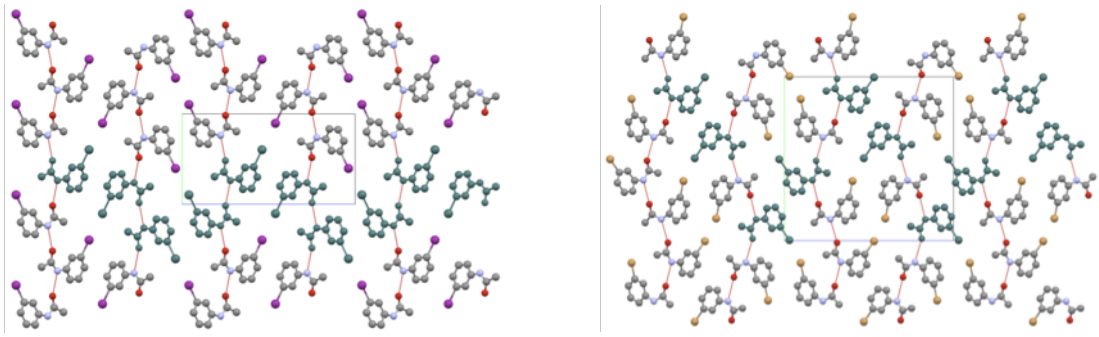


Figure 4.40: Layer **L3** as viewed parallel to stack direction in **Me:m-I** (left) and **Me:m-Br(o)** (right).

The direction of propagation of this row of stacks now coincides with one unit cell axis in **Me:m-I** as is the case in **Me:m-Br(o)**. However, in contrast to **Me:m-Br(m)** only one of the independent molecules in the ASU of **Me:m-Br(o)** is involved in the construct. To illustrate this point the packing of the two polymorphs is shown again in Figure 4.41 but with the different colours indicating the two types of molecules in the ASU.

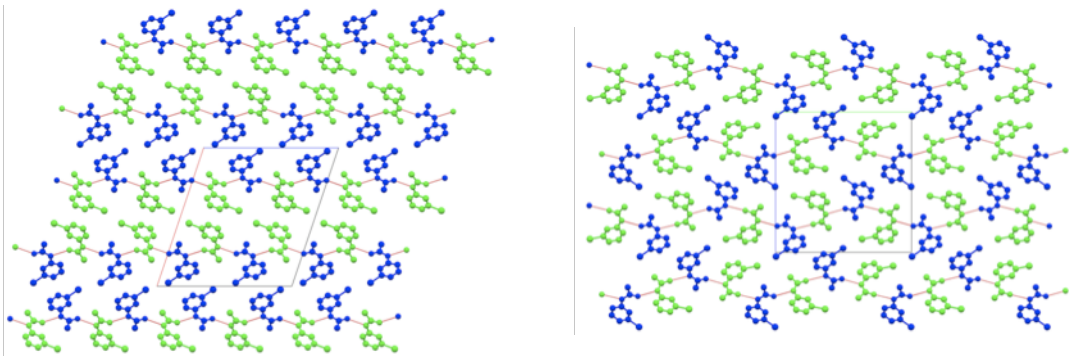


Figure 4.41: Packing in **Me:m-Br(m)** (left) and **Me:m-Br(o)** (right) with colour coding according to independent components of the ASU.

The combination of Figures 4.40 and 4.41 hence reveals that actually the packing of one component of the ASU in **Me:m-Br(o)** is equivalent to the packing mode of **Me:m-I** and its analogues. However, in **Me:m-Br(o)** this assembly is expanded so as to accommodate the second component of the ASU. The overall packing of **Me:m-Br(o)** and **Me:m-Cl** can thus be understood as an interlacing of two lattices of the **Me:m-I** type. The crystal structure of the polymorph **Me:m-Br(m)** on the other hand is built up through alternating layers of one component of the ASU.

Finally, layer **L4** is constructed via packing of the HB chain $\mathbf{C1_{[g]}}$ in rows so that the wedges are placed next to each other (rather than stacked on top of each other). The construct is displayed in Figure 4.42 for the three structures as viewed along the HB chain.

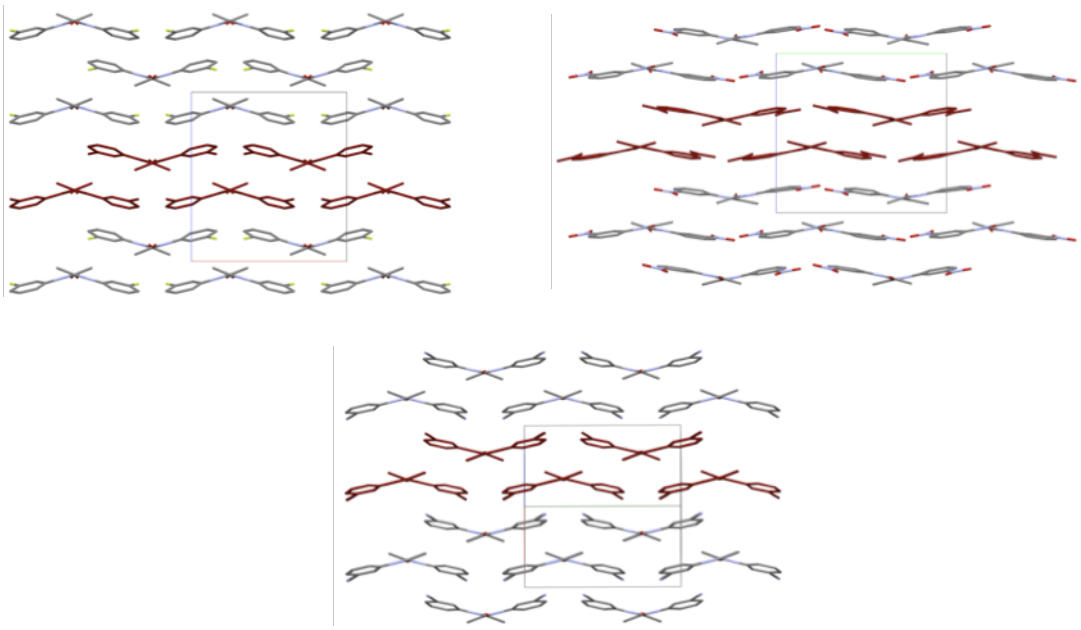


Figure 4.42: Layer **L4** in **Me:m-F** (top left), **Me:m-NO_2** (top right) and **Me:m-CN** (bottom).

In this arrangement neighbouring HB chains are related by 2_1 screw symmetry so that a corrugated layer occurs. XPac identified **L4** in **Me:m-F** and its packing analogue **Me:m-NO$_2$** as well as in **Me:m-CN**.

The difference between the crystal structure of **Me:m-CN** and **Me:m-F/Me:m-NO$_2$** hence manifests itself in the packing of these layers. In **Me:m-CN** neighbouring layers **L4** are stacked with 2_1 screw symmetry whereas in **Me:m-F** stacking of **L4** occurs with glide symmetry and pseudo glide symmetry in **Me:m-NO$_2$**, respectively. Hence the alternation of the direction of the HB chains between the layers varies between the two crystal assemblies.

4.B.2.5. The structure of Me:m-NHCOMe

Me:m-NHCOMe crystallises in an assembly that has no similarity with any of the other fourteen meta acetanilides. The shape of the molecule is quite different from the other molecules of the other compounds in the series. One of the amide groups (N1, C7, O1) is twisted out of the phenyl ring plane by approximately 52˚ whereas the ring plane and the amide plane are nearly co-planar in most of the crystal structures of the other compounds. In comparison the largest deviation is a ca 22˚ twist in one of the two molecules in the ASU of **Me:m-Br(o)**. The second amide group comprised of N2, C9, O2 (*c.f.* Ortep in Chapter 4.A.) is rotated out of the phenyl ring plane by approximately 30˚. Both amide groups engage in hydrogen bonding so that the amide group containing N1 interacts with the group of N2 and vice versa. An inversion HB dimer is established (shown in Figure 4.43), but in contrast to $\mathbf{D2}_{[\bar{1}]}$ this dimer is not planar. Instead the hydrogen bonding direction is perpendicular to the phenyl ring plane, which is facilitated through the twist of the amide groups. Two further hydrogen bonds are established so that each **Me:m-NHCOMe** molecule is involved in four hydrogen bonds. These additional hydrogen bonds point in the opposite direction to the dimer HBs and form an infinite helix with 2_1 screw symmetry effectively propagating along

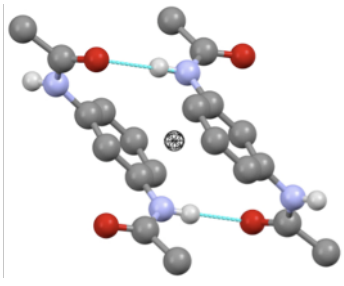


Figure 4.43: HB dimer in **Me:m-NHCOMe**. Location of inversion centre indicated by black circle.

the width of the molecule. This helix is depicted in Figure 4.44.

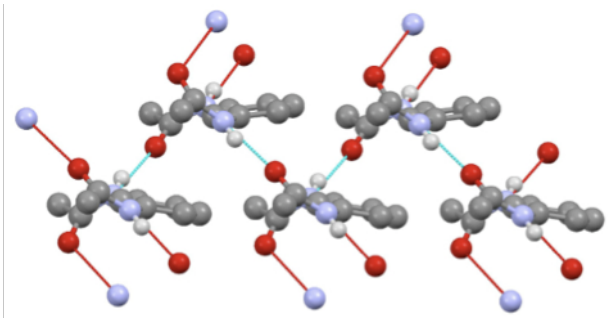


Figure 4.44: HB helix **Me:m-NHCOMe** propagating from left to right. Hanging contacts above and below the helix indicate direction of HB dimer formation.

The interplay of the helical and dimer arrangement results in a HB net parallel to the bc plane.

The resulting overall structure is layered as shown in Figure 4.45 and neighbouring nets are related by 2_1 screw symmetry.

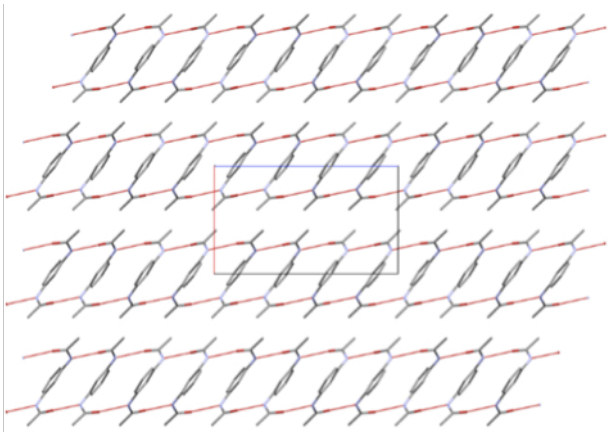


Figure 4.45: Packing of HB nets in **Me:m-NHCOMe**. HB contacts are drawn in red.

4.B.3. Structural Relationships in Me:p-X

The structural similarity in the para-acetanilides as determined with XPac is summarised in the diagram in Figure 4.46 overleaf. Remarkably, out of 35 crystal structures only two, namely **Me:p-OCOMe(m1)** and **Me:p-COOMe** are unique and share no SC with any of the other structures in the set. Four para acetanilides crystallise in polymorphic forms. The monoclinic and orthorhombic structures of **Me:p-Me** have already been reported in the literature. Also **Me:p-Br(o)** was previously determined, whilst CSF provided the new monoclinic form, **Me:p-Br(m)**. In the case of **Me:p-OCOMe** the literature indicates that three polymorphs were observed during manufacture, but only the polymorph **Me:p-OCOMe(m1)** was characterised crystallographically[11]. During this study a different crystal structure **Me:p-OCOMe(m2)** was obtained also in the monoclinic crystal setting. Finally, the polymorphs of paracetamol, **Me:p-OH**, have been the focus of man studies and hence the crystal structures of the thermodynamic monoclinic polymorph, **Me:p-OH(m)**, and the kinetic orthorhombic polymorph, **Me:p-OH(o1)**, are well known. The crystal structure of a third metastable polymorph has been elusive for a long time. Recently a structure solution derived from powder diffraction and theoretical calculations was reported[12]. The structure **Me:p-OH(o2)** is in the orthorhombic setting and contains two molecules in the ASU (Z'=2). Given the importance of paracetamol, this structure although not derived from single crystal X-ray diffraction was also included in the XPac analysis of the para acetanilides.

The similarity between the crystal structures of the para acetanilides can be classified into six groups of analogous structures, five 2D constructs and seven 1D common features in accordance with Figure 4.46.

The para series exhibits greater structural variety than the ortho and meta acetanilides, but it should be noted that there are more than twice as many crystal structures in the para series as in the ortho and meta sets. The number of constructs in the para series does not scale with the library size however, and the structural similarity in this series is not as complex as one would expect for this size of library[13].

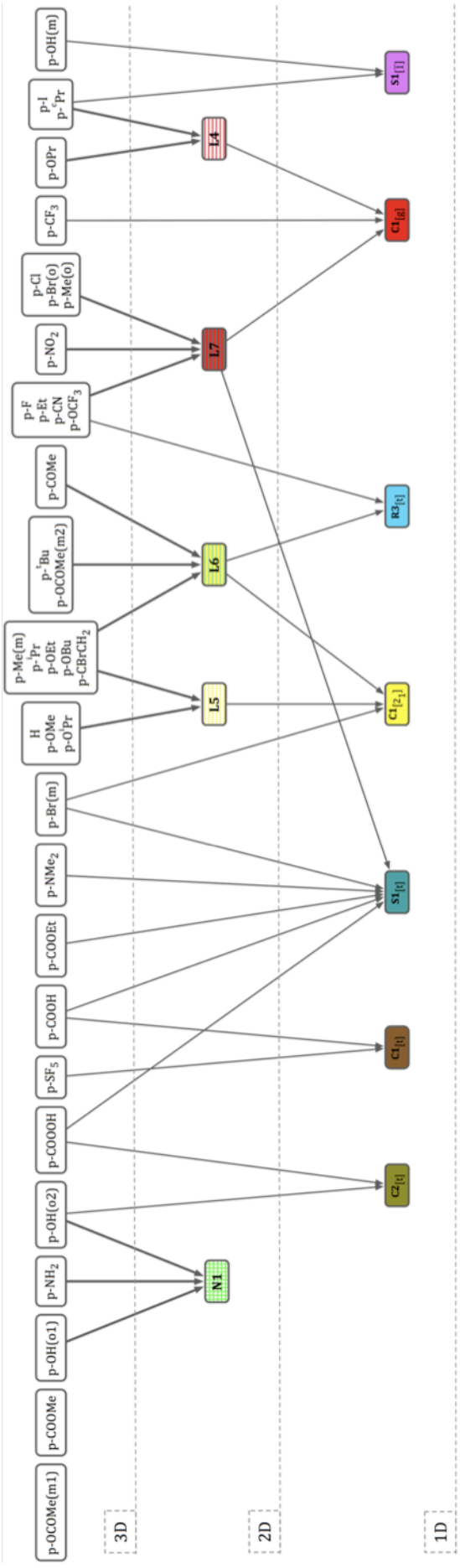


Figure 4.46: Structural similarity plot for para-substituted acetanilides.

At this stage it should be noted that the structural similarity plot experienced the longest development phase and even though the number of structures has steadily risen the structural relationships have been surprisingly robust and few in number.

The similarity plots of the initial set of 13 structures and the plot of a set of 24 para acetanilide structures is shown for comparison in Appendix 4-5. Whilst the similarity in the original concerned 7 constructs (5 1D and 2 2D), addition of 11 para acetanilide structures resulted in further 4 constructs (1 and 3 additional 1D and 2D constructs respectively). Upon reaching the current library size of 35 para acetanilide crystal structures only one new 1D SC was identified. This shows two aspects of this work. On one hand a presented structural similarity plot is never finished. As the crystal structures of new compounds or polymorphic modifications are obtained these structures need to be added to the structural comparison and hence the derivation of a similarity plot can be a seen as a working process. On the other hand the findings of this work demonstrate that the structural relationships in the series of para substituted acetanilides are very robust and similar packing arrangements are found in the crystal structures of quite different compounds. The packing features obtained through the XPac analysis are discussed in more detail in the following sections. Thereby the Aufbau principle of the para acetanilide crystal structures will be explained from the "bottom up", i.e. starting with the 1D features, it will be shown how most crystal structures can be perceived as alternative assemblies of these 1D constructs. Corresponding base vectors can be found in Appendix 4-2. The total lattice energies are recorded in Appendix 4-3 for the assessment of the importance of the pairwise interaction energies on the crystal stability.

4.B.3.1. 1D Similarity

The seven 1D SCs in the para acetanilides can be divided into constructs with hydrogen bonding and those without. The chains **C1** and **C2** belong to the former group whilst the features $\mathbf{S1_{[t]}}$, $\mathbf{S1_{[\bar{1}]}}$and $\mathbf{R3_{[t]}}$ are constructs without strong intermolecular interactions between adjacent molecules.

4.B.3.1.a SCs without strong intermolecular interaction: $S1_{[t]}$, $S1_{[\bar{1}]}$ and $R3_{[t]}$

The constructs $\mathbf{S1_{[t]}}$ and $\mathbf{S1_{[\bar{1}]}}$ are offset stacks of molecules which propagate with the depth of the molecules. $\mathbf{S1_{[t]}}$ was already discovered in the meta acetanilides. As can be seen from Figure 4.46 this is quite a common packing motif with thirteen occurrences corresponding to eight unique crystal structures in the para set. Furthermore it should be recalled that $\mathbf{S1_{[t]}}$ was the most commonly observed SC in the meta series, hence this feature deserves a closer inspection. Considering the molecular dimensions of the acetanilides it can be concluded that the general shape of the molecule is that of a 2D object (*c.f.* Figure 4.47) where the depth of the molecule is much smaller than the other dimensions.

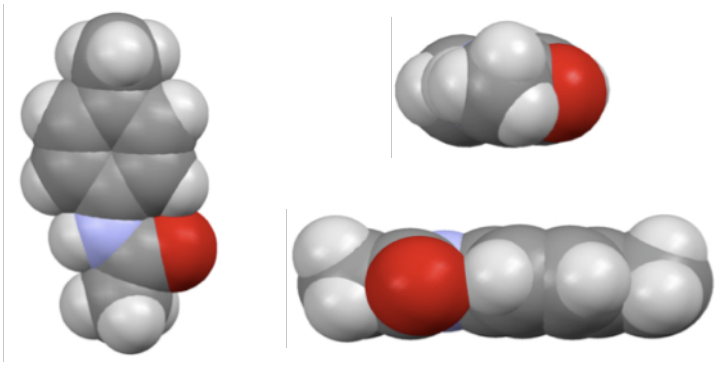
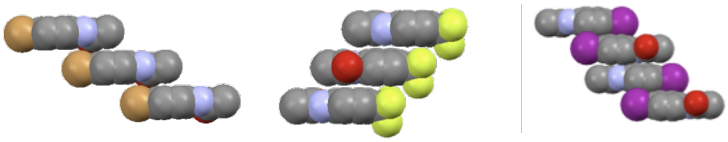


Figure 4.47: General shape of para-acetanilides shown parallel to the depth, length and width of **Me:m-Me(o)**.

Assembling such objects into infinite 1D features with crystallographic symmetry can be achieved in a number of ways. For example, rows of the molecules could be constructed propagating with the length or width or a linear combination of the two dimensions. Molecular stacks could also be formed, which is especially a common form of intermolecular assembly for aromatic systems (*c.f.* Chapter 1.A.3.). A stack of this type will yield the smallest exposed molecular surface compared with rows. Allen *et al*. have demonstrated with their box model that an arrangement that minimises this surface is preferred over large exposed molecular surfaces[14]. Furthermore as was seen in Section A.3. of Chapter 1 stacks involving π-systems are generally offset rather than parallel. After these considerations the

detection of the molecular stacks can be easily rationalised. Neighbouring molecules in $\mathbf{S1_{[t]}}$ and $\mathbf{S1_{[\bar{1}]}}$ are related by translation and inversion, respectively. $\mathbf{S1_{[t]}}$ can be found in the following crystal structures: **Me:p-X** with **X = F, Cl, Br(m), Br(o), OCF_3, Me(o), Et, CN, NO_2, NMe_2, COOH, COOEt** and **COOOH.** This makes the translational stack one of the most commonly occurring constructs in the series of para acetanilide structures. $\mathbf{S1_{[\bar{1}]}}$ on the other hand, is only observed in three structures: **Me:p-I**, **Me:p-^{c}Pr** and **Me:OH(m)**. Figure 4.48 displays the two stacks for **Me:p-Br(m)** and **Me:p-I**.



(a) **(b)** **(c)**

Figure 4.48: (a) $\mathbf{S1_{[t]}}$ stack in **Me:p-Br(m)**. (b) Stacking of **Me:p-CF_3** molecules. (c) SC $\mathbf{S1_{[\bar{1}]}}$ as found in **Me:p-I**.

Also shown in Figure 4.48 is a third molecular stack as it is found in **Me:p-CF_3**. However, this stack possesses glide symmetry and no further para acetanilide structure contains this type of stack hence it is not listed in the similarity plot. Table 4.17 lists the offsets of stack $\mathbf{S1_{[t]}}$.

Structure	d_{ring} [Å]	d_{offset} [Å]	$d_{X\cdots ring}$ [Å]
Me:p-F	3.428	3.088	3.464 / 3.443[a)]
Me:p-Cl	3.402	5.628	3.655[b)]
Me:p-Br(m)	3.634	5.650	3.728[b)]
Me:p-Br(o)	3.436	5.776	3.554[b)]
Me:p-Me(o)	3.427	5.571	3.566[b)]
Me:p-Et	3.439	3.032	3.505 / 3.435[a)]
Me:p-CN	3.464	1.671	n/a
Me:p-NO_2	3.405	1.232	n/a
Me:p-NMe_2	3.373	6.710	n/a
Me:p-OCF_3	3.552	2.743	3.522 / 3.531[a)]
Me:p-COOH	3.163	6.066	3.792[b)]
Me:p-COOEt	3.261	7.790	n/a
Me:p-COOOH	3.527	3.590	3.545 / 3.599[c)]

Table 4.17: d_{ring} = distance between neighbouring ring planes in $\mathbf{S1_{[t]}}$; d_{offset} = offset of phenyl ring with respect to neighbouring ring; $d_{X\cdots ring}$ = distance between substituent **X** and ring centroid in neighbouring phenyl ring. [a)] Distance between ring centroid and substituent **X** / neighbouring amide nitrogen atom. [b)] Distance between ring centroid and acetamide methyl carbon atom. [c)] Distance between ring centroid and O3 / amide carbon atom.

The largest offset occurs in **Me:p-COOEt** followed by **Me:p-NMe$_2$**. In these structures the phenyl rings do not overlap nor are the substituents in close enough proximity to the phenyl ring for any interaction with the π-system. The next group has offsets ranging from 5.571 to 6.066 Å. In these structures there is some overlap between the phenyl ring and the adjacent acetamido methyl group with the separation being smaller than the sum of the van der Waals radii of the constituent atoms (3.77 Å). This indicates that there may be some C-H···π interaction present in the stacks $\mathbf{S1_{[t]}}$ in the compounds **Me:p-X** with **X** = **Cl**, **Br**, **Me(o)** and **COOH**. Smaller offsets ranging between 2.743 and 3.590 Å occur in **Me:p-F**, **Me:p-Et**, **Me:p-OCF$_3$** and **Me:p-COOOH**. In these structures the π-system is enclosed by the substituent **X** on one side and by an adjacent amide group on the other side. However the distance between the ring centroid and the respective enclosing atoms is slightly greater than the respective sum of the van der Waals radii of the constituent atoms. This might be the result of the balance between possible interactions on both sides of the phenyl ring and any steric hindrance associated with smaller intermolecular separation. Lastly the smallest offsets occur in **Me:p-CN** and **Me:p-NO$_2$** where the shift of the phenyl rings with respect to each other is so small it does not permit any overlap with substituents or the amide residue of the molecule. This arrangement comes closest to a parallel assembly of the phenyl rings in the stack $\mathbf{S1_{[t]}}$.

The inversion symmetry of stack $\mathbf{S1_{[\bar{1}]}}$ results in an overlap of amide group with the neighbouring phenyl ring. The distances between the ring centroid and the amide nitrogen atom increase in the order: **Me:p-OH(m)** (3.318 Å), **Me:p-I** (3.380 Å), **Me:p-^{c}Pr** (3.545 Å) and are close to the sum of the van der Waals radii (3.32 Å) of the constituent atoms. Given the inversion symmetry both phenyl ring and amide group of one molecule are involved in the overlap described above.

The respective pairwise interaction energies are recorded in Table 4.18. The intermolecular energies reveal that stack $\mathbf{S1_{[\bar{1}]}}$ is associated with a much greater energy gain than $\mathbf{S1_{[t]}}$. Presumably the antiparallel overlap of the molecules within the stack presents a favourable geometry for maximising the dispersive forces between the molecules.

Structure	$r_{A\text{-}B}$ [Å]	E_{tot}	E_{coul}	E_{pol}	E_{disp}	E_{rep}
$S1_{[t]}$						
Me:p-F	4.614	-15.7	-0.7	-3.2	-27.8	16.0
Me:p-Cl	6.528	-12.4	-2.3	-2.4	-21.8	14.1
Me:p-Br(m)	6.718	-13.6	-3.8	-3.2	-21.7	15.1
Me:p-Br(o)	6.721	-12.3	-4.4	-3.5	-24.7	20.2
Me:p-Me(o)	6.541	-14.3	-8.8	-4.9	-25.5	25.0
Me:p-Et	4.585	-21.5	-4.7	-5.7	-40.2	29.1
Me:p-CN	3.846	-13.7	4.6	-5.5	-42.9	30.1
Me:p-OCF_3	4.489	-16.1	-1.8	-3.4	-30.8	16.3
Me:p-COOH	6.841	-15.9	-4.2	-2.6	-20.0	10.9
Me:p-COOEt	8.445	-21.2	-10.7	-4.9	-25.8	20.1
$S1_{[\bar{1}]}$						
Me:p-OH(m)	3.493	-31.5	-14.6	-6.3	-39.8	29.2
Me:p-^{c}Pr	4.546	-31.4	-8.4	-3.9	-36.3	17.2

Table 4.18: Intermolecular interaction energies for the two stacks. Energies are given in kJ·mol^{-1}.

The interaction energies of stack **$S1_{[t]}$** vary from -12.3 to -21.5 kJ·mol^{-1} and fall into the energy range expected for interactions involving the π-system as acceptor[15]. All interactions in the stack are of dispersive nature as the values in Table 4.18 show. It should be noted that 'double overlap' of the phenyl ring with both substituent and amide group does not result in significant energy stabilisation overall, although the dispersive energies are enhanced in these cases (*c.f.* Table 4.17 an 4.18). Interestingly, although there is no direct molecular overlap between the stacked molecules in **Me:p-COOEt**, the overall energy is one of the highest of the stacks **$S1_{[t]}$**.

Construct **$R3_{[t]}$** is a simple row of molecules propagating along the length of the molecule and with translational symmetry only. This structural feature is present in the twelve crystal assemblies **Me:p-X** where **X = F**, **OCF_3**, **Me(m)**, **Et**, **^{i}Pr**, **^{t}Bu**, **CN**, **OEt**, **OBu**, **COMe**, **C_2H_2Br** and **OCOMe(m2)**. Figure 4.49 depicts the construct in two orientations for **Me:p-Me(m)**.

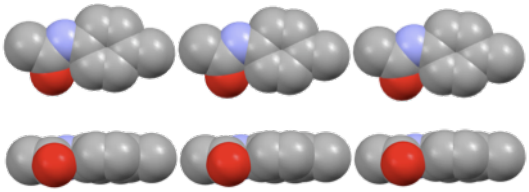


Figure 4.49: R3$_{[t]}$ as found in **Me:p-Me(m)**. Top: View parallel to the depth of the molecule. Bottom: View parallel to the width of the molecule.

The frequency of occurrence demonstrates that **R3$_{[t]}$** can be regarded as another robust motif for the intermolecular aggregation of the para acetanilide molecules. Table 4.19 contains the pairwise intermolecular energies.

Structure	$r_{A\text{-}B}$ [Å]	E_{tot}	E_{coul}	E_{pol}	E_{disp}	E_{rep}
Me:p-Me(m)	11.678	-2.3	-0.4	-0.3	-5.5	3.9
Me:p-OEt	13.622	-3.0	-1.2	-0.4	-5.7	4.4
Me:p-OBu	15.881	-2.8	-1.0	-0.3	-4.3	2.7
Me:p-CBrCH$_2$	12.812	-5.2	-2.2	-0.6	-6.0	3.6
Me:p-OCOMe(m2)	13.877	-1.7	0.5	-0.2	-4.0	2.1
Me:p-COMe	12.900	0.5	2.3	-0.5	-4.1	2.9
Me:p-F	10.978	-2.8	-2.0	-0.6	-3.7	3.5
Me:p-Et	12.718	-1.9	-0.2	-0.4	-3.9	2.5
Me:p-CN	12.505	-8.1	-6.6	-1.7	-3.5	3.7
Me:p-OCF$_3$	12.972	-3.2	-2.3	-0.3	-3.2	2.6

Table 4.19: Intermolecular interaction energies for the row **R3$_{[t]}$**. Energies are given in kJ·mol^{-1}.

From the energy values it can be deduced that the molecules in **R3$_{[t]}$** are too far apart for there to be any significant intermolecular interaction. In fact the closest coordination spheres of **Me:p-^{i}Pr** and **Me:p-^{t}Bu** do not include the molecules in **R3$_{[t]}$** so that intermolecular energies could not be obtained in these cases. This demonstrates that construct **R3$_{[t]}$** is established on purely geometrical grounds and hence cannot be structure determining for the para acetanilide crystal structures. Instead it may be the result of the interplay of other more energetically favourable intermolecular assemblies confirming the similar geometrical arrangement of the molecules in the crystal.

In summary, the stacks $\mathbf{S1_{[t]}}$ as well as the row $\mathbf{R3_{[t]}}$ are to be considered as secondary packing motifs compared with any hydrogen bonded features since the intermolecular energies are of weak strength. Instead they arise because of the close balance of weak and dispersive forces. However, the reoccurrence of certain motifs across a range of structures implies that these arrangements are not entirely random but constitute some form of packing preference due to efficient space filling. The interaction energies of $\mathbf{S1_{[\bar{1}]}}$ on the other hand are more in the range of hydrogen bond interactions and may play a greater role in directing the crystal assembly. This stack is only observed in three structures though and hence the combination of other geometrical intermolecular assemblies offer alternatives that are energetically as favourable as stack $\mathbf{S1_{[\bar{1}]}}$.

4.B.3.1.b Hydrogen bonded features: $C1_{[t]}$, $C1_{[g]}$, $C1_{[2_1]}$ and $C2_{[t]}$

The hydrogen bonded chains $\mathbf{C1_{[t]}}$, $\mathbf{C1_{[g]}}$, $\mathbf{C1_{[2_1]}}$ were already detected as constructs in the ortho and meta series, although only two of the three chains were observed at the time in the individual series. These HB chains occur frequently in the para series and facilitate amide-amide hydrogen bonding. In terms of HB connectivity the chains are of the same type but of different geometry due to inherent symmetry. In line with this view, the geometry of one chain can be transformed into another via rotation of the constituent molecules about the HB chain.

The chain $\mathbf{C1_{[t]}}$ is the least common of the HB chains of type 1 with only two occurrences in **Me:p-COOH** and **Me:p-SF_5**. Neighbouring molecules in the chain are related by translational symmetry only. In order to facilitate the HB interaction in this intermolecular assembly the molecular backbone is twisted, so that the phenyl ring plane does not coincide with the mean plane of the amide group. Furthermore in **Me:p-COOH** the substituent engages in the well known hydrogen bond dimer between carboxylic acid groups. This dimer links neighbouring chains $\mathbf{C1_{[t]}}$ into an overall ladder motif (Figure 4.50).

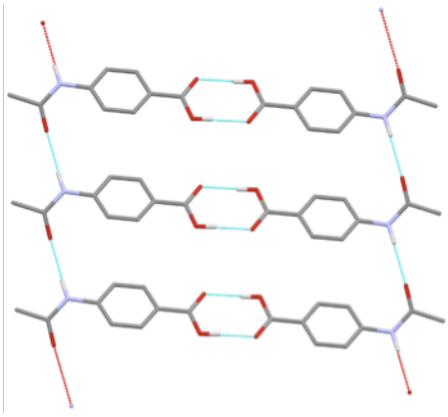


Figure 4.50: Hydrogen bonding in **Me:p-COOH**: chain $\mathbf{C1_{[t]}}$ and inversion carboxylic acid dimer.

The chains $\mathbf{C1_{[g]}}$ and $\mathbf{C1_{[2_1]}}$ are very common and each SC occurs in 12 para acetanilide structures. Neighbouring molecules within $\mathbf{C1_{[g]}}$ are related by glide symmetry leading to the wedge-like (or V) shape already mentioned in the discussion of the structural similarity of the ortho and meta series. $\mathbf{C1_{[2_1]}}$ possesses 2_1 screw symmetry along the HB direction. Hence chains with varying degrees of corrugation are formed. Figure 4.51 displays an example for each of the HB chains **C1** as viewed along and perpendicular to the direction of propagation of the feature.

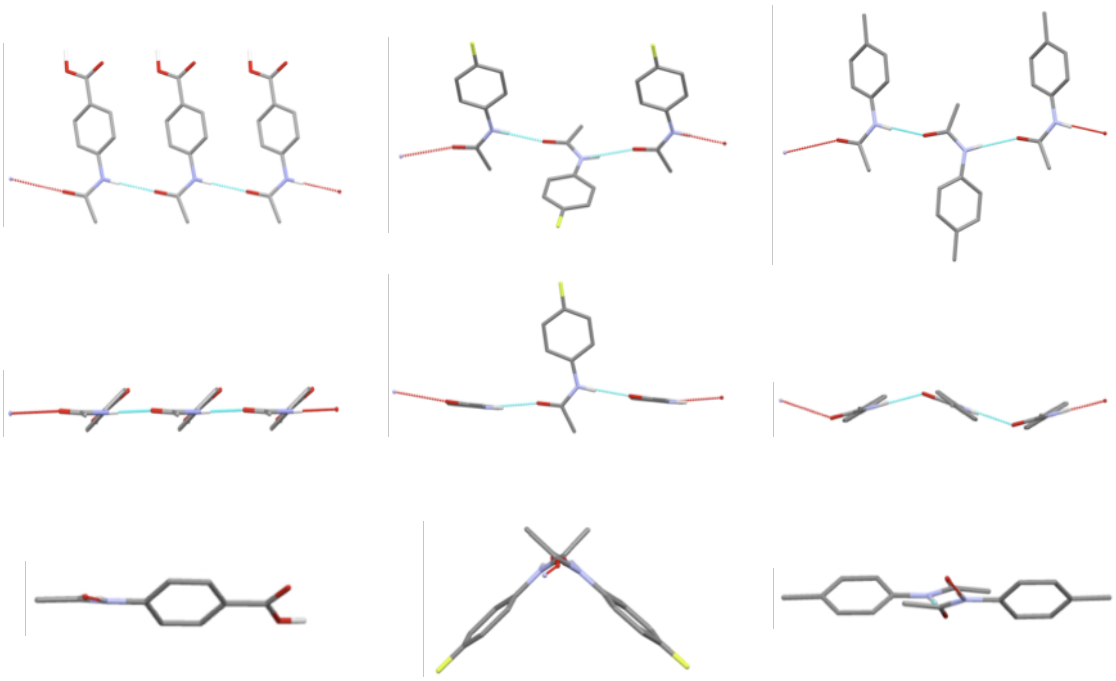
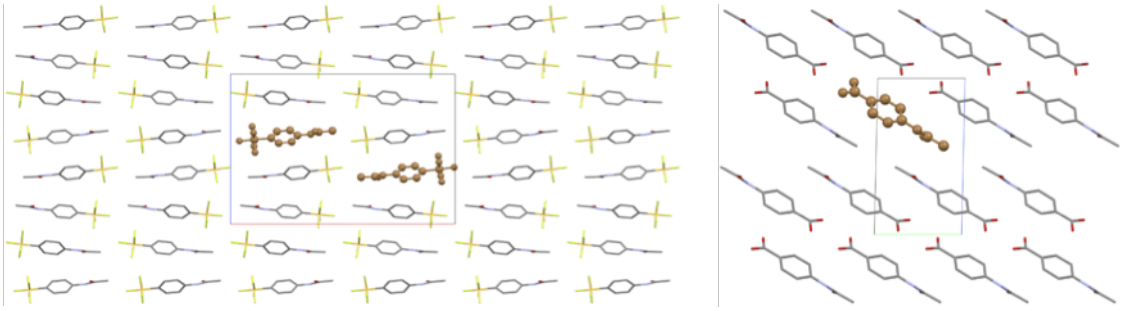


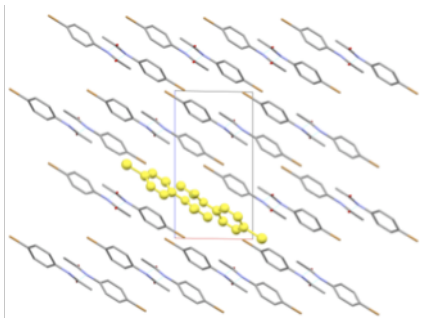
Figure 4.51: The three HB motifs of type **C1** as viewed from the top (top), side-on (middle) and along the HB (bottom). $\mathbf{C1_{[t]}}$ shown for **Me:p-COOH** (left), $\mathbf{C1_{[g]}}$ in **Me:p-F** (middle) and $\mathbf{C1_{[2_1]}}$ as found in **Me:p-Me(m)** (right).

Figures 4.52 to 4.54 display the chains $\mathbf{C1_{[t]}}$, $\mathbf{C1_{[2_1]}}$ and $\mathbf{C1_{[g]}}$, respectively, as they occur in the unique structures of the para acetanilides. In the figures the viewing direction is along the hydrogen bond (i.e. into the plane of the paper).

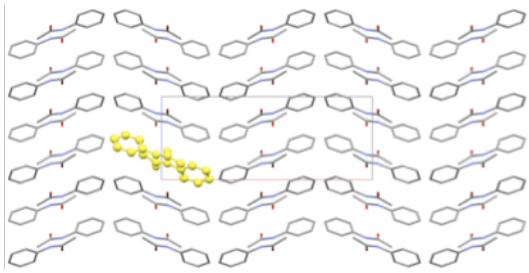


Me:p-SF_5 **Me:p-COOH**

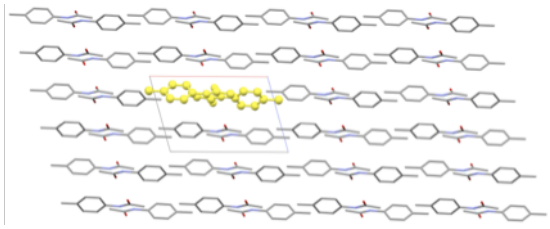
Figure 4.52: The chain $\mathbf{C1_{[t]}}$ as it occurs in the crystal packing of **Me:p-COOH** and **Me:p-SF_5**. The HB chain is highlighted in brown for each component of the ASU and viewed parallel to the HB direction.



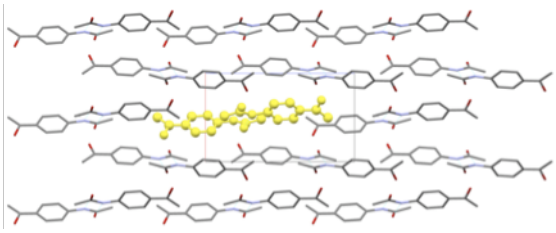
Me:p-Br(m)



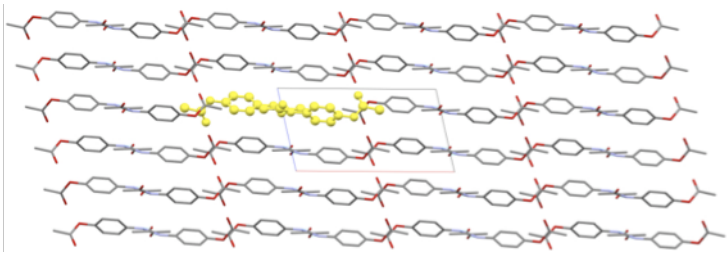
Me:H



Me:p-Me(m)

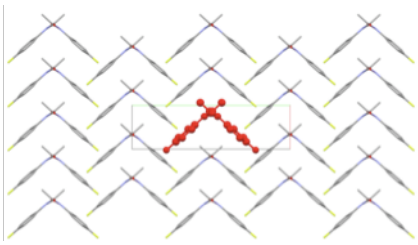


Me:p-COMe

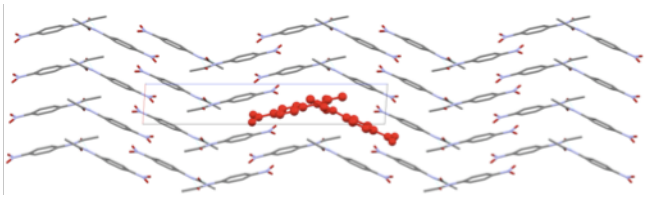


Me:p-OCOMe(m2)

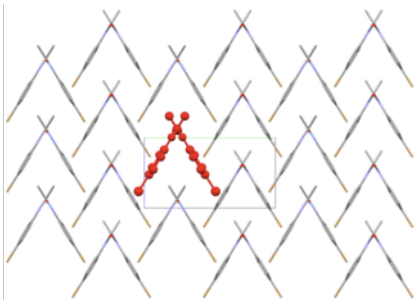
Figure 4.53: The chain $\mathbf{C1}_{[2_1]}$ as it occurs in the packing of the unique structures **Me:p-Br(m)**, **Me:H**, **Me:p-Me(m)**, **Me:p-COMe** and **Me:p-OCOMe(m2)**. Viewed along the HB direction.



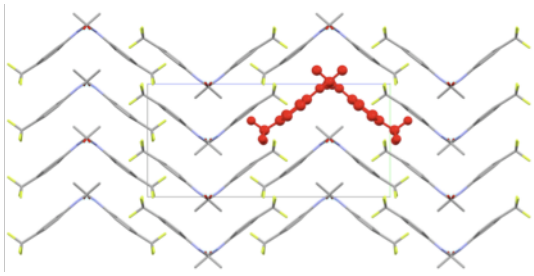
Me:p-F



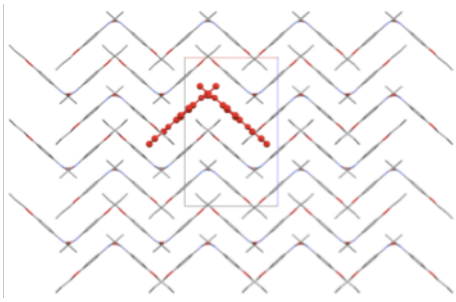
Me:p-NO_2



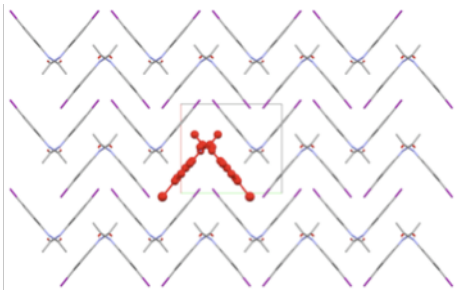
Me:p-Br(o)



Me:p-CF_3



Me:p-OPr



Me:p-I

Figure 4.54: The chain **C1**$_{[g]}$ as it occurs in the packing of the unique structures **Me:p-X** with **X** = **F**, **NO_2**, **Br(o)**, **CF_3**, **OPr**, **I**. Viewed along the hydrogen bond.

In contrast to **C1**$_{[t]}$ the geometry of the intermolecular hydrogen bonding between the molecules in chains **C1**$_{[g]}$ and **C1**$_{[2_1]}$ does not impose any conformational restraints on the molecular backbone. However, inspection of the molecular conformation indicates some correlation between the molecular structure and the geometry of the **C1** hydrogen bonding.

The molecular variable found to differ between the HB chains is the dihedral angle θ between the phenyl ring and amide group mean planes. Table 4.20 contains the angular data as well as the hydrogen bond parameters for the three chains **C1**$_{[t]}$, **C1**$_{[g]}$ and **C1**$_{[2_1]}$.

Structure	θ [°]	$r_{N\cdots O}$ [Å]	$\angle_{N-H\cdots O}$ [°]
$C1_{[t]}$			
Me:p-SF_5	35.596/37.133	3.082/3.092	159/155
Me:p-COOH	39.255	2.949(3)	176(3)
$C1_{[g]}$			
Me:p-F	3.322	2.799(2)	171(2)
Me:p-Cl	3.761	2.834(2)	172
Me:p-Br(o)	5.22	2.864(3)	162(2)
Me:p-I	16.282	2.866(2)	171(3)
Me:p-CF_3	7.45	2.841(4)	162
Me:p-OCF_3	5.312	2.862(9)	157
Me:p-Me(o)	5.662	2.903	174
Me:p-Et	5.019	2.868(4)	153
Me:p-^{c}Pr	9.531	2.907	173
Me:p-CN	6.195	2.829(18)	156
Me:p-NO_2	3.791/12.896	2.958(9)/2936(9)	170/171
Me:p-OPr	10.265	2.799	161
$C1_{[2_1]}$			
Me:H	16.142	2.914	171
Me:p-Br(m)	18.998	2.879(2)	173(2)
Me:p-Me(m)	16.616	2.904(2)	174(2)
Me:p-^{i}Pr	14.074	2.881(1)	173(1)
Me:p-^{t}Bu	22.576	2.840(2)	171(2)
Me:p-OMe	17.813	2.854(1)	177(1)
Me:p-OEt	27.68	2.922(1)	175(1)
Me:p-OiPr	19.677	2.869	175
Me:p-OBu	25.52	3.000(1)	166(1)
Me:p-OCOMe(m2)	10.149	2.846(2)	173(2)
Me:p-COMe	15.535	2.919(7)	175
Me:p-$CBrCH_2$	19.651	2.984	176

Table 4.20: Dihedral angles and HB parameters for the para acetanilides containing the HB chains **C1**.

Plotting the distance between the HB donor and acceptor versus the dihedral angle θ (Figure 4.55) does not result in a strong correlation, but trends in the data are clearly discernible. The plot in Figure 4.55 shows that the largest dihedral angles (θ > 30°) are observed in chain $\mathbf{C1_{[t]}}$ together with the longest hydrogen bonds.

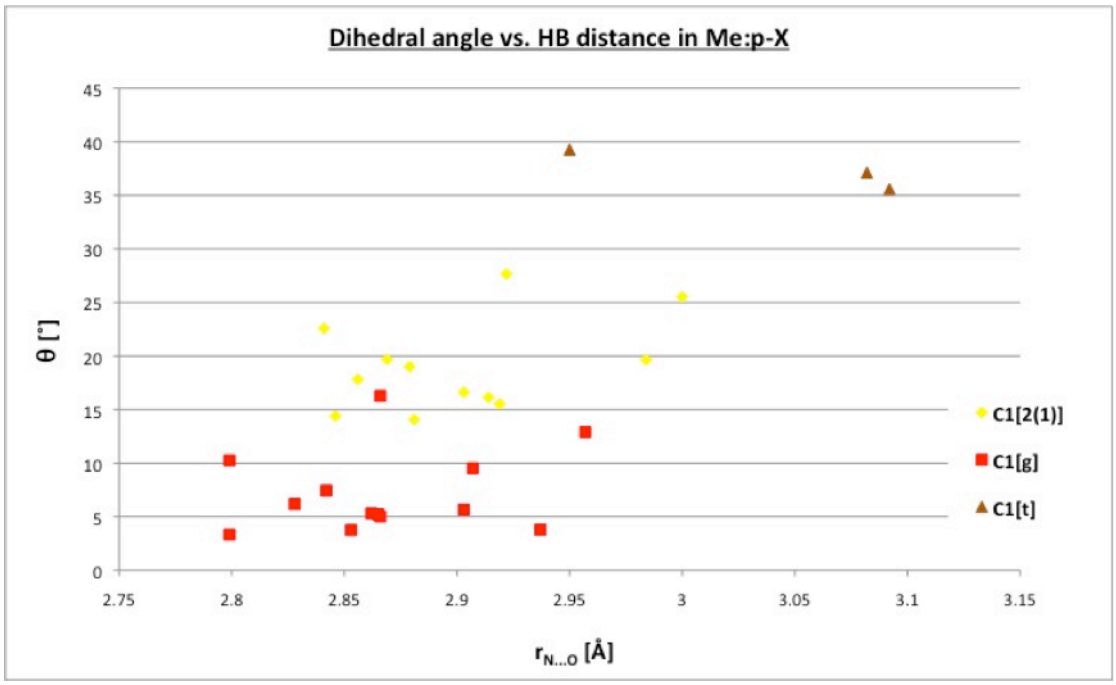


Figure 4.55: Dihedral angle plotted against the HB distance in para acetanilides. Colour coding according to HB chain motif.

A medium range of dihedral angles ($12 < \theta < 30^{\circ}$) and HB lengths is found in $\mathbf{C1_{[2_1]}}$ whilst the smallest dihedral angles ($0 < \theta < 12^{\circ}$) and shortest hydrogen bonds occur in $\mathbf{C1_{[g]}}$. With respect to the dihedral angle there is one "outlier": the molecule **Me:p-I** has a dihedral angle of approximately 16° which lies in the medium range, but the HB chain geometry is that of $\mathbf{C1_{[g]}}$. One of the two independent molecules of the ASU in **Me:p-NO$_2$** has a dihedral angle of ca 12° and lies hence at the borderline. Furthermore there is a much less clear correlation between the HB distances and the HB chain geometry. Values between 2.84 and 2.96 Å are observed both in $\mathbf{C1_{[g]}}$ and $\mathbf{C1_{[2_1]}}$. Hence the plot in Figure 4.55 can only be regarded as indicator of trends rather than the confirmation of definite correlation. However, the observation of these trends prompted the analysis of the conformational freedom of the acetanilide molecules, which is discussed in Chapter 5.

Energetically the HB chains are quite similar as the intermolecular interaction energies in Table 4.21 reveal. The highest energy was calculated for $\mathbf{C1_{[t]}}$ and the dispersive energy indicates that the close proximity of neighbouring aromatic

systems in **Me:p-COOH** is energetically favourable. This is caused by the chain geometry and was also observed in the construct $\mathbf{C1_{[t]}}$ present in the ortho series.

Structure	$r_{A\text{-}B}$ [Å]	E_{tot}	E_{coul}	E_{pol}	E_{disp}	E_{rep}
$C1_{[t]}$						
Me:p-COOH	5.021	-42.5	-40.3	-13.2	-28.6	39.5
$C1_{[g]}$						
Me:p-F	5.640	-31.6	-47.3	-18.9	-18.4	53.1
Me:p-Cl	5.349	-34.9	-43.5	-16.4	-19.7	44.7
Me:p-Br(o)	5.348	-34.9	-41.9	-14.8	-19.5	41.4
Me:p-CF_3	6.551	-31.5	-41.4	-15.9	-16.8	42.6
Me:p-OCF_3	7.088	-31.0	-39.9	-15.0	-17.9	41.9
Me:p-Me(o)	5.696	-34.1	-39.3	-14.0	-18.4	37.6
Me:p-Et	7.354	-32.5	-40.6	-15.1	-17.5	40.7
Me:p-^{c}Pr	6.449	-35.3	-39.2	-14.2	-19.4	37.5
Me:p-OPr	8.287	-30.8	-47.4	-18.3	-18.2	53.1
Me:p-CN	6.478	-27.5	-27.7	-11.1	-18.0	29.3
$C1_{[2_1]}$						
Me:H	5.798	-36.6	-41.3	-15.4	-22.9	43.0
Me:p-Br(m)	5.961	-33.4	-44.1	-18.3	-25.0	53.9
Me:p-Me(m)	6.917	-36.1	-41.1	-15.1	-20.6	40.7
Me:p-^{i}Pr	8.706	-36.8	-45.5	-16.6	-21.7	47.1
Me:p-^{t}Bu	9.223	-33.8	-44.3	-17.4	-20.8	48.7
Me:p-OMe	7.187	-37.6	-47.1	-18.1	-23.1	50.7
Me:p-OEt	8.474	-37.2	-39.8	-13.9	-17.2	33.7
Me:p-O^{i}Pr	9.151	-35.8	-44.8	-17.7	-20.8	47.5
Me:p-OBu	10.418	-36.4	-34.0	-11.6	-18.2	27.5
Me:p-COMe	7.693	-38.7	-29.6	-9.6	-19.1	19.6
Me:p-OCOMe(m2)	8.280	-35.4	-46.0	-17.8	-21.1	49.5
Me:p-$CBrCH_2$	7.351	-38.2	-37.6	-13.5	-21.0	33.9

Table 4.21: Intermolecular interaction energies for the HB chains of type **C1**. Energies are given in kJ·mol^{-1}.

However, it should be kept in mind that it was only possible to compute the interaction energy for one para compound with $\mathbf{C1_{[t]}}$; hence it would be advantageous to base any interpretation concerning the stability of the three chains on this one calculated value. The hydrogen bond energies in $\mathbf{C1_{[g]}}$ and $\mathbf{C1_{[2_1]}}$ are very similar and there is only a very weak indication that chain $\mathbf{C1_{[2_1]}}$ is the slightly more stable arrangement of the two **C1** chains, but again this should not be over-interpreted.

One further HB feature was identified by the XPac analysis: $\mathbf{C2_{[t]}}$, which occurs in two para acetanilide structures, namely in **Me:p-COOOH** and in the meta-stable paracetamol polymorph **Me:p-OH(o2)**. In this construct hydrogen bonding is established between the substituent and the amide group in a head-to-tail interaction as displayed in Figure 4.56.

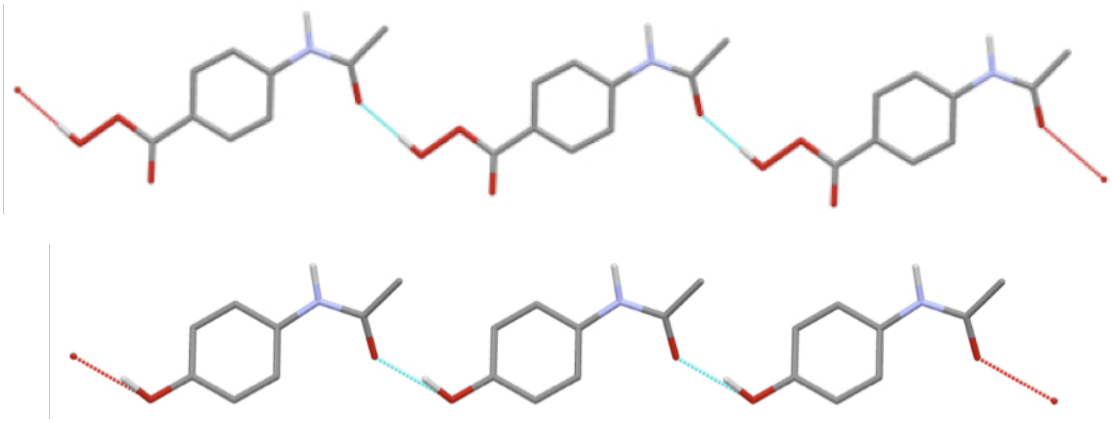


Figure 4.56: The chain $\mathbf{C2_{[t]}}$ as formed by **Me:p-COOOH** (top) and **Me:p-OH(o2)** (bottom).

It can be seen that the amide oxygen atom acts as the HB acceptor with the substituent OH group assuming the role of HB donor. The associated graph sets are $\mathbf{C_1^1(11)}$ and $\mathbf{C_1^1(9)}$ for **Me:p-COOOH** and **Me:p-OH(o2)** with the substituents causing the difference in the pattern degree (i.e. repeat length of the chain). Both independent molecules in the ASU of **Me:p-OH(o2)** form separate chains $\mathbf{C2_{[t]}}$. Actually a SC of the same type has already been discussed in the meta series, namely construct $\mathbf{C2_{[g]}}$. These two HB features are of the same type, i.e. they involve the same HB acceptor and donor groups, but they differ in their geometry and hence symmetry: one has translational symmetry whereas the other is consistent with glide symmetry. However, they can be transformed into one another by rotation of the molecules about the hydrogen bond and are hence assigned the same type.

Analysing the occurrence of all HB SCs it can be deducted that hydrogen bonding between amide groups is an important energetic interaction. This interaction is robust even if stronger HB acceptors than the amide group are introduced at the para position on the phenyl ring. It is noteworthy that big acceptor groups, such as

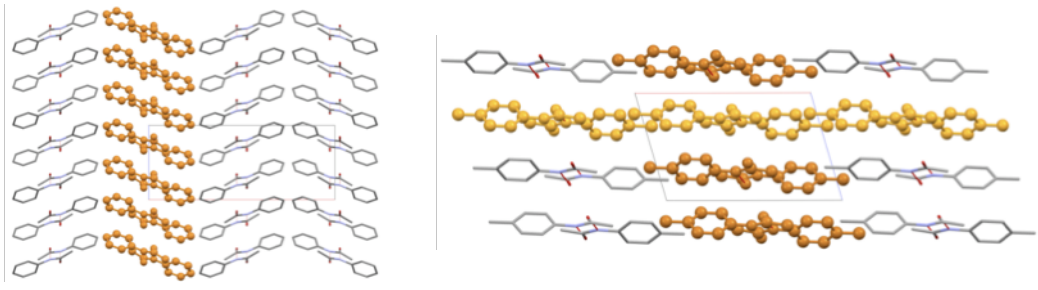
COOMe and **COOEt** prevent the type **C1** hydrogen bonding acting as HB acceptor instead of the amide group. However, as seen from the polymorphs **Me:p-OCOMe(m1)** and **Me:p-OCOMe(m2)** bulky substituents with HB acceptor functionality can also be accommodated in hydrogen bonds of type **C1**. The presence of alternative hydrogen bond donors on the other hand disrupts the amide-amide hydrogen bond in every instance except for **Me:p-COOH**.

4.B.3.2. 2D assemblies

Of the five 2D constructs identified with XPac only the net **N1** cannot be obtained via a combination of lower dimensionality features. The layers **L5** and **L6** contain the amide-amide hydrogen bond $\mathbf{C1}_{[2_1]}$ whereas $\mathbf{C1}_{[g]}$ is part of the layers **L7** and **L4**. In fact all four 2D layers can be considered as 2D assemblies of the 1D HB chains $\mathbf{C1}_{[2_1]}$ and $\mathbf{C1}_{[g]}$ and will be discussed as such. It should be noted that **L4** is also present in the meta series.

4.B.3.2.a Layered packing of $\mathbf{C1}_{[2_1]}$

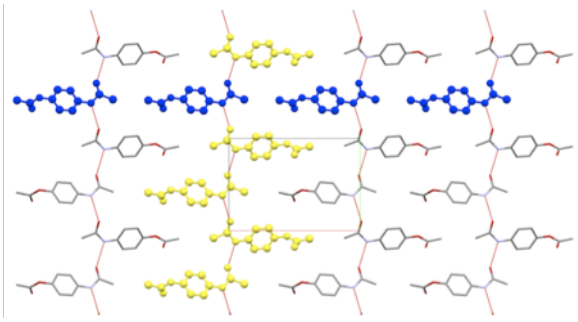
The layers **L5** and **L6** present two alternatives for the close packing of the HB chain $\mathbf{C1}_{[2_1]}$. On one hand the chains assemble along their depth (Figure 4.57) whereby neighbouring chains are related by glide symmetry in **L5**.



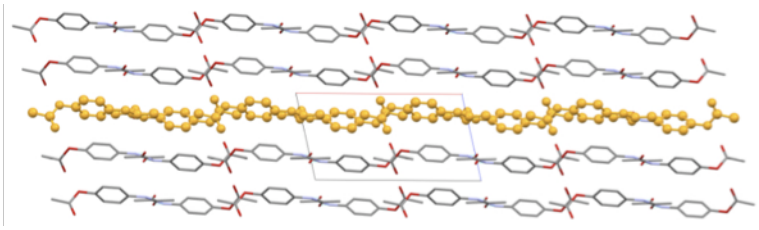
Me:H **Me:p-Me(m)**

Figure 4.57: The layer **L5** in **Me:H** and **Me:p-Me(m)**. Layer **L6** is also shown in **Me:p-Me(m)**.

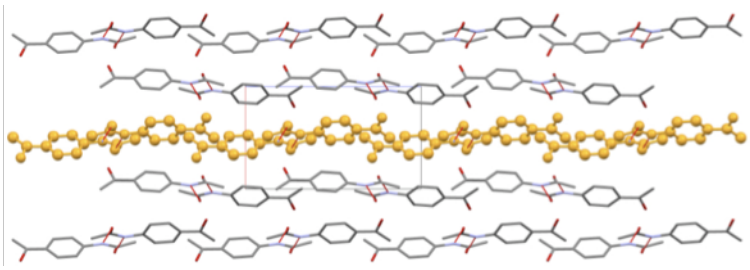
On the other hand layer **L6** is constructed by combining $\mathbf{C1}_{[2_1]}$ with the row $\mathbf{R3}_{[t]}$. This means the chains are translated along their width as displayed in Figure 4.58.



$\mathbf{C1}_{[2_1]}$ & $\mathbf{R3}_{[t]}$



Me:p-OCOMe(m2)



Me:p-COMe

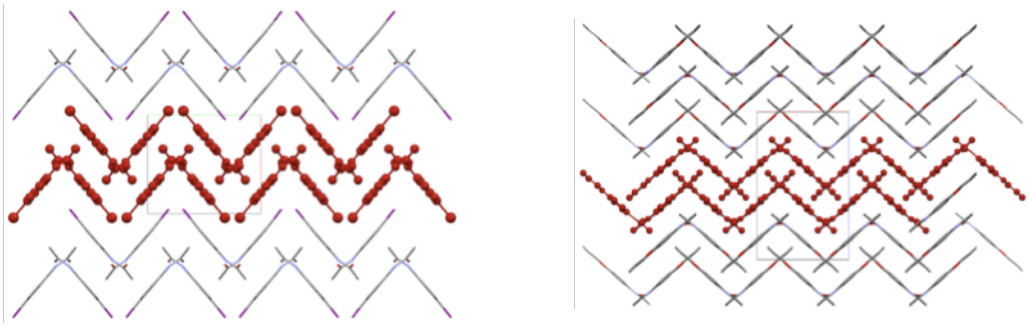
Figure 4.58: The composition of layer **L6** (top) and **L6** in the structures of **Me:p-OCOMe(m2)** and **Me:p-COMe**.

Each layer occurs in seven crystal structures. In the case of **L5** these seven structures can be classified into two groups of packing analogues. These essentially differ in the assembly of **L5** into the 3D crystal structures. In **Me:H** and its analogues neighbouring layers are related by glide symmetry whilst in the structures of the **Me:p-Me(m)** type the layer is simply translated in the third dimension. The crystal structures containing **L6** on the other hand resemble three individual structure types. **Me:p-Me(m)** and its packing analogues belong to one of them as well. In fact both layers **L5** and **L6** coexist in **Me:p-Me(m)** and its crystal structure analogues producing the overall 3D assembly (*c.f.* Figure 4.57). In

Me:p-COMe the layers **L6** are stacked with 2_1 symmetry, whilst in the crystal structures of **Me:p-^{t}Bu** adjacent layers **L6** are related by inversion symmetry.
Interestingly the layers are observed in a wide variety of para acetanilides with the substituents ranging from hydrogen in the unsubstituted acetanilide parent compound to acetoxy in **Me:p-OCOMe(m2)**.

4.2.3.2.b The layers L4 and L7

The assembly of $\mathbf{C1_{[g]}}$ is comparable to that of $\mathbf{C1_{[2_1]}}$. In **L4** the wedge-shaped chains $\mathbf{C1_{[g]}}$ are arranged with 2_1 screw symmetry between adjacent chains forming a corrugated layer as shown in Figure 4.59.



Me:p-I **Me:p-OPr**

Figure 4.59: The layer **L4** in **Me:p-I** and **Me:p-OPr**.

The layer **L7** is a combination of $\mathbf{C1_{[g]}}$ and $\mathbf{S1_{[t]}}$. The 2D feature is depicted in Figure 4.60. In this way the wedges are stacked on top of each other. **L4** was only detected in three para acetanilide structures, whereas **L7** was found in eight crystal structures. The packing of layer **L4** is achieved through translation in **Me:p-I** and **Me:p-^{c}Pr**, whereas adjacent layers **L4** are related by 2_1 screw symmetry in **Me:p-OPr**.
The assembly of layer **L7** in a 3D crystal structure is afforded through glide symmetry in **Me:p-F** and its packing analogues. In the structures of the type of **Me:p-Cl** neighbouring layers **L7** are related by 2_1 screw symmetry and in **Me:p-NO_2** the layers pack with inversion symmetry.

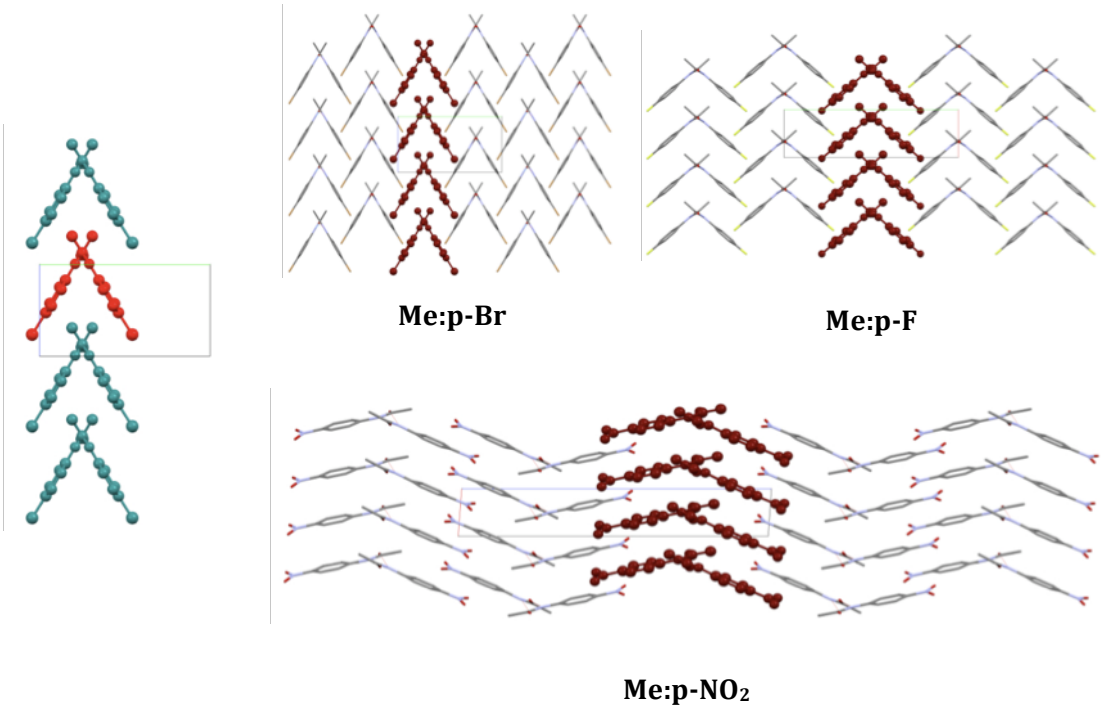


Figure 4.60: The composition of **L7**: **C1**$_{[g]}$ and **S1**$_{[t]}$ (far left). The layer **L7** in **Me:p-Br** (top left), **Me:p-F** (top right) and **Me:p-NO$_2$** (bottom).

As for **L5** and **L6** the layers **L4** and **L7** exist in a variety of para acetanilide structures. However, as can be seen from the similarity plot the halo acetanilides and generally those compounds with smaller substituents seem to preferentially assemble in layers **L4** and **L7** than in the layers **L5** and **L6**.

4.B.3.2.c The net N1

The final 2D SC identified with XPac is the net N1. This feature was found in **Me:p-OH(o1)**, **Me:p-NH$_2$** and **Me:p-OH(o2)**. The nets are slightly corrugated and are constructed through hydrogen bonding along the length and width of the acetanilide molecule. The nets are depicted in Figure 4.61 for the three structures. Within the net every molecule is involved in four hydrogen bonds with neighbouring molecules as can be seen from Figure 4.61. Both the amide group and the substituent act as HB donor as well as acceptor. In **Me:p-NH$_2$** this means that one of the amino hydrogen atoms is not involved in the HB within the net.

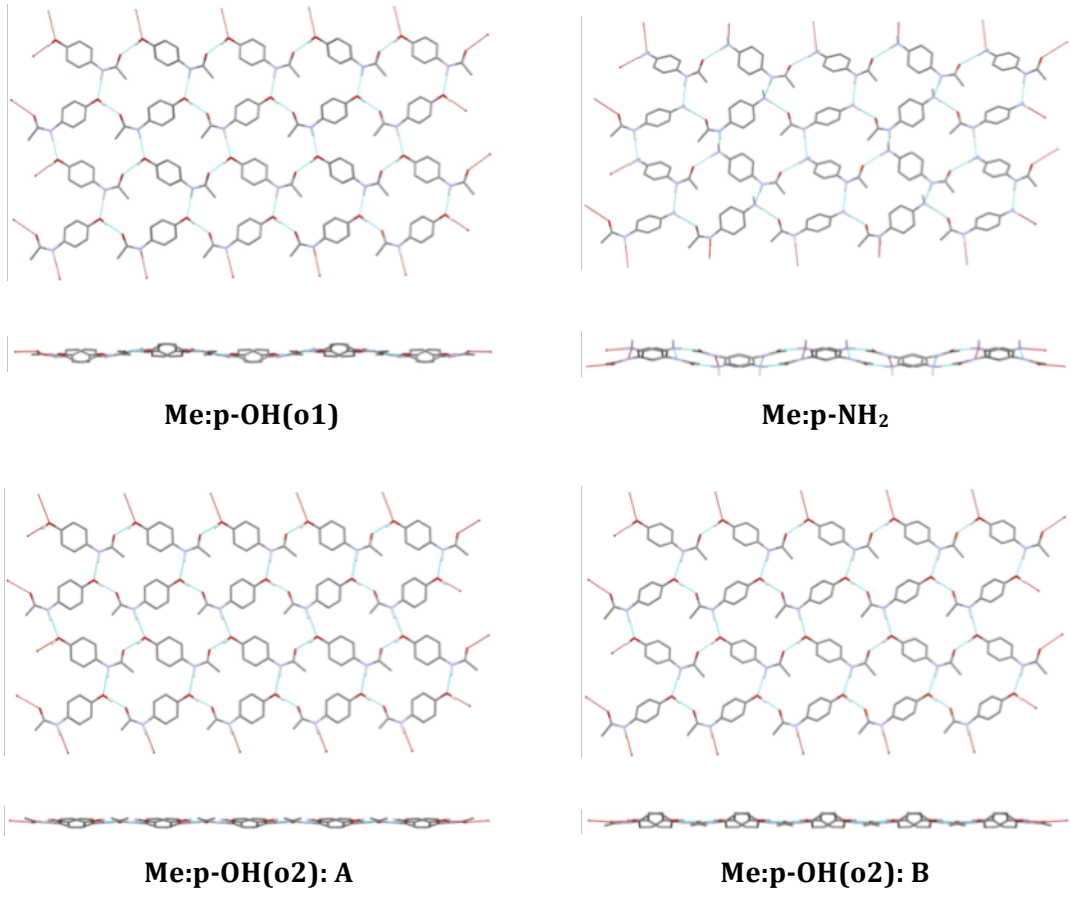
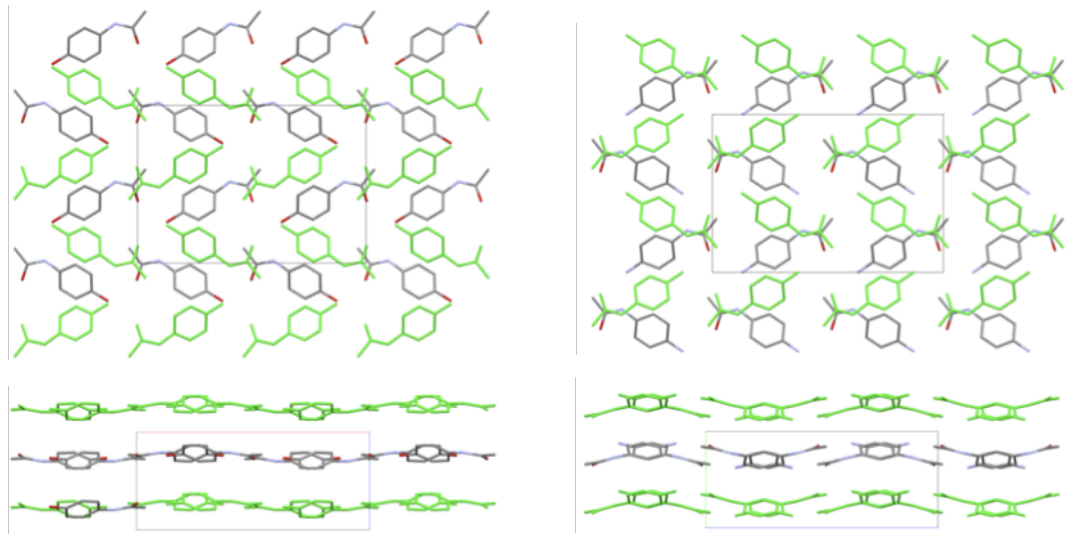


Figure 4.61: The HB net **N1** in **Me:p-OH(o1)**, **Me:p-NH$_2$** and **Me:p-OH(o2)**. Viewed from top and side-on. **Me:p-OH(o2): A/B** refers to the two independent molecules A and B in the ASU.

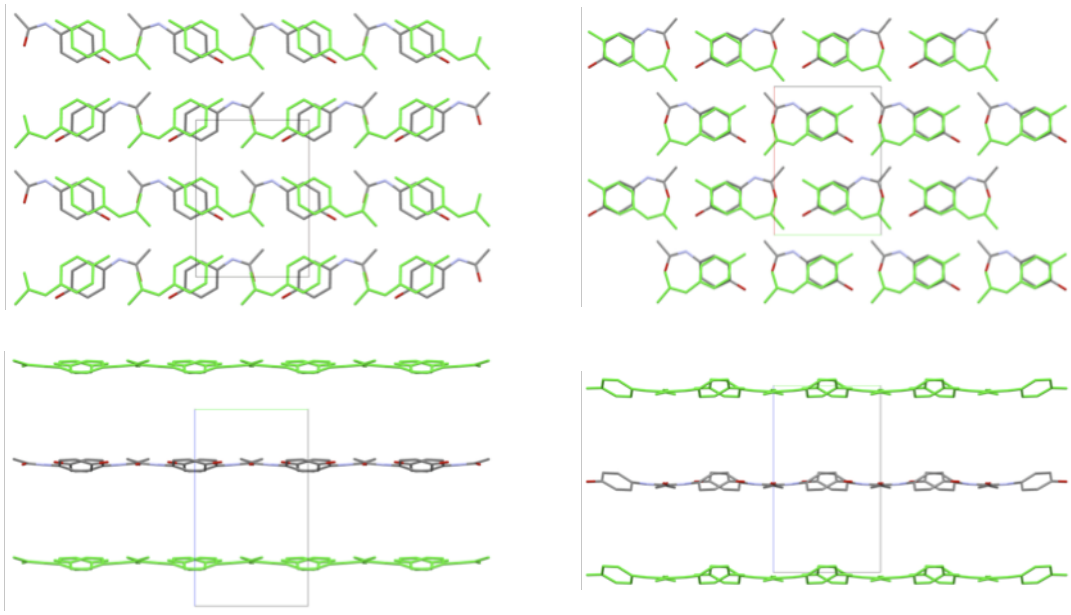
However this atom engages in a long hydrogen bond of approximately 3.28Å with the amide oxygen in an adjacent net. Hence all HB donors are saturated. Since the lone pair of the amino nitrogen lies outside the plane of the amino hydrogen atoms, the corrugation of **N1** is strongest in **Me:p-NH_2.** It should be noted that the structure of the meta-stable **Me:p-OH(o2)** contains two independent molecules in the ASU and each component of the ASU assembles in a separate net **N1**.

The hydrogen bonding along the width of the molecules has glide symmetry in all three structures. Neighbouring molecules in the head-to-tail HB chains however, are related by glide symmetry in **Me:p-OH(o1)** and **Me:p-NH** and by translation in **Me:p-OH(o2)**, respectively. This difference arises since the orientation of the molecules with respect to the unit cell axis is different in **Me:p-OH(o2)** compared with the other two structures. Hence one of the **N1** base vectors equates to a double unit cell axis in **Me:p-OH(o2)**. The three structures differ in the packing of the nets as demonstrated in Figure 4.62 overleaf.

The reference net is shown in green and layers related by translation are drawn in the same colour. In **Me:p-OH(o1)** and **Me:p-NH_2** neighbouring nets are related via glide planes perpendicular to the net plane. The hydrogen bonding between adjacent nets in **Me:p-NH_2** results in a shift of the nets with respect to each other differing from that in **Me:p-OH(o1)**. Stacking of **N1** in **Me:p-OH(o2)** is achieved by an increase in Z' so that pair of nets are generated.



Me:p-OH(o1) **Me:p-NH_2**



Me:p-OH(o2): A **Me:p-OH(o2): B**

Figure 4.62: Packing of HB net **N1** in **Me:p-OH(o1)**, **Me:p-NH_2** and **Me:p-OH(o2)**. Viewed from top and side-on. Layers related by translation are shown in same colour. **Me:p-OH(o2): A/B** refers to the two independent molecules A and B in the ASU.

Adjacent pairs of this kind are then related by glide symmetry. The arrangement of the pairs of nets due to Z' = 2 is illustrated in Figure 4.63.

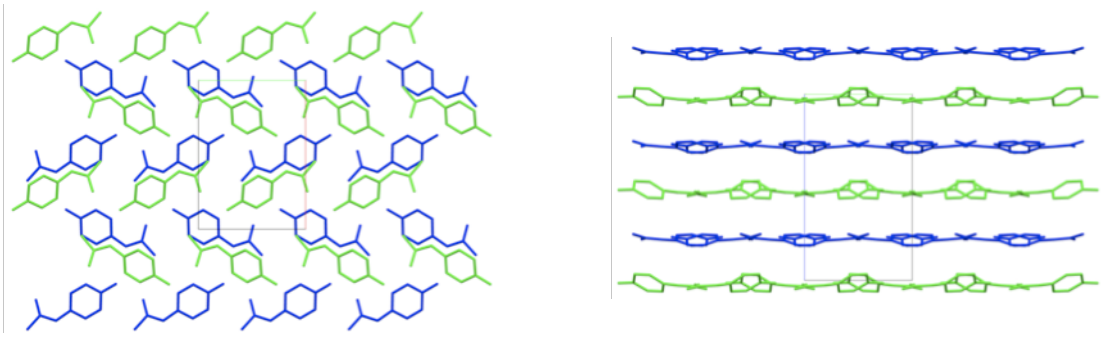


Figure 4.63: Packing of **N1** due to Z' = 2 in **Me:p-OH(o2)**. The independent components of the ASU are coloured green and blue, respectively.

4.B.3.3. Packing in 3D

In the para acetanilide library there are six groups of crystal structure packing analogues. Two groups comprise two structures: **Me:p-^{t}Bu** and **Me:p-OCOMe(m2)** form group 1, whereas **Me:p-I** and **Me:p-^{c}Pr** belong to group 2. Groups 3 and 4 contain three crystal structures each, namely **Me:H**, **Me:p-OMe**, **Me:p-O^{i}Pr** and **Me:p-Cl**, **Me:p-Br(o)**, **Me:p-Me(o)**, respectively. The four structures **Me:p-F**, **Me:p-Et**, **Me:p-CN** and **Me:p-OCF$_3$** form group 5 and group 6 contains the five structures of **Me:p-Me(m)**, **Me:p-^{i}Pr**, **Me:p-OEt**, **Me:p-OBu** and **Me:p-CBrCH$_2$**. The corresponding unit cell parameters are listed in Table 4.22. From the table is can be seen that the unit cells are very similar for groups 2 and 4, whereas in the other groups at least one unit cell parameter deviates from similarity quite significantly. This can be rationalised when considering the substituents in the respective structures. The substituents in groups 2 and 4 are comparable in size whilst in the other groups very different substituents are found both in terms of their shape and size.

For example the acetanilides with **O^{i}Pr** and **OMe** crystallise in the same crystal structure as the unsubstituted parent compound **Me:H** and the longest cell axis varies by approximately 12 Å. In this case 3D similarity was established through combination, since 12 Å exceeds the XPac detection limits. The XPac analysis confirmed packing analogy in **Me:H** and **Me:p-OMe** as well as in **Me:p-O^{i}Pr** and **Me:p-OMe**, but only detected 2D similarity between **Me:H** and **Me:p-O^{i}Pr**. However, from the combination of these results it must follow that **Me:H** and **Me:p-O^{i}Pr** are also packing analogues, which was confirmed by visual inspection of the crystal structures. The unit cell packing diagrams in Figure 5.64 reveal the origin of the large difference in the unit cell dimension.

The distance between neighbouring stacks of the HB chain $\mathbf{C1_{[2_1]}}$ is increased in **Me:p-O^{i}Pr** to accommodate the *iso*-propoxy substituent.

The unit cell parameter variation is not as pronounced in groups 1, 5 and 6, which can be seen as the result of the different substituent sizes and shapes.

Structure	a [Å]	b [Å]	c [Å]	β [˚]	Space Group
Group 1					
Me:p-^{t}Bu	16.417(1)	9.616(1)	7.199(1)	97.32(1)	$P2_1/c$
Me:p-OCOMe(m2)	13.877(1)	9.577(1)	7.442(1)	103.02(1)	
Group 2					
Me:p-I	9.578(1)	10.305(1)	9.426(1)	109.29(1)	$P2_1/c$
Me:p-^{c}Pr	9.557(4)	10.980(3)	9.641(2)	106.34(3)	
Group 3					
Me:H	19.509(11)	9.364(8)	7.778(10)	90	Pbca
Me:p-OMe	9.117(1)	7.495(1)	24.663(1)	90	
Me:p-O^{i}Pr	9.301(2)	7.649(2)	31.394(6)	90	
Group 4					
Me:p-Cl	9.698(1)	12.387(1)	6.576(1)	90	$Pna2_1$
Me:p-Br(o)	9.773(1)	12.504(1)	6.721(1)	90	
Me:p-Me(o)	9.899(5)	12.956(6)	6.541(4)	90	
Group 5					
Me:p-F	4.614(1)	17.042(1)	9.562(1)	93.38(1)	Cc
Me:p-Et	4.585(1)	21.395(1)	9.431(1)	90.00(1)	
Me:p-CN	3.846(1)	22.190(5)	9.448(2)	90.05(2)	
Me:p-OCF_3	4.488(1)	22.168(3)	9.528(1)	90.05(1)	
Group 6					
Me:p-Me(m)	11.678(1)	9.490(1)	7.419(1)	106.52(1)	$P2_1/c$
Me:p-^{i}Pr	14.005(1)	9.349(1)	7.810(1)	90.71(1)	
Me:p-OEt	13.243(1)	9.586(1)	7.478(1)	103.35(1)	
Me:p-OBu	10.516(1)	9.803(1)	10.903(1)	95.69(1)	
Me:p-$CBrCH_2$	11.151(6)	9.237(5)	10.142(5)	93.39(1)	

Table 4.22: Unit cell parameters for packing analogues in the para acetanilide series.

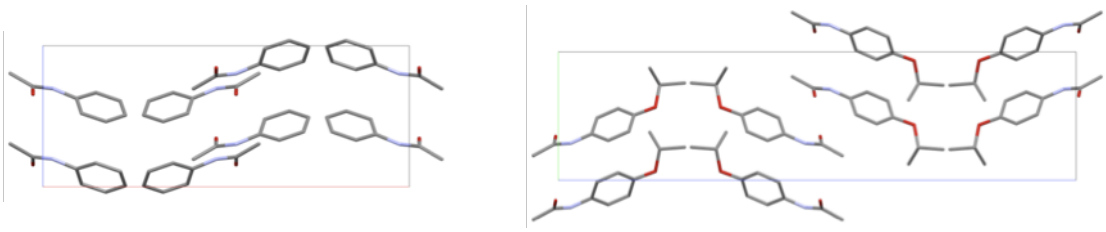


Figure 4.64: Unit cell content if **Me:H** (left) and **Me:p-O^{i}Pr** (right) as viewed along the HB direction.

It should be noted that none of the groups of packing analogues contain a para acetanilide with a HB donor substituent. Also not many compounds with substituents with HB acceptor functionality are found in the groups. Instead the packing analogues mainly contain alkyl-, halo- and alkoxy-substituents, showing that the variety of crystal structures is to some extend dependent on the intermolecular hydrogen bond network that can be established. Of course unique crystal structures can also be found for acetanilides with substituents without HB functionality as is the case for **Me:p-Br(m)** and **Me:p-CF$_3$**, but compared with the number of structures crystallising as packing analogues such occurrences are rare. Overall it is remarkable that 19 para acetanilides crystallise in only 6 unique crystal structures, especially given the variety of substituents – who would expect bromoethylene is comparable with a methyl group! The underlying principle however is that the packing of the para acetanilides occurs in a certain robust way that goes unmodified by a large number of quite different substituents. By seeing the crystal structure as an assembly of the lower dimensionality SCs rather than individual molecules, it can be understood that the possibilities of close packing of these features is more limited than the assembly of separate molecules. Most para acetanilides contain one of the HB chains **C1**, for which packing in two dimensions is restricted to two to three options. As these layers then assemble in a 3D crystal structure, again the number of possibilities for close packing is limited to a maximum of three options. Hence it appears that for the para acetanilides with substituents without HB donor functionality it is not so much the kind of substituent that determines the overall crystal structure but rather the possibilities of closely packing the 1D building blocks formed through amide-amide hydrogen bonds.

4.B.4. The influence of the position of substitution

The final XPac comparison of all 64 acetanilide crystal structures was aimed at elucidating what influence the position of substitution **s** has on the packing behaviour of these types of compounds. The findings from this systematic search are summarised in the structural similarity plot in Figure 4.65 overleaf.

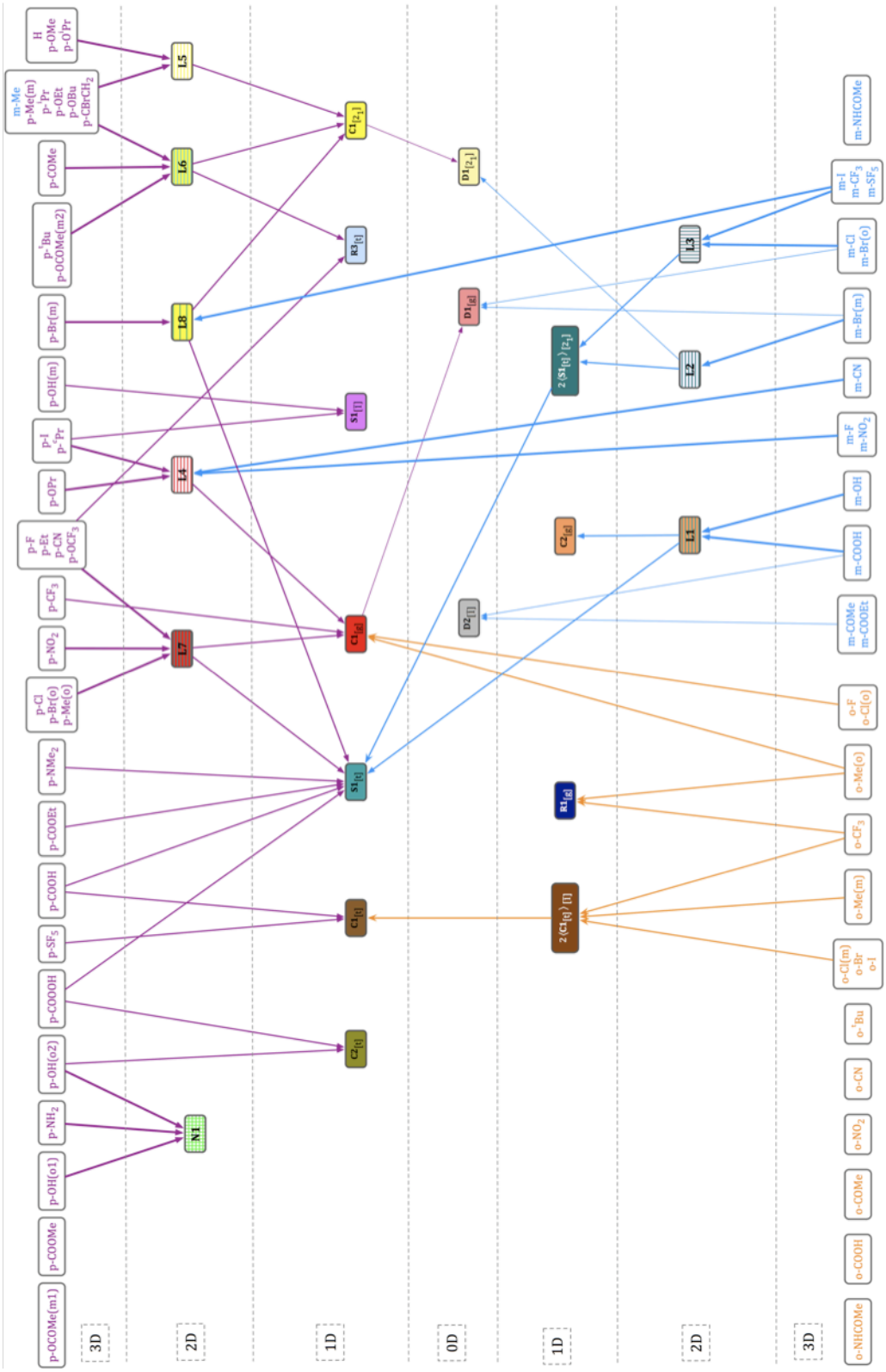


Figure 4.65: Structural similarity plot for the full set of mono-substituted acetanilides. Crystal structure entries and arrows are colour encoded according to position of substitution: ortho (orange), meta (aquamarine), para (purple).

In this plot the similarity diagrams of all three series are combined. For clarity one SC was omitted: **R2$_{[t]}$** present in the meta series was not included in the plot (*c.f.* Figure 4.22). The colour coding of the arrows in the structural similarity plot in Figure 4.65 was introduced so that constructs common to structures in different series can be identified at a glance. To derive the SCs a structure contains, the arrows should be traced from tail to head. If arriving at a SC box, which only has arrowheads aimed at it, then this forms the terminal point of this particular path.

In total XPac identified 43 unique crystal structure assemblies or types amongst which are 12 groups of packing analogues. Approximately one third of the acetanilides crystal structures belong to one of these twelve groups. But between the ortho-, meta- and para-series 3D similarity was only observed once: **Me:m-Me** was found to crystallise in the same structure type as **Me:p-Me(m)**.

Furthermore the XPac analysis reveals that the similarity between the ortho-, meta- and para-series is largely restricted to the HB chains **C1$_{[t]}$**, **C1$_{[g]}$** and **C1$_{[2_1]}$**, which occur in a total of 41 acetanilide structures. Indeed the ortho series has no further SCs in common with any of the structures in the meta and para acetanilides. The amide-amide hydrogen bonding certainly plays a major role in the assembly of all acetanilides irrespective of the position of substitution. Including the dimers of type **D1** and the helical HB chain in **Me:o-^{t}Bu** a total of 46 acetanilide structures – more than two thirds of the whole family – are based on amide-amide hydrogen bonding. The remaining 18 structures correspond to acetanilide modifications in which the substituent **X** has HB donor or acceptor functionality.

A related construct occurs in some para- and meta-structures; the layer **L4** is present in **Me:p-I**, **Me:p-^{c}Pr**. **Me:p-CN** as well as in **Me:m-CN**, **Me:m-F** and **Me:m-NO$_2$**, respectively. None of the other 2D constructs was identified outside the individual series. Instead a new 2D SC was found by XPac, which is constructed by combining chain **C1$_{[2_1]}$** with stack **S1$_{[t]}$** as shown in Figure 4.66. This layer **L8** was detected in **Me:p-Br(m)** and **Me:m-CF$_3$** and its packing analogues.

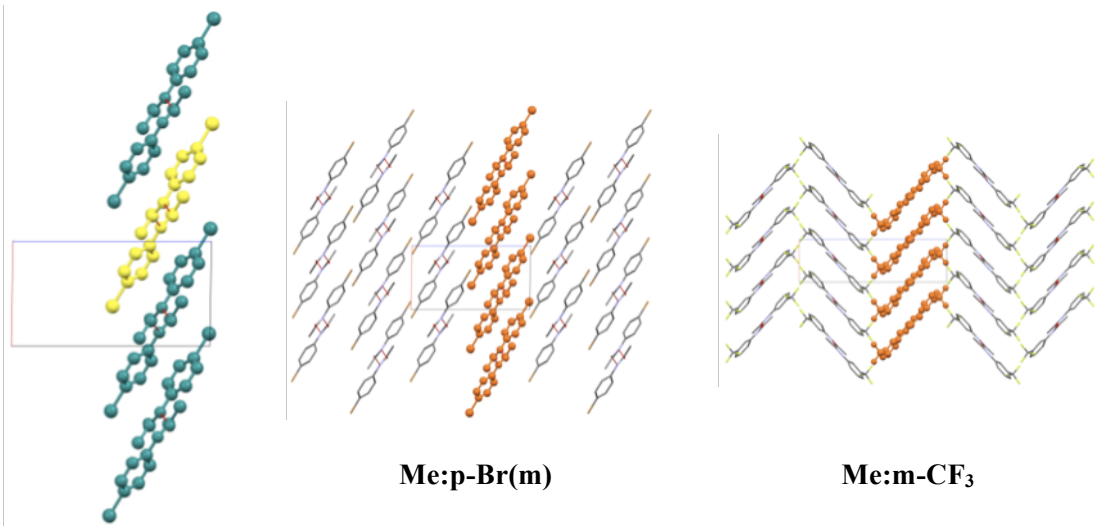


Figure 4.66: The layer **L8** composed of chain $\mathbf{C1}_{[2_1]}$ and stack $\mathbf{S1}_{[t]}$ (far left). The layer in **Me:p-Br(m)** and **Me:m-CF_3**.

The most prominent and only other single SC observed in both meta and para acetanilides is the translational stack $\mathbf{S1}_{[t]}$ occurring in 21 structures in total. In the preceding sections the stack was analysed in terms of the pairwise interaction energy and it was shown that this molecular aggregate is mainly based on weak intermolecular interactions. However, as supported by the frequency of occurrence this particular arrangement of molecules must present a preferred geometry for the assembly of the acetanilides. It was also discussed in previous sections that the major energy gain arises from intermolecular hydrogen bonding especially with respect to the amide-amide HB chains of type **C1**. Hence it can be concluded that the primary driving force for the aggregation of acetanilide molecules in the crystal is that of involving the amide group in hydrogen bonding. The close packing of the phenyl ring in a stacked system is preferred but this packing motif is secondary to the hydrogen bonding and in many cases the stack $\mathbf{S1}_{[t]}$ seems incompatible with the amide-amide HB chains as the similarity plot (Figure 4.65) shows.

Overall the establishment of the amide-amide HB chains seems to be little dependent on the position of substitution on the phenyl ring, although the crystal structures of the ortho series are affected by any intramolecular interactions between the amide group and the substituent **X**. The stacking of the acetanilides along their depth only seems to prevail in the crystal structures of the meta- and

para-substituted compounds. However, it should be noted that the dihedral angle between the mean planes of the phenyl ring and amide group is very large in the ortho-acetanilide structures containing **C1**$_{[t]}$. Hence neighbouring phenyl rings are already 'stacked' within the chain excluding the possibility for a purely translational stack such as **S1**$_{[t]}$.

4.C. References

[1] Chris Frampton was the industrial supervisor at Pharmorphix Ltd.

[2] G. Perlovich, T. Volkova, A. Bauer-Brandl, *Journal of Thermal Analysis and Calorimetry,* **89**(3), 767-774 (2007).

[3] T. Beyer, G. Day, S. Price, *Journal of the American Chemical Society,* **123**(21), 5086-5094 (2001)

[4] M.-A. Perrin, M. A. Neumann, H. Elmaleh, L. Zaske, *ChemComm*, 3181-3183 (2009).

[5] C. F. Macrae, I. J. Bruno, J. A. Chisholm, P. R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. van de Streek, P. A. Wood, ***"Mercury CSD 2.0 - New Features for the Visualization and Investigation of Crystal Structures"*** in *Journal of Applied Crystallography*, **41**, 466-470 (2008).

[6] D. A. Fletcher, R. F. McMeeking, D. Parkin, ***"The United Kingdom Chemical Database Service"*** in *Journal of Chemical Information and Computing Science*, **36**, 746-749 (1996).

[7] A. Gavezzotti, ***"Molecular Aggregation"***, Oxford University Press, Oxford, UK (2007).

[8] Taking into account the different radii of the elements.

[9] The structures of **Me:m-Cl**, **Me:m-Br(o)**, **Me:m-SF_5** and **Me:m-NO_2** can be found under the CSD refcodes GISPOO, HOTDOK, OBEJUA and UGUHEJ, respectively. For the corresponding references *c. f.* Appendix 2-2.

[10] Note that the interaction energies could not be calculated for Z' > 1 structures, **Me:p-COOOH** and **Me:s-I** as mentioned in Chapter 3.

[11] D. Bhalla, J. Lalla, *Drug Development and Industrial Pharmacy,* **16**(1), 115-135 (1990).

[12] M.-A. Perrin, M. Neumann, H. Elmalch, L. Zaske, *ChemCommun*, 3181-3183 (2009).

[13] a) In a set of 25 carbamazepine compounds 9 1D and 6 2D constructs were found as described in T. Gelbrich, M. B. Hursthouse, *CrystEngComm*, **8**, 448-460 (2006); b) In a series of 37 sulfonamides XPac identified 12 1D and 8 2D common packing features as presented by T. Gelbrich in a poster at the Spring Meeting of the British Crystallographic Association (2007); The analysis of 50 disubtsituted

chalcones resulted in 14 1D and 13 2D SCs as summarised in G. Tizzard, *"Charaterisation and Investigation of Structural Relationships of 4, 4'-disubstituted Chalcones"* (2008).

[14] E. Pidcock, W. S. Motherwell, *Crystal Growth & Design*, **4**(3), 611-620 (2004).

[15] A. Gavezzotti, ***"Molecular Aggregation"***, Oxford University Press, Oxford, UK (2007).

CHAPTER 5: CONFORMATIONAL ANALYSIS OF MONO-SUBSTITUTED ACETANILIDES

5.A. Introduction

In the previous chapter the variations in the crystal structures of 64 mono-substituted acetanilides **Me:s-X** were analysed. In doing so the influence of the type of substituent **X** and the position of substitution **s** on intermolecular interactions and the crystal assembly was probed. In this chapter the molecular structure is examined more closely; in particular the influence of **X** and **s** on the conformational preferences of the molecule in the solid state is investigated.

The acetanilide shape is mainly defined by the core of the molecule. The aromatic system and the amide group are planar moieties and together they form a conjugated system. Hence in principle the acetanilide core is expected to assume a flat shape, with the depth of the molecule being much smaller than its width and length. However, rotation about the bond linking the amide group and the phenyl ring is possible and depends on the bond strength, which can vary due to conjugation. This bond C_{ph}–N_{am} is shown in Figure 5.1 along with the torsion angle α.

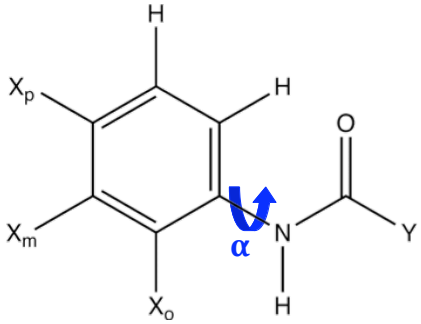


Figure 5.1: The torsion angle alpha.

Further variation to this basic shape is introduced with the substituent **X** at position **s**.

When the compound library was initially designed the following was assumed:

a) due to conjugation the acetanilides should be generally flat,
b) most of the substituents **X** are small and rotationally symmetric or pseudosymmetric so that the depth of the molecule should be least affected by the introduction of **X**,
c) the larger and rotationally asymmetric substituents **X** are expected to coincide with the phenyl ring plane due to conjugation with the π-system,
d) α should be unaffected by para- and meta-substitution as the substituents are too far removed from the amide group and the respective molecule should thus remain planar,
e) bulky ortho-substituents should cause a twist about the C_{ph}–N_{am} bond because of steric hindrance.

As for the first point, it was seen in Chapter 4 that most of the acetanilides do not assume a planar configuration in the crystal structure at all. For some acetanilides conformational polymorphism was observed where α appeared to be the only molecular variable. Furthermore some correlation between α and the three amide-amide hydrogen bonding motifs was noticed for the para acetanilides. These observations prompted the further study of the molecular conformation of the acetanilides with respect to the substitution pattern. The conformational freedom of the acetanilides was thus assessed by computation of the potential energy surface (PES) for the rotation about α in 10° increments. These calculations were preformed in the gas phase as outlined in Chapter 3. The PES was computed for a complete, i.e. 360° rotation about α, to probe for inequality of the two sides of the molecule with respect to the phenyl ring plane rather than assuming the molecule to be symmetrical with respect to the phenyl ring plane. The values calculated in this way are all listed in Appendix 5-1. In the subsequent sections the results of the theoretical conformational analysis (CA) are presented and discussed. The PES of the acetanilides **Me:s-NHCOMe** will not be considered in the discussion since the distinction between substituent and core moiety is not possible in this molecule resulting in two indistinguishable torsion angles. This is demonstrated in Figure 5.2 using **Me:m-NHCOMe** as example. The PES

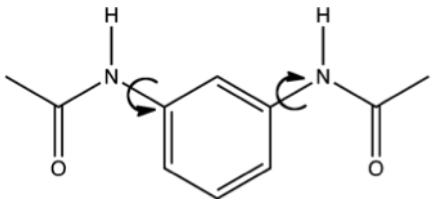


Figure 5.2: The two equivalent torsions in **Me:m-NHCOMe**.

would hence have to be calculated as a 2-dimensional map with both C_{ph}-N_{am} bonds as variables, but such a map is then no longer directly comparable to those compounds with only one rotatable C_{ph}-N_{am} bond.

5.B. Conformational analysis of ortho-substituted acetanilides

Maps of the PES estimated for the ortho-substituted acetanilides are shown in Figures 5.3 to 5.6. Figure 5.3 displays the PES maps for the halogenated ortho acetanilides whilst Figure 5.4 shows the maps for **X = Me**, **CF_3** and **^{t}Bu**, i.e. for bulkier and HB inactive groups. Figure 5.5 depicts the plots of the PES for the substituents with HB acceptor functionality, i.e. **X = CN**, **NO_2** and **COMe**, and Figure 5.6 contains the PES for the substituent with HB donor functionality, i.e. **X = COOH.**

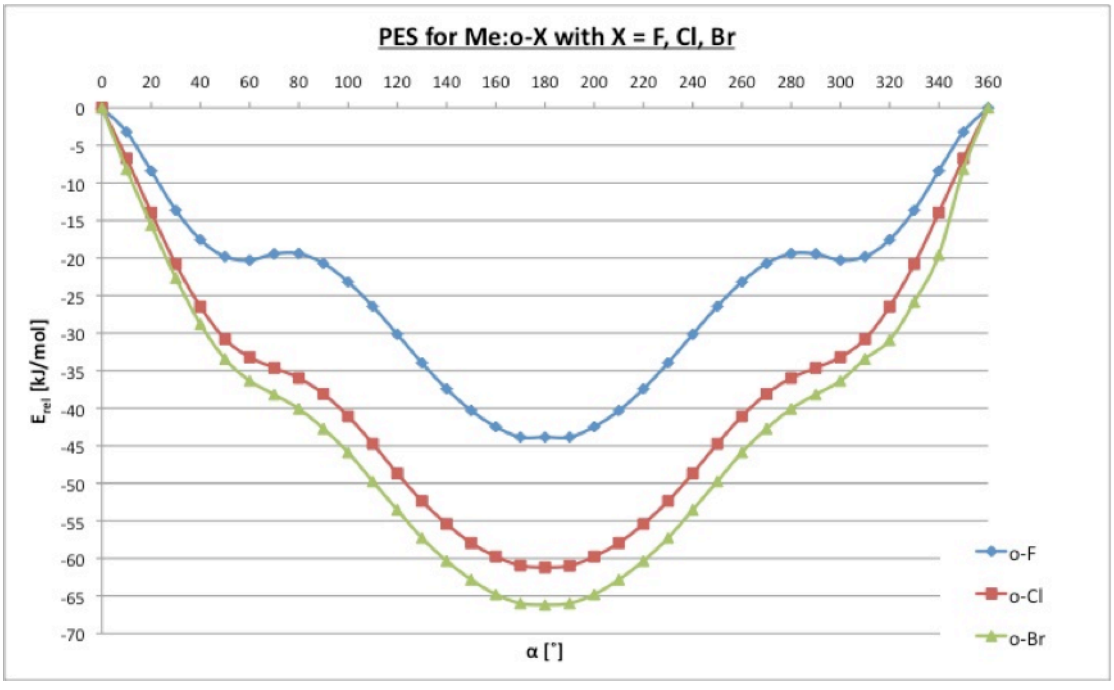


Figure 5.3: Potential energy surfaces for the rotation about the torsion angle α for **Me:o-F**, **Me:o-Cl**, **Me:o-Br**.

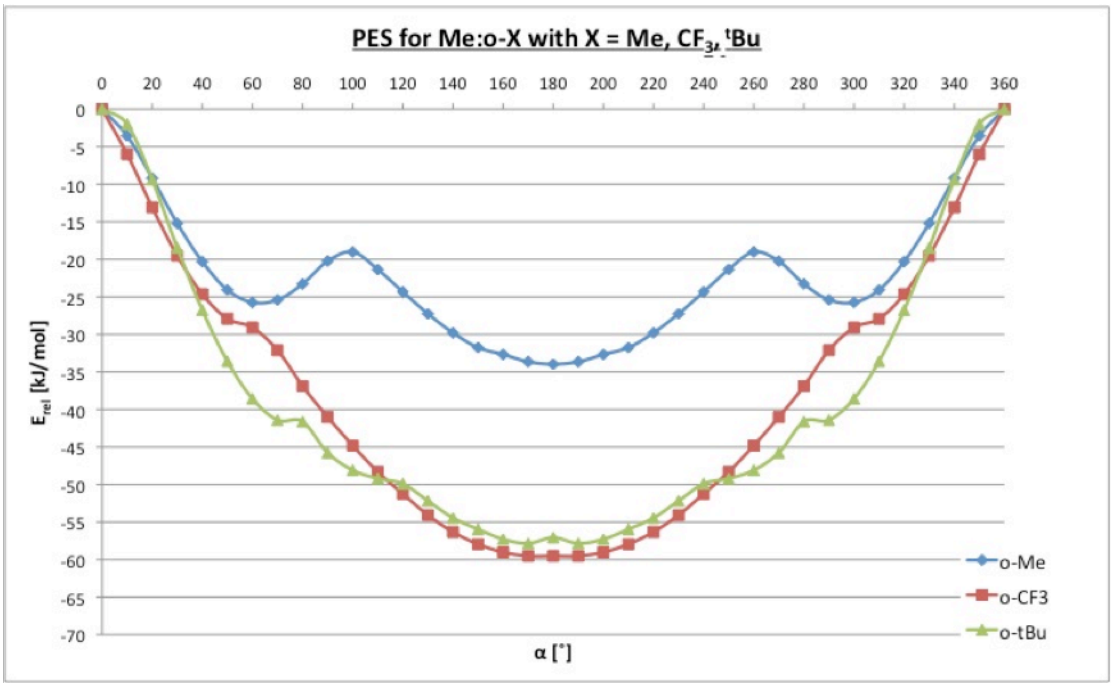


Figure 5.4: Potential energy surfaces for the rotation about the torsion angle α for **Me:o-Me**, **Me:o-CF_3**, **Me:o-^{t}Bu**.

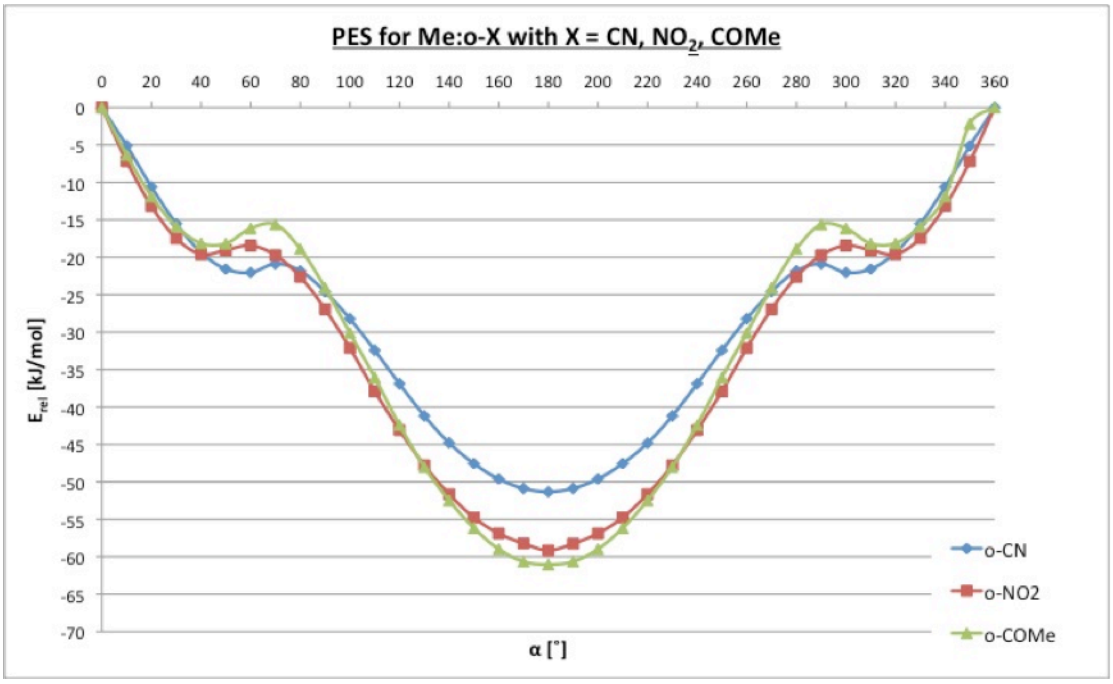


Figure 5.5: Potential energy surfaces for the rotation about the torsion angle α for **Me:o-CN**, **Me:o-NO_2**, **Me:o-COMe**.

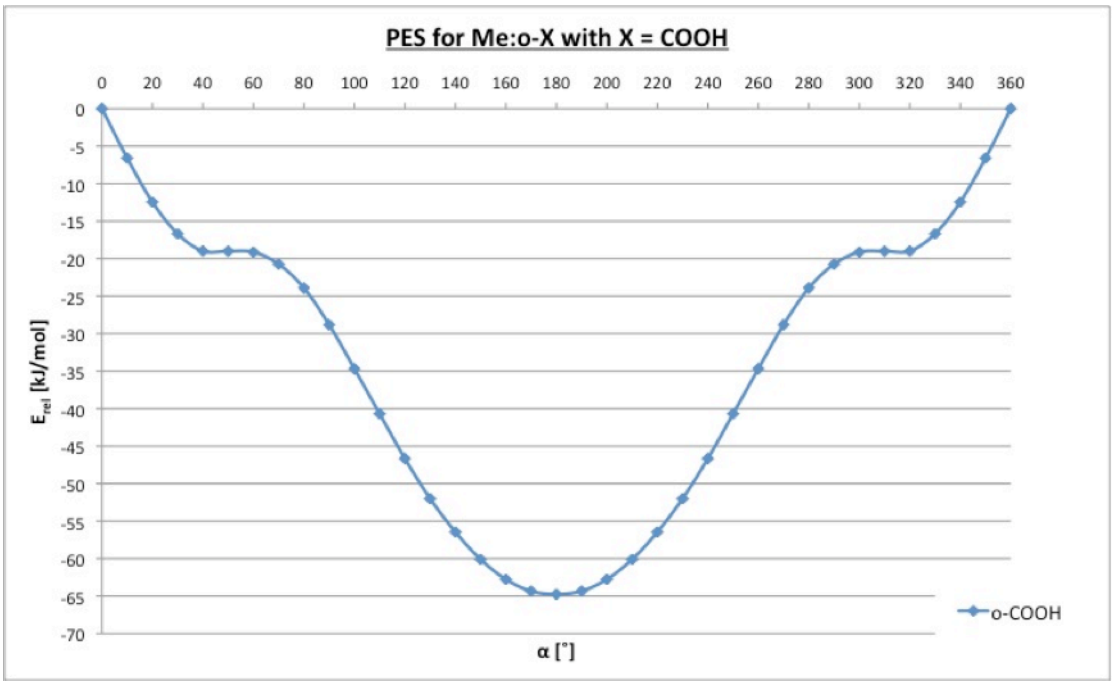


Figure 5.6: Potential energy surface for the rotation about the torsion angle α for **Me:o-COOH**.

The maps demonstrate that the PES is symmetrical with respect to α = 180° and the conclusion can thus be drawn that both sides of the phenyl ring plane and hence the molecule are equivalent. Hence only the area 0 ≤ α ≤ 180° will be further considered.

The shape of the PES is quite similar for all ortho-substituted acetanilides. The global minimum occurs at a torsion angle of α = 180°. This conformation corresponds to a fully planar molecule where the amide oxygen is positioned *anti* to the ortho-substituent. The calculations also show that the conformation of α = 0°, i.e. the planar *syn* arrangement of **X** and amide oxygen, is energetically least favourable for the ortho acetanilides. These two different configurations are displayed in Figure 5.7.

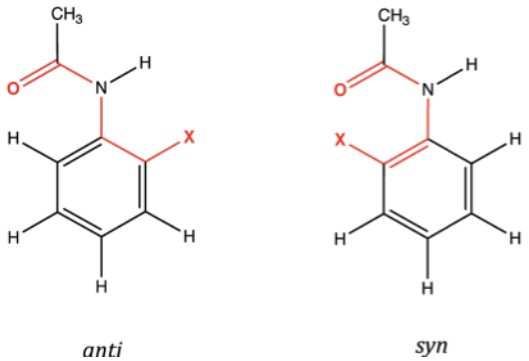


Figure 5.7: The *anti* and *syn* configurations as obtained by 180° rotation about α.

In addition to the global minimum at α = 180° the PES exhibits shoulders or even small local minima in the regions 40° < α < 100° and 260° < α < 320° although these are energetically far removed from the global minima.

The relative energies at the global minimum are listed in Table 5.1. These energies vary from -66.175 kJ·mol^{-1} to -33.976 kJ·mol^{-1}.

Molecule	α_{theo} [°]	E_{theo} [kJ·mol^{-1}]
Me:o-F	180	-43.853
Me:o-Cl	180	-61.185
Me:o-Br	180	-66.175
Me:o-CF$_3$	170	-59.515
Me:o-Me	180 (60*)	-33.976 (-25.745*)
Me:o-^{t}Bu	170	-57.844
Me:o-CN	180	-51.324
Me:o-NO$_2$	180	-59.145
Me:o-COMe	180	-61.027
Me:o-COOH	180	-64.746

Table 5.1: Torsion angles α_{theo} and energies E_{theo} at the global minimum are listed for the ortho series. *Local minimum values.

The PES of **Me:o-Me** and **Me:o-F** exhibit the shallowest minima whilst the minima of **Me:o-Br** and **Me:o-COOH** are the deepest of the ortho acetanilides. In fact the minimum energies are similar to within 10 kJ·mol^{-1} for all substituents except for **X = F**, **Me** and **CN**. The geometry of the substituent **CN** is unfavourable for intramolecular hydrogen bonding compared with the other substituents with HB acceptor functionality explaining the higher (less stable) minimum energy. The substituents **Me** and **F** on the other hand are not regarded as HB partners and hence their energies are the least stable. The size of the methyl group may impose steric hindrance hence resulting in the lowest energy of all ortho acetanilides. Interestingly both **Cl** and **Br** are stabilising substituents indicating the existence of a halogen bond as detailed in Chapter 1.A.3. This concept if supported since the conformational energies of **Me:o-Cl** and **Me:o-Br** are very similar to the energies of the molecules with classical strong intramolecular hydrogen bonds such as **Me:o-COOH**, **Me:o-COMe** and **Me:o-NO$_2$**. Furthermore the global minimum energy of **Me:o-CF$_3$** also falls into this range and is comparable to that of **Me:o-NO$_2$**. This may be ascribed to the size and electrostatic nature of the substituent **CF$_3$**.

In contrast to the other ortho acetanilides, the PES of **Me:o-Me** contains two pronounced local minima at α = 60° and α = 300°, which are approximately 8 kJ·mol^{-1} higher in energy than the global minimum at α = 180°. This shows that for **Me:o-Me** a non-planar conformation is associated with a relatively small energy penalty compared with the other ortho acetanilides making the

observation of two different conformers likely. This behaviour can be understood as a consequence of the size of the substituent.

The torsions angles found in the crystal structures of the ortho acetanilides are recorded in Table 5.2 together with the deviation of the solid state torsion angles from α = 180˚. The table also contains the difference between the energy at the global minimum and at the experimental value of α. Since experimental energies are not available the energy of the experimental conformer was estimated from the PES, whereby the theoretical value of α closest to the experimental torsion angle was chosen and the corresponding energy used in the energy difference estimation.

Structure	α_{exp} [˚]	$\Delta\alpha(\lvert\alpha_{theo} - \alpha_{exp}\rvert)$ [˚]	$\Delta E(E_{exp}\text{-}E_{theo})$ [kJ·mol^{-1}]
Me:o-F	145.85	34.15	3.540
Me:o-Cl(m)	138.56	41.44	5.766
Me:o-Cl(o)	143.75	36.25	5.766
Me:o-Br	138.47	41.53	5.842
Me:o-CF_3	128.22	41.78	5.440
Me:o-Me(m)	129.63	50.37	6.721
Me:o-Me(o)	65.92	114.08 (5.92*)	8.574 (0.343*)
Me:o-^{t}Bu	114.79	55.21	7.941
Me:o-CN	142.39	37.61	6.530
Me:o-NO_2	151.83	28.17	4.377
Me:o-COMe	164.75	15.25	2.053
Me:o-COOH	174.02	5.98	0.406

Table 5.2: α_{exp} = experimental torsion angle; Δα = absolute difference between α_{exp} and theoretical optimum torsion angle α_{theo}; ΔE = difference between energy of α_{exp} and α_{theo}; *local minimum value.

Table 5.2 shows that all the experimental values of α deviate from the theoretical optimal angle of 180˚. It should be emphasised here that the theoretical PES was calculated for the isolated molecule with optimised geometry in the gas phase. Any intermolecular effects that arise from the crystal assembly are neglected in such computations. Hence the deviation of experimental solid state torsion angles from the theoretical values demonstrates that the molecular conformation is affected by the forces acting in the crystal. In the ortho-acetanilides intramolecular hydrogen bonding between the substituent and the amide group is expected to favour a planar molecular conformation over non-planar arrangements. Indeed the smallest deviation from the gas phase optimum conformation occurs in **Me:o-COOH** with a

difference of 5.98˚ followed by **Me:o-COMe** with a deviation of 15.25˚. In the gas phase these differences in torsion angles are associated with less than 2 kJ·mol^{-1} in energy. This energy difference is comparable with the free energy available to a molecule at room temperature[1] and is thus negligible. The largest deviation from α = 180˚ occurs in **Me:o-Me(o)** with 114.08˚, but even this substantial change in α would only result in an increase in energy by 8.57 kJ·mol^{-1} over the optimum conformation in the gas phase. Furthermore the torsion angle in **Me:o-Me(o)** lies within 6˚ and 0.343 kJ·mol^{-1} of the local minimum at α = 60˚. The experimental observation of conformational polymorphism is thus in line with the PES, which predicts the likelihood of two alternative conformations for the compound. As was shown in Chapter 4.B. the intermolecular forces acting in a crystal can certainly compensate for such an energy penalty hence conformationally less energetic molecules can still be incorporated in a stable crystal structure.

Regarding the acetanilides without intramolecular hydrogen bonding the experimentally observed conformations can be directly linked to intermolecular interactions in the crystal. For example the torsion angle in **Me:o-CN** amounts to 142.39˚ allowing for intermolecular hydrogen bonding with the amide group of a neighbouring molecule. If α were 180˚ hydrogen bonding would not be possible due to steric hindrance. This applies to all the ortho-acetanilides without intramolecular hydrogen bonding and accounts for the deviations of the experimental torsion angles from the theoretical optimum conformation.

5.C. Conformational analysis of meta-substituted acetanilides

The conformational analysis for the meta acetanilides yielded the PES maps displayed in Figures 5.8 to 5.11 according to the following grouping:

- halogenated meta acetanilides, i.e. **X = F**, **Cl**, **Br** → Figure 5.8
- bulkier, HB inactive substituents, i.e. **X = Me**, **CF_3**, **SF_5** → Figure 5.9
- substituents with HB acceptor functionality, i.e. **X = CN**, **NO_2**, **COMe**, **COOEt** → Figure 5.10
- substituents with HB donor functionality, i.e. **X = OH**, **COOH** → Figure 5.11

As for the ortho acetanilides the PES maps of the meta compounds have mirror symmetry with respect to 180˚ and hence only the region between 0 and 180˚ will be discussed further.

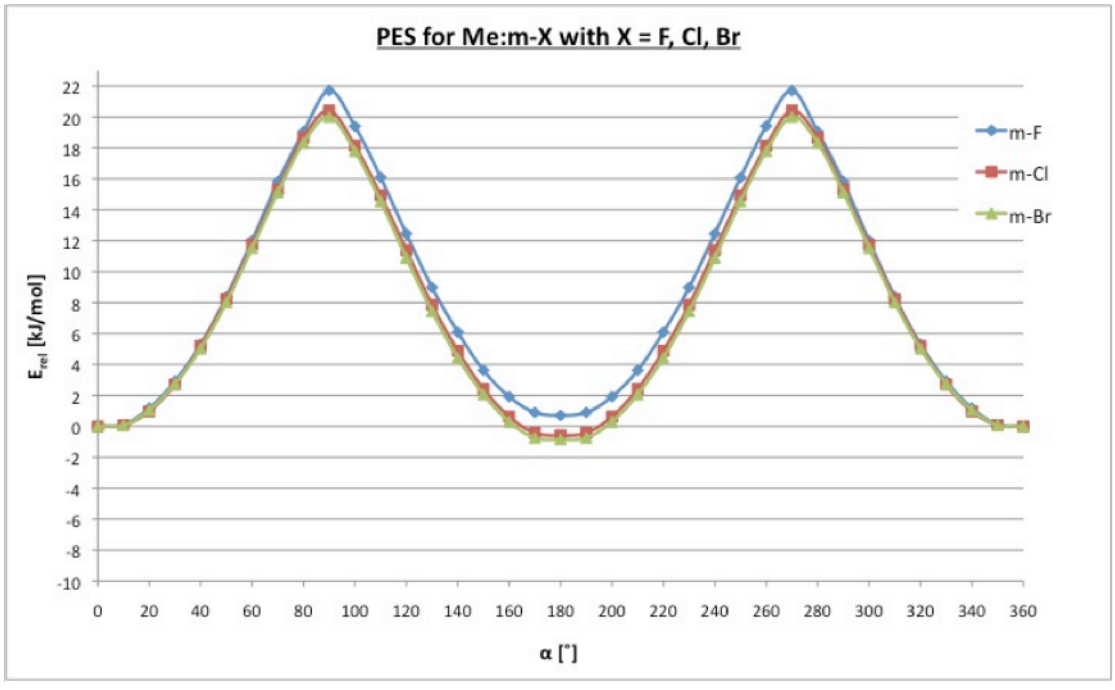


Figure 5.8: Potential energy surfaces for the rotation about the torsion angle α for **Me:m-F**, **Me:m-Cl**, **Me:m-Br**.

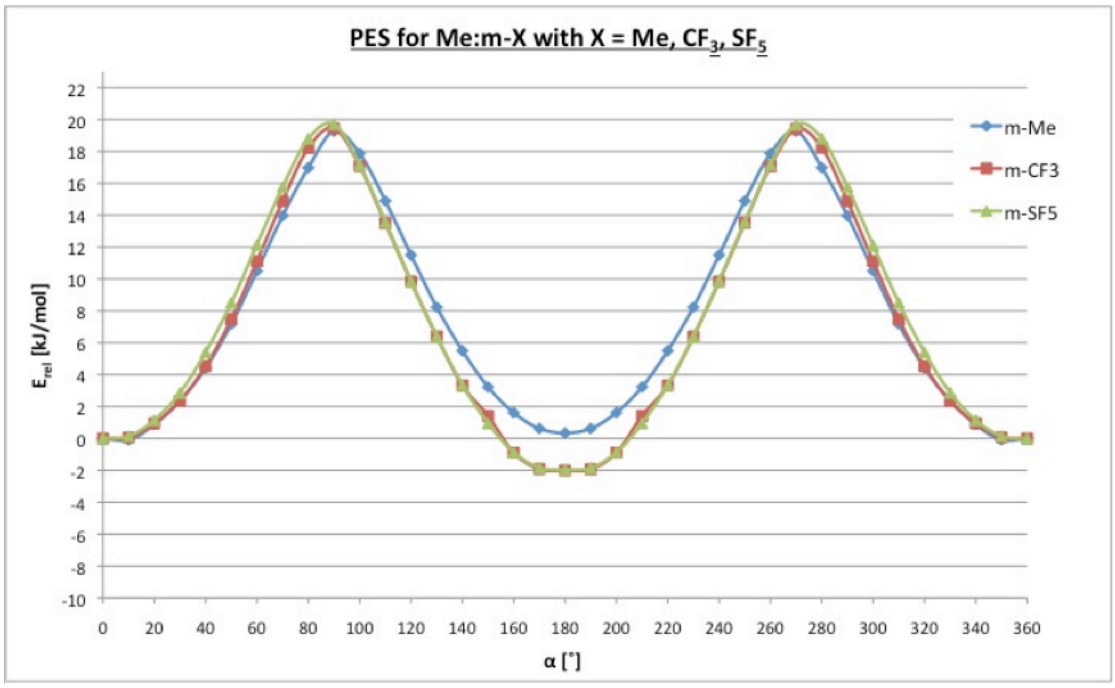


Figure 5.9: Potential energy surfaces for the rotation about the torsion angle α for **Me:m-Me**, **Me:m-CF_3**, **Me:m-SF_5**.

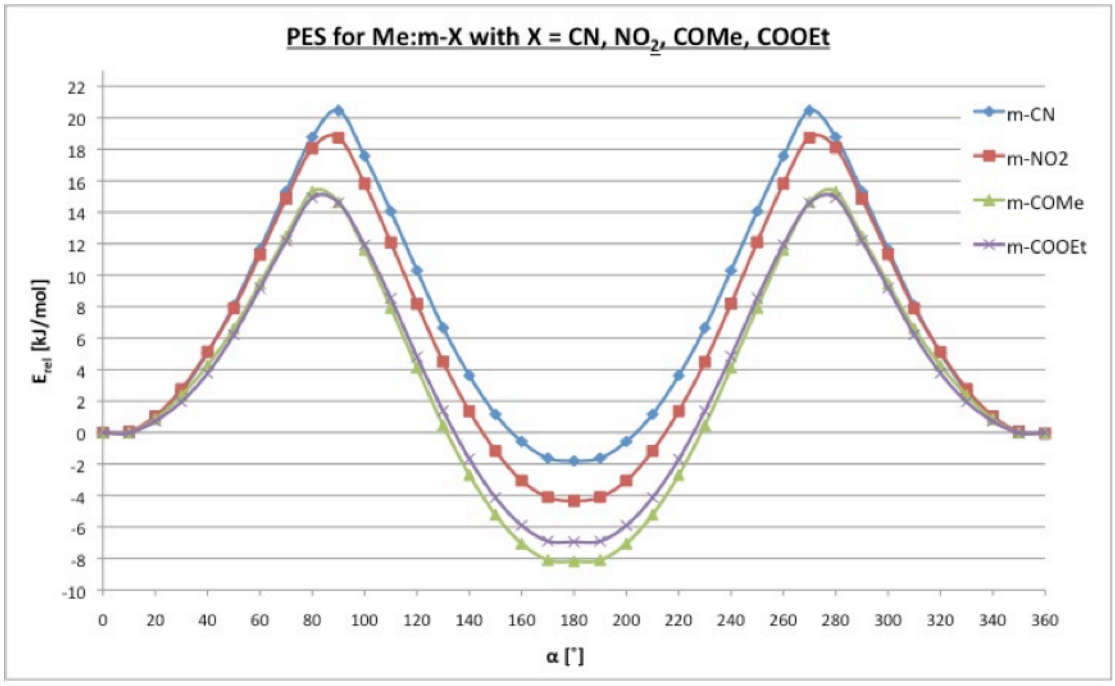


Figure 5.10: Potential energy surfaces for the rotation about the torsion angle α for **Me:m-CN**, **Me:m-NO_2**, **Me:m-COMe**, **Me:m-COOEt**.

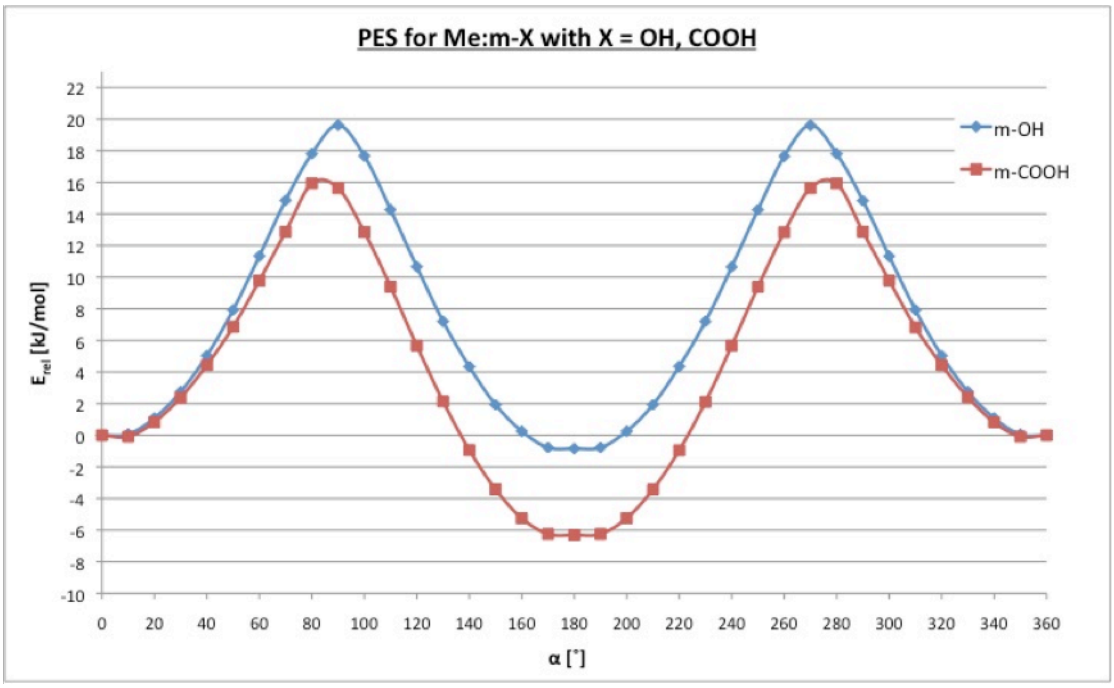


Figure 5.11: Potential energy surfaces for the rotation about the torsion angle α for **Me:m-OH**, **Me:m-COOH**.

In contrast to the PES of the ortho series the curves exhibit minima at α = 0° and α = 180° as well as a maximum at α = 90° or α = 80°. These maxima constitute rotational barriers between the energetic minimum conformations, and their relative energies vary between approximately 15 and 22 kJ·mol^{-1}. A listing of the relative energy values at the minima and maxima is provided in Table 5.3.

Molecule	$\alpha_{theo,\,min}$	$\alpha_{theo,\,max}$	$E_{theo,min}$	$E_{theo,max}$	$\Delta E_{theo}(E_{max}\text{-}E_{min})$
Me:m-F	0 (180)	90	0 (-0.701)	21.710	21.710 (21.009)
Me:m-Cl	180 (0)	90	-0.578 (0)	20.412	20.990 (20.412)
Me:m-Br	180 (0)	90	-0.844 (0)	20.000	20.844 (20.000)
Me:m-CF$_3$	180 (0)	90	-2.005 (0)	19.447	21.452 (19.447)
Me:m-SF$_5$	180 (0)	90	-1.963 (0)	19.704	21.667 (19.704)
Me:m-Me	0 (180)	90	0 (-0.335)	19.268	19.268 (18.934)
Me:m-CN	180 (0)	90	-1.800 (0)	20.464	22.264 (20.464)
Me:m-NO$_2$	180 (0)	90	-4.344 (0)	18.736	23.080 (18.736)
Me:m-OH	180 (0)	90	-0.842 (0)	19.619	20.460 (19.619)
Me:m-COMe	180 (0)	80	-8.178 (0)	15.311	23.489 (15.311)
Me:m-COOH	180 (0)	80	-6.309 (0)	15.928	22.237 (15.928)
Me:m-COOEt	180 (0)	80	-6.950 (0)	14.913	21.863 (14.913)

Table 5.3: $\alpha_{theo,min}/\alpha_{theo,max}$ = torsion angle at theoretical global minimum/maximum; $E_{theo,min}/E_{theo,max}$ = energies at theoretical global minimum/maximum; ΔE_{theo} = rotational barrier between global minimum and maximum. Values in parenthesis refer to local minima. Energies and angles are given in kJ·mol^{-1} and °, respectively.

From the values in the table it can be seen that global minimum occurs at α = 180˚ and the local minimum at α = 0˚ for all compounds except for **Me:m-F** and **Me:m-Me**. The global and local minimum is found at α = 0˚ and α = 180˚, respectively, for these two compounds. However, overall the energy between the global and local minimum conformations differs by less than 10 kJ·mol^{-1}. As was discussed for the ortho series, the conformers of α = 180˚ and α = 0˚ correspond to the planar *anti* and *syn* configurations between the meta-substituent and the amide oxygen atom.

Table 5.3 also contains the heights of the rotational barriers for the transition from the global to the local minimum and vice versa. The barriers are all similar within approximately 4 and 6 kJ·mol^{-1} for the former and latter case, respectively. The difference between the energy maximum and minimum values represents the penalty associated with the least favourable conformation, which is found at 90˚ for all meta acetanilides. In comparison with the ortho series these values are approximately three times smaller meaning the optimum conformations of the meta acetanilides are less stable overall than the ortho acetanilides. This can be ascribed to the stabilising effect of intramolecular interactions between the ortho substituents and the amide group. The meta substituents on the other hand are too far removed from the amide group for any intramolecular interactions to take effect on the overall energy of the molecule.

However, from Figures 5.8 to 5.11 it can be seen that the minimum at α = 180˚ becomes increasingly deeper than the minimum at α = 0˚as the substituents increase in size or become more electron withdrawing. This effect is strongest for **Me:m-COMe** where the energy difference between the conformations α = 180˚ and α = 0˚ is 8.18 kJ·mol^{-1} followed by **Me:m-COOEt** (6.95 kJ·mol^{-1}), **Me:m-COOH** (6.31 kJ·mol^{-1}) and **Me:m-NO$_2$** (4.34 kJ·mol^{-1}). For the other structures this energy difference amounts to less than 2 kJ·mol^{-1}. It should be mentioned that the substituents **COMe**, **COOEt** and **COOH** are oriented so that the HB acceptor atom is closest to the amide hydrogen atom in the optimised structures as shown in Figure 5.12. These results indicate that there is some long-range influence of the meta substituent on the torsion angle α, which causes energetic non-equivalence of the *anti* and *syn* configurations in the gas phase.

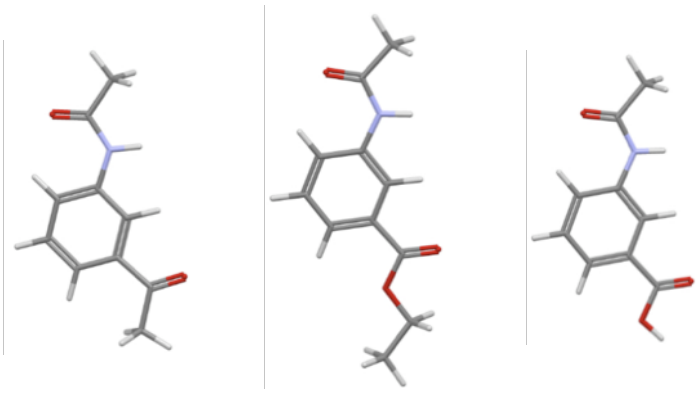


Figure 5.12: The meta substituent orientation is shown for **COMe** (left), **COOEt** (middle) and **COOH** (right).

This influence is most pronounced if the substituent contains an "exposed" oxygen atom such as in a carbonyl group or in the nitro substituent and is probably electrostatic in nature. The origin of this effect may lie in the optimisation of the charge distribution over the molecule so as to minimise the molecular dipole.

The torsion angles measured in the crystal structures of the meta series are listed in Table 5.4.

Structure	α_{exp} [°]	$\Delta\alpha(\lvert\alpha_{theo} - \alpha_{exp}\rvert)$ [°]	$\Delta E(E_{exp}\text{-}E_{theo})$ [kJ·mol^{-1}]
Me:m-F	15.03	15.03 (164.97)	1.173 (0.473)
Me:m-Cl: A	19.04	160.96 (19.04)	1.542 (0.964)
Me:m-Cl: B	339.64	159.64 (20.36)	1.544 (0.966)
Me:m-Br(m): A	3.3	176.7 (3.3)	0.844 (0.000)
Me:m-Br(m): B	14.65	165.35 (14.65)	0.913 (0.069)
Me:m-Br(o): A	21.49	158.51 (21.49)	1.894 (1.050)
Me:m-Br(o): B	337.83	157.83 (22.17)	1.894 (1.050)
Me:m-CF_3	10.27	169.73 (10.27)	2.055 (0.051)
Me:m-SF_5	14.56	165.44 (14.56)	2.102 (0.138)
Me:m-Me	165.87	165.87 (14.13)	0.622 (0.287)
Me:m-CN	11.15	168.85 (11.15)	1.875 (0.075)
Me:m-NO_2: A	10.55	169.45 (10.55)	4.409 (0.065)
Me:m-NO_2: B	356.45	176.45 (3.55)	4.278 (-0.066)
Me:m-NO_2: C	171.61	8.39 (171.61)	0.238 (-4.106)
Me:m-NO_2: D	181.23	1.23 (178.77)	0.000 (-4.344)
Me:m-COMe	168.76	11.24 (168.76)	0.081 (-8.097)
Me:m-COOEt	169.24	10.76 (169.24)	0.055 (-6.895)
Me:m-OH	172.22	7.78 (172.22)	0.066 (-0.775)
Me:m-COOH	172.96	7.04 (172.96)	0.064 (-6.246)

Table 5.4: α_{exp} = experimental torsion angle; $\Delta\alpha$ = absolute difference between α_{exp} and theoretical optimum torsion angle α_{theo}; ΔE = difference between energy of α_{exp} and α_{theo}. Values in parenthesis correspond to local minima. The notation ": A" indicates the number of molecules in the ASU.

As can be seen from the table, the molecules in the solid state almost all assume the *syn* configuration – according to the theoretical calculations the less favourable conformation compared to the *anti* arrangement. However, all experimental torsion angles are within 25˚ of one of the two calculated minimum conformations and the difference to the closest energy minimum does not exceed 1 $kJ \cdot mol^{-1}$.

Hence the meta molecules assume a more planar conformation in the solid state than in the ortho-series. The reason for this may be the greater distance between the substituent and the amide group so that any steric effects are reduced in the meta series. Interestingly the molecules in **Me:m-COMe**, **Me:m-COOEt** and **Me:m-COOH** assume the theoretical energetically more stable *anti* configuration also in the crystal assembly. **Me:m-NO_2** crystallises with four molecules in the ASU, two of which are in the *syn* and two in the *anti* configuration (*c.f.* picture of ASU in Chapter 4.A.). This Z'=4 structure is reproducible and efforts to produce a Z'=1 structure have so far been fruitless. Hence **Me:m-NO_2** may present a case of a 'frustrated' molecule for which the mixture of the two configurations is the most stable solid state arrangement. This case also demonstrates that both configurations can coexist rather than exclude each other, e.g. due to orientation preferences occurring during synthesis. Hence for the other meta acetanilides it may be concluded that the intermolecular assembly in the crystal determines which configuration is incorporated in the solid state structure. It could of course be also argued that the gas phase calculations show that the *anti* and *syn* conformers are energetically very similar and hence polymorphism may be likely for these compounds. Although a polymorph screen was not the objective of this study, **Me:m-Br** was found to be at least dimorphic with both polymorphs being Z'=2 structures. However, in both crystal structures both molecules in the ASU assume the *syn* configuration.

5.D. Conformational analysis of para-substituted acetanilides

The PES maps of the para acetanilides are displayed in Figures 5.13 to 5.18 using the grouping as outlined in Table 5.5. Due to the large number of compounds two styles were chosen to show similarities and differences more clearly.

X	Figure
H, Me, Et, ^{i}Pr, ^{c}Pr, ^{t}Bu	5.13
F, Cl, Br, I, CF_3, SF_5	5.14
CN, NO_2, OCF_3, OCOMe	5.15
NH_2, NMe_2	5.16
COMe, COOMe, COOEt, COOH, COOOH, $CBrCH_2$	5.17
OH, OMe, OEt, OPr, O^{i}Pr, OBu	5.18

Table 5.5: Listing of the Figures 5.12 to 5.17 containing the PES for the para acetanilides.

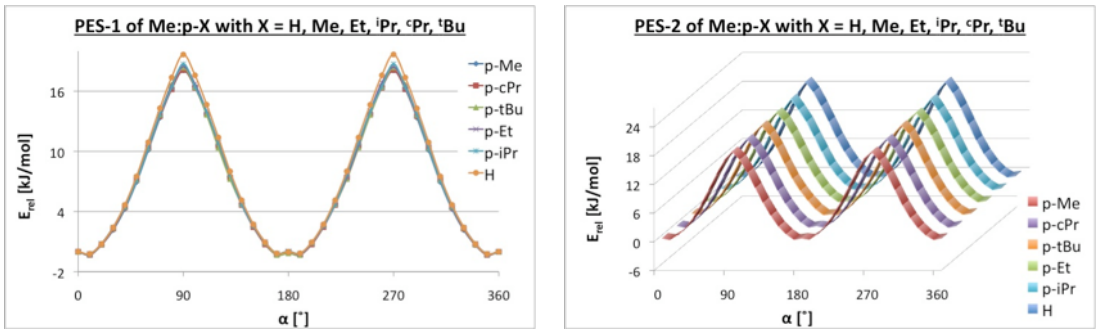


Figure 5.13: Electrostatic potential energy surfaces for the rotation about the torsion angle α for **Me:p-H**, **Me:p-Me**, **Me:p-Et**, **Me:p-^{i}Pr**, **Me:p-^{c}Pr**, **Me:p-^{t}Bu**. The ordering in the graph on the right hand side is in ascending energy at the maximum.

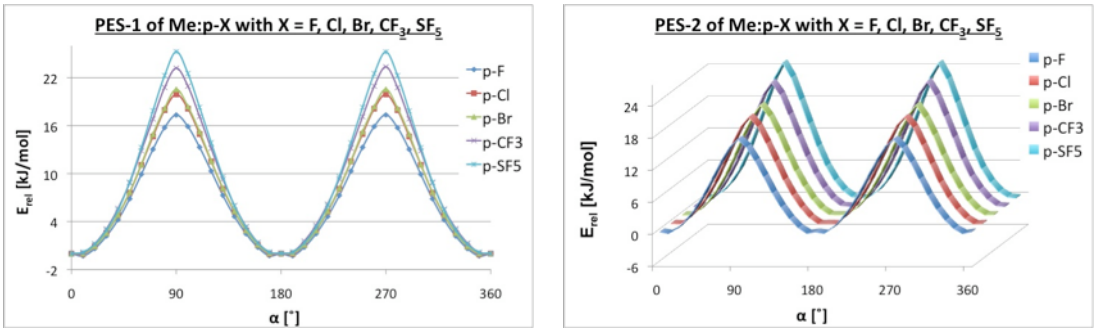


Figure 5.14: Electrostatic potential energy surfaces for the rotation about the torsion angle α for **Me:p-F**, **Me:p-Cl**, **Me:p-Br**, **Me:p-CF_3**, **Me:p-SF_5**. The ordering in the graph on the right hand side is in ascending energy at the maximum.

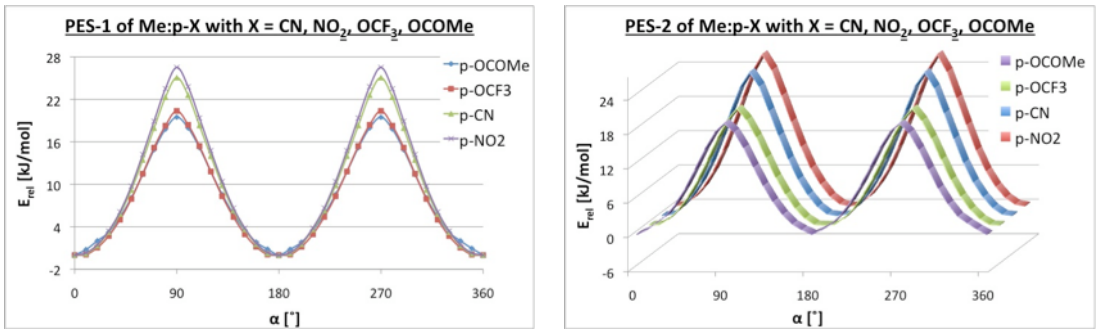


Figure 5.15: Electrostatic potential energy surfaces for the rotation about the torsion angle α for **Me:p-CN**, **Me:p-NO$_2$**, **Me:p-OCF$_3$**, **Me:p-OCOMe**. The ordering in the graph on the right hand side is in ascending energy at the maximum.

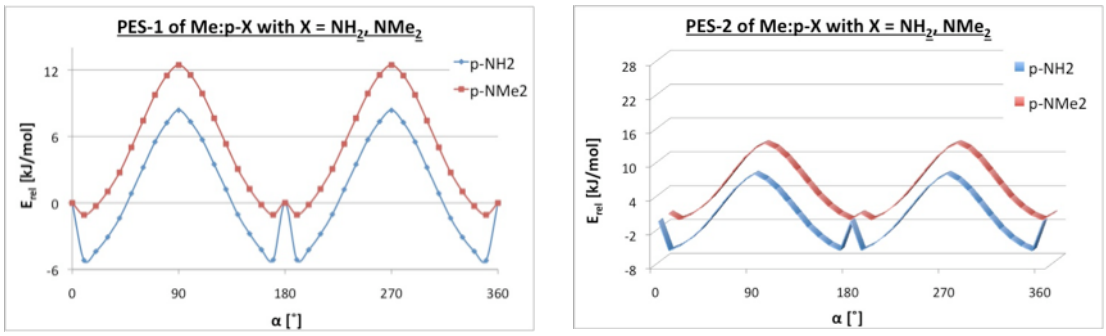


Figure 5.16: Electrostatic potential energy surfaces for the rotation about the torsion angle α for **Me:p-NH$_2$**, **Me:p-NMe$_2$**.

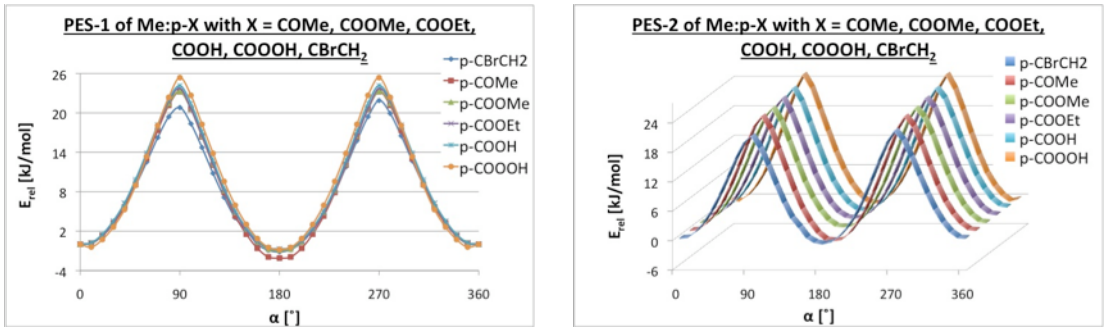


Figure 5.17: Electrostatic potential energy surfaces for the rotation about the torsion angle α for **Me:p-COMe**, **Me:p-COOMe**, **Me:p-COOEt**, **Me:p-COOH**, **Me:p-COOOH** and **Me:p-CBrCH$_2$**.

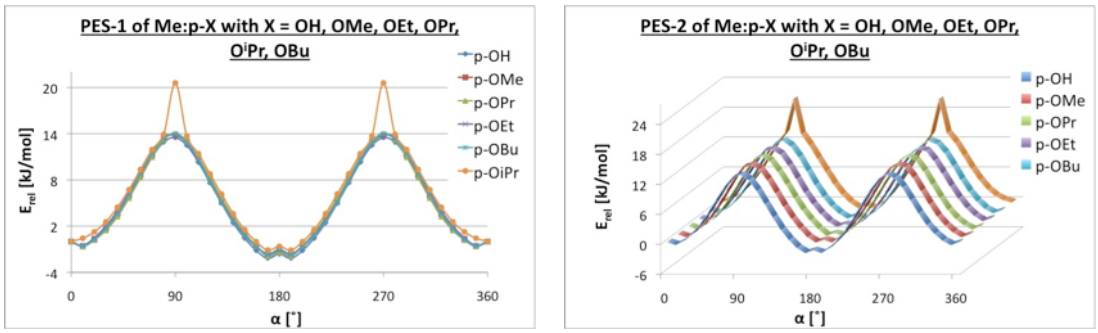


Figure 5.18: Electrostatic potential energy surfaces for the rotation about the torsion angle α for **Me:p-OH**, **Me:p-OMe**, **Me:p-OEt**, **Me:p-OPr**, **Me:p-O^{i}Pr** and **Me:p-OBu**.

Again the surfaces have mirror symmetry with respect to α = 180°. For most para acetanilides the curves between α = 180° and α = 0° are also symmetrical with respect to α = 90° and the conformations of α = 180° and α = 0° are energetically equivalent. This means that no distinction can be made between a *syn* and *anti* configuration as was the case in the ortho and meta series. The calculations confirm this for the compounds shown in Figures 5.13, 5.14, 5.15 and 5.16. The substituents of these compounds can be broadly classified as alkane, amino or halogen derivatives. In addition **Me:p-CN**, **Me:p-NO$_2$** and **Me:p-OCOMe** also fall into this category. These substituents are either rotationally symmetric or their conformation is consistent with the mirror symmetry of the phenyl ring (the mirror plane being perpendicular to the ring plane) so that the sides of the phenyl ring are equivalent. For example, in **Me:p-OCOMe** the substituent mean plane lies perpendicular to the phenyl ring mean plane hence coinciding with the mirror plane of the aromatic system as illustrated in Figure 5.19.

The PES maps of the remaining para acetanilides exhibit energy differences between the α = 180° and α = 0° conformers (*c.f.* Figure 5.17 and 5.18). It is notable that this applies to all compounds with alkoxy substituents and carbonyl[2] derivatives as well as **Me:p-OH** and **Me:p-CBrCH$_2$**. For these compounds the molecular structure is slightly more stable if α = 180°. Inspection of the optimised structures shows that this arrangement can be interpreted as an *anti* configuration and arises since the

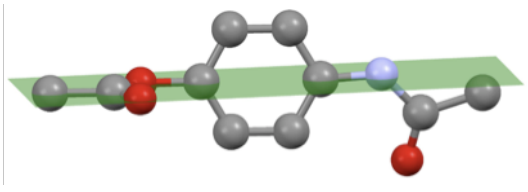


Figure 5.19: Molecular conformation of **Me:p-OCOMe**. The mirror plane dissecting the phenyl ring and substituent is shown in green.

mirror symmetry of the phenyl ring is broken due to the substituent. In the compounds with **X = COMe**, **COOMe**, **COOEt**, **COOH** and **COOOH** the amide oxygen is positioned on the opposite side of the molecule in relation to the substituent carbonyl oxygen (*anti)*. In the para alkoxy acetanilides and **Me:p-OH** the amide oxygen resides in the *anti* position to the lone pairs of the substituent oxygen atom. Figure 5.20 displays the α = 0˚ and α = 180˚ conformers for **Me:p-OMe** and **Me:p-OMe** to illustrate the difference in configuration.

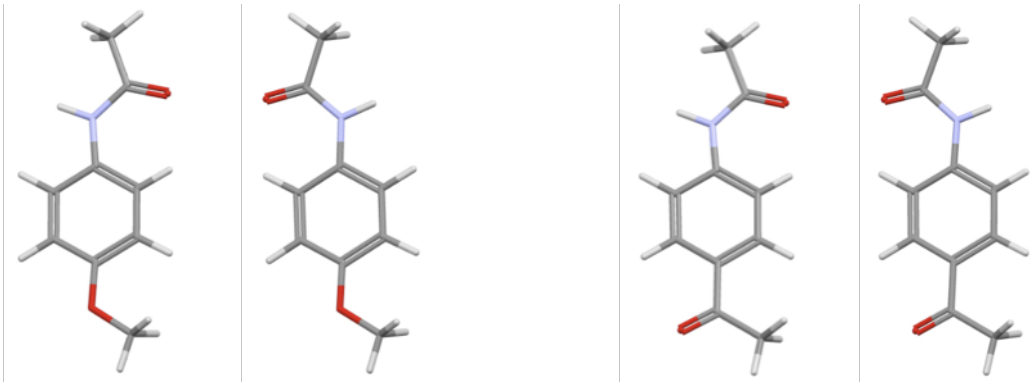


Figure 5.20: The *anti* (α=180˚) and *syn* (α=0˚) configurations are shown pair wise for **Me:p-OMe** (left) and **Me:p-COMe** (right).

It should be mentioned that the substituent mean plane is coplanar with the phenyl ring in these eleven structures. The situation is similar for **Me:p-CBrCH$_2$**, although in this structure the substituent mean plane is rotated out of the phenyl ring plane by approximately 48˚ so that the bromine atom is located above the ring as shown in Figure 5.21.

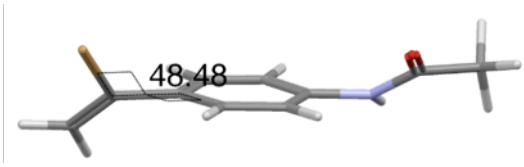


Figure 5.21: The torsion angle between substituent and phenyl ring is shown for **Me:p-CBrCH$_2$**.

Overall the PES curves of the para series are similar to those of the meta acetanilides. They exhibit minima in the regions of α = 0˚ and α = 180˚ as well as maxima at α = 90˚. However, the global minimum is often found at α = 10˚ or α = 170˚, respectively, which is most pronounced in the curves of **Me:p-NH$_2$** and **Me:p-NMe$_2$** (Figure 5.16). Table 5.6 contains the relative energies of the minimum regions and the maxima of the PES curves.

Molecule	$\alpha_{theo,min}$	$\alpha_{theo,max}$	$E_{theo,min}$	$E_{theo,max}$	$\Delta E_{theo}(E_{max}-E_{min})$
Me:H	10 (170)	90	-0.204 (-0.171)	19.664	19.868 (19.834)
Me:p-F	10 (170)	90	-0.295 (-0.270)	17.325	17.621 (17.596)
Me:p-Cl	10 (170)	90	-0.069 (-0.043)	19.912	19.981 (19.955)
Me:p-Br	0 (10)	90	0.000 (0.010)	20.553	20.553 (20.543)
Me:p-CF_3	0 (180)	90	0.000 (0.034)	23.197	23.197 (23.164)
Me:p-SF_5	0 (180)	90	0.000 (0.029)	25.253	25.253 (25.224)
Me:p-OCF_3	0 (180)	90	0.000 (0.001)	20.379	20.379 (20.377)
Me:p-Me	10 (170)	90	-0.295 (-0.275)	18.061	18.356 (18.336)
Me:p-Et	10 (170)	90	-0.222 (-0.188)	18.471	18.692 (18.658)
Me:p-^{i}Pr	10 (170)	90	-0.251 (-0.170)	18.725	18.976 (18.895)
Me:p-^{c}Pr	10 (170)	90	-0.323 (-0.291)	18.115	18.438 (18.406)
Me:p-^{t}Bu	170 (10)	90	-0.163 (-0.230)	18.380	18.543 (18.610)
Me:p-CN	180 (0)	90	-0.005 (0.000)	25.071	25.076 (25.071)
Me:p-NO_2	180 (0)	90	-0.013 (0.000)	26.524	26.537 (26.524)
Me:p-NH_2	10 (170)	90	-5.195 (-5.143)	8.374	13.569 (13.517)
Me:p-NMe_2	10 (170)	90	-1.077 (-1.064)	12.442	13.519 (13.506)
Me:p-OH	170 (180)	90	-2.168 (-1.626)	13.489	15.657 (15.115)
Me:p-OMe	170 (180)	90	-1.634 (-1.121)	13.825	15.459 (14.946)
Me:p-OEt	170 (180)	90	-1.691 (-1.138)	13.803	15.494 (14.941)
Me:p-OPr	170 (180)	90	-1.901 (-1.387)	14.035	15.936 (15.422)
Me:p-O^{i}Pr	170 (180)	90	-1.112 (-0.628)	20.592	21.705 (21.220)
Me:p-OBu	170 (180)	90	-1.642 (-1.005)	13.956	15.598 (14.961)
Me:p-OCOMe	0 (180)	90	0.000 (0.046)	19.481	19.481 (19.435)
Me:p-COMe	180 (170)	90	-2.086 (-1.927)	23.228	25.314 (25.155)
Me:p-COOH	180 (170)	90	-0.684 (-0.516)	24.154	24.839 (24.670)
Me:p-COOMe	180 (170)	90	-0.886 (-0.729)	23.302	24.189 (24.031)
Me:p-COOEt	180 (170)	90	-0.732 (-0.521)	23.753	24.484 (24.273)
Me:p-COOOH	180 (170)	90	-0.793 (-0.472)	25.373	26.166 (25.845)
Me:p-$CBrCH_2$	180 (170)	90	-1.155 (-0.900)	20.808	21.963 (21.709)

Table 5.6: $\alpha_{theo,min}/\alpha_{theo,max}$ = torsion angle at theoretical global minimum/maximum; $E_{theo,min}/E_{theo,max}$ = energies at theoretical global minimum/maximum; ΔE_{theo} = rotational barrier between global minimum and maximum. Values in parenthesis refer to local minima. Energies and angles are given in kJ·mol^{-1} and °, respectively.

It can be seen that for 16 compounds the α = 10° or α = 170° is energetically more favourable than α = 0° or α = 180°, which constitutes the global minimum for 10 compounds. However, the energy differences are very small with generally less than 1 kJ·mol^{-1}. The greatest variation occurs in the amino derivatives with 5.14 kJ·mol^{-1} (**Me:p-NH_2**) followed by 1.08 kJ·mol^{-1} (**Me:p-NMe_2**). In these particular cases the amino nitrogen assumes a trigonal planar configuration at α = 0° and α = 180° whereas the amino group geometry can be described as a capped trigonal pyramid with the lone pair in the apical position in all other optimised conformers. Given the energy difference in the PES it must be concluded

that the planar geometry is energetically unfavourable over the pyramidal configuration. Such variation in geometry was not observed for any of the other structures, except for **Me:p-O^{i}Pr** at α = 90°. In the optimised structure the substituent assumes as different geometry relative to the phenyl ring as illustrated in Figure 5.22 causing a sharp spike in energy at the maximum (*c.f.* Figure 5.18).

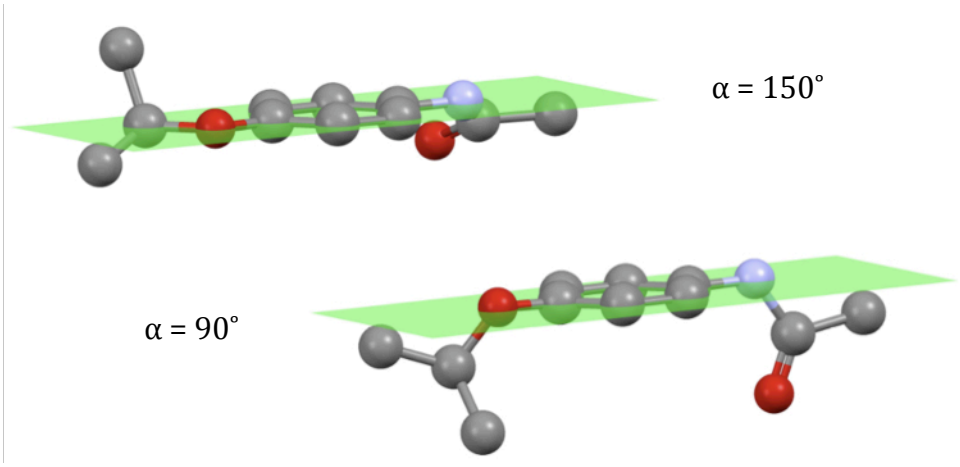


Figure 5.22: The change in the substituent position with respect to the phenyl ring as going to the 90° conformer observed in **Me:p-O^{i}Pr** is shown. The phenyl ring mean plane is included in green. Hydrogen atoms were omitted for clarity.

The slight preference of the α = 10° or α = 170° conformers suggests the conjugation of amide group and phenyl ring of the molecular backbone causes energetic strain in the fully planar geometry. However it should be emphasised that the energies involved concern only a fraction of 1 kJ·mol^{-1} and are hence even less indicative of any substituent effects compared with the meta and ortho series. The maximum energy values range from 8.37 kJ·mol^{-1} (**Me:p-NH$_2$**) to 26.52 kJ·mol^{-1} (**Me:p-NO$_2$**). The rotational barriers (Table 5.6) for the transition from the global to the local minimum closely match those for the reverse transition with the largest differences occurring for the acetanilides with alkoxy or carbonyl derivatives as already mentioned above. Compared with the meta series the rotational barriers vary over a wider range indicating that the level of conjugation between phenyl ring and amide group depends on the nature of the para substituent. The substituent influences the electron distribution in the phenyl π-system and hence the conjugation of the whole molecular backbone via inductive and mesomeric effects. The Hammett constants have been devised as measures of

these effects in order to relate reaction rates with substituent properties[3]. Figure 5.23 displays the plot of the theoretical rotational barriers versus the Hammett constants of the para substituents[4].

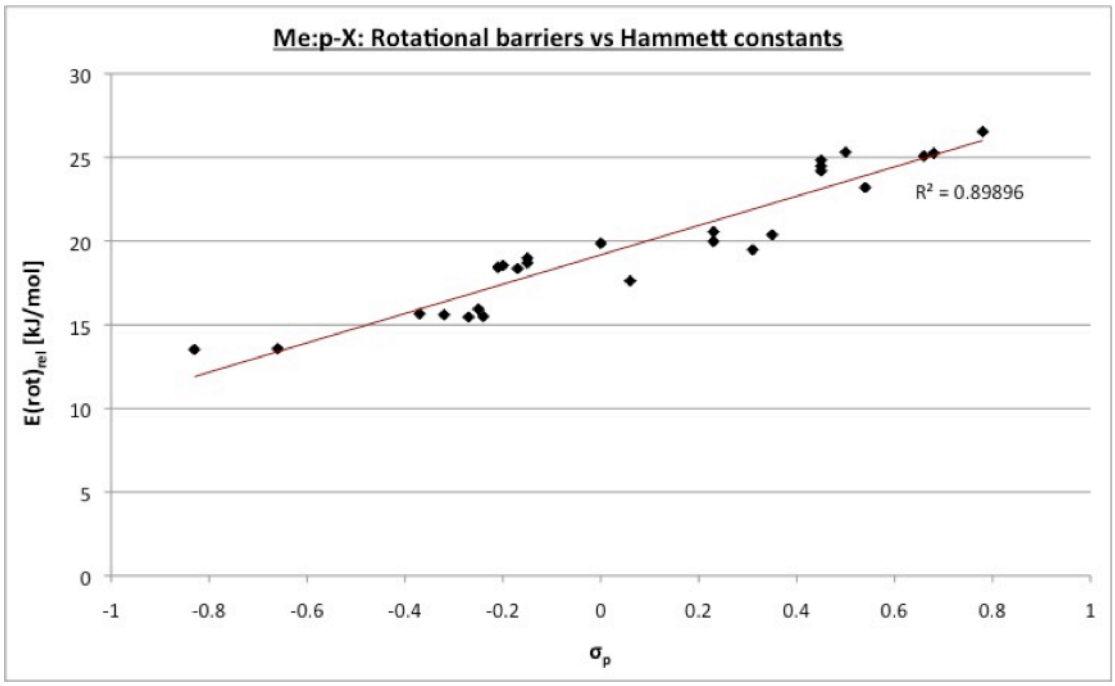


Figure 5.23: The relative rotational barriers are plotted versus the Hammett constants for the para-substituted acetanilides.

It can be seen that the two quantities are positively correlated, hence the more electron withdrawing the substituent the higher the rotational barrier. This in turn means that the bond between the amide group and the phenyl ring becomes less rotationally flexible.

Table 5.7 lists the values of the torsion angles observed in the crystal structures.

Structure	θ	α_{exp}	$\Delta\alpha(\lvert\alpha_{theo} - \alpha_{exp}\rvert)$	$\Delta E(E_{exp}-E_{theo})$
Me:H	16.142	162.70	152.7 (7.30)	1.153 (1.120)
Me:p-F	3.322	176.30	166.3 (6.30)	0.273 (0.248)
Me:p-Cl	3.761	175.74/178.22	165.74/168.22 (5.74/8.22)	0.068/0.068 (0.042/0.042)
Me:p-Br(m)	18.998	164.34	164.34 (154.34)	1.109 (1.099)
Me:p-Br(o)	5.22	174.63	174.63 (164.63)	0.050 (0.040)
Me:p-CF_3	7.45	173.11/179.41	173.11/179.41 (6.89/0.59)	0.174/0.034 (0.140/0.000)
Me:p-SF_5: A	35.596	142.68	142.68 (37.32)	5.974 (5.945)
Me:p-SF_5: B	37.113	146.35	146.35 (33.65)	3.285 (3.255)
Me:p-OCF_3	5.312	174.92/175.28	174.92/175.28 (5.08/4.72)	0.081/0.001 (0.079/0.000)
Me:p-Me(m)	16.616	163.51	153.51 (6.49)	1.155 (1.135)
Me:p-Me(o)	5.662	173.64	163.64 (3.64)	0.020 (0.000)
Me:p-Et	5.019	173.75/173.88	163.75/163.88 (3.75/3.88)	0.034/0.034 (0.000/0.000)
Me:p-^{i}Pr	14.074	167.94	157.94 (2.06)	0.081 (0.000)
Me:p-^{c}Pr	9.531	169.87	159.87 (0.13)	0.032 (0.000)
Me:p-^{t}Bu	22.576	154.77	15.23 (144.77)	2.836 (2.903)
Me:p-CN	6.195	174.42/177.10	5.58/2.90 (174.42/177.10)	0.258/0.000 (0.253/-0.005)
Me:p-NO_2: A	3.791	169.18	10.82 (169.18)	0.375 (0.362)
Me:p-NO_2: B	12.896	178.91	1.09 (178.91)	0.000 (-0.013)
Me:p-OCOMe(m1)	14.394	168.31	168.31 (11.69)	0.841 (0.796)
Me:p-OCOMe(m2)	10.149	164.55	164.55 (15.45)	1.824 (1.779)

Table 5.7: θ = dihedral angle between phenyl ring and amide group mean planes; α_{exp} = experimental torsion angle; Δα = absolute difference between α_{exp} and theoretical optimum torsion angle α_{theo}; ΔE = difference between energy of α_{exp} and α_{theo}. Values in parenthesis correspond to local minima. The notation ": A" indicates the number of molecules in the ASU. Energies and angles are given in $kJ \cdot mol^{-1}$ and °, respectively.

Structure	φ	α_{exp}	$\Delta\alpha(\lvert\alpha_{theo} - \alpha_{exp}\rvert)$	$\Delta E(E_{exp}\text{-}E_{theo})$
Me:p-NH₂	34.689	146.67	136.67 (23.33)	2.398 (2.346)
Me:p-NMe₂: A	5.766	136.48	126.48 (33.52)	4.156 (4.142)
Me:p-NMe₂: B	44.691	173.70	163.70 (3.70)	0.013 (0.000)
Me:p-OH(m)	20.487	158.85	11.15 (21.15)	1.000 (0.459)
Me:p-OH(o1)	16.729	165.99	4.01 (14.01)	0.000 (-0.541)
Me:p-OH(o2): A	36.051	149.82	20.18 (30.18)	2.607 (2.066)
Me:p-OH(o2): B	25.243	154.21	15.79 (25.79)	2.607 (2.066)
Me:p-OMe	17.813	160.68	9.32 (19.32)	1.046 (0.533)
Me:p-OEt	27.68	152.55	17.45 (27.45)	2.554 (2.001)
Me:p-OPr	10.265	168.77	1.23 (11.23)	0.000 (-0.514)
Me:p-O^{i}Pr	19.677	161.18	8.82 (18.82)	1.084 (0.599)
Me:p-OBu	25.52	155.79	14.21 (24.21)	0.960 (0.323)
Me:p-COMe	15.535	164.21/174.39	15.79/5.61 (5.79/4.39)	1.483/0.159 (1.325/0.000)
Me:p-COOH	39.255	142.82	37.18 (27.18)	5.789 (5.621)
Me:p-COOMe	9.821	170.84	9.16 (0.84)	0.158 (0.000)
Me:p-COOEt	14.138	165.07	14.93 (4.93)	0.211 (0.000)
Me:p-COOOH	4.002	176.97	3.03 (6.97)	0.000 (-0.320)
Me:p-CBrCH₂	19.651	160.92	19.08 (9.08)	1.174 (0.920)

Table 5.7 *continued.*

It should be noted that the distinction between the α = 0˚ and α = 180˚ conformers is not possible for the compounds with rotationally symmetric substituents. The atoms chosen to define the torsion angle are displayed in Figure 5.24 and hence the values are closer to 180˚ than to 0˚ in α. In order to gain a better picture of the distortion from planarity the dihedral angle θ between the mean planes of the phenyl ring and the amide group is recorded in Table 5.7 also.

From the values in the table it can be seen that most experimental torsion angles correspond to the PES minima of the isolated molecule and lie within 20° of this minimum energy. Since these are close to α = 0° or α = 180° it can be concluded that most para-acetanilides assume a near planar conformation in the solid state.

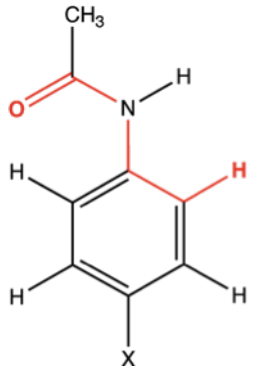


Figure 5.24: The atoms defining the experimental torsion angle.

However, a few structures do not follow this trend. The biggest deviations from planarity occur in **Me:p-SF$_5$**, **Me:p-NH$_2$**, in one of the two molecules in the ASU of **Me:p-NMe$_2$** and **Me:p-OH(o2)** as well as in **Me:p-COOH** with the associated dihedral angles of 35.60°/37.11°, 34.69°, 44.70°, 36.05° and 39.26°, respectively. Accordingly the relative energy difference to the minimum conformer is largest for these structures (*c.f.* Table 5.7). This deviation from the optimum conformation cannot be related to any inductive or mesomeric effects and is hence most likely a consequence of intermolecular forces acting in the crystal structure.

Table 5.8 summarises the configuration of the substituent with respect to the amide oxygen atom for the acetanilides with substituents that are not rotationally symmetric. The calculations show that the *anti* configuration is the more stable than the *syn* arrangement, however, the largest energy difference between α = 0° and α = 180° only amounts to 2.08 kJ·mol^{-1} for **Me:p-COMe**. In the solid state structures both configurations occur with similar frequency. Hence the intermolecular forces in the crystal are sufficient to accommodate the incorporation of the less stable conformer and the energetic difference between the *anti* and *syn* configurations are not sufficient to achieve a preference in the molecular structure in the crystal.

Structure	***Configuration***
Me:p-OH(m)	*syn*
Me:p-OH(o1)	*syn*
Me:p-OH(o2)	*syn/syn*
Me:p-OMe	*syn*
Me:p-OEt	*anti*
Me:p-OPr	*syn*
Me:p-OiPr	*anti*
Me:p-OBu	*anti*
Me:p-COMe	*syn*
Me:p-COOH	*anti*
Me:p-COOMe	*anti*
Me:p-COOEt	*anti*
Me:p-COOOH	*syn*
Me:p-CBrCH$_2$	*anti*

Table 5.8: Listing of the configuration between substituent and amide group.

5.E. Conformational differences in polymorphic forms

It was noted during this study that some of the polymorphs encountered differed in their molecular conformation with respect to the torsion angle α. Table 5.9 summarises the observed torsion angles for all polymorphs encountered during this study. Also collated are the differences in torsion angles and the equivalent theoretical energies.

Structure	α_{exp} [°]	$\Delta\alpha_{exp}$ [°] *ΔE [kJ·mol^{-1}]*				
Me:o-Cl(m)	138.56	5.19 *0*				
Me:o-Cl(o)	143.75					
Me:o-Me(m)	129.63	63.71 *1.853*				
Me:o-Me(o)	65.92					
Me:m-Br(m): A	3.30	11.35 *0.069*	18.19 *1.050*	25.47 *1.050*		
Me:m-Br(m): B	14.65				6.84 *0.981*	36.82 *0.981*
Me:m-Br(o): A	21.49	43.66 *0*				
Me:m-Br(o): B	337.83					
Me:p-Br(m)	164.34	10.29 *1.059*				
Me:p-Br(o)	174.63					
Me:p-Me(m)	163.51	10.13 *1.135*				
Me:p-Me(o)	173.64					
Me:p-OH(m)	158.85	7.14 *1.000*	9.03 *1.607*	4.64 *1.607*		
Me:p-OH(o1)	165.99				16.17 *2.607*	11.78 *2.607*
Me:p-OH(o2): A	149.82	4.39 *0*				
Me:p-OH(o2): B	154.21					
Me:p-OCOMe(m1)	168.31	3.76 *0.983*				
Me:p-OCOMe(m2)	164.55					

Table 5.9: α_{exp} = experimental torsion angles; $\Delta\alpha_{exp}$ = difference in torsion angles between polymorphs; ΔE = theoretical energy difference between polymorph conformations. The energy values are enclosed by asterisks. The notation ": A" refers to the independent molecules in the ASU if Z' > 1.

As the conformational analysis described in the preceding three sections of this chapter has shown the associated energy differences in Table 5.9 are small. Although changes in α are observed between the polymorphs, almost all conformations remain within the energy well of the global minimum on the PES as was confirmed in the previous sections. Only in the case of **Me:o-Me** the two conformations present in the monoclinic and orthorhombic crystal structure, respectively, correspond to different energy minima on the PES. The torsion angles of the molecules in **Me:p-Me(m)** and **Me:p-Me(o)** hence differ by over 60° - the

largest variation in the torsion angles compared with the other compounds. The next largest difference in torsion angles is actually found between the independent molecules in the ASU of **Me:m-Br(o)** and between the conformations in **Me:m-Br(o)** and those in its polymorph, which is also a Z'=2 structure. However, the energy differences are again small and all observed torsions are close to the theoretical global minimum. A brief comment regarding the unusually large values of α = 337.83° in **Me:m-Br(o): B** (i.e. the second molecule in the ASU). Usually the conformers of α = 0° and α = 360° are indistinguishable and in crystal structures with an inversion centre or mirror plane only the values of α = 0° to 180° are unique[5]. However, the crystal structure **Me:m-Br(o)** do not contain any such symmetry element causing a relation between the torsion angle and the direction of rotation about α. Hence with the phenyl ring mean plane as reference then in one molecule of the ASU in **Me:m-Br(o)** the amide group is located above the plane and in the other molecule it is positioned below the plane as demonstrated in Figure 5.25.

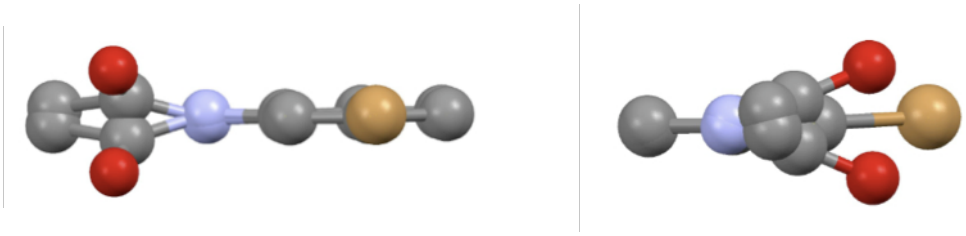


Figure 5.25: Two perspectives of the overlay of the two independent molecules in the ASU.

In **Me:m-Br(m)** on the other hand, due to the presence of an inversion centre such distinction is not possible. Nevertheless the two conformers in the ASU of **Me:m-Br(o)** only correspond to the approximately 22° deviation from the theoretical global minimum value, the fluctuation simply occurs on either side of the minimum within the same energy well.

In terms of the polymorphism observed the conformational analysis has revealed that the energy differences between the polymorphic forms are small. They are of an order that is comparable with the weakest intermolecular interactions and hence the different conformers of the same compound can easily be incorporated into different crystal structures.

5.F. References

[1] The energy available at room temperature is obtained as RT = 2.436 kJ·mol^{-1} for T = 293K.

[2] The substituent is connected to the phenyl ring via the carbon atom of the carbonyl group.

[3] C. Hansch, A. Leo, R. W. Taft, *Chemical Reviews*, **91**, 165-195 (1991).

[4] Hammett constants could not be found for **X = CBrCH$_2$** and **COOOH**.

[5] This means the molecules are symmetric with respect to α = 180˚ hence 350˚ ≡ 10˚ or 330˚ ≡ 30˚ and so forth.

CHAPTER 6: THE BIGGER PICTURE

As already mentioned in Chapter 2, the study of the acetanilide crystal structures forms part of a bigger project. Since the systematic analysis of the structural similarities in the acetanilides revealed that certain constructs appeared to be very robust and predominant, the compound library was soon extended to include other types of acylanilides. By changing the substituent on the amide group the influence of the substituent on the 'other end' of the core molecule could be assessed. As the synthesis and crystal structure matrices introduced in Chapter 2 show, this approach produced a great wealth of compounds and crystallographic data. The detailed discussion of the structural similarities in this extended compound library goes beyond the scope of this work, but a few aspects will be briefly highlighted in the following sections. They do not only concern structural similarity in the sense as it was discussed in Chapter 4 for the acetanilides. Undertaking such a big study has also enabled the observation of trends in crystallisation behaviour and correlation between crystallographic phenomena and the type of acylanilide. The size of the compound library has furthermore attracted interest within the crystallographic community and a few collaborations have since been started. A summary of these will also be provided in this chapter. Overall this chapter is intended as an overview of the variety of research topics that a systematic study of this size provides and summarises further developments and future work associated with this project. The associated crystallographic data will be made publicly available through the eCrystals archive in due course.

6.A. Structural similarity in para-bromo acylanilides, Y:p-Br

The inclusion of different residues on the amide side of the molecule allows the comparison of crystal structures along the rows of the compound matrices introduced in Chapter 2. Such a comparison then provides information about the influence the type of acylanilide has on the crystal packing. As an example, one complete row is chosen corresponding to the series of para-bromo acylanilides **Y:p-Br**, where **Y**= **H**, **Me**, **Et**, **Pr**, **^{t}Bu**, **CF_3**, **OMe**, **OEt**, **NH_2**. The resulting structural similarity plot is shown in Figure 6.1. For this particular comparison a COSP of the following 5 atoms was chosen: the amide group nitrogen, carbon and oxygen atoms as well as carbon 1 and 4 (para) of the phenyl ring.

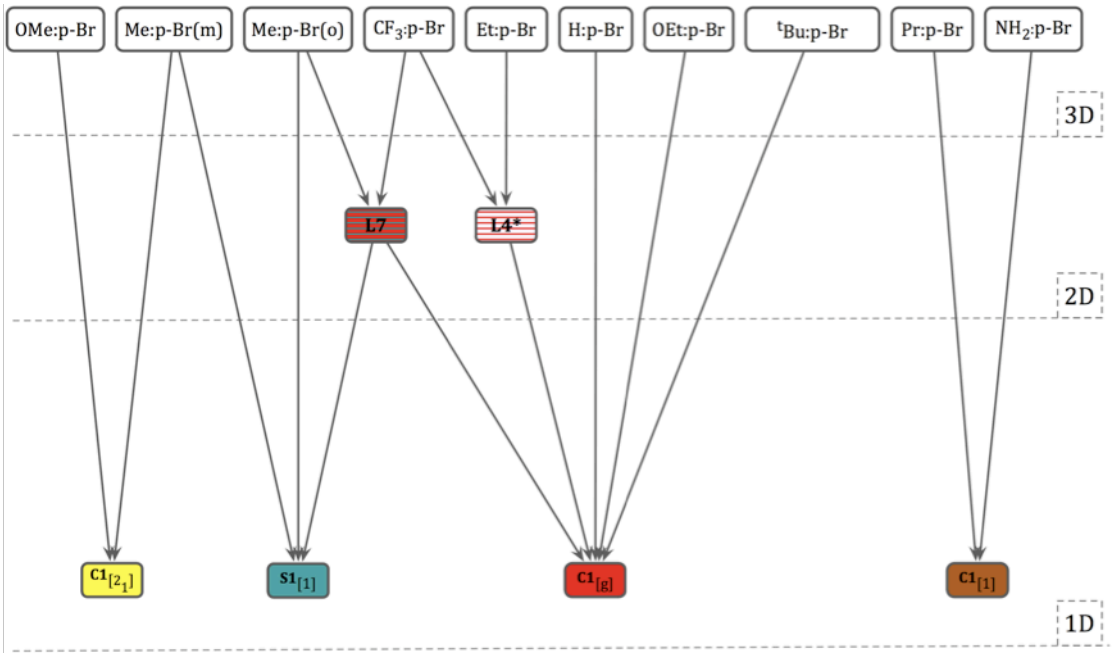


Figure 6.1: Structural similarity plot for ten para-bromo acylanilide crystal structures.

As XPac confirmed, only SCs familiar from the acetanilide series are observed in this small set. All compounds exhibit amide-amide hydrogen bonding in one of the common three HB geometries: **C1$_{[t]}$**, **C1$_{[2_1]}$** or **C1$_{[g]}$**. The former two chains each occur in two structures only, whereas **C1$_{[g]}$** is present in the remaining 6 structures. This construct can clearly form irrespective of the type of acylanilide.

The "longest" amide substituent however, is accommodated in the translational chain **C1**$_{[t]}$. In this case the geometry of chain **C1**$_{[t]}$ places the alkyl chains of **Pr:p-Br** alongside each other whilst separating them from the phenyl rings resulting in a flat chain. The close packing of these planar objects is more easily achievable than the V-shape typical of the glide symmetry in **C1**$_{[g]}$. Figures 6.2 to 6.4 display the **C1** chains as they occur in the para-bromo acylanilides.

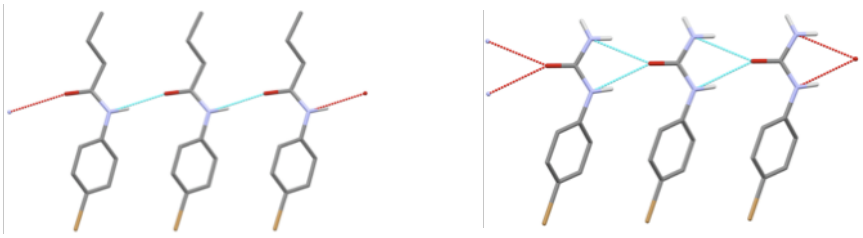


Figure 6.2: Chain **C1**$_{[t]}$ in **Pr:p-Br** (left) and **NH$_2$:p-Br** (right).

It is noteworthy that the hydrogen bonding of para-bromo phenylurea, **NH$_2$:p-Br**, can be classified as the chain **C1**$_{[t]}$, but it can be seen from Figure 6.2 that this hydrogen bond is bifurcated. Given the additional hydrogen bonding opportunities the NH_2 group provides, the phenyl ureas form a separate class within the acylanilides. Nevertheless the geometry of this particular hydrogen bond is consistent with the translational chain **C1**$_{[t]}$.

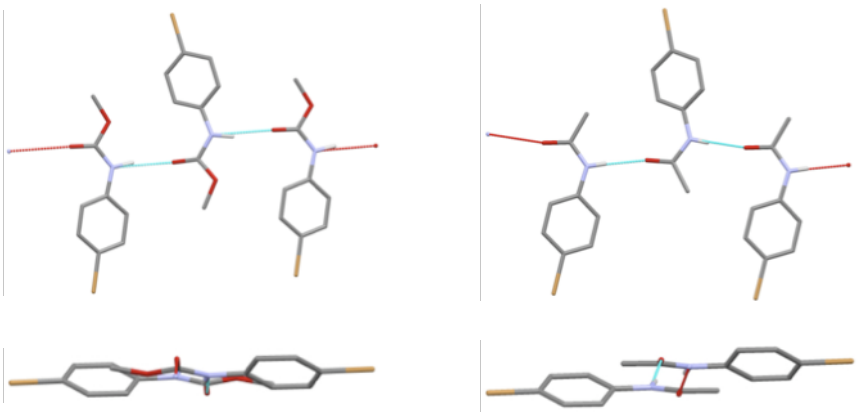


Figure 6.3: Chain **C1**$_{[2_1]}$ in **OMe:p-Br** (left) and **Me:p-Br(m)** (right). The viewing direction is perpendicular (top) and parallel (bottom) to the HB direction.

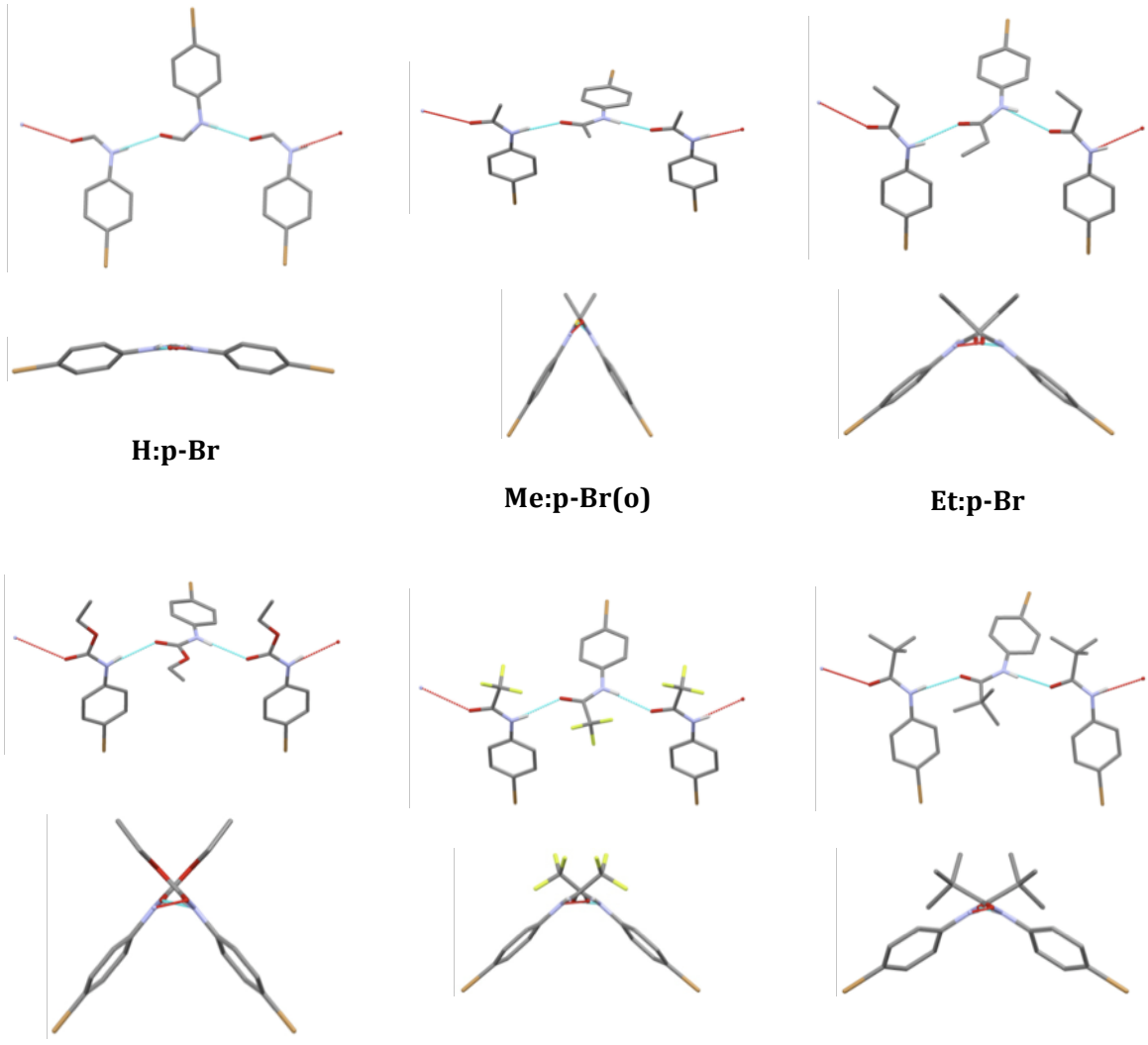
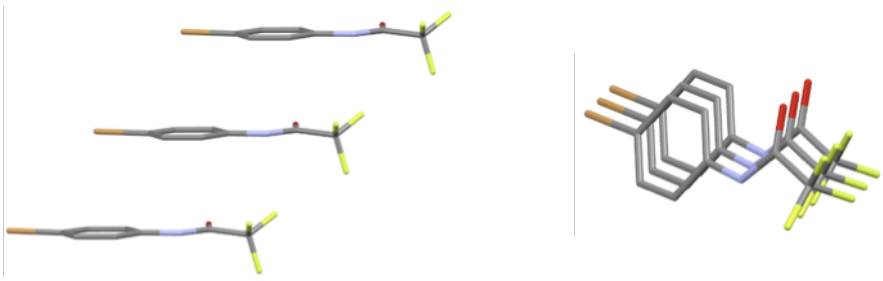
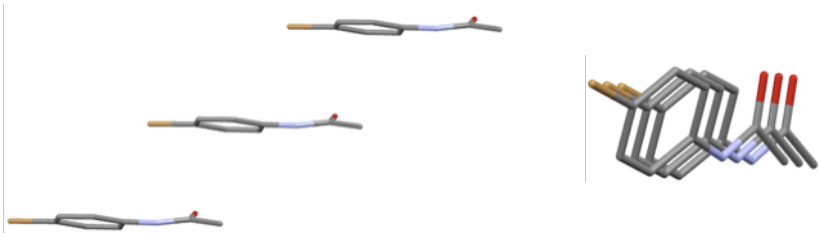


Figure 6.4: Chain **C1**$_{[g]}$ in **Y:p-Br**. Two perspectives are shown in each case: viewed perpendicular (top) and parallel (bottom) to the HB direction.

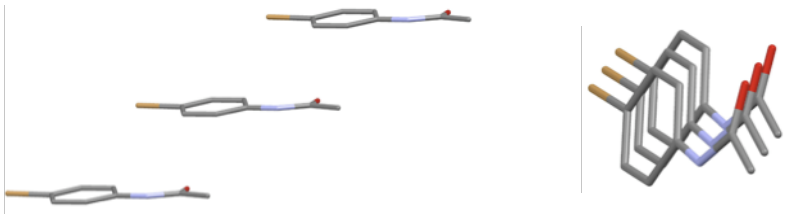
Figure 6.4 clearly shows the varying span of the V-shape characteristic for the HB chain **C1**$_{[g]}$ with the widest V occurring in **H:p-Br**. The only other 1D construct identified in this subset is the translational stack **S1**$_{[t]}$, which was found in **CF$_3$:p-Br** and the two polymorphs of **Me:p-Br**. Again this feature has already been discussed in Chapter 4.B. Figure 6.5 shows the stack for the three structures.



CF_3:p-Br



Me:p-Br(m)



Me:p-Br(o)

Figure 6.5: Stack **$S1_{[t]}$** in **CF_3:p-Br**, **Me:p-Br(m)** and **Me:p-Br(o)** each shown as viewed side on (left) and along the stack direction (right).

The combination of stack **$S1_{[t]}$** and **$C1_{[g]}$** produces the 2D layer **L7**, which was also found in the para acetanilide series. Here it is observed in **CF_3:p-Br** and **Me:p-Br(o)** and the layer is depicted for the former compound in Figure 6.6.

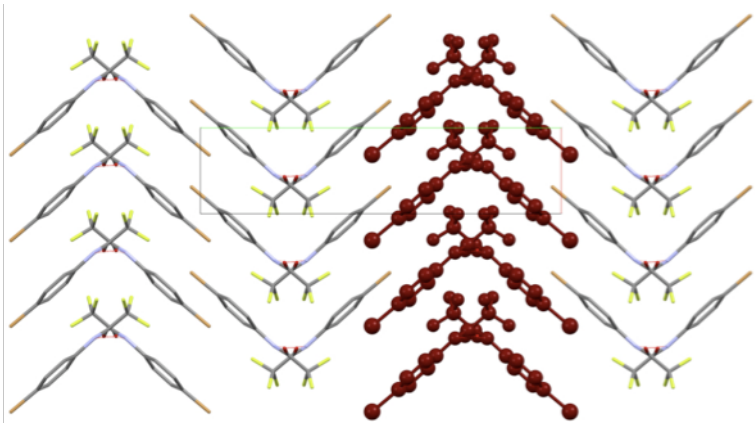


Figure 6.6: Layer **L7** in the crystal structures of **CF_3:p-Br** highlighted in red.

Finally, XPac identified layer **L4*** as common to **CF_3:p-Br** and **Et:p-Br**. The layer is highlighted in the packing diagrams in Figure 6.7 for both structures. This layer is closely related to layer **L4** introduced in Chapter 4 hence the similar identifier.

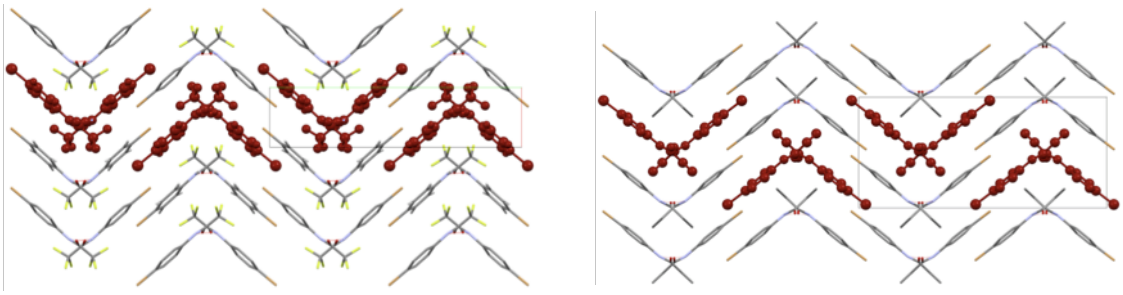


Figure 6.7: Layer **L4*** in the crystal structures of **CF_3:p-Br** (left) and **Et:p-Br** (right) highlighted in red.

Recalling from the previous discussion, layer **L4** contains the HB chain **$C1_{[g]}$** and neighbouring chains are related by 2_1 screw symmetry. In contrast to layer **L4** the spacing between neighbouring HB chains is widened in layer **L4***. Hence adjacent layers **L4*** are more interwoven than in the structures containing **L4**.

6.B. Persistence of HB motifs $C1_{[t]}$, $C1_{[2_1]}$ and $C1_{[g]}$ in the acylanilides

The predominance of the three HB chains $\mathbf{C1_{[t]}}$, $\mathbf{C1_{[2_1]}}$ and $\mathbf{C1_{[g]}}$ in the acetanilide library raised the question on how robust and transferrable these chains are to exchanging the amide substituent. Also of interest was to probe which end of the molecule has the greater influence on the HB geometry – the phenyl or amide end. For this reason the crystal structures were assessed on the presence of the three HB chains. This was done immediately after the basic crystal structure was established. The results of this search are summarised in Figures 6.8 to 6.10.

ORTHO	H:o-X	Me:o-X	Et:o-X	Pr:o-X	^{t}Bu:o-X	CF_3:o-X	OMe:o-X	OEt:o-X	NH_2:o-X
Y / X	H	CH_3	C_2H_5	C_3H_7	$C(CH_3)$	CF_3	OCH_3	OC_2H_5	NH_2
F									
Cl									
Br									
I									
CH_3									
CF_3									
CN									
NO_2									
COOH									
COMe									
NHCOY									
OH									
^{t}Bu									

Figure 6.8: Summary of occurrence of HB chains **C1** in ortho acylanilides. Colour coding is as follows: brown – $\mathbf{C1_{[t]}}$, yellow – $\mathbf{C1_{[2_1]}}$, red – $\mathbf{C1_{[g]}}$. Stars indicate that HB chain is part of more complex HB network.

META	H:m-X	Me:m-X	Et:m-X	Pr:m-X	^{t}Bu:m-X	CF_3:m-X	OMe:m-X	OEt:m-X	NH_2:m-X
X \ Y	H	CH_3	C_2H_5	C_3H_7	$C(CH_3)$	CF_3	OCH_3	OC_2H_5	NH_2
F									
Cl									
Br									
I									
CF_3									
CH_3									
C_2H_5									
CN									
NO_2									
NH_2									
OH									
OCH_3									
COOH									
COOMe									
COOEt									
COMe									
OCOMe									
NHCOY									
SF_5									

Figure 6.9: Summary of occurrence of HB chains **C1** in meta acylanilides. Colour coding is as follows: brown – $\mathbf{C1}_{[t]}$, yellow – $\mathbf{C1}_{[2_1]}$, red – $\mathbf{C1}_{[g]}$. Starred entries in **NH_2:m-X** column indicate that HB chain is part of more complex HB network. Two coloured starred entries elsewhere symbolise a mixture of HB chain motifs due to Z' > 1.

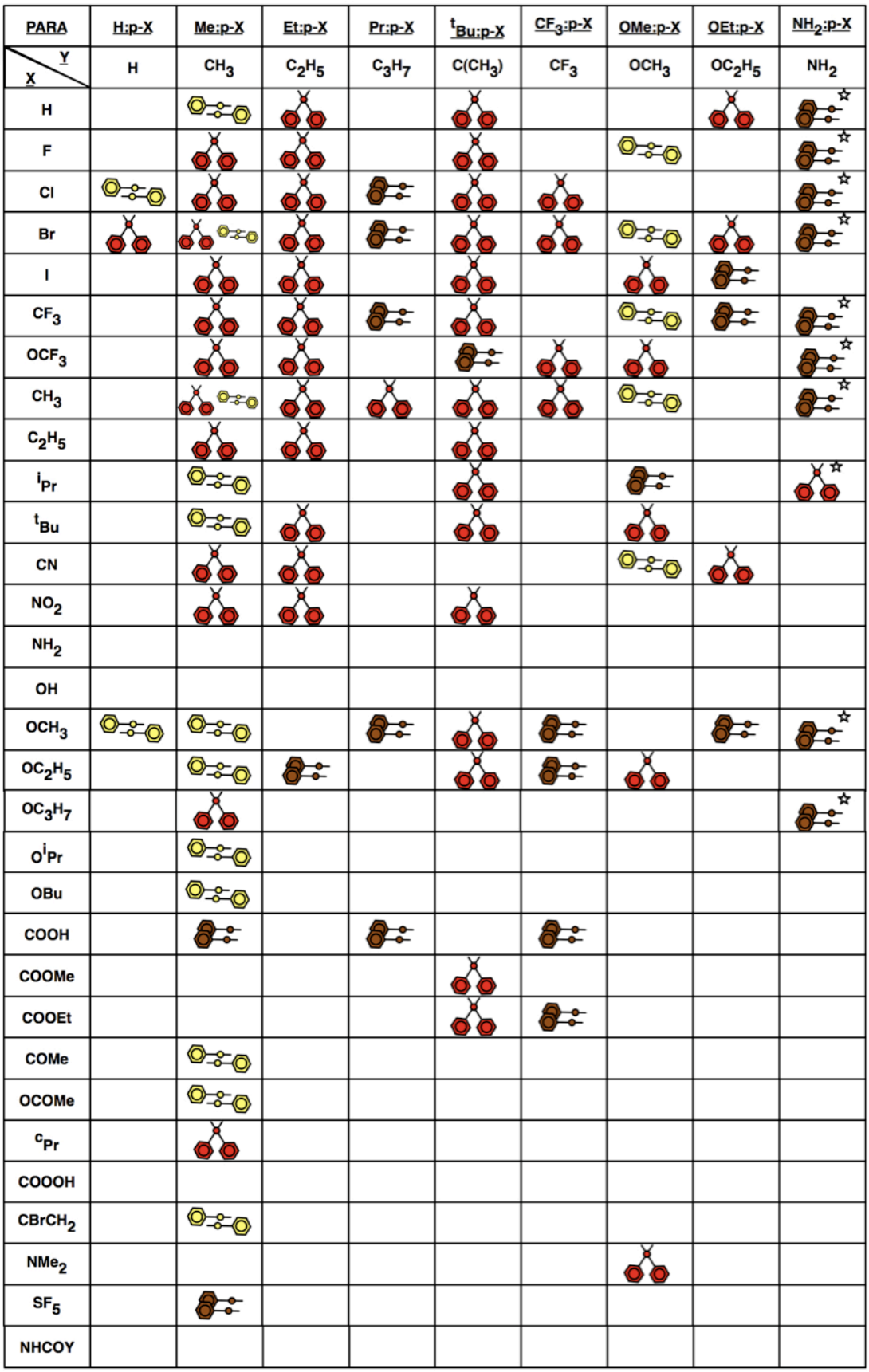


Figure 6.10: Summary of occurrence of HB chains **C1** in paraa acylanilides. Colour coding is as follows: brown – $\mathbf{C1}_{[t]}$, yellow – $\mathbf{C1}_{[2_1]}$, red – $\mathbf{C1}_{[g]}$. Stars indicate that HB chain is part of more complex HB network.

The symbols used in the figures show the basic geometry of the chains **C1** and are colour encoded according to the structural similarity plots, i.e. $\mathbf{C1_{[t]}}$ is shown in brown, $\mathbf{C1_{[g]}}$ symbols are drawn in red and $\mathbf{C1_{[2_1]}}$ are highlighted in yellow. The entries in the columns of **NH_2:s-X** are starred since the chains of type **C1** are part of a more complex HB network. Furthermore in the meta diagram some entries are combinations of the different geometries and colours (further highlighted with a star). These refer to structures with Z' = 2. Due to pseudosymmetry the HB chains in these structures are formed of dimers with the respective geometry and symmetry. For comparison the cases of **Me:m-Br(m)** and **Me:m-Br(o)** were discussed in Chapter 4.B.

From Figures 6.8 to 6.10 it can be seen that the distribution of the three HB chain motifs in the ortho acylanilides differs quite significantly from the meta and para compounds. The amide-amide HB observed most in the ortho acylanilides is the translational chain $\mathbf{C1_{[t]}}$, followed by $\mathbf{C1_{[g]}}$ and then $\mathbf{C1_{[2_1]}}$. In chain $\mathbf{C1_{[t]}}$ the intermolecular interaction is enhanced due to the overlap of substituent and adjacent phenyl ring as was shown in the discussion of the ortho acetanilides in Chapter 4.B. This favourable geometry is of course directly linked to the position of substitution on the phenyl ring explaining the increased occurrence of $\mathbf{C1_{[t]}}$ in the ortho acylanilides. In these compounds the position of substitution seems to be the primary driving force for the formation of the HB motif. There is little indication that the type of phenyl substituent **X** has any major effect on the HB geometry. However there are two groups of acylanilides in which the amide substituent breaks the HB geometry pattern. In the ortho trimethylacetanilides (pivalomides) **^{t}Bu:o-X**, the amide-amide hydrogen bonding geometry is almost exclusively that of chain $\mathbf{C1_{[2_1]}}$. In this series the substituent on the amide group appears to have a directing effect on the HB geometry. If incorporated in chain $\mathbf{C1_{[t]}}$ the *tert*-butyl groups would cause steric clashes between neighbouring molecules. Instead, in the 2_1 screw symmetry chain $\mathbf{C1_{[2_1]}}$ C-H···π interactions are facilitated between the amide substituent **^{t}Bu** and the adjacent phenyl ring, hence enhancing the intermolecular association in $\mathbf{C1_{[2_1]}}$. The ortho phenylureas **NH_2:o-X**, form the other group in which there is a clear preference of a different HB chain geometry than that of $\mathbf{C1_{[t]}}$. In the ortho phenylureas the HB chain $\mathbf{C1_{[g]}}$ is mainly observed.

An explanation of this behaviour is not so readily available however. The urea moiety is engaged in a more complex intermolecular hydrogen bonding network due to the additional HB donors in the molecule. Hence a direct comparison with the other acylanilides and their hydrogen bonding is not appropriate. It is however interesting that the familiar HB chain geometries are also observed as part of the more complex hydrogen bonding of these compounds. The same applies to the meta and para phenylureas.

In the other meta and para acylanilides the prevailing HB motif is that of chain **C1$_{[g]}$**. The V-shape geometry of this chain is clearly compatible with most compounds irrespective of the substitution on the phenyl ring and on the amide group. Especially the trimethylacetanilides **^{t}Bu:m-X** and **^{t}Bu:p-X** as well as the propionanilides **Et:m-X** and **Et:p-X** form the HB chain **C1$_{[g]}$** in almost all instances. Outside of the acetanilides **Me:m-X** and **Me:p-X** the HB chain **C1$_{[2_1]}$** is little observed. Only in the methylcarbamates **OMe:m-X** and **OMe:p-X** can chain **C1$_{[2_1]}$** be regarded as a common feature, but there is no clear pattern of occurrence with respect to the phenyl substituent **X** and its inductive or mesomeric effects. Such relationships are not really discernible for any of the acylanilide families. The para acetanilides **Me:p-X** are perhaps the only exception. But even in this set the structural data only provides an indication that compounds with electron withdrawing substituents **X** mainly form the HB chain **C1$_{[g]}$**, whereas chain **C1$_{[2_1]}$** is established in compounds with electron donating substituents **X**. There are exceptions and the polymorphic forms of **Me:p-Br** and **Me:p-Me** clearly thwart this rule of thumb. Hence the 'correlation' can only be understood as an indication of tendencies. Furthermore there are still plenty of gaps in the matrices in Figure 6.9 and 6.10, which need filling in before any definite rules can be established or discarded.

A brief comment regarding the formanilides **H:s-X**: the respective columns in Figures 6.8 to 6.10 are sparsely populated despite many more crystal structures having been determined for this compound family. The distinct lack of the three HB chains of type **C1** arises since the formanilides frequently crystallise with the amide group in the *cis* configuration. In those cases the intermolecular hydrogen bonding does not yield infinite chains but often catemeric rings are formed. Hence the formanilides present another independent group in terms of their structural

similarities and comparison with the other acylanilides is dependent on the amide group configuration.

6.C. Hydrogen bond propensities in the acylanilides

When probing the acylanilides for the presence of the HB chains of type **C1** it was noticed that phenyl substituents **X** with HB donor functionality disrupt the amide-amide hydrogen bonding in almost every instance. The only exception to this behaviour was observed when **X** = **COOH**, in which case complementary hydrogen bonding was found. In **Me:p-COOH**, **Pr:p-COOH** and **CF_3:p-COOH** the amide groups interact in the HB chain **$C1_{[t]}$** and the carboxylic acid groups form an inversion HB dimmer. In this way neighbouring chains **$C1_{[t]}$** are linked through the dimer and an overall ladder motif results.

On the other hand compounds containing substituents **X** with HB acceptor ability were often found to maintain the amide-amide hydrogen bond. One particularly interesting case is that of the polymorphic forms **Me:p-OCOMe(m1)** and **Me:p-OCOMe(m2)**. In the former crystal structure the HB network is based on amide-acetyl interactions whilst in the latter the HB chain **$C1_{[2_1]}$** persists.

Hence the question arises what the 'competing power' of these substituents is and how the amide-amide hydrogen bond can be ranked in comparison with the substituent-amide hydrogen bonding. For these reasons a collaboration was started with Peter Galek from the Cambridge Crystallographic Data Centre. This collaboration is aimed at calculating the HB propensity for the acylanilides with additional HB donor and acceptor groups using the predictive algorithm developed by Galek *et al*[1]. For these calculations the CSD was harvested for structures with similar chemical functionality as the acylanilides under investigation. This target set was subsequently analysed for the presence of intermolecular hydrogen bonding providing the parameters for the HB propensity models. Finally these models were applied to predict the likelihood of hydrogen bonding between the various HB functional groups in the acylanilides.

The preliminary results for two acylanilides are briefly presented here.

6.C.1. HB propensity in Paracetamol

The groups competing for hydrogen bonding in Paracetamol, **Me:p-OH**, are the amide group of the molecular backbone and the substituent hydroxyl group. In the crystal structures of the polymorphs hydrogen bonding was only observed between the hydroxy and amido groups, but in principle hydrogen bonding is also possible involving only amide or hydroxy groups. The propensity for all these combinations was estimated by Peter Galek with his predictive method. Table 6.1 summarises the results from these calculations.

HB donor	HB acceptor	π	+/– [a]	Me:p-OH(m)	Me:p-OH(o1)
N of amide	O of amide	0.971	0.085	×	×
O of hydroxy	O of amide	0.937	0.038	✓	✓
N of amide	O of hydroxy	0.428	0.103	✓	✓
O of hydroxy	O of hydroxy	0.247	0.027	×	×

Table 6.1: Hydrogen bonding propensities for two polymorphs of paracetamol.

The predictions rank the amide-amide hydrogen bond as the most probable HB interaction in Paracetamol, yet this hydrogen bond is not observed in any of the known crystal structures of **Me:p-OH**. Instead the hydroxyl group acts as the HB donor for the amide oxygen and as HB acceptor for the amide nitrogen as shown in Figure 6.11.

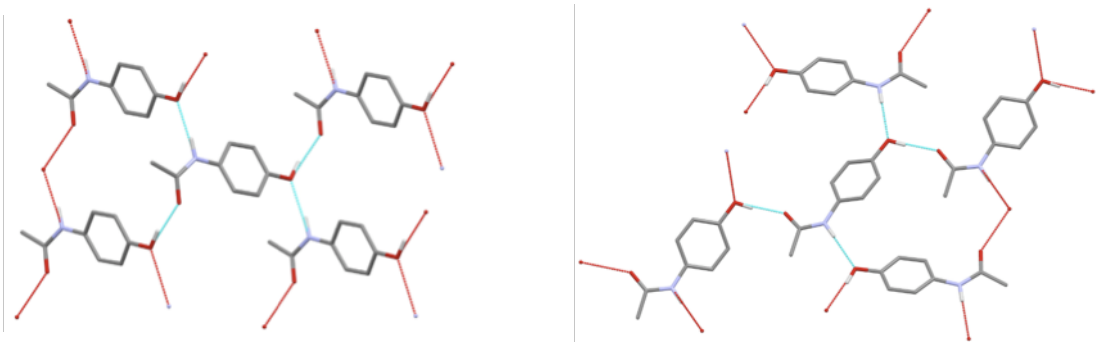


Figure 6.11: HB network in the paracetamol polymorphs shown for **Me:p-OH(m)** on the left and **Me:p-OH(o1)**on the right hand side.

The first of these interactions is also highly probable according to the predictions but the propensity of the second interaction is ranked much lower. Further work

in this respect should be aimed at understanding why the highly likely amide-amide hydrogen bond is not observed in these structures.

6.C.2. Hydrogen propensity in the polymorphs of Me:p-OCOMe

The second case is also concerned with a member from the para acetanilide series. In the case of **Me:p-OCOMe** the acetyl and amide groups are potential hydrogen bonding partners and the propensity for all possible interactions was estimated by Peter Galek. Table 6.2 lists the results from this prediction.

HB donor	HB acceptor	π	+/–	Me:p-OCOMe(m1)	Me:p-OCOMe(m2)
N of amide	O of amide	0.793	0.082	×	✓
N of amide	O of acetyl	0.262	0.146	✓	×
N of amide	O of acetyl	0.032	0.183	×	×

Table 6.2: Hydrogen bonding propensities for **Me:p-OCOMe(m1)** and **Me:p-OCOMe(m2)**.

As for Paracetamol, the amide-amide hydrogen bond is estimated as the most probable amongst the possible interactions. In fact its propensity is much higher than any of the other interactions. Two polymorphic crystal structures are known of **Me:p-OCOMe**. The structure of **Me:p-OCOMe(m1)** has been published previously and it exhibits the less probable HB interaction between the amide nitrogen HB donor and the carbonyl oxygen HB acceptor of the acetyl group. During this work the second polymorph **Me:p-OCOMe(m2)** was discovered and in this structure hydrogen bonding is established between the amide groups only forming the chain $\mathbf{C1}_{[2_1]}$. According to the propensity predictions this interaction is the more probable and hence the discovery of the polymorph is not very surprising. However, calculation of the intermolecular interaction energies reveals that the amide-amide hydrogen bond is slightly less stable at -35.4 kJ·mol^{-1} than the amide-acetyl HB interaction with -37.7 kJ·mol^{-1}.

The alternative HB connectivity is shown in Figure 6.12 for both polymorphs.

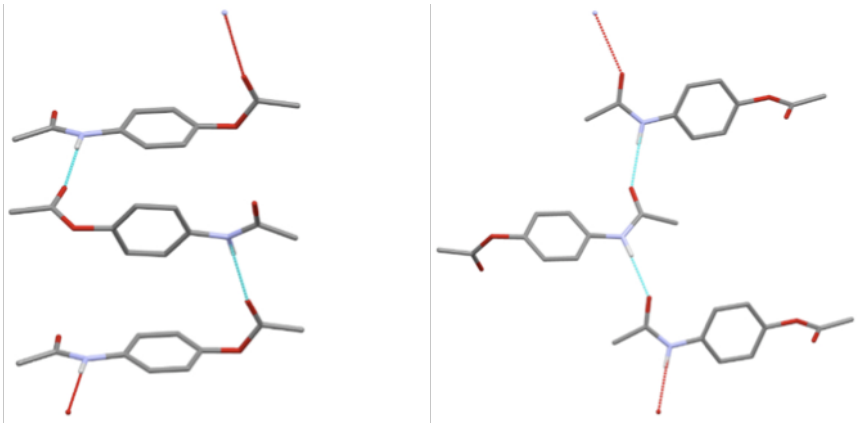


Figure 6.12: Hydrogen bonding in **Me:p-OCOMe(m1)** (left) and **Me:p-OCOMe(m2)** (right).

A third polymorphic form has been alluded to in the literature but no crystal structure is known to date[2]. Further investigations concerning this compound should be aimed at analysing the thermodynamic characteristic of the polymorphs.

6.D. Crystallisation trends of the acylanilides

The samples for this work were all synthesised from the respective aniline in an acylation reaction and then crystallised mainly from ethanol. The syntheses and crystallisations were all performed by Terry Threlfall over the course of one year. Given the size of acylanilide compound library this was a truly gargantuan effort. However, as a 'side product' of the sample preparation it was possible to observe certain trends in the crystallisation behaviour of the different compounds. These observations are summarised in a recent publication by Threlfall *et al*[3] and the reader is referred to this paper for a more thorough discussion of the findings. Here the key trends are only summarised briefly. During this project it became apparent that there is some correlation between crystallisation time and the type of acylanilide produced. Crystallisation time in this context is to be understood as the time it took for the first visible crystals to appear. A distinction could be made based on the position of substitution **s** on the phenyl ring as well as based on the type of acylanilide, i.e. the amide substituent **Y**. It was found that the ease of crystallisation decreased as going from para to meta with the ortho acylanilides having intermediate crystallisation times. In terms of the type of acylanilide it was observed that the crystallisation time increased with decreasing spherical character of the residue **Y**. In other words, as the conformational flexibility of the amide substituent **Y** increased the crystals took longer to form. This is reflected to some extend in the compound matrices presented in Chapter 2. In fact some of the butyrylanilides **Pr:s-X** and ethylcarbamates **OEt:s-X** are still crystallising.

The nature of the phenyl substituent **X** appeared to have less of an impact on the crystallisation time. However, it was noticed that crystallistion slowed down if **X** has hydrogen bond functionality.

It should be mentioned that this study was not aimed at investigating the crystallisation behaviour in the slightest and hence no appropriate experiments were performed to confirm any relation between the conformational freedom of the acylanilide and the crystallisation time. Future efforts will be directed at monitoring the onset of crystallisation and measuring half-zone widths to get accurate values for any correlation studies.

6.E. Other crystallographic phenomena

During the data collection process a number of crystallographic phenomena were observed, which can be linked to the type of acylanilides under investigation. For example the crystal hardness was noticed to reduce as the flexibility of the amide substituent **Y** increased. This goes in line with the crystallisation trends mentioned in the previous section. Generally the single crystals of the butyrylanilides and ethylcarbamates were found to be much softer than those of the acet-, trimethylacet- and trifluoroacetanilides. Also the crystals could not be cut easily but instead splaying was observed. In some cases liquid crystalline phases were observed under the polarizing microscope where cutting of the crystal had been attempted. Furthermore some of the crystal structures of these acylanilides were modulated. Figure 6.13 displays a precession image clearly showing the modulation in the compound **Pr:p-Cl**.

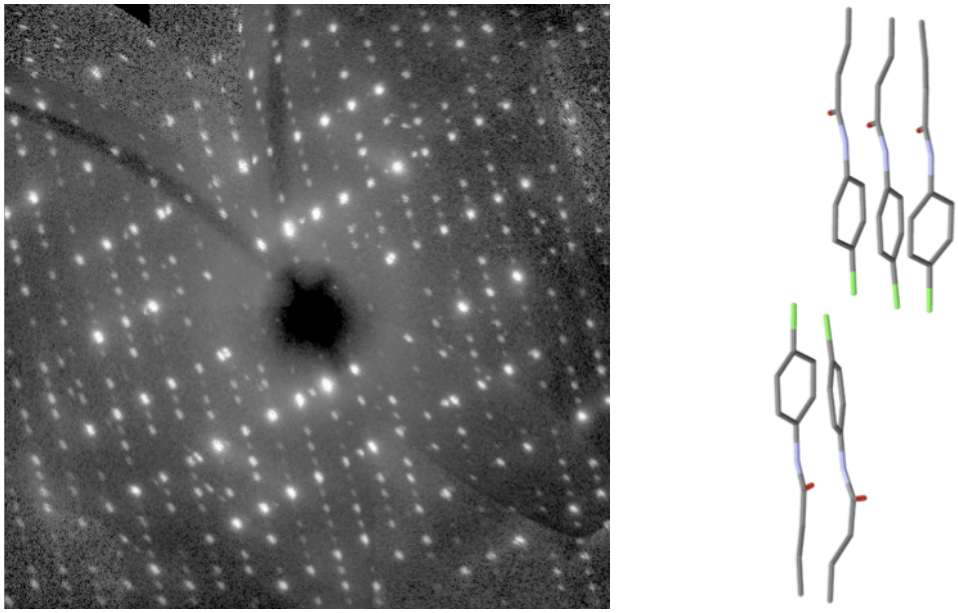


Figure 6.13: Precession image of the modulated crystal structure of **Pr:p-Cl**. The content of the ASU is also shown.

This shows that the crystal structures of these compounds are not only slow to form, but they contain irregularities to the periodic assembly and their stability is low.

Twinning was often encountered for the carbamates and formanilides indicating packing problems in these compounds. Figure 6.14 shows a typical precession image of a twinned sample.

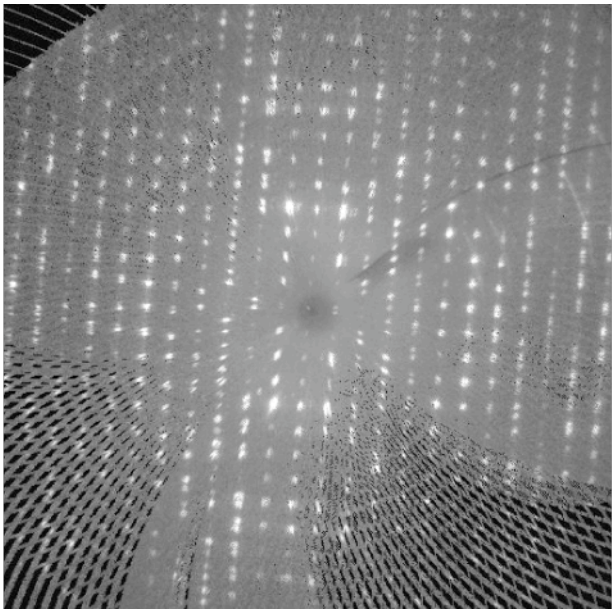


Figure 6.14: Precession image of a twinned sample.

The formanilides in particular produced several structures with more than one independent molecule in the ASU. This increase in Z' was frequently accompanied by cis/trans isomerism of the amide group in the formanilides meaning that both isomers were found in the crystal. The propionanilides also produced several Z' = 2 structures. Here the molecules in the ASU differed in the conformation of the amide substituent as shown in Figure 6.15 overleaf.

Lastly, disorder was observed in several structures, but particularly those compounds containing a fluorinated phenyl ring or those crystallising in space group Cc appeared prone to disorder of the phenyl ring or amide group, respectively. A precession image containing diffuse scattering indicative of the disorder in the crystal is displayed for **Me:p-Cl** in Figure 6.16 also overleaf.

As was the case for the crystallisation trends, these observations were a side product of the project and the sheer number of crystal structure determinations. Future work would be concerned with both crystallographic and thermodynamic experiments to probe the link between the type of compound and irregularities in the crystal structure.

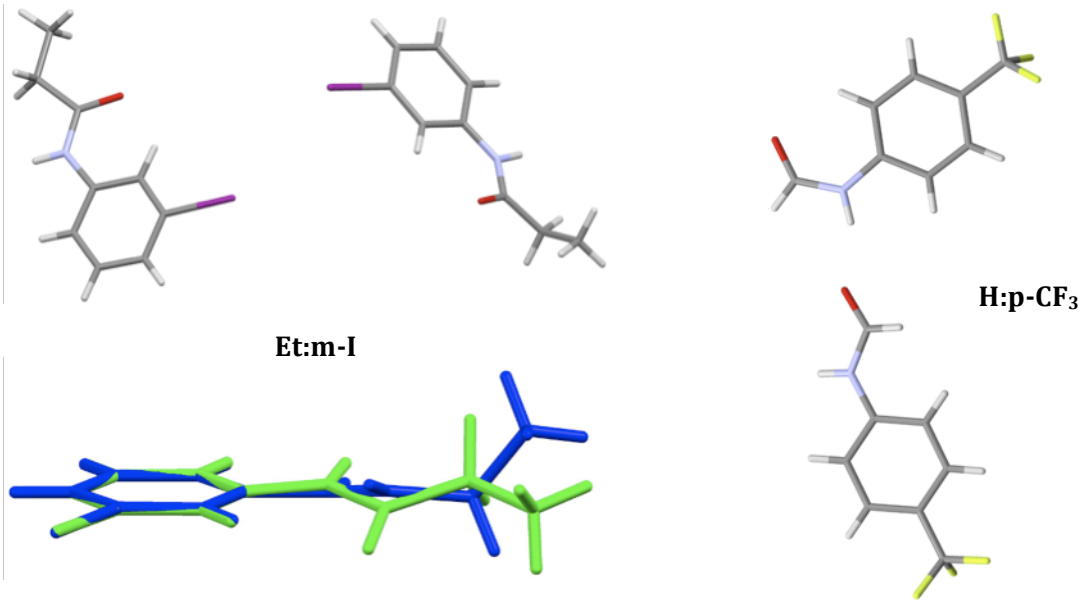


Figure 6.15: Two examples of Z' = 2 structures. ASU shown for **Et:m-I** (left) and **H:p-CF_3** (right). Difference in amide chain conformation in **Et:m-I** shown as molecular overlap in green and blue (bottom left).

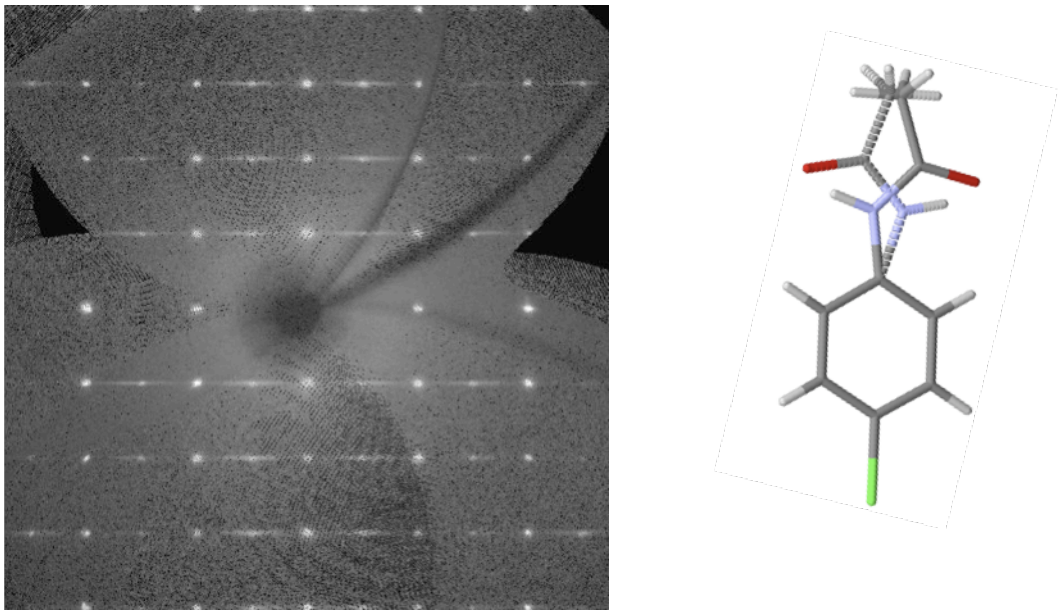


Figure 6.16: Precession image containing diffuse scattering and the corresponding disordered structure of **Me:p-Cl** (disordered part drawn with dashed bonds).

6.F. Future work

A few possibilities for further work on the acylanilides have been mentioned in the preceding sections, but this list can be easily extended. Again, it is the size and diversity of the compound library that lends itself to and inspires a variety of different studies. In the following a few suggestions for future work are briefly summarised.

1) Future work will be concerned with the structural similarity of the acylanilides other than the acetanilide series. Crystal structure comparisons need to be carried out not only for each individual acylanilide family but also across the series as was demonstrated in Section A of this chapter.

2) Gaps in the compound matrix remain to be filled in. Some of the reactions failed to solidify or yield the desired compound. In other cases samples were produced as amorphous or powder material. Efforts should be directed at producing single crystals by modifying the synthesis step and the crystallisation process.

3) The compound library can still be extended to include other substituents **X** on the phenyl ring or **Y** on the amide group. This should be done systematically so that the understanding of the influence of the respective substituents on the crystal packing and on hydrogen bonding behaviour can be increased.

4) Polymorphism has been observed in several acylanilides albeit only by serendipity. The results from the analysis of the amide-amide hydrogen bonding and the HB propensity calculations suggest that polymorphism may be much more widely spread in these compounds and a polymorph screen should be performed. Furthermore existing samples of polymorphs should be used for phase transition tests and thermal analysis.

5) The robustness of crystal structure types with respect to substitution could be assessed with crystal structure predictions. Such work could also aid the search for polymorphs. For example there have been indications that para-chloro acetanilide may exist in two polymorphic modifications in analogy to para-bromo and para-methyl acetanilide. However, so far all attempts at finding a monoclinic polymorph have been unsuccessful. A full crystal structure prediction would produce a lattice energy landscape on which potential polymorphic forms are

identifiable. Such alternative crystal structures can then be compared with already known structure types from the acetanilide set to probe for similarities. Furthermore the energy landscape would provide an indication as to the likelihood of observing such a polymorph under standard conditions.

6) The influence the substituent **X** has on the conjugation between the phenyl ring and the amide group is not yet well understood. Analysis of the electronic structure may elucidate how the electron density is distributed across the molecule, which in turn may provide more information on the conformational preferences of the acylanilides. The electron density can be determined with high resolution single crystal X-ray diffraction and charge density data sets have been collected for a number of acylanilides and currently await processing.

Of course this list is by no means exhaustive and many more studies could be devised for the acylanilides. However, they offer a good starting point for the more detailed study of structural similarities and any correlation of molecular and crystal structure.

6.G. References

[1] P. T. A. Galek, L. Fábián, W. D. S. Motherwell, F. H. Allen, N. Feeder, *Acta Crystallographica*, **B63**, 768-782 (2007).

[2] D. Bhalla, J. Lalla, *Drug Development and Industrial Pharmacy,* **16**(1), 115-135 (1990).

[3] M. B. Hursthouse, L. S. Huth, T. L. Threlfall, *Organic Process Research & Development*, **13**, 1231-1240 (2009).

CHAPTER 7: SUMMARY & CONCLUSIONS

This work was aimed at gaining a better understanding of the Aufbau principle of crystal structures of molecular compounds. For this reason a large array of closely related mono-substituted acylanilides was synthesized and the corresponding crystal structures determined. To date the compound library comprises 255 crystal structures, of which 215 were obtained during the course of this project (the rest being harvested from the CSD). The influence of the molecular substitution pattern on the crystal assembly was studied in detail for a subset of 64 mono-substituted acetanilides. Of these 27 new crystal structures were presented as part of this thesis. The crystallographic data for these and any redetermined structures have been made publicly available on the eCrystals archive.

The study of the acetanilides entailed crystal structure comparisons to establish structural similarity relationships. The XPac algorithm was used for the identification of common packing patterns. According to the position of substitution on the phenyl ring the acetanilide family can be further divided into three series: ortho, meta and para, populated with 14, 15 and 35 structures, respectively. Similarity searches were performed for each individual set as well as for the whole family of acetanilides. Lattice energy calculations were used to assess the stability of the 1D common packing features. Furthermore the energy calculations provided insight into the importance of these features in the crystal assembly.

The structural similarity search revealed that acetanilides without additional HB functionality form amide-amide HB chains of three geometries. Their interaction energy is in the expected range of the classical strong hydrogen bond and they account for approximately one quarter to one third of the total lattice energy. The translational chain $\mathbf{C1_{[t]}}$ is found most frequently in the ortho acetanilides, where intermolecular adhesion is enhanced due to substituent-π interaction. This chain does not occur in the meta series and was only detected in two instances in the

para acetanilides. Instead the chains $\mathbf{C1_{[2_1]}}$ and $\mathbf{C1_{[g]}}$ persist in these two series. Amongst the para acetanilides a trend is discernible for certain substituents to prefer one of these two chains. Small, spherical or highly electron withdrawing substituents seem to align towards the HB chain with glide symmetry ($\mathbf{C1_{[g]}}$), whereas acetanilides with alkane and alkoxy substituents seem to preferentially adopt the 2_1 screw symmetry chain $\mathbf{C1_{[2_1]}}$. There are exceptions to this rule, in particular **Me:p-Br** and **Me:p-Me** are able to form both chains resulting in polymorphic modifications. The energy differences between the two chains were too small to allow a conclusive assignment of the more stable arrangement.

The robustness of these three amide-amide HB chains was put to test by introducing substituents with additional HB functionality. It was found that HB donors disrupt the amide hydrogen bonding in every instance, except for **p-COOH**, which engages in a carboxylic acid HB dimer interaction. Interestingly HB propensity calculations estimate that amide-amide hydrogen bonding should still be the most probable interaction even if competing HB donors are present in the structure.

On the other hand HB acceptors had different effects on the intermolecular hydrogen bonding. Especially small HB acceptors did not seem to affect the interaction between the amide groups. Interestingly the polymorphs of **Me:p-OCOMe** contain the two alternative HB interactions possible for the compound: the amide-amide HB chain $\mathbf{C1_{[2_1]}}$ is observed in **Me:p-OCOMe(m2)** whilst amide-acetoxy hydrogen bonding is present in **Me:p-OCOMe(m1)**. Although HB propensity calculations predict a higher likelihood for the amide-amide interaction, lattice energy calculations assign a slightly higher energy gain to the amide-acetoxy hydrogen bond. However, the differences are small, making the observation of polymorphism understandable.

In addition to the amide-amide hydrogen bonds another construct was observed for many meta and para acetanilides – the translational stack $\mathbf{S1_{[t]}}$. This geometrical arrangement is driven by the assembly of the aromatic part of the molecule. The frequency of occurrence of this feature demonstrates that the stacked packing of π-systems is an important factor in the crystal assembly. Energy calculations rank this geometrical arrangement as a secondary packing feature, with interaction energies worth approximately one third of that of the amide

hydrogen bonds. However, it was also noted that, in particular, the combination of stack $\mathbf{S1_{[t]}}$ and the HB chain $\mathbf{C1_{[2_1]}}$ appears as unfavourable. This may be due to the molecular conformation of the phenyl ring in the HB chain, where the twisted conformation may be incompatible with face-to-face stacking of the phenyl rings.

During the similarity search in the para-acetanilide series, indications of a correlation of molecular conformation with the amide-amide HB chain geometry became apparent. Hence for compounds with dihedral angles greater than 30˚ the translational chain $\mathbf{C1_{[t]}}$ was the preferred geometry, whilst the compounds with dihedral angles smaller than 12˚ were found to adopt the glide symmetry chain $\mathbf{C1_{[g]}}$. The 2_1 screw symmetry chain $\mathbf{C1_{[2_1]}}$ was observed for those compounds with intermediate dihedral angles. In order to assess conformational preferences, a theoretical conformational analysis was performed for the acetanilides. The results showed rotational barriers could only be related to mesomeric and inductive substituent effects for the para series. Intramolecular hydrogen bonding in the ortho series leads to a significant stabilization of the molecular conformation in comparison with the meta and para series. Generally the theoretical calculations predicted a planar conformation as the most energetically stable for all three series. In the ortho and meta series this corresponds to the *anti* geometry of the substituent with respect to the amide oxygen atom (i.e. $\alpha = 180^{\circ}$). This distinction is not directly transferrable to the para series due to the mirror symmetry of the phenyl ring. In the case of para amino acetanilide, the theoretical calculations showed that the energy penalty associated with a slight deviation from planarity is off-set by relaxation of the substituent into its preferred geometry. For compounds with alkoxy substituents it was found that the conformations corresponding to $\alpha = 0^{\circ}$ and $\alpha = 180^{\circ}$ are not equivalent as for the rest of the para acetanilides. Instead the calculations show that the configuration in which the lone pairs of the amide and substituent oxygen are positioned on opposite sides of the molecule (*anti*) is energetically more stable.

In the crystal structures the acetanilide molecules were found to be non-planar throughout. Deviations from planarity were largest in those ortho acetanilides where intramolecular hydrogen bonding was not possible. Whilst the largest torsion angle in **Me:o-^{t}Bu** can be ascribed to steric effects of the substituent, the distortion in the other ortho compounds is stabilised by intermolecular

interactions in the crystal. Experimentally, both *syn* and *anti* configurations were found in the meta acetanilides, but this variation could not be correlated with any substituent effects. Again, the molecules are non-planar in the crystal and it is proposed that the molecular conformation is affected by the intermolecular assembly and results from the maximisation of intermolecular interactions.

The molecular conformation of the para acetanilides was found to deviate the least from planarity, although large torsion angles were observed in a few compounds. The correlation of Hammett constants with the rotational barriers indicates that the conjugation throughout the whole molecular backbone is most effective in these compounds, i.e. the mesomeric and inductive properties of the para substituent influence the conformation at the other end of the molecule. For the para alkoxy derivatives the experimental configurations vary between *syn* and *anti* without a clear link to any substituent effects. Again it is concluded that intermolecular interactions may have some influence on the molecular conformation in the crystal.

Although the experimental torsion angles were found to deviate from the theoretical optimum, they could still be considered as located within the energy wells calculated. Energy differences between the conformations in the crystal structures and the theoretical minima turned out to be negligible or small. Hence it can be concluded that the overall stability of the crystal is not dominated by molecular conformation. The computation of the pairwise interaction energies has shown that the intermolecular assembly can easily compensate for any energy penalties associated with molecular conformation. The conformational analysis has in so far not provided deeper insight into the correlation of molecular structure with hydrogen bonding features. However, the theoretical results and the comparison with experimental findings indicate conformational flexibility of the acetanilides and indicate that polymorphism may be much more widely spread in this family of compounds than so far observed.

Overall this study has shown that the crystal structures of the mono-substituted acetanilides are dominated by intermolecular hydrogen bonding. In the structures where amide-amide hydrogen bonding persist, the crystal structure can be understood as the 3D assembly of these one-dimensional features, which limits the structural diversity considerably. Hence many acetanilides with seemingly very

different substituents are found to assume analogous packing. The three amide-amide HB chain geometries are in principle interconvertible through rotation of adjacent molecules about the hydrogen bond. Whether this conversion is possible in the solid state or only in the liquid phase, is a question that remains to be addressed in the future. Some of the polymorphic systems could certainly be used for such an investigation since they prove that one compound can assume alternative HB chain geometries.

Whilst the amide-amide HB chains can be understood as primary packing features, the off-set stacks are considered as secondary constructs. In contrast to the HB chains, the stack geometry constitutes efficient packing of differently shaped molecules. If the molecular conformation is favourable, the stacked geometry is combined with the HB chains, but the energetic gain of the stack can be achieved by alternative intermolecular arrangements.

The importance of the amide-amide HB chains for the assembly of anilide derivates has become even clearer when considering the whole library of mono-substituted acylanilides. Again all three geometries are present in most compounds, but it was found that the type of acylanilide has a greater effect on the HB geometry than the phenyl ring substituents. The V-shaped HB chain $\mathbf{C1_{[g]}}$ is observed in most structures showing that this feature provides efficient packing for many varied molecular shapes.

So far this study has but scratched the surface of the packing behaviour of the mono-substituted acylanilides. Whilst the identification of clear preferences for intermolecular association and the limited structural diversity is encouraging, the structural similarity search needs to be extended to the full compound library in order to obtain a better understanding of the role of the molecular structure in the crystal assembly. With such a wealth of samples and crystallographic data a large variety of investigations and experiments can be derived for the future exploration of these compounds.