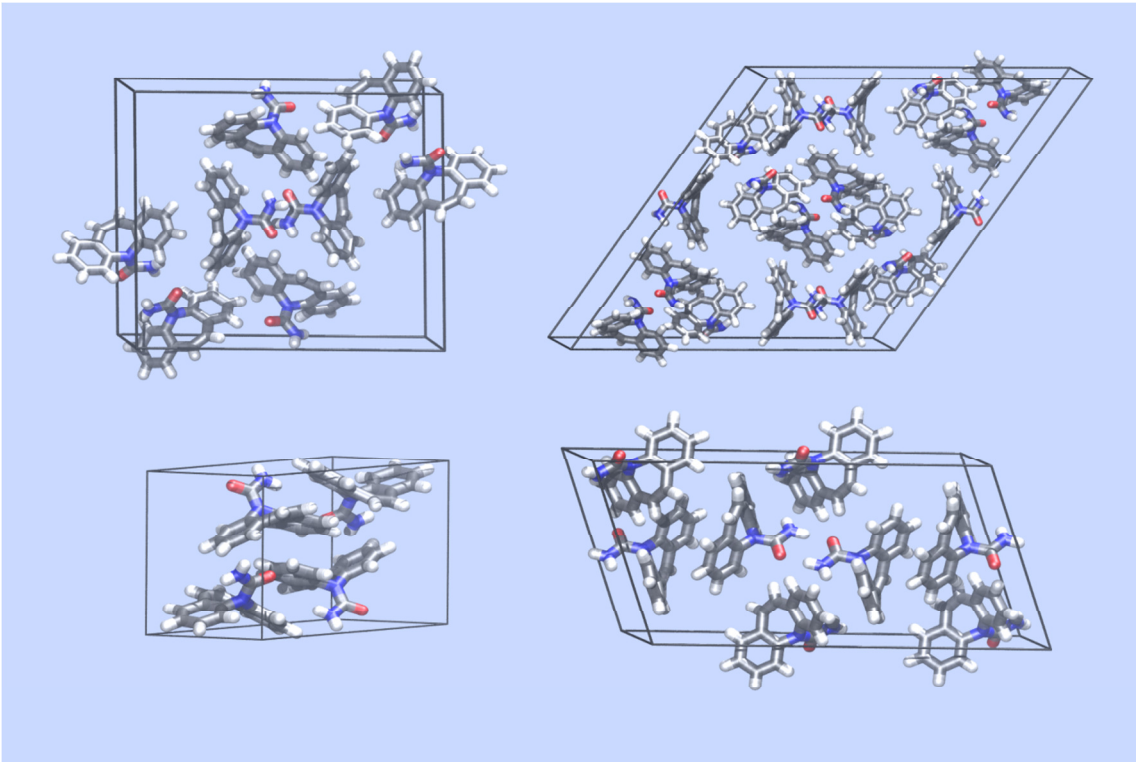

Evaluation of crystalline solid properties

Experimental and theoretical comparison of polymorphs, solvates and co-crystals

Master's Thesis within the Applied Physics Masters Programme

ALEXANDER LILJEBLAD

Chalmers University of Technology, Gothenburg
AstraZeneca R&D, Mölndal
Sweden, 2015

Evaluation of crystalline solid properties
Experimental and theoretical comparison of polymorphs, solvates and co-crystals
ALEXANDER C.E. LILJEBLAD

Department of Applied Physics
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 10 00

In collaboration with AstraZeneca AB
SE-431 83 Mölndal
Sweden
Telephone: +46 (0)31-776 10 00

Cover: Crystal structures of four polymorphs (I, II, III & IV) of the drug molecule carbamazepine used in this project. Elements are carbon (grey), hydrogen (white), nitrogen (blue) and oxygen (red) See section 2.1 for more information.

Printed at Chalmers Reproservice
Gothenburg, Sweden 2015

MASTER'S THESIS

Evaluation of crystalline solid properties
Experimental and theoretical comparison of polymorphs, solvates and co-crystals

Master's Thesis within the *Applied Physics* programme

ALEXANDER LILJEBLAD

SUPERVISORS
Boissier C., Broo A., Lindkvist M., Nilsson Lill S.
AstraZeneca R&D, Mölndal

EXAMINER
Aleksandar Matic
Chalmers University of Technology

Department of Applied Physics
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2015

Evaluation of crystalline solid properties
Experimental and theoretical comparison of polymorphs, solvates and co-crystals

ALEXANDER C.E. LILJEBLAD
Department of Applied Physics
Chalmers University of Technology

ABSTRACT

Being able to predict solubility has been a long standing goal for the pharmaceutical industry. The aim of this project was to evaluate whether melting temperatures and solubilities of drug molecules could be predicted with high accuracy. Crystals (polymorphs, co-crystals and solvates) of indomethacin and carbamazepine were produced experimentally and characterized by X-ray diffraction, differential scanning calorimetry (melting temperature) and thermal gravimetric analysis. Dissolution rates in water were determined by absorbance measurements. Surface morphology and size distribution were investigated by scanning electron microscopy and laser diffraction measurements. To assess differences in accuracy between simulation methods, simulations were performed on a benchmark set ($C19_{RT}$) consisting of 19 organic molecular crystals. Crystalline geometries and lattice energies of the benchmark set were evaluated. The Grimme (D3) dispersion method was found to give the highest accuracy. A correlation between sublimation free energy and melting temperature was found. Entropy simulations were performed to obtain Gibbs' free energies. Correlations linking entropy and zero point energy to the unit cell volume were found. Solvation energies were calculated by simulating hydration free energies utilizing a thermodynamic cycle. Solubilities were then calculated from the solvation free energies of the benchmark set. For the simulations of indomethacin and carbamazepine crystals, the correlations found for entropy and zero point energy were used. Melting temperatures were predicted from simulated sublimation free energies using the previously found correlation and compared to generated experimental data. A weak correlation was found. Calculated solubilities of the indomethacin and carbamazepine crystals did not fully agree with experimental data, indicating the need for development of more accurate simulation methods.

Keywords: polymorph, co-crystal, solvate, multi-component crystal, solubility, dissolution, indomethacin, carbamazepine, active pharmaceutical ingredient, drug substance, DFT, simulation, benchmark

ABBREVIATIONS

API	Active Pharmaceutical Ingredient
BCS	Biopharmaceutics Classification System
BJ	Becke-Johnson
CSD	Cambridge Structural Database
$C19_{RT}$	Data set consisting of 19 organic molecular crystals
CBZ	Carbamazepine
DFT	Density Functional Theory
DSC	Differential Scanning Calorimetry
FDA	US Food and Drug Administration
GGA	Generalized Gradient Approximation
IND	Indomethacin
LDA	Local Density Approximation
LSDA	Local Spin Density Approximation
MM	Molecular Mechanics
PBE	Perdew-Burke-Ernzerhof exchange-correlation functional
PSD	Particle Size Distribution
QM	Quantum Mechanics
$RMSD_{15}$	Root Mean Standard Deviation (15 molecules)
SEM	Scanning Electron Microscopy
TGA	Thermal Gravimetric Analysis
TS	Tkatchenko-Scheffler
vDW-DF2	van der Waals Density Functional 2
XRPD	X-ray Powder Diffraction
ZPE	Zero point Energy

CONTENTS

1. INTRODUCTION

Medicines have been a very important factor for public health and welfare during the last centuries. When the first modern medicines were discovered the concept of clinical trials did not exist, therefore making the drug development process very quick. An example would be when Dr. Edward Jenner during the 18th century discovered a vaccine against smallpox. He inoculated his caretakers' son with cowpox under the hypothesis that it might induce immunization against smallpox. He then tried to infect the boy with smallpox several times. All infection attempts were unsuccessful, allowing him to draw the conclusion that his vaccine worked. As this example also shows, there were no regulations or ethical frameworks for what could be tested on human patients. Patient safety has improved since, with the most advances being made in the 20th century. This has however increased the time necessary for developing drugs.[1]

Drug development during the early years of the pharmaceutical industry was often focused on analyzing well known medicinal herbs in order to find a therapeutic agent which could be extracted and synthesized. This technique is still used today, although less popular.[2] Drug discovery today (disregarding the case of antibodies or proteins produced by biological systems) relies on large scale screening. A target of high interest is chosen (*e.g.* receptor, protein etc.) and matched against a library of substances. The library may contain on the order of 10^5 different compounds. The most potent molecules from such screenings are selected for further analysis. These molecules, termed new molecular entities (NME), often exhibit low aqueous solubility. This represents a major challenge during drug development, since low solubility results in low uptake rates for the molecule. A drug molecule having high solubility may allow faster development of its delivery system (pill, liquid, inhaler etc.), since focus can be diverted from solubility issues.[1]

There exist a couple of different strategies for increasing the solubility of an active pharmaceutical ingredient (API). These strategies includes salt creation,[3] crystallization in different polymorphic states,[4] co-crystallization (crystallization with another molecular entity into a joint crystal),[5] utilizing amorphous phases among other techniques. To experimentally determine all possible salts, polymorphs or co-crystals for an API and subsequently select suitable candidates with desirable solubility is next to impossible. As a way of aiding or guiding the experiments, computer modeling (*in silico*) can be done. The main problem is that simulation of various properties results in varying degrees of accuracies. Computational modeling and simulation of the screening process in an accurate way would save both time and resources as well as increase the scientific knowledge allowing informed project decisions to be made.

1.1 AIM

In silico studies of solubility are mainly focused on obtaining an absolute value, as accurate as possible [6, 7]. The aim of this project is directed towards determining solubilities (relative and absolute) as well as dissolution rates for a system of polymorphs and multi-component (here defined as solvates and co-crystals) crystals. Possible connections between solubility, dissolution rate, melting temperature, lattice energies and sublimation enthalpies will be explored. The end goal is to be able to predict physical properties such as melting temperatures and solubility with a high accuracy using the determined connections. To achieve this, experimental and simulated

(from first principles) data will be generated for two APIs (indomethacin and carbamazepine) that are subsequently evaluated and compared.

2. BACKGROUND

In order to understand the different techniques and methods and what differentiates them from each other, it is important to understand the underlying theory. A deep knowledge of the theory allows for greater understanding of limitations in the results obtained.

2.1. MODEL SYSTEMS

The molecules that were chosen for this study are indomethacin and carbamazepine, due to the facts that they can form multi-component crystals and are well studied in literature.

The drug indomethacin (IND) is a non-steroidal anti-inflammatory drug discovered in 1963 and used against pain, fever, stiffness and even swelling.[8] It works by inhibiting the production of prostaglandin, a hormone which regulates - among many other functions - fever and inflammation. The structure of indomethacin can be seen in Figure 1. Interest in the drug has seen a surge due to discovery of different polymorphs, co-crystals and solvates. This creates the possibility to explore what effect polymorphism and solvates have on various properties, for example processability, stability and dissolution rate.[9]

Carbamazepine (CBZ) is an anticonvulsant drug and is primarily used for treating seizures. Like IND, interest in CBZ has increased which can be attributed to the fact that it has been reported in numerous polymorphs, solvates and co-crystals forms. The number of confirmed unique multi-component crystals of CBZ is approaching 100. The structure of the drug can be seen in Figure 2.

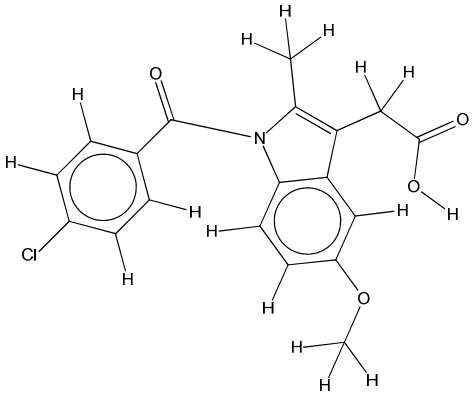

Figure 1. Structure of indomethacin.

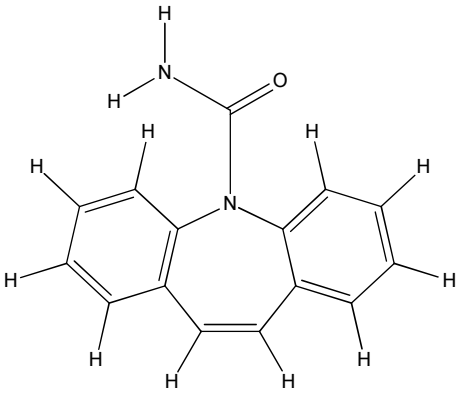

Figure 2. Structure of carbamazepine.

2.1.1. Crystal structures

For a crystalline material it is important to know the crystal structure. A crystal is made up of identical unit cells, defined by three lattice parameters (a, b, c) and three angles (α, β, γ). The unit cell creates a three-dimensional volume in which a number of molecules can be placed.

Lattice, or cohesive energy, is a measure of how much energy is gained when a crystal is formed. In order to account for different conformers and multi-component systems, the lattice energy is defined by eq 1,

$$E_{lattice} = \frac{E_{crystal-} \sum_i^n c_i E_{vac,i}}{\sum_i^n c_i} \quad (1)$$

where the summation is over all different molecular species n, c_i the number of molecules of species i in the unit cell and $E_{vac,i}$ the vacuum energy of molecule i. Sublimation energy is defined as the negative of the lattice energy.

2.1.2. Thermodynamical properties

In this project, a number of different thermodynamical properties are used.

2.1.2.1. Entropy

The entropy S of a system can be defined from Boltzmann's equation (eq 2), where Ω is the number of micro states that are associated to the macro states of the system and k_b is the Boltzmann constant. Entropic contributions for a molecule are translation, rotation and vibration.

$$S = -k_B \ln \Omega \quad (2)$$

If a molecule is in a crystal lattice, only vibrations are possible since the molecule has a fixed positions. For a perfect crystal at T=0 K, there exists only one state thus the entropy of the crystal is zero. As temperature increases, lattice defects and molecular motions are introduced. A collective molecular motion in a crystal is defined as a phonon. A phonon can be seen as a superposition of a set of different vibrational modes. Since vibrations are the only molecular contribution to the entropy of a crystal, it is very important to be able to calculate the phonons accurately.

2.1.2.2. Enthalpy

The enthalpy H of a system is the sum of the internal energy U and the product of pressure P and volume V. The temperature dependence of the enthalpy is given by $H = U - 2RT$. The enthalpy must also be corrected by addition of the zero point energy (ZPE) which is the ground state energy of the system.

2.1.2.3. Free energy

The free energy of a system is the useful energy, available to perform work. It exists two different definitions of free energy, The Helmholtz free energy and the Gibbs' free energy.

The Helmholtz free energy, A, is defined for an isothermal system, as $A = U - TS$ where T is the temperature. Gibbs' free energy G is defined for an isobaric and isothermal system, as $G = H - TS$. Free energy in this project is defined by Gibbs' free energy.

2.1.2.4. Phases

A region of a material that has identical physical properties can be defined as a region of the same phase. A material can exist in many different phases, for example in glass or liquid phase.

Phases can be thermodynamically stable, metastable or unstable. Since all processes strive to minimize the free energy, the stable phases exist where the global free energy is minimal. Metastable phases are defined as having a local free energy minimum. This allows the phase to exist for a period of time but will eventually (after infinite time) transform to the stable phase. All other phases of materials are unstable and will transform to stable or metastable phases. During transition between phases, physical properties undergo changes. Some of the affected properties are volume, specific heat and magnetization.

If a material exists in different phases as a solid, it is polymorphic. Polymorphs can have different properties.[10]

Multi-component crystals (as defined in this project) consist of at least two distinct molecules, and can be divided into solvates or co-crystals. The definition of a co-crystal chosen for this project [11], is that the bulk material of the co-crystal former must be solid at room temperature. If it is liquid at room temperature, the resulting crystal will then be referred to as a solvate.

2.1.3. Solubility

The definition of solubility is the amount of substance that can be dissolved in a solvent, resulting in a saturated solution at constant temperature and pressure. In a pharmaceutical context, the solubility is often expressed as µg/mL. To compare solubilities of different compounds, the *intrinsic* solubility (S_0) is often used. It is defined as the equilibrium solubility measured at a pH corresponding to an uncharged compound

When calculating solubility for a crystalline material, it is quite difficult to model the free energy of solvation in an accurate way. The often used method is instead to calculate it indirectly by using a thermodynamic cycle as can be seen in Figure 3.[7]

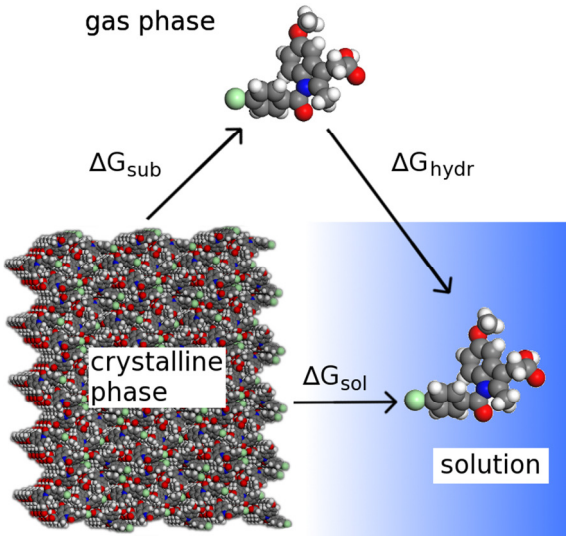

Figure 3. Schematic picture over the thermodynamic cycle utilized in the solubility calculations. ΔG_{sub} is the sublimation energy, ΔG_{hydr} the hydration energy and ΔG_{sol} the solvation energy. All energies mentioned are Gibbs' free energies.

The cycle, explained by eq 3, will be utilized in this project.

$$\Delta G_{sol} = \Delta G_{sub} + \Delta G_{hydr} \quad (3)$$

Equation 4 explains how the free energy of solvation is related to the equilibrium solubility S.[6]

$$S = \frac{P_0}{RT} \cdot \exp\left(\frac{-\Delta G_{sol}}{RT}\right) \quad (4)$$

Pharmaceutical drugs are classified based on their solubility and permeability, using the so-called Biopharmaceutics Classification System (BCS).[12] The BCS system consists of four classes, of which class II (low solubility and high permeability) and IV (low solubility and low permeability) are the most interesting from a solubility perspective. This classification system is extensively used by the US Food and Drug Administration (FDA).

2.1.4. Dissolution rate

Dissolution of a material is driven mainly by two phenomena, cavity creation and diffusion. In order for one solid molecule to dissolve (leave the ordered structure of the solid), a molecule sized cavity must be created in the solvent. For a pure solvent this can be quite limiting and to overcome this, different types of additives are often introduced. The additives break up the ordered solvent structure and induce a low level of stress in the system, thus reducing the energy required to create the cavity. After a molecule from the solid has been inserted into the cavity it needs to diffuse away from the surface to allow other molecules to dissolve.[13]

Diffusion is governed by Fick's first law. A generalized version of Fick's first law is given by eq 5,

$$\boldsymbol{J} = -D\nabla \mathrm{c} \quad (5)$$

where **J** is the molar flux, D the diffusion coefficient and ∇c the concentration gradient.[1] It states that the larger the concentration gradient is, the larger the resulting molar flux will be. In order to describe transient diffusion, Fick's second law must be used (eq 6).

$$\frac{\partial c}{\partial t} = D\nabla^2 c \quad (6)$$

The concentration profile for a 1D stationary case can be seen in Figure 4. Here the concept of boundary layers is introduced. Boundary layers build up when solvated molecules remain close to the surface. The layer thickness is related to the diffusion coefficient of the solute. If convection is applied, the boundary layer will be reduced.[10]

[1] $\nabla = \frac{d}{dx} + \frac{d}{dy} + \frac{d}{dz}$

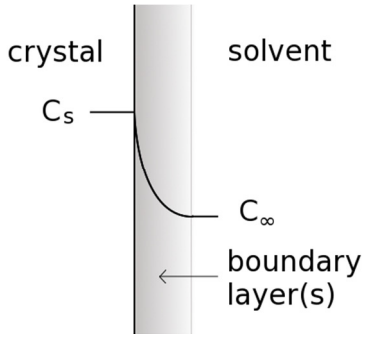

Figure 4. Concentration profile of solute during dissolution. Diffusion occurs from the surface, through the boundary layer(s) to the solvent. The solute concentration is highest near the surface (C_s) and decreases until the saturation concentration C_∞ of the solution has been reached.

In 1897, Noyes and Whitney formulated an equation that can be used to determine the dissolution rate. The Noyes-Whitney equation (eq 7) relates the dissolution rate to the properties of both the solid and solvent, where D is the diffusion coefficient of the solute, A the surface area of the solid, C_s concentration of solid in the diffusion layer, C solute concentration in bulk medium and L length of the diffusion layer.[13]

$$\frac{dm}{dt} = \frac{DA(C_s - C)}{L} \tag{7}$$

The concentration of solute near the crystal is the solubility of the material in the particular solvent. If the dissolution takes place in a very large container, such that the bulk concentration of the solute is identically equal to zero, so called sink conditions are valid. This is a very important concept since if sink conditions are valid, the initial dissolution rate will be linearly proportional to the solubility as well as surface area. The surface area directly influences the dissolution rate and is therefore important to know in order to quantitatively analyze the dissolution process. If the surface area is known, the initial dissolution for different systems can be assumed to only differ by their solubilities, making it possible to compare them to each other. This can be done since the diffusion coefficient and layer thickness are assumed to be almost identical for most samples evaluated.

2.2. EXPERIMENTAL BACKGROUND

In order to understand the techniques used for characterization of the substances, brief explanations will be given in this section.

2.2.1. X-ray Diffraction

A scattering or diffraction experiment can be conducted by using a variety of sources, electrons, neutrons, photons etc. It is important that the wavelength of the radiation matches the characteristic length of the material to be studied; otherwise structural information might be lost. The wavelength of X-ray radiation is between 0.01 and 10 nanometers, which correspond well to the distances of atomic bonds and lattice parameters and is why it is normally chosen for structural characterization.

Incoming X-rays on a crystal will be diffracted if Bragg's law (eq 8) is fulfilled:

$$n\lambda = 2d_{hkl}\sin\theta \quad (8)$$

where n is an integer dependent on scattering order, λ radiation wavelength, d_{hkl} is the distance between lattice planes with Miller index *hkl* (*hkl* are integer numbers) and θ the angle between incoming radiation and the diffracted planes.
During an experiment, θ is varied by stepwise elevation of the detector and the tube. The scattered intensity will peak for certain values corresponding to Miller planes which fulfill Bragg's law.[14] A typical example of an X-ray diffractogram can be seen in Figure 5.

In this project X-ray diffraction (XRPD) measurements have been performed on crystalline powders for identifying the crystal form of the samples.

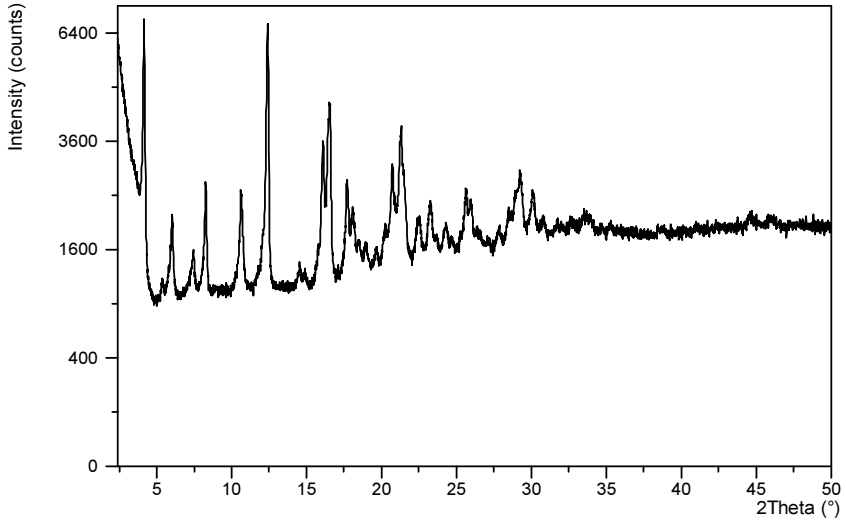

Figure 5. Typical XRPD pattern, showing intensity (photon count) of the scattered radiation as a function of scattering angle 2θ for 1.54Å CuK_α X-rays. Each peak corresponds to a Miller plane.

2.2.2. Thermal Gravimetric Analysis

Thermal gravimetric analysis (TGA) is an old technique with a relative straightforward concept. By measuring the weight of the sample while it is being heated, mass loss due to evaporation of solvent can be tracked.[15] The result can then be visualized as a temperature – mass loss (percentage) graph, Figure 6. By comparing the actual loss with a theoretical mass loss (for a specific solvent) the molar ratio between the main molecule and solvent can be determined.

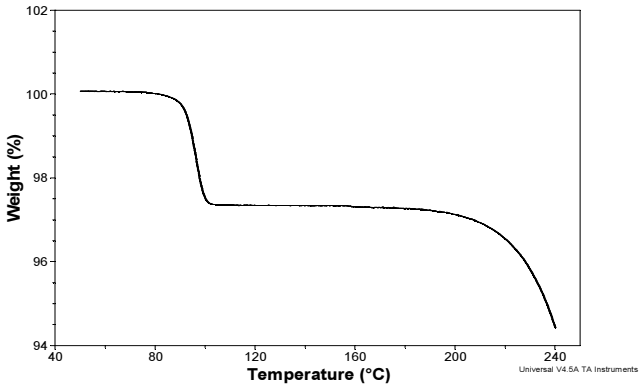

Figure 6. Typical visualization of a TGA scan, weight percentage is plotted against temperature. It can clearly be seen that solvent is evaporating from the sample at an early stage. After the evaporation, composition of the sample is stable until high temperature is reached where it starts to decompose.

2.2.3. Differential Scanning Calorimetry

Understanding when phase transitions occur in a sample is of great importance in order to characterize it, but also to be able to handle it correctly during storing and manufacturing. Differential Scanning Calorimetry (DSC) uses changes in specific heat to find phase transitions. It works by simultaneously heating the sample and a reference while monitoring the heat required for the sample to obtain the same temperature as the reference.[16]

The results are visualized in a heat flow versus temperature graph. A typical DSC scan can be seen in Figure 7. Here the thermal events are marked. In the graph, a local minimum corresponds to an endothermic event, *e.g.* melting of the crystal or evaporation of solvent. Likewise, a local maximum corresponds to an exothermic phase transition, *e.g.* recrystallization.

By analyzing the DSC scan, melting points (T_m) and glass transition temperatures (T_g) can be

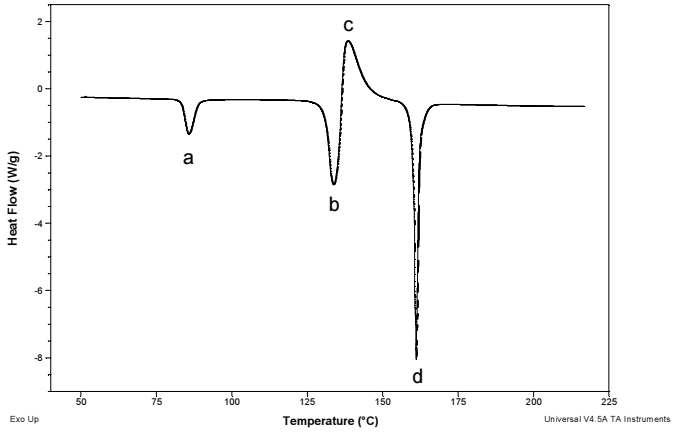

Figure 7. DSC thermogram showing four thermal events. (a), (b), (d) are endothermic events, for example melting points, while (c) is an exothermic event, for example recrystallization.

found. The energy required for the phase transition can also be determined from the scan, but it requires a more thorough analysis. By integrating over the peak of interest, the energy for that specific phase transition is obtained, for example giving values of ΔH_{fusion} or $\Delta H_{recrystallization}$.

2.2.4. Solubility and dissolution rate measurements

One way of measuring the solubility of a material in a solvent is by absorbance measurements, since a molecule will absorb light of certain wavelengths. The absorbed wavelengths depend on the functional groups and general structure of the molecule. The absorbance A_λ (dependent on wavelength λ) is related to the concentration c by the Beer-Lambert law (eq 9), where I_0 and I are the submitted and measured light intensity respectively, ε the molar absorption coefficient and l the length of the beam path in the sample.

$$A_\lambda = \log_{10} \frac{I_0}{I} = \varepsilon \, l \, c \tag{9}$$

As stated, the absorbance of the sample is wavelength dependant, so in order to interpret the results correctly it is important to select the wavelengths that are absorbed by the molecule of interest.[17] Dissolution rate of a sample can be determined by continuously monitoring the absorbance as it is being dissolved.

2.2.5. Scanning Electron Microscopy

The Scanning Electron Microscope (SEM) works by scanning a sample using electrons. The short wavelength of electrons allows the resolution to greatly surpass optical microscopes, reaching resolutions as low as 1 nm. One of the biggest drawbacks is that the electrons will gradually destroy the sample. As a way of countering this, samples are usually coated with a thin layer of gold which also increases the measured signal. For each point in the scan, an electron beam is focused on the surface. As electrons from the beam hit the surface, they interact with the sample. Operating voltage for a typical SEM is around 5 keV. The electron beam is focused using magnetic lenses.[18]

One way for incoming electrons to interact with the sample is by scattering with a valence electron. For the incoming electron this results in a slight path change, while the valence electron will be ejected from the atom and called a secondary electron (SE). The energy of such an electron is usually lower than 50eV. Since SE:s have such low energies, they can only escape from a very shallow region close to the surface. In other words, the number of detected SE:s depends strongly on the topography of the surface (see Figure 8). Images formed by these electrons therefore have a very high topographic resolution. Image formation is done by collecting the scattered electrons using detectors and analyzing the intensities for each point in the grid.[18, 19]

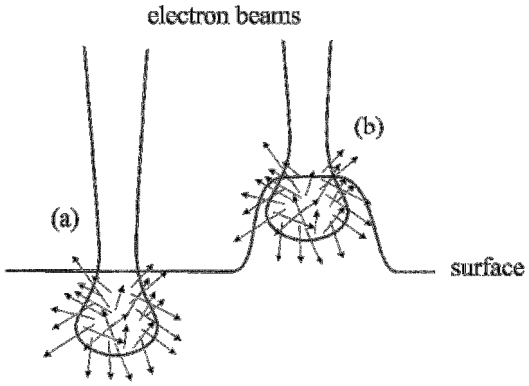

Figure 8. Schematic drawing over how surface detail is generated in a SEM. As the electron beam hits the surface, SE electrons are generated from the so-called interaction volume. On a flat surface (a) only few electrons escape. If instead the beam is focused somewhere near an edge as in (b), more electrons can escape. This effect, called edge-contrast, is what creates contrast in surface topography imaging.

The topography images in this project will be taken using an SE detector. The length of different features on the sample can be determined by analyzing the image. If the sample consists of a powder, careful analysis makes it possible to determine a distribution of particle sizes.

2.2.6. Particle sizing

The sizes of particles in a powder are never identical. When the particle size is to be determined, the distribution of sizes is what is wanted. One way to determine this is by using laser diffraction. The technique is based on the Fraunhofer diffraction theory, which proportionately relates the scattered light intensity to the particle size. A small particle deviates the laser beam to a greater extent than a large particle. Using accurate detectors and sophisticated software algorithms, the particle size distribution (PSD) can be determined. One major drawback of this technique is that the particles, regardless of actual shape, are assumed to be spherical.[20]

2.3. SIMULATION PRINCIPLES

Simulations will be carried out using methods based on density functional theory as well as molecular mechanics.

2.3.1. Quantum mechanics

A quantum mechanical (QM) system is described by the Schrödinger equation and can be associated to its quantum state $|\varphi>$. Solving the equation yields the energy of the system E as eigenvalues to the Hamiltonian $\hat{H}$. The Hamiltonian can, for a system consisting of N electrons and M nucleus, be written as[2]

$$\hat{H} = \hat{T} + \hat{V}_{ee} + \hat{V}_{ext} = -\frac{1}{2}\sum_{i=1}^{N}\nabla^2 + \sum_{\mathrm{i}=1}^{\mathrm{N}}\sum_{\mathrm{i}<j}^{\mathrm{N}}\frac{1}{|\mathbf{r}_{ij}|} - \sum_{\mathrm{i}=1}^{\mathrm{N}}\sum_{\mathrm{A}=1}^{\mathrm{M}}\frac{\mathrm{Z_A}}{|\mathbf{R}_A - \mathbf{r}_{ij}|} \qquad (10)$$

[2] In Rydberg atomic units

where $\hat{T}$ is the kinetic energy, $\hat{V}_{ee}$ the electron-electron interaction energy and $\hat{V}_{ext}$ the potential energy between electrons and nucleus. The individual contributions are then further expanded where ∇^2 is the kinetic energy, $\mathbf{r}_i$ the coordinates of electron *i*, $\mathbf{R}_A$ the coordinates and Z_A the charge of nucleus A. Equation 10 is only possible to solve for the hydrogen atom. For bigger atoms (or systems of atoms) approximations are needed. One of the most common is the Born-Oppenheimer approximation, where the movement of the nucleus is neglected. This can be done since the mass of a nucleon is of the order 10^3 times higher than the mass of an electron.

2.3.2. Density Functional Theory

In 1964, Hohenberg and Kohn proved that the electron density could be used instead of wave functions when solving the Schrödinger equation.[21] They showed that a quantum system is uniquely described by its electron density $n(\mathbf{r})$ but also that the variational principle can be applied for electron densities (eq 11):

$$E_0 = \min_{\mathrm{n}} \ (E[n]) = \min_{\mathrm{n}} \ (T[n] + U_C[n] + E_{ncl}[n] + \int n V_{Ne} d\boldsymbol{r}) \qquad (11)$$

Here T is the kinetic energy, U_C the Coulomb energy, E_{ncl} are a non classical energy term arising from self-interactions, anti-symmetric and correlational effects. The last term in eq 11 is the periodic potential arising from the lattice structure. One important condition for the electron density is that integration over the system must yield exactly the number of electrons.

If we consider a hypothetical system consisting of only non-interacting electrons, the kinetic energy (T_s) of such a system can be calculated exactly. The kinetic energy difference between this hypothetical system and the corresponding real system is unknown and will be referred to as T_C. T_C, along with E_{ncl} cannot be calculated exactly and are grouped together to form the exchange-correlation functional E_{XC}, resulting in the following expression (eq 12):

$$E[n] = (T_s[n] + U_C[n] + E_{XC}[n] + \int n V_{Ne} d\boldsymbol{r}) \qquad (12)$$

Still assuming non-interacting electrons, an effective potential can be introduced. This potential arises from the nuclei, averaged over the system, allowing creation of a complete non-interacting system. The Hamiltonian operator for such a system can, starting from eq 10, be expressed in form of eq 13:

$$\hat{H}_s = -\frac{1}{2}\sum_i^N \nabla^2 + \sum_i^N V_s(\boldsymbol{r}_i) \qquad (13)$$

According to the Pauli principle, electrons must have anti-symmetric wave functions that are orthogonal to each other. The wave function of a single electron is denoted by φ_i. This allows an operator to be defined, the Kohn-Sham one electron operator, which acts on individual orbitals. It is defined by eq 14:

$$\hat{f}^{KS} = -\frac{1}{2}\nabla^2 + \mathrm{V_s}(\mathbf{r}) \quad , \quad \hat{f}^{KS}\varphi_\mathrm{i} = \epsilon_i \varphi_\mathrm{i} \qquad (14)$$

In order to couple the non-interacting system to the real one, the potential V_s must be chosen wisely. It should be chosen so that the non-interacting electron density is equal to the real electron density. Using eq 10 as a starting point, an expression for the necessary potential can be obtained. This potential, V_{eff}, is defined by eq 15:

$$V_s(\boldsymbol{r}) = V_{eff}(\boldsymbol{r}) = \int \frac{n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} d\boldsymbol{r}' + V_{XC}[n] - \sum_A^M \frac{Z_A}{|\boldsymbol{R}_A - \boldsymbol{r}|} \quad (15)$$

Equation 14 and 15 form a set that must be solved self-consistently (iteratively) for $V_{eff}(\boldsymbol{r})$. When the solution has converged the energy of the system can be calculated.

2.3.2.1. Electronic representation

An important part of quantum mechanical simulations is how the electron density is represented; it can be done by using localized orbitals or plane-waves. The localized orbitals method uses a basis set, unique for each individual element, centered on the atom. The basis set consists of a linear combination of (usually) Gaussian-type functions. The minimal requirement is that each orbital is represented with one function. An example of such a minimal basis set is STO-3G. In order to describe chemical bonding, polarization functions are added. Increasing the basis set or adding polarization functions increases the time required for calculations. Localized orbitals describe molecular systems with high accuracy, but a drawback of the representation is their inherent basis set superposition error (BSSE). The BSSE contribution to the energy can be calculated and corrected for *a posteriori*.[22]

Another way of representing the electron density is to use a plane-wave basis set. A plane-wave basis set consists of a set of wave functions, mutually orthogonal. They are an efficient choice when used in systems with periodic boundary conditions. A cutoff energy is used as the maximum energy for the waves. A higher cutoff value means higher quality of the simulations at the cost of increased computational demand. Plane-waves are frequently used in combination with pseudopotentials. A pseudopotential is a modified atomic potential in which the core electrons are included, which results in that the plane-waves only represent the valence electron. This is justified by the fact that chemical bonding is exclusive to valence electrons, as the core electrons are much tighter bound to the nucleus. The approximation results in a less complex description of the system which increases the computational performance with respect to time. There are a few different implementations: *ultrasoft, norm-conservin*g or *on-the-fly* pseudopotentials. Norm-conserving was the first developed pseudopotential type, requiring the norms of the one-electron and the all-electron wave functions to be identical. The norm criterion caused convergence issues in some cases which gave rise to ultrasoft pseudopotentials, where the norm criterion is relaxed. Most pseudopotentials are determined explicitly one single time and then used in many instances. On-the-fly pseudopotentials are generated individually for each simulation.[22, 23]

2.3.2.2. The Exchange-Correlation functional

The E_{XC}-functional contains, as stated earlier, everything that cannot be calculated exactly. The energy contribution from the functional is small compared to the other terms in Equation 12, but is chemically relevant so it cannot be neglected. In order to construct an implementable

algorithm it is necessary to approximate the form of the functional. Development of different functionals has been in process since DFT first emerged.[24]

The first usable functional to be constructed was LDA[25], eq 16. It is based on the theory of the homogenous electronic gas (or jellium) in which the electrons can move freely over the periodic potential arising from the nucleus.

$$E_{XC}^{LDA} = \int n(\boldsymbol{r})\epsilon_{XC}(n)d\boldsymbol{r} \tag{16}$$

An improved version was soon introduced by regarding the spin of the electrons as well. This created the Local Spin Density Approximation (LSDA),[26] where $\epsilon_{XC}(n) \rightarrow \epsilon_{XC}(n_\uparrow, n_\downarrow)$. The implementation resulted in increased accuracy of the results. However, the DFT method was still only applicable (regarding accuracy) to metallic systems.

Using the LSDA functional as a starting point, the concept of a GGA functional was created (eq 17) introducing dependence of the spin density gradients. This greatly increased the complexity of the algorithm thus increasing the time required for calculations substantially. However, DFT could now be used with a high degree of accuracy for determining molecular geometries and properties.

$$E_{XC}^{GGA} = \int n(\boldsymbol{r})\epsilon_{XC}(n_\uparrow, n_\downarrow, \boldsymbol{\nabla}\mathrm{n}_\uparrow\boldsymbol{\nabla}n_\downarrow)d\boldsymbol{r} \tag{17}$$

In this project a GGA functional called PBE[27] will be used. It is one of the most common functionals in use and has been tested thoroughly on varying systems with results of high accuracy.[24]

2.3.2.3. Dispersive forces

Dispersive forces are intermolecular, weak, long range forces. They arise as electrons can spontaneously create regions of high or low electron density as they oscillate around the molecule causing a local dipole moment. This phenomenon is called instantaneous dipoles, and they can polarize nearby molecules effectively inducing dipoles in them. The forces between such dipoles are called dispersive forces. The individual contribution of dispersive energy from each dipole pair is very small, but summation over the system results in a large energy for many materials. For some certain classes dispersive forces are utterly important, for example noble gas crystals and non-polar molecules with high molecular weight (polymers). Another case for which the dispersive forces are important is crystals of organic molecules.

One of the inherent major drawbacks of native DFT is the fact that it is not able to describe dispersive interactions. In order to increase the accuracy and allow more diverse system to be investigated it is therefore of paramount importance to be able to correctly describe dispersive forces. There exists two different way of implementing dispersive forces into DFT, implicitly and explicitly. An implicit method tries to design E_{XC} so that the dispersive forces are taken into account during each step of the calculations. The implementation is usually done by first principles, without addition of empirical data. An example of an implicit functional is vDW-DF2.[28] This does however greatly increase the calculation time. Explicit methods are, in contrast, based on empirical data and exist in many different forms. Common for all these methods is that the dispersive forces are added at the end of each calculation step.

The most common dispersion corrections (also used in this project) are Tkatchenko-Scheffler (TS) [29] and Grimme (revision D2 [30] and D3 [31, 32]) methods. The TS and D3 method are similar due to the fact that the correction coefficients of both methods are dependent on the local environment of the atoms. In the D3 method, the single most important contribution to the dispersive energy is the two-body term between atom A and atom B which can be seen in eq 18:

$$E_{disp}^{(2)} = -\frac{1}{2}\sum_{A\neq B}\sum_{n=6,8} s_n \frac{C_n^{AB}}{r_{AB}^n + f(R_0^{AB})^n} \tag{18}$$

Here s_n is a constant depending on which functional that is used and $C_{6,8}^{AB}$ are the 6th and 8th order dispersion coefficients respectively. In order to prevent a singularity that could arise if the distance r_{AB} becomes too small, a damping function $f(R_0^{AB})$ is added. It is possible to choose between two different functions, Becke-Johnson (BJ) or zero damping. Both s_n and exact coefficients of the damping functions $f(R_0^{AB})^n$ are dependent on which functional that is used. The performance of both damping schemes has been evaluated in this project.

2.3.3. Molecular mechanics

Molecular mechanics (MM) is a well used method in molecular calculations. It uses atoms as the smallest building block, instead of electrons as in DFT. This causes all quantum effects to be incorporated into the parameters. The system is effectively transformed to a classical mechanical system where atoms are treated as spheres and bonds as springs of varying lengths and stiffness. The stiffness is given by the force constant of the spring and since the electrons are omitted from the system, must be given explicitly. The spring constant is different for different atoms and their local environment, the bond of a C atom can for example be given as sp^2- or sp^3-type.

In MM methods, the molecular energy is divided into different contributions (eq 19).

$$E_{MM} = E_{stretch} + E_{bend} + E_{torsional} + E_{vdw} + E_{electronic} + E_{cross} \tag{19}$$

The total energy is then minimized using a force field. The field is parameterized so that each energy term has different constants for each atom type. [22]

In this project the force field COMPASSII has been used due to the capability of modeling crystalline systems.

2.3.4. Quantum mechanics vs. molecular mechanics

When deciding which method to use, careful considerations must always be taken as to what kind of simulated data (and its accuracy) that is needed. Comparing quantum and molecular mechanics methods with each other, it is clear that they both offer advantages as well as drawbacks. The major advantage of MM methods is the very quick calculations. An example of the timescale difference between the method is that calculations taking only seconds to perform with MM can take takes days using QM. Another advantage is that the results can be of extremely high accuracy if parameterization is done correctly. If however the parameterization is done poorly or done for a system with no resemblance to the evaluated one, low accuracy is a major drawback.

The major advantage of DFT is the computational accuracy. Using high quality settings, certain properties can be determined with even higher accuracy than its experimental

counterpart.[22] Choosing optimal settings requires deep knowledge of the system. The functional in DFT must be chosen carefully, but it is also necessary to reflect on how the electrons should be represented – by plane waves or localized orbitals. Considering MM methods, the most challenging question can be deciding which force field should be used. In some cases existing force fields are not sufficient, thus requiring a customized field to be developed.

3. METHODS

In order to reach the aim of the project, reliable experimental and theoretical data must be produced. The techniques and methods used for generating the data will be explained in this section.

3.1. EXPERIMENTAL METHODS

For the experimental part, one of the main goals of the project was preparation of polymorphs, solvates and co-crystals of indomethacin and carbamazepine. The second step was characterization of the materials produced. In order to achieve this, a number of different techniques were used: X-ray powder diffraction (XRPD), Thermal gravimetric analysis (TGA) and Differential scanning calorimetry (DSC). These techniques, when used together, are often enough to determine which crystal form the material has. The last step was determination of solubility and dissolution rates for the different compounds.

3.1.1. Preparation

The preparation of all polymorphs and multi-component crystals were done in AstraZeneca's facilities in Mölndal (Sweden) unless otherwise stated. All samples were prepared in batches of roughly 100 mg.

3.1.1.1. Indomethacin

Bulk powder (γ-polymorph) of indomethacin was ordered from CHEMTRONICA. The α-polymorph was prepared by precipitation from a hot ethanol solution as cold distilled water was added, as described in the literature.[33]

A set of different solvents were used to screen for indomethacin solvates. Hot saturated solutions of indomethacin were allowed to slowly evaporate to form the solvate.[9, 34]

One co-crystal of indomethacin was prepared by mixing indomethacin and saccharin (Aldrich) in a molar ratio of 1:1 with ethyl acetate as solvent. The solution was allowed to slowly evaporate forming indomethacin-saccharin co-crystals.[35]

The attempted different crystal forms of indomethacin are summarized in Table 1.

Table 1. Summary of all material used in the attempts to prepare the different crystal forms of indomethacin

Prepared indomethacin crystal forms		
Crystal form	**Coformer**	**Solvent**
α-polymorph	-	Ethanol (Kemetyl)
γ-polymorph	-	-
Acetone solvate	-	Acetone (Scharlau)
tert-Butanol solvate	-	*tert*-Butanol (Sigma)
Diethyl ether solvate	-	Diethyl ether (Riedel-de Haën)
Ethanol solvate	-	Ethanol
Ethyl acetate solvate	-	Ethyl acetate (Sigma-Aldrich)
Methanol solvate	-	Methanol (Sigma-Aldrich)
2-Propanol solvate	-	2-propanol (Sigma)
Saccharin co-crystal	Saccharin (Aldrich)	Ethyl acetate
Tetrahydrofuran solvate	-	Tetrahydrofuran (Sigma)

3.1.1.2. Carbamazepine

Bulk CBZ powder was manufactured by Hawkins Inc.. CBZ polymorph I was prepared by heating CBZ to 170 °C for 3 h. Polymorph II was prepared by dissolution of CBZ in hot ethanol, followed by slow evaporation and cooling to 5 °C. Polymorph III was prepared by dissolution in ethanol at room temperature followed by rapid cooling.[36]

The procedure for preparing CBZ solvates was the same as for indomethacin solvates, saturated hot solutions were allowed to evaporate slowly and form crystals.[37, 38]

Co-crystals were prepared by dissolving (either in a 1:1 or a 1:0.5 molar ratio) in a suitable solvent as indicated in the literature.[37, 39, 40]

The attempted different crystal forms of carbamazepine crystal forms are summarized in Table 2.

Table 2. Summary of all material used in the attempts to prepare the different crystal forms of carbamazepine.

Prepared carbamazepine crystal forms		
Crystal form	**Coformer**	**Solvent**
Polymorph I	-	-
Polymorph II	-	Ethanol
Polymorph III	-	Ethanol
Acetic acid solvate	-	Acetic acid
Acetone solvate	-	Acetone
Dimethylsulfoxide (DMSO) solvate	-	DMSO (Fluka Analytical)
Glutaric acid co-crystal	Glutaric acid (Aldrich)	Acetonitrile (Fischer scientific)
Malonic acid co-crystal	Malonic acid (Aldrich)	Acetonitrile (Fischer scientific)
N,N-dimethylformamide solvate	-	*N,N*-dimethylformamide (Sigma-Aldrich)
Nicotinamide co-crystal	Nicotinamide (Fluka Analytical)	50/50vol% DMSO/methanol
Saccharin co-crystal	Saccharin	Ethanol
Salicylic acid co-crystal	Salicylic acid (Scharlau)	Acetonitrile (Fischer scientific)
Succinic acid co-crystal	Succinic acid (Sigma)	33/66vol% Ethyl acetate/Ethanol
Terephtalaldehyde co-crystal	Terephtalaldehyde (Fluka Analytical)	Methanol

3.1.2. Sample characterization

Characterization of the samples was done with X-ray diffraction, Thermal gravimetric analysis and Differential scanning calorimetry.

For X-ray diffraction, samples were mounted on single silicon crystal (SSC) wafer mounts and powder X-ray diffraction was recorded with a Theta-Theta PANalytical X'Pert PRO diffractometer (wavelength of X-rays 1.5418 Å nickel-filtered Cu radiation, voltage 45kV and filament emission 40 mA). Automatic variable divergence and antiscatter slits were used and the samples were rotated during measurement. Samples were scanned from 2 - 50° 2θ using a 0.013° step width and a 115.77 second count time using a PIXCEL detector (active length 3.35° 2θ).

Analysis of possible solvent residues was performed with a TGA Q500 from TA Instruments. Samples (~3 mg) were placed in aluminum sample holders, TA Instrument, inside platinum pans. The procedure used for the experiments was a linear ramp, heating rate 10 °C/min, up to a final temperature of 250 °C.

Differential scanning calorimetry measurements were performed with a DSC Q2000 from TA Instruments. Pans and lids, Tzero from TA Instruments, were individually weighed before sample addition in order to increase the sensitivity of the experiments. Inert gas (N_2) was used with a flow of 40 ml/min. Modes used were either linear ramp (heating rate of 10 °C/min) or modulated ramp (modulation cycle 45 s, modulation temperature ±3 °C, overall heating rate 5 °C/min).

3.1.3. Dissolution

Dissolution experiments were carried out with µDISS profiler from Pion coupled to a HAAKE B3 thermal bath unit for temperature control. The µDISS system has 6 identical probes allowing for simultaneous measurements. Each probe continuously measures the UV absorbance of the medium. The setup of the instrument can be seen in APPENDIX D.

3.1.3.1. Reference curve

In order to measure the dissolution rate, a reference curve must first be created that shows the relationship between absorbance and solute concentration. For this, a stock solution of high concentration was created by dissolving 1 mg of bulk powder in 25 mL 95% EtOH. Sonication was performed with a Transsonic T700 sonicator from ELMA to ensure complete dissolution. Six µDISS vials (one for each probe) were prepared with 20 mL of dissolution media from which a blank spectrum was collected. A small amount of the stock solution was pipetted into the µDISS vials with subsequent collection of the spectrum. At least three spectra were measured for each concentration. Careful considerations were taken in order to ensure that no air bubbles were present near the probe cavity. The procedure was repeated until eight different measurement points were established. The obtained reference curve was required to have an R^2 fit higher than 0.99 for ensuring a correct linear relationship between the concentration and absorbance. The linear relationship is assumed not to be affected by the presence of solvent, due to its low maximum volume concentration of <5%.

By examining how the absorbance changes over the entire measured spectrum, a specific wavelength range can be chosen that is individual for the two substances. Using the selected range, the relationship between the absorbance and solute concentration is calculated by the software AuPro. For an example of how the reference curve is visualized, see APPENDIX D.

3.1.3.2. Measurements

Samples were weighed and prepared as triplicates in 20 mL µDISS vials. For each measurement series, the bulk material was present as control in three vials. Precisely 20.0 mL of dissolution media was prepared for each vial. The media was chosen so that the molecule was uncharged. As the experiment started, the dissolution media was added to a new vial every 20 seconds. After completion the results were exported for analysis. Exported values are time and concentration. Measured time was adjusted so that each experiment started at $t = 0$. The concentration was normalized to show dissolution in percentage of starting material.

The dissolution media chosen for all experiments was 0.01M HCl/NaCl buffer (pH 1.2) to ensure charge neutrality of all molecules.

3.1.4. Particle morphology and size distribution

Surface morphology was examined using a scanning electron microscope (Quanta 200, FEI). The analysis was performed in high vacuum mode with a variable spot size and acceleration voltage of 20.0 kV with an Everhart-Thornley detector. SEM micrographs were obtained using the software xT Microscope Control. Samples were prepared by applying powder onto double sided carbon black tape (carbon tabs, 12mm, Agar Scientific), attached to a sample holder (aluminum specimen stubs, 0.5", Agar Scientific). Excess powder was removed to achieve a monolayer. Samples were then sputter coated with gold (Cressington 108auto) using a current of 20 mA for 150 s.

Particle sizing were performed using a Malvern Mastersizer 3000 with the Aero S dry powder dispersion unit. Fraunhofer diffraction method was chosen with background measurement time of 8 s, sample measurement time of 65 s , feed rate of 18% and obscuration range 0.1-10%. All samples were ensured to be homogenous by shaking the container for a period of 20 seconds.

Specific surface area (area per weight) was calculated from the determined diameter using eq 20, where D is the particle size and ρ sample density.

$$A_{specific} = \frac{\pi D^2}{\frac{\pi D^3 \rho}{6}} = \frac{6}{D\rho} \tag{20}$$

3.2. SIMULATION METHODS

The procedures for obtaining all the energies sought in this project are time-consuming. Crystal energies must first be calculated followed by energies for gas phase molecules. After these energies are obtained, lattice energies and sublimation enthalpies of the crystal can be calculated using Eq. 1. Entropy contributions for both the crystalline and gas phase as well as hydration energies must also be simulated. With all the individual contributions, the free energy of solvation can be calculated using the thermodynamic cycle explained earlier.

3.2.1. Simulation software

Simulations were performed in Materials Studio 7.0 using the modules CASTEP[41], DMol^3[42, 43] and Forcite. The program Jaguar[44, 45] of the Schrödinger suite was also used. Starting from the default CASTEP settings, some changes were made in order to improve convergence which was justified by in-house knowledge and literature. The changes included

cell compressibility (set to "medium") and pseudopotentials (set to "on-the-fly"). The functional PBE was used throughout the DFT calculations unless otherwise stated.

The program DFT-D3 developed by Grimme was used in order to implement D3 dispersion.

3.2.2. Energy simulation

Simulations of the internal energy were done on the crystalline structures obtained from CSD, after geometry optimization.

The vacuum energy, the energy of a gas phase molecule, can be a bit more complex to calculate correctly. This is the case when using CASTEP since it uses plane waves as electronic representation and requires a unit cell for the calculations. In order to simulate the vacuum, the molecule of interest is placed in a large empty unit cell. The dimensions of the cell should be large enough to make sure that repeated units do not interact with each other. A very large unit cell (compared to the molecular dimension) gives a correct vacuum energy, but substantially increases the computational time. The influence of the unit cell size on the vacuum energy accuracy was investigated.

All the simulated energies were corrected with the enthalpy correction term described in section 2.1.2.2.

3.2.3. Entropy simulation

Entropy of each crystal or molecule was calculated with the "phonon" property in the CASTEP module. Finite displacement method was used instead of linear response to allow more flexibility regarding settings. The output of the entropy calculations is a graph, from which the entropy is taken at 298K. This can be seen in Figure 9.

As the entropy of each system had been calculated, the zero point energy (ZPE) of the system could be extracted. This energy needs to be added to the internal energy in order to achieve correct results.

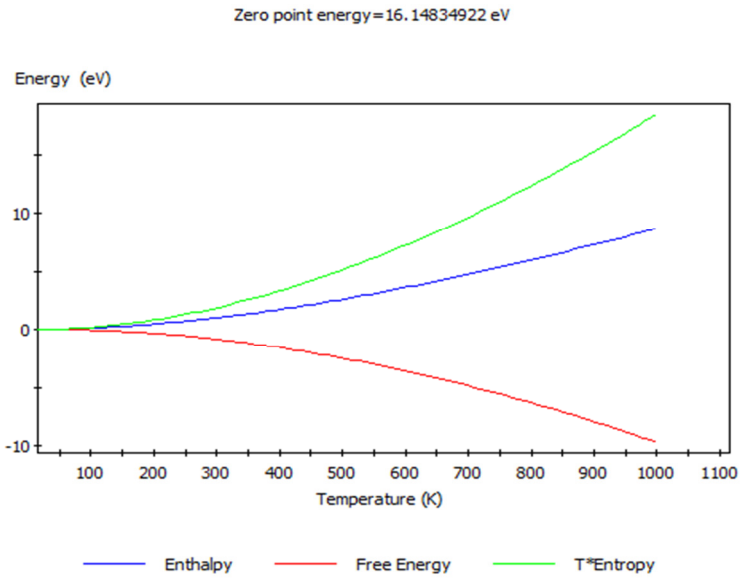

Figure 9. CASTEP output of entropy calculations. The value of the entropy is multiplied with the temperature. The entropy contribution is taken at 298K. The zero point energy is given as a value (top).

3.2.4. Hydration energy simulation

Hydration energy was calculated to complete the thermodynamic cycle mentioned previously. To assess different hydration models, three different models were tested against each other. The models are COSMO,[46] PBF and SM6/8[47-49]. In these models, the water surrounding the molecule is modeled as a continuous dielectric media. This causes individual water molecules to be neglected, removing possible orientational solvent-solute interactions and effects. A more rigorous way that includes possible orientational effect, is to use a finite system of water molecules and allow the molecule to relax in the presence of the water molecules.

COSMO is an embedded functionality in the DMol^3 module of Materials studio. Molecular geometries previously optimized by CASTEP were assumed to be accurate enough as starting points and used as input for the DMol^3 hydration calculations. The unit cells of the isolated vacuum molecules were removed and geometry optimizations were performed with the solvent effect added.

PBF, SM6 and SM8 are modules in the Jaguar program. The procedure used for PBF and SM6 was identical to the COSMO one. Due to some internal limitations in the SM8 method, Jaguar is unable to perform geometry optimization. This causes the SM8 hydration energy to be calculated on CASTEP optimized geometries.

3.2.5. Benchmark set

A benchmark set consisting of organic molecular crystals, $C19_{RT}$,[50] was selected after a literature study. The benchmark set consisted of 19 crystals with experimentally determined lattice parameters and lattice energies. Crystal structures of the compounds were found searching the Cambridge Structural Database [51] (CSD) with unique codes available from the benchmark set. The search was performed using the program Conquest.[52] Crystal structural data was found for all molecules except adamantane and cyanamide (lack of data for these two molecules lead to exclusion from the benchmark set).

The structures were visualized in Mercury [53] and afterwards exported to Material Studio 7.0. Using CASTEP and Forcite in Material Studio a selected number of crystal structures were geometry optimized (allowing both lattice parameters and atomic positions to relax) with different quality settings. Comparison of experimental data to the different quality settings were done in order to decide which settings corresponded to the best balance between simulation time and accuracy.

The determined quality setting (see APPENDIX A for the different settings) was used to optimize the crystal geometries with TS and Grimme dispersion scheme, as well as with the force field COMPASSII. Vacuum energies of the molecules were simulated by taking a molecule from the crystal, isolating it in vacuum and allowing the atomic positions to relax. The optimized structures (crystal and vacuum) were exported to Mercury. Using the DFT-D3 software suite, a correction term was calculated in order to obtain energies at the D3 level of theory.

3.2.6. Indomethacin and Carbamazepine

Crystalline structures for polymorphs, co-crystals and solvents for the respective molecules were found searching CSD with a name query. Each hit was inspected and non-relevant structures (*e.g.* salts) were removed. The remaining structures were imported to Materials Studio for geometry optimizations, with the settings established during the benchmark. The optimized

structures were exported for DFT-D3 calculation of more accurate dispersion energies. By using eq 1, sublimation energies of the multi-component crystals were calculated. Hydration energies and entropy contributions were then simulated which made it possible to calculate the solubility.

4. RESULTS & DISCUSSION

Experimental results are first presented, followed by simulated results and thereafter exploration of possible correlations between experimental and simulated data.

4.1. EXPERIMENTAL STUDIES

The experimental results will be categorized under the two different molecules.

4.1.1. Indomethacin

Indomethacin polymorphs, solvates and cocrystals were produced using the procedures described in section 3.1.1.1. X-ray characterization were performed on the samples. Unique crystal diffractograms can be seen in Figure 10. Analysis of DSC and TGA data was performed (see APPENDIX B for figures) in order to ensure that preparation was successful. Table 3 summarizes the analysis of the crystal forms of indomethacin. All data was compared to available literature.[33-35]

Table 3. Summary of the XRPD, DSC and TGA analysis of the indomethacin crystal forms. Numbers in parenthesis indicate the theoretical weight loss for a hemi-solvate. Results for all unsuccessfully prepared crystal forms will be omitted from the report.

Wanted crystal form	Successful preparation	T_m (DSC) [K]	TGA weight loss [%]
α-polymorph	Yes	153.3	<0.1
γ-polymorph	Yes	160.0	<0.1
Acetone solvate	No	-	
tert-Butanol hemi-solvate	Yes	101.0	8.8 (9.4)
Diethylether solvate	No	-	
Ethanol solvate	No	-	
Ethyl acetate solvate	No	-	
Methanol hemi-solvate	Yes	83.3	4.9 (4.3)
2-Propanol solvate	No	-	
Saccharin co-crystal	Yes	182.8	<0.1
Tetrahydrofuran solvate	No	-	

Dissolution experiments were performed on the different available indomethacin forms and the results can be seen in Table 4.

Since dissolution rate is closely tied to the surface area, it is imperative to determine the surface area of the powder. To get an overview of the morphology of the different crystal forms (as well as decision basis for which method to use for determining the particle size distribution), SEM pictures were taken (Figure 11). The morphology and size of all successfully prepared indomethacin crystal forms can be seen in APPENDIX B.

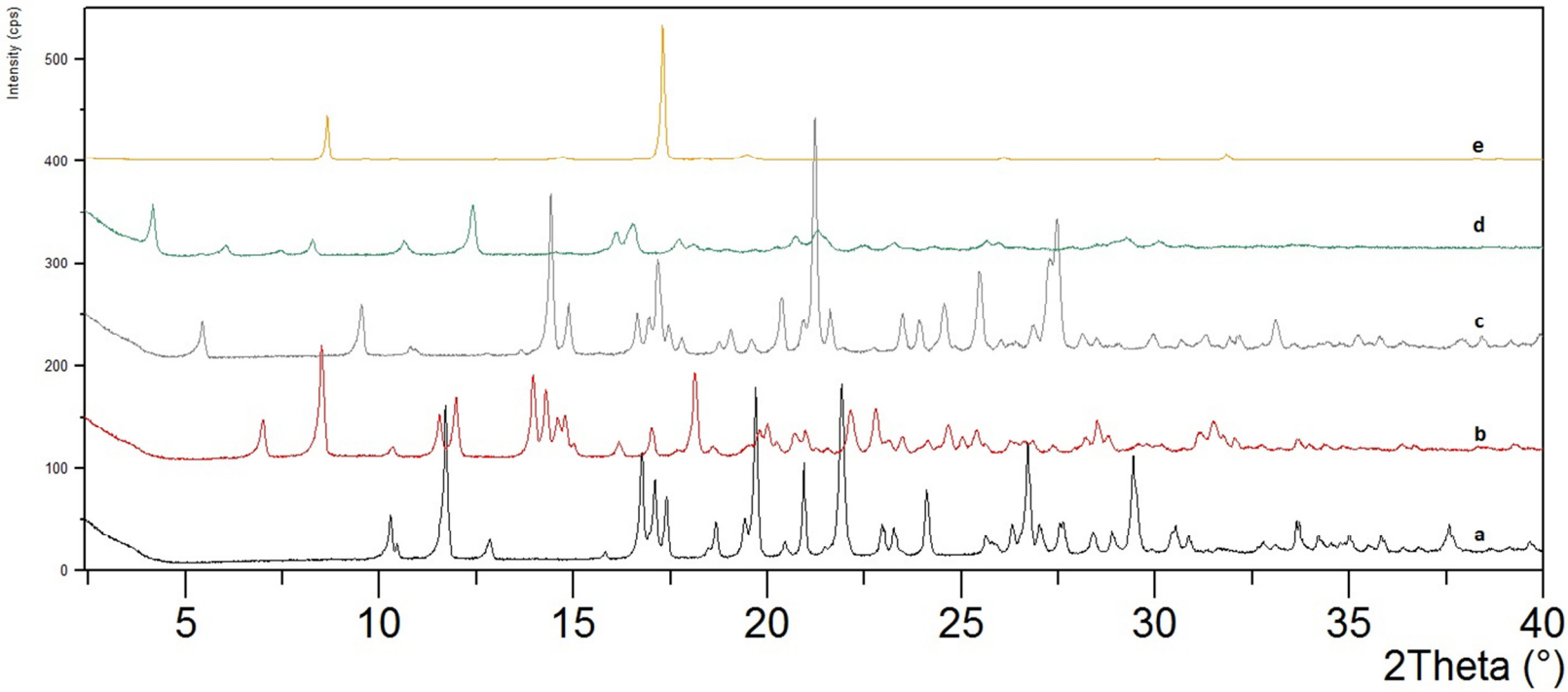

Figure 10. X-ray diffractograms of the successfully prepared indomethacin crystals. From bottom: γ-polymorph (a), α-polymorph (b), saccharin co-crystal (c), *tert*-butanol hemi-solvate (d) and at the top methanol hemi-solvate (e). The lack of detail in the top diffractogram is due to orientational effects of the indomethacin methanol hemi-solvate crystals.

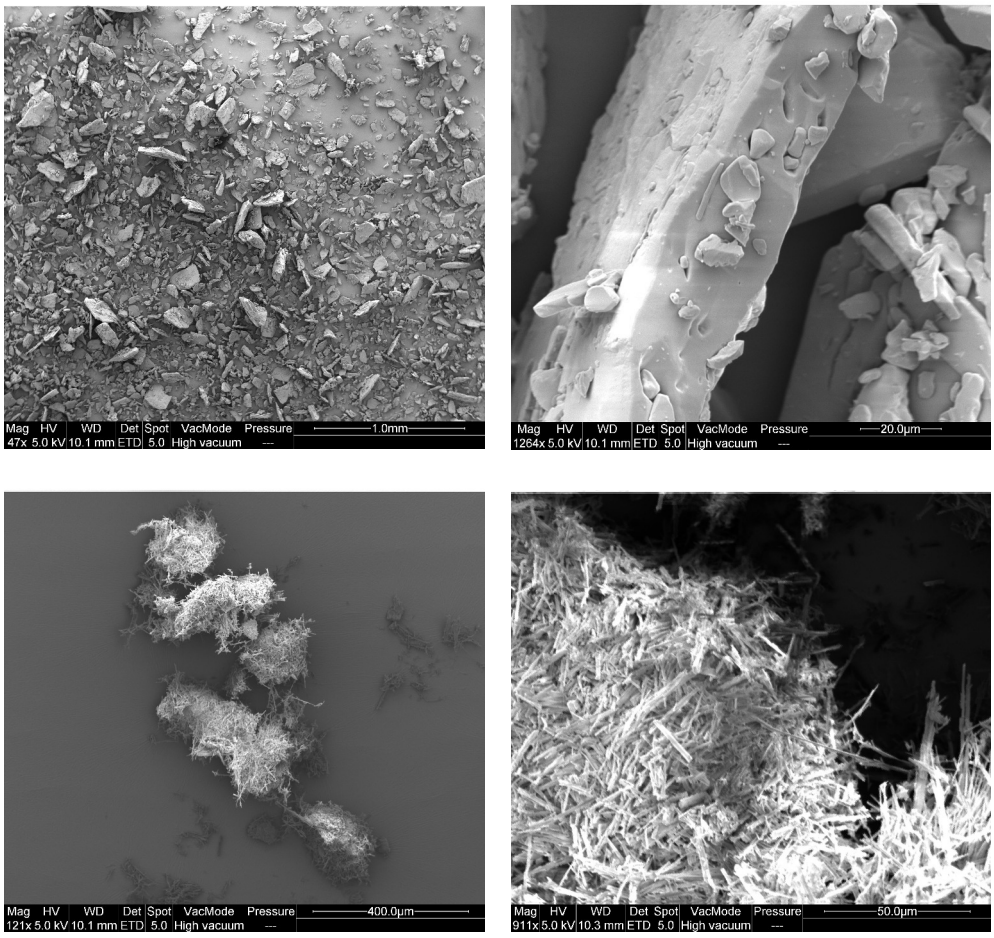

Figure 11. SEM morphology overview of the polymorphs of indomethacin: γ-polymorph (top) and α-polymorph (bottom). N.B. the varying length scales and morphology. The α–polymorph can be seen to consist of primary particles with fairly homogenous sizes, but they are prone to forming large aggregates or agglomerates (secondary particles).

The morphology and size can be seen to vary greatly between samples which cause a major problem during the dissolution rate measurements. A powder consisting of small particles (primary particles) will have tremendously larger surface area available for dissolution. However, even if the powder consists of small primary particles, they can agglomerate into larger particles (secondary particles), which greatly reduces the solvent available surface area. A more similar morphology of all samples could possibly be achieved by developing the crystallization protocols, but this might take months for a single substance. Crystallization protocol development was therefore not done in this project. Another alternative would have been to use some form of grinding or milling to achieve more homogeneity between the samples. Such techniques might induce amorphous materials and were not used in the project.

Using the SEM pictures as a starting point, the particle size distribution was determined using different pressures. The distributions of γ- and α-polymorph can be seen in Figure 12. It can easily be seen that determining a true value of the particle sizes is far from trivial. The γ-polymorph shows an ideal pattern, where the particles are deagglomerated as the pressure

increases. Even though deagglomeration is successful, it is still a challenge to determine the true size of the primary particles. Since the experimental method assumes spherical particles, the reported size is the mean of the particles' length scales. This adds a great uncertainty to the experiments and makes it hard to determine the true size unless the exact morphology can be

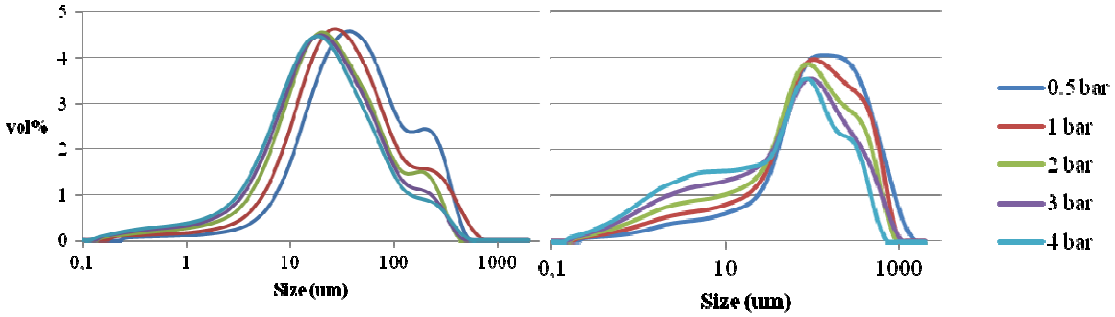

Figure 12. PSD of indomethacin γ-polymorph (left) and α-polymorph (right). Looking at the γ-polymorph, intensity of the peak at ~200µm decreases as the pressure is increased. This suggest successful de-agglomeration. Turning the attention to the α-polymorph, the shape of the curve changes substantially as the pressure is increased. Looking more closely around 10µm, a peak emerges as the pressure is increased, which could be due to deagglomeration.

examined. Looking at the data obtained for the α-polymorph, the situation is even more complex. Deagglomeration occurs as the distribution shifts towards lower sizes, but a peak is emerging for sizes 0.5 - 20 µm. This effect can be due to primary particles that break free as deagglomeration occurs (which would be ideal), but it could also be destruction of the primary particles. In order to correctly understand what happens, the previously taken SEM micrographs were studied with the PSD results in focus. It was concluded that the most probable cause of the peak buildup was due to deagglomeration. See APPENDIX B for the PSDs for the multi-component indomethacin crystals. D_{50} values for the particle sizes for the different samples were taken at 3 bars, except for the methanol solvate from which it was taken from 1 bar. Particle sizes, along with surface adjusted dissolution rates for indomethacin crystals are summarized in Table 4.

Table 4. Summary of determined particle sizes, dissolution rates and the corresponding surface independent dissolution rates for indomethacin crystal forms. It can clearly be seen that the surface area plays a big role in the dissolution rate.

Indomethacin crystal form	Particle size (D_{50}) [µm]	Dissolution rate [nmol/s]	Surface adjusted dissolution rate [nmol/cm^2 s]
α-polymorph	21.1	0.6	10.0
γ-polymorph	61.5	0.8	10.2
tert-Butanol solvate	4.5	1.45	2.2
Methanol solvate	6.8	1.2	5.2
Saccharin cocrystal	179.0	0.4	21.3

The PSD method used is not ideal; the optimal one would have been to use a method measuring the distribution in a saturated liquid dispersion. This would have given a more accurate particle size for the powder. The reason for using a dry dispersion method is due to lack of material.

During the preparation stage, the required amount for all subsequent characterizations was underestimated, causing a lack of material during the particle size distribution determinations.

4.2. CARBAMAZEPINE

Carbamazepine crystal forms were prepared and subsequently analyzed by XRPD, DSC and TGA. Based on the data analyzed (see APPENDIX C for all figures), it was concluded that preparation resulted in ten unique crystal forms, summarized in Table 5. In Figure 13 the XRPD patterns for all the successfully prepared crystal forms of CBZ can be seen.

Table 5. Outcome of the crystal form characterizations performed with XRPD, DSC and TGA. Numbers in parenthesis indicate either the theoretical weight loss for a mono/hemi-solvate or the temperature range in which no weight loss was observed. Results for unsuccessfully prepared crystal forms will be omitted from the report.

Wanted crystal form	Successful preparation	T_m (DSC) [K]	TGA weight loss [%]
Polymorph I	Yes	190.7	<0.1
Polymorph II	No	-	-
Polymorph III	Yes	174.7	<0.1
Acetic acid mono-solvate	Yes	133.3	19.7 (20.3, mono)
Acetone solvate	No	-	-
Dimethylsulfoxide hemi-solvate	Yes	90.5	17.25 (19.8, hemi)
Glutaric acid co-crystal	Yes	124.9	<0.1
Malonic acid co-crystal	Yes	107.0	26.5
N,N-dimethylformamide solvate	No	-	-
Nicotinamide co-crystal	No	-	-
Saccharin co-crystal	Yes	174.2	<0.1 (25-160 °C)
Salicylic acid co-crystal	Yes	158.7	<0.1 (25-140 °C)
Succinic acid co-crystal	Yes	169.8	<0.1
Terephtalaldehyde co-crystal	Yes	63.1	32.2

Surface analysis of the different crystal forms of CBZ were performed with SEM and a selection of the different morphologies are presented in Figure 14. See APPENDIX C for micrographs of all the surfaces. Just as the case for indomethacin, surface morphology varies greatly between the samples. The discussion on indomethacin morphologies therefore applies on the samples of carbamazepine as well.

Particle sizes, dissolution rates and the corresponding surface adjusted dissolution rate from PSD and uDISS experiments are summarized in Table 6. All D_{50} values were taken at 3 bars except for malonic acid co-crystal for which it was taken at 4 bars.

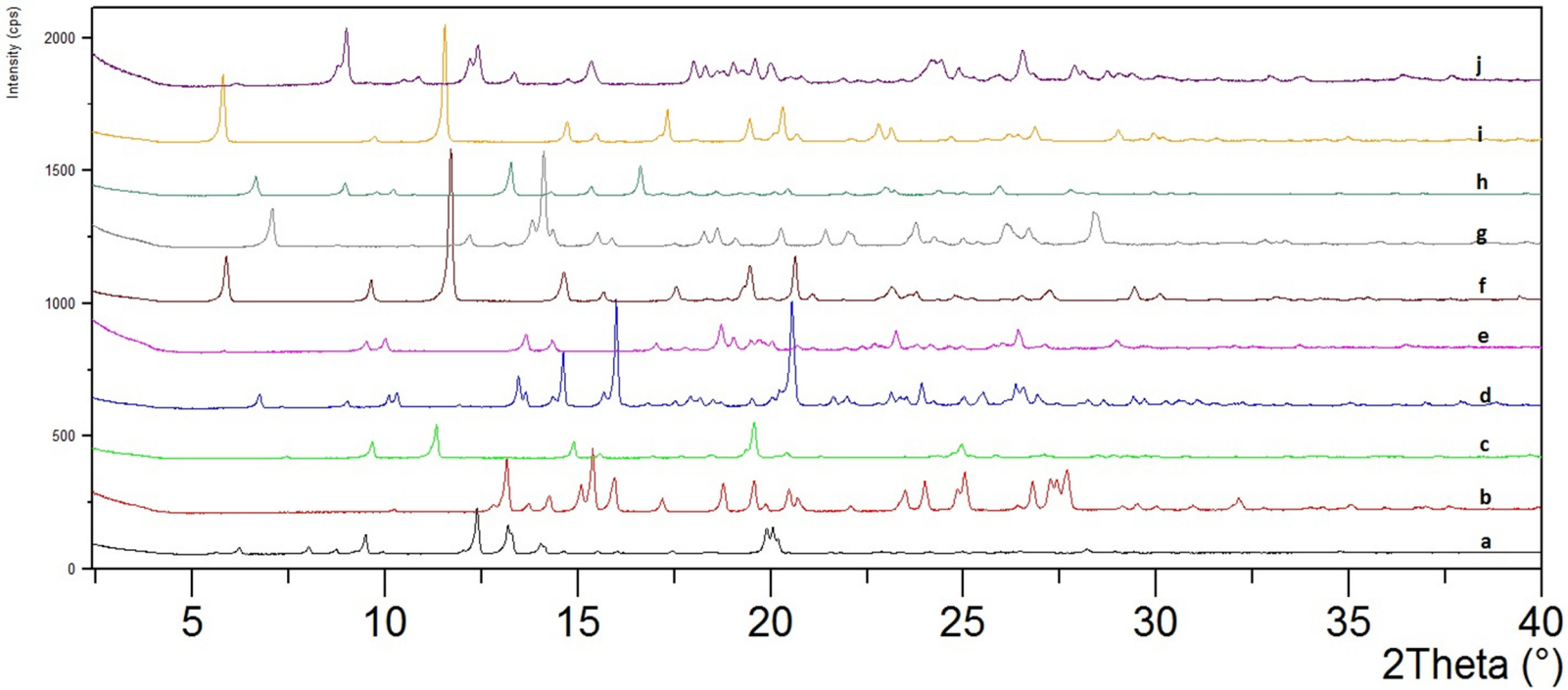

Figure 13. Combined X-ray diffractograms for all successfully synthesized polymorphs, solvates and co-crystal of carbamazepine. Starting from the bottom, data is shown for Polymorph I (a), Polymorph III (b), Acetic acid mono-solvate (c), Dimethylsulfoxide hemi-solvate (d), Glutaric acid co-crystal (e), Malonic acid co-crystal (f), Saccharin co-crystal (g), Salicylic acid co-crystal (h), Succinic Acid co-crystal (i) and Terephtalaldehyde co-crystal (j).

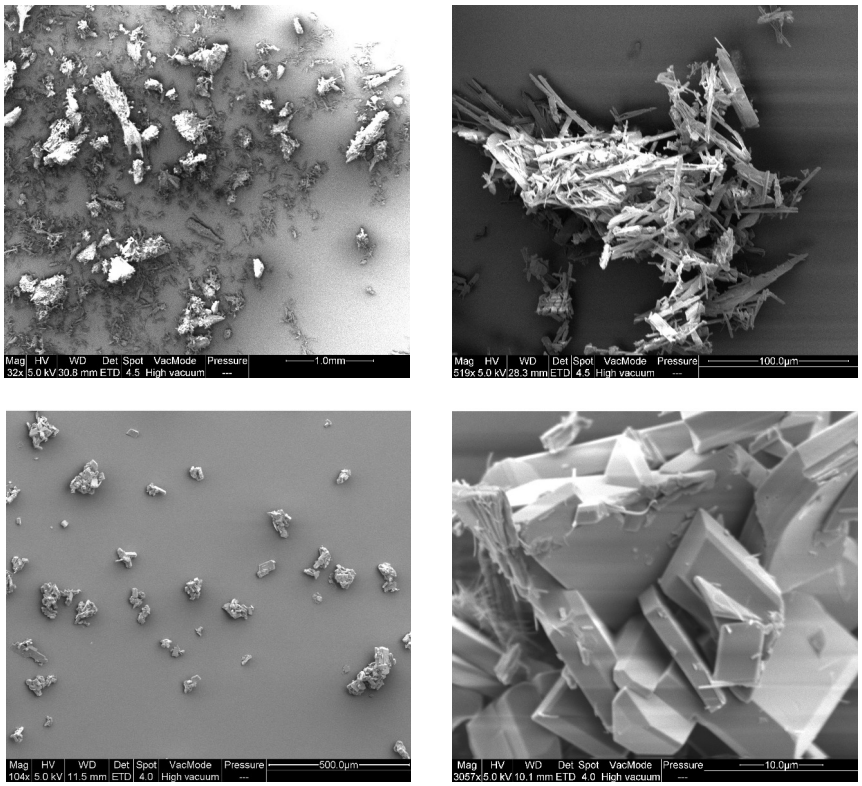

Figure 14. SEM morphology pictures of carbamazepine glutaric acid co-crystal (top) and carbamazepine saccharin co-crystal (bottom) N.B. the varying length scales.

Table 6. Summary of determined particle sizes, dissolution rates and the corresponding surface independent dissolution rates for carbamazepine crystal forms. It can clearly be seen that the surface area plays a big role in the dissolution rate. Particle size could not be determined for DMSO due to lack of material.

Carbamazepine crystal form	Particle size (D_{50}) [µm]	Dissolution rate [nmol/s]	Surface adjusted dissolution rate [nmol/cm^2 s]
Polymorph I	20.6	1.8	6.9
Polymorph III	3.8	4.9	2.6
Acetic acid mono-solvate	10.2	1.5	2.9
Dimethylsulfoxide hemi-solvate	-	0.1*	-
Glutaric acid co-crystal	12.5	0.5	1.4
Malonic acid co-crystal	149.0	0.1	3.1
Saccharin co-crystal	32.1	1.0	4.6
Salicylic acid co-crystal	11.2	0.8	1.6
Succinic acid co-crystal	45.7	0.7	7.7
Terephtalaldehyde co-crystal	212.0	0.5	22.7

*It is possible that the sample is not dry enough due to the high boiling point of dimethylsulfoxide which might affect the dissolution measurements.

4.3. IN SILICO STUDIES

The module CASTEP was first benchmarked to find optimal settings, followed by testing on the $C19_{RT}$ set. Using the established settings, simulations on the polymorphs, co-crystals and solvates of indomethacin and carbamazepine were carried out.

4.3.1. Simulation parameters

Benchmarking the CASTEP module of Materials Studio 7.0 was done using a selection of the $C19_{RT}$ benchmark set. Different quality settings (*medium, fine* and *ultra-fine*) were tested (see APPENDIX A for a complete description of the different quality settings) and the optimal was determined to be *"fine"*. This was based on the fact that "*medium"* caused convergence issues and *"ultra-fine"* significantly increased calculation time but only offered very limited accuracy gains. The electronic cutoff was changed to a custom value of 550 eV for consistency throughout the set.

The impact of vacuum cell size on the calculated energy was tested (Figure 15) and a separation of at least 20-25 Å between repeated molecules was chosen as optimal. Based on this, a cubic vacuum cell with lattice lengths of 30 Å was used throughout the vacuum energy simulations.

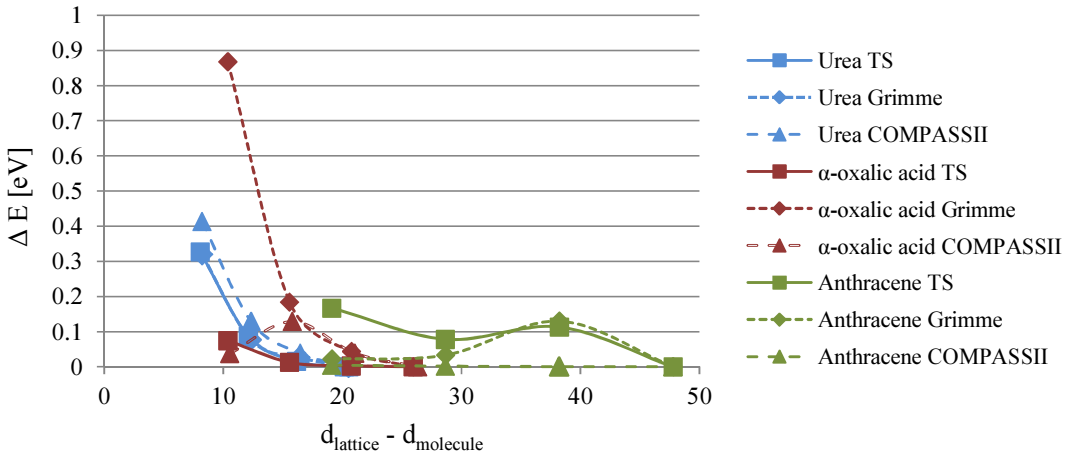

Figure 15. Impact of cell size on relative vacuum energy for a set of three molecules. The molecules investigated were of varying dimension to cover as many cases as possible. The energy of the largest vacuum cell for each of the three different molecules was taken as the baseline (the correct energy), effectively using it as a reference energy. From the appearance of the graph, adding at least 20-25 Å to the molecule length was regarded as optimal.

4.3.2. $C19_{RT}$ benchmark set

Using the settings described in section 3.2 and "*fine*" quality setting, geometry optimization and energy calculations were performed on the benchmark set $C19_{RT}$. The obtained geometries were compared to experimental data (Figure 16 and Table 7) to see how the different methods compared to each other. The $RMSD_{15}$ values in the table were calculated by Mercury. The performance is similar for TS and Grimme dispersion correction methods, with TS generating slightly better geometries. Even though TS produced better geometries overall, there is not a

clear systematic trend in the results which would be desirable. The Grimme geometries can be seen to show a small systematic trend. These two aspects were taken into account as the methods were deemed equal for simulation of molecular geometries. The geometries obtained from COMPASSII showed very large discrepancies leading to the decision to consider the method inadequate for the project aims.

Table 7. Summary of percentage deviation of lattice parameter ratios and unit cell volumes as well as $RMSD_{15}$ values evaluated with Mercury. MAE is mean absolute error, and MUE is the mean unsigned error.

	a/c	**b/c**	**V_{cell}**	**$RMSD_{15}$**
TS - MAE	2.01	2.89	4.40	-
TS - MuE	0.41	1.30	-0.63	0.13
Grimme - MAE	3.22	3.19	6.06	-
Grimme - MuE	0.66	0.66	-4.21	0.17
CompassII - MAE	5.76	7.23	5.64	-
CompassII - MuE	2.61	2.65	-4.38	0.37

Calculated lattice energies of the benchmark set can be seen in Figure 17. The energies calculated by Grimme can clearly be seen to be in better agreement with experimental data than TS. Based on this, the decision was made to use Grimme dispersion for the future simulations.

Looking at a summary of the geometries obtained using COMPASSII (Table 7), it can clearly be seen that they are not optimal. The energy is however better than those obtained using TS dispersion scheme. A hypothesis that the energies calculated by COMPASSII might be better if they were calculated on more “correct” geometries was tested. Results of this can be seen in Figure 28 (APPENDIX A). The lattice energies calculated by COMPASSII became worse as it was applied to geometries obtained by TS or Grimme methods. Based on the data, hybrid calculations like these (using COMPASSII energy calculations on DFT optimized geometries) are not likely to offer any benefits.

The implemented version of Grimme dispersion in Materials Studio 7.0 is D2. In order to test accuracy, the newer (D3) correction was tested in a separate program. Calculated energies for the different versions of Grimme dispersion corrections can be seen in Figure 18 and Table 8. With a MAE of 6.89 kJ/mol for D3 used with BJ damping, the method was deemed accurate and was determined to be used on the real systems of indomethacin and carbamazepine crystals.

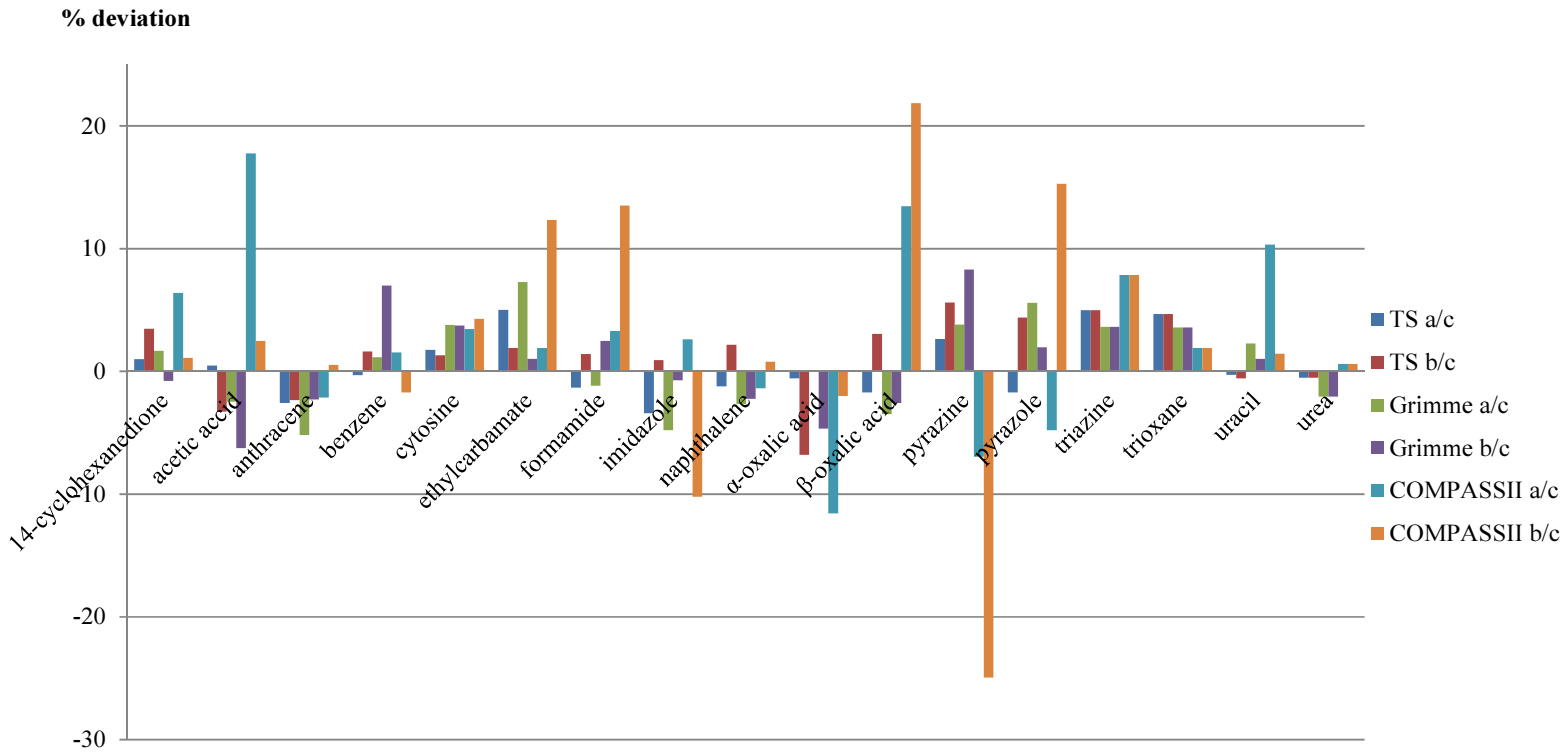

Figure 16. Lattice parameter ratios (a/c and b/c) for the different methods (TS, Grimme D2 and COMPASSII) as percentage deviation compared to experimental data.

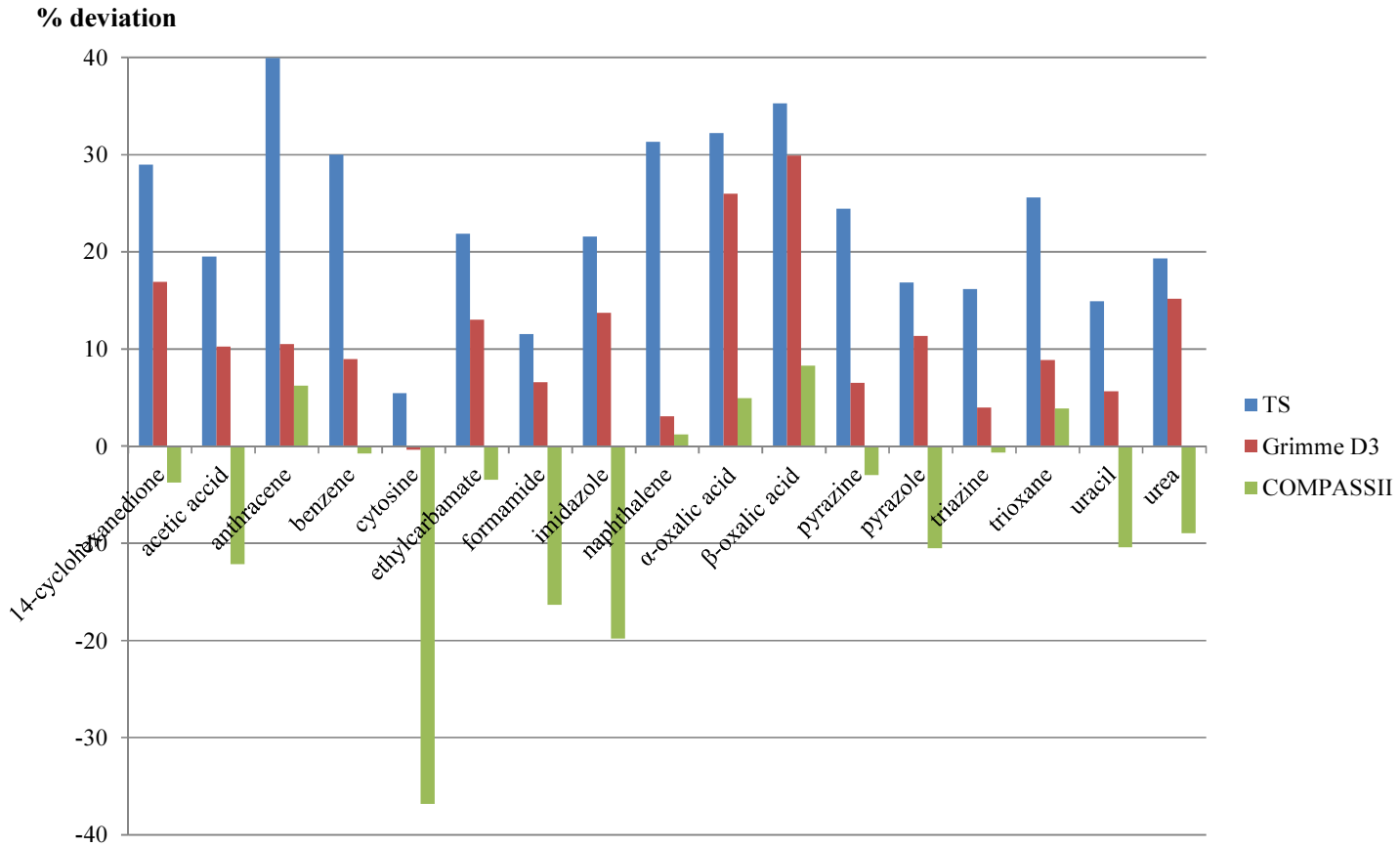

Figure 17. Percentage deviations for simulated compared to experimental lattice energies.

These results from the benchmarking (using the D3 method) are in overall agreement with Grimme et al.[54] as the lattice energies of cytosine and urea are under- and overestimated respectively. However, the lattice energy of α- and β-oxalic acids are gravely overestimated, in contrast to earlier results. The error might originate from the fact that in that study, [54] geometries were optimized with the D3 dispersion, as opposed to these results which only uses D3 dispersion as an energy correction term.

A summary of the results from the different Grimme dispersion corrections can be found in Table 8. When used with Becke-Johnson damping the lattice energies are within ±6.89 kJ/mol from the experimental values, which is close to the experimental variability.

Table 8. Comparison of the different versions of Grimme dispersion scheme (D2& D3) using both zero (Z) damping and Becke-Johnson (BJ) damping for the sum(S) and mean (M) of the absolute (A) and unsigned (U) errors. MD stands for max deviation. Deviations are calculated as the difference between experimental and simulated values.

	Absolute deviations [kJ/mol]			**% deviations**		
	D2	**D3 - Z**	**D3 - BJ**	**D2**	**D3 - Z**	**D3 - BJ**
SAE	164.87	210.43	117.09	190.77	245.44	127.08
MAE	9.7	12.38	6.89	11.22	14.44	7.48
SuE	163.54	210.43	72.63	189.98	245.44	87.76
MuE	9.62	12.38	4.27	11.18	14.44	5.16
MD	28.65	29.07	22.59	29.89	30.39	23.57

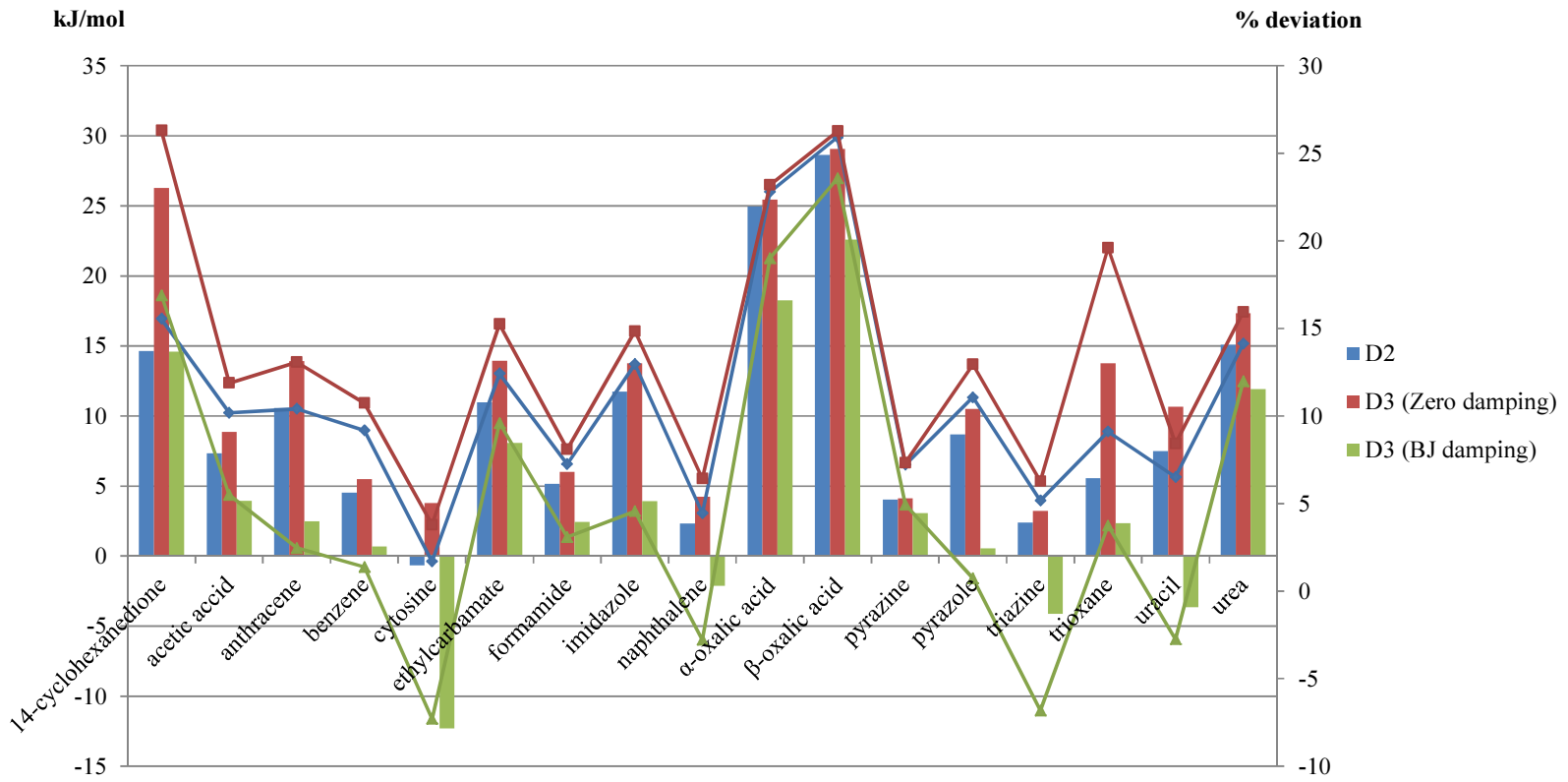

Figure 18. Absolute deviations (left side, bars) and percentage deviation (right side, lines) for the different versions of Grimme dispersions. (Overall, D3 with Becke-Johnsson (BJ) damping appears to perform best.)

Simulations were performed on the benchmark set to obtain the entropy contribution to the internal energy. In Figure 19, calculated entropy is plotted against the crystal volume where a distinct linear relationship is seen with the form $S = 0.1857V_{cell} + 8.2168$. Increasing the unit cell volume corresponds to an increase in the entropy. This effect might be explained by the fact that molecules in a crystal need to be in fairly close proximity to each other to ensure crystal stability. A volume increase of the unit cell is therefore translated to an increase in number of atoms/molecules. This allows more molecules to vibrate, thus increasing the entropy.

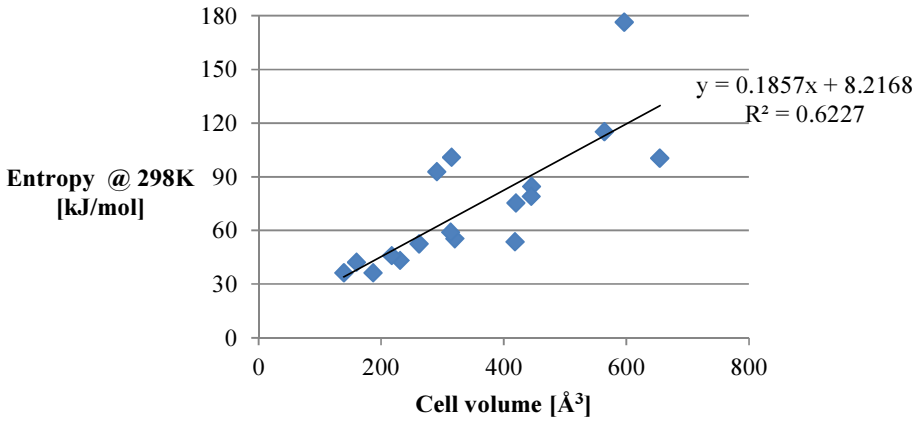

Figure 19. Crystal entropy plotted against unit cell volume for the benchmark set.

From the free energy calculations the zero point energy (ZPE) was extracted. In Figure 20, the ZPE values are plotted against the unit cell volume. A very good relationship $ZPE = 2.3547V_{cell}$ was found. These two correlations, along with the enthalpy correction term $H = U - 2RT$ make it possible to estimate the temperature contribution towards Gibbs free energy without performing tedious entropy calculations.

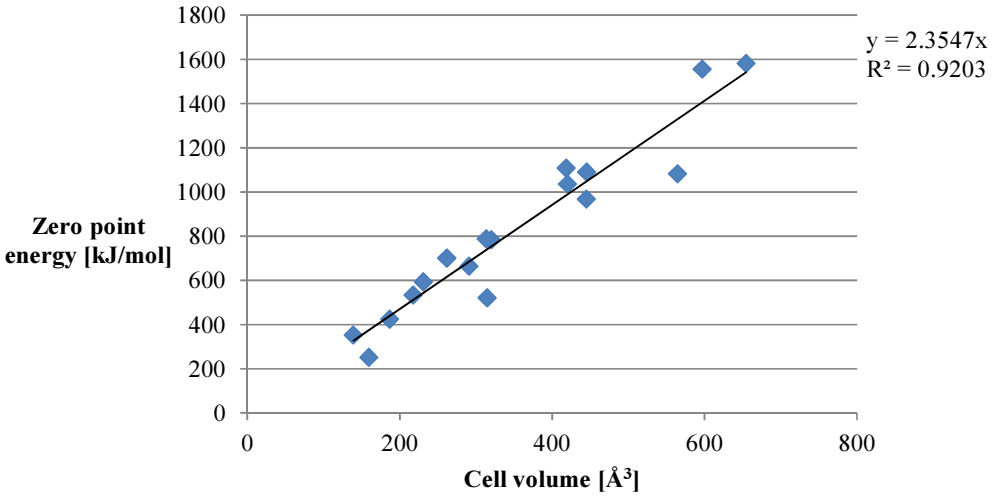

Figure 20. Zero point energy plotted as a function of unit cell volume for the benchmark set.

Using the simulated entropy contributions, the lattice free energy was calculated. In Figure 21 the sublimation free energy is visualized against the sublimation energy. The entropy contribution can be seen to change the results somewhat, but there is a strong correlation.

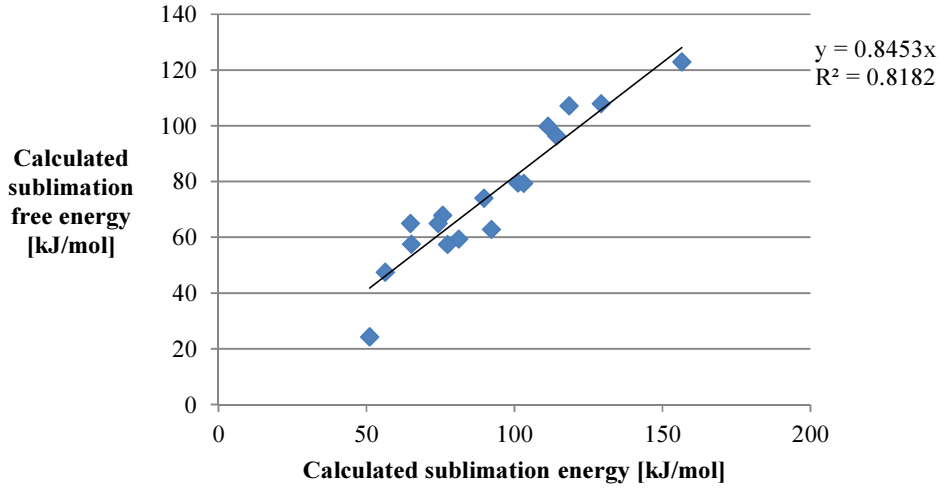

Figure 21. Gibbs' free energy plotted against the energy of the lattice for the benchmark set. A clear correlation can be seen.

Melting temperature should depend on the lattice energy since a high melting temperature should correspond to a larger amount of energy required to break the bonds in the crystal. The relationship between melting temperature and sublimation energy was examined for the $C19_{RT}$ benchmark set in Figure 22. A clear trend can be seen with $T_m = 3.0436 \cdot E_{sub} + 111.88$.

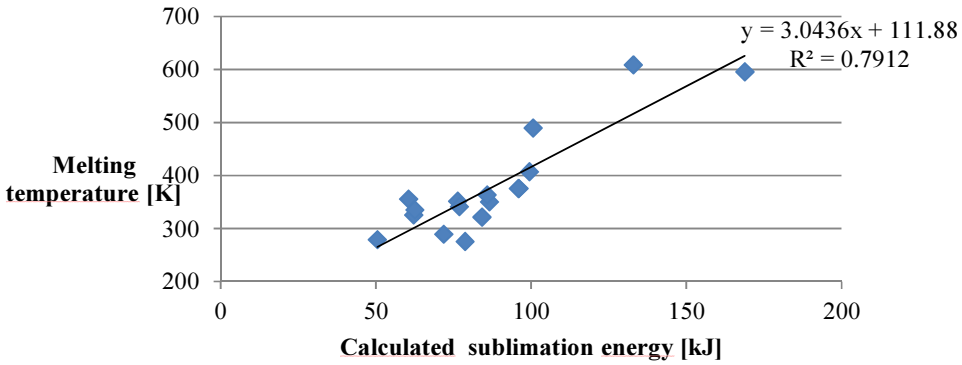

Figure 22. Plot of melting temperatures as a function of sublimation energies for the $C19_{RT}$ benchmark set.

No accurate literature values of hydration energies were found for the benchmark set. However, literature values of solubilities were found for six of the molecules. Using eq 4 the solubilities of these molecules of the benchmark set was calculated. In Figure 23 experimental and calculated solubilities are compared to each other. A correlation can be seen to exist, even though it is based on only a small number of crystals. It would be interesting if this correlation

would hold if more molecules of the benchmark were added (*i.e.* if more solubilities could be found in the literature).

After all simulations were performed on the benchmark set, simulations were then performed on indomethacin and carbamazepine crystals using the found correlations.

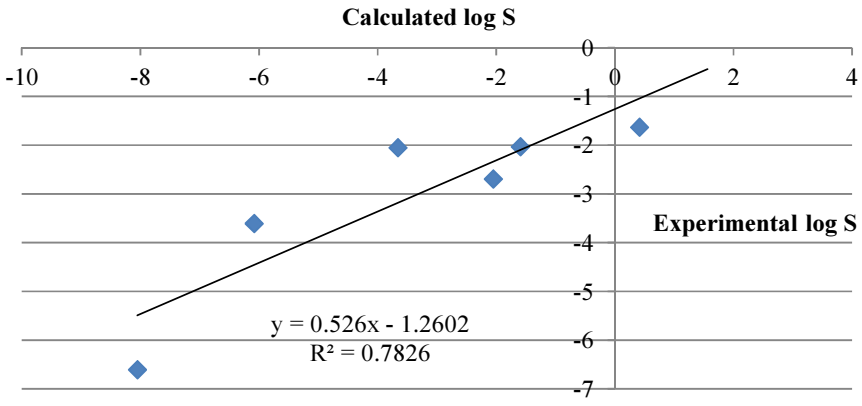

Figure 23. Experimental solubility as a function of calculated solubility for a selection of the benchmark $C19_{RT}$ set. The unit of the solubility is in mol/L.

4.3.1. Indomethacin & carbamazepine

Modeling of the substances was done with the settings established from the benchmarking set.

4.3.1.1. Lattice energy, entropy and hydration energy

Lattice energies were first calculated for both indomethacin and carbamazepine, and corrected with D3-BJ dispersion. As the entropy of the crystals was being calculated, it became clear that the chosen method was very time consuming. A simulation on one of the crystals took almost a month to complete. The optimal way would have been to change the method and rerun all calculations, but that was not feasible from a time perspective. Instead, the linear relationship $S = 0.1857V_{cell} + 8.2168$ was used for predicting the entropy contribution, even though extrapolation was necessary due to larger unit cells. Afterwards, entropy of the gas-phase molecules was calculated.

The different methods used for simulating the hydration energies were tested. In Figure 24 the different methods are compared to each other. It can be seen that the methods follow the same trend for almost all crystal forms, but the results are varying for the different polymorphs. Looking more closely at the SM6/8 results, the energy difference for the two carbamazepine conformers are roughly 10 kJ/mol. This is alarming since these two conformers are mirror images of each other, and as such their energy should be identical. Looking at either of the two other methods, the trends are similar. The varying result between indomethacin conformers can be explained easier due to the fact of different molecular configuration. This shows the importance of simulating data for the correct conformer.

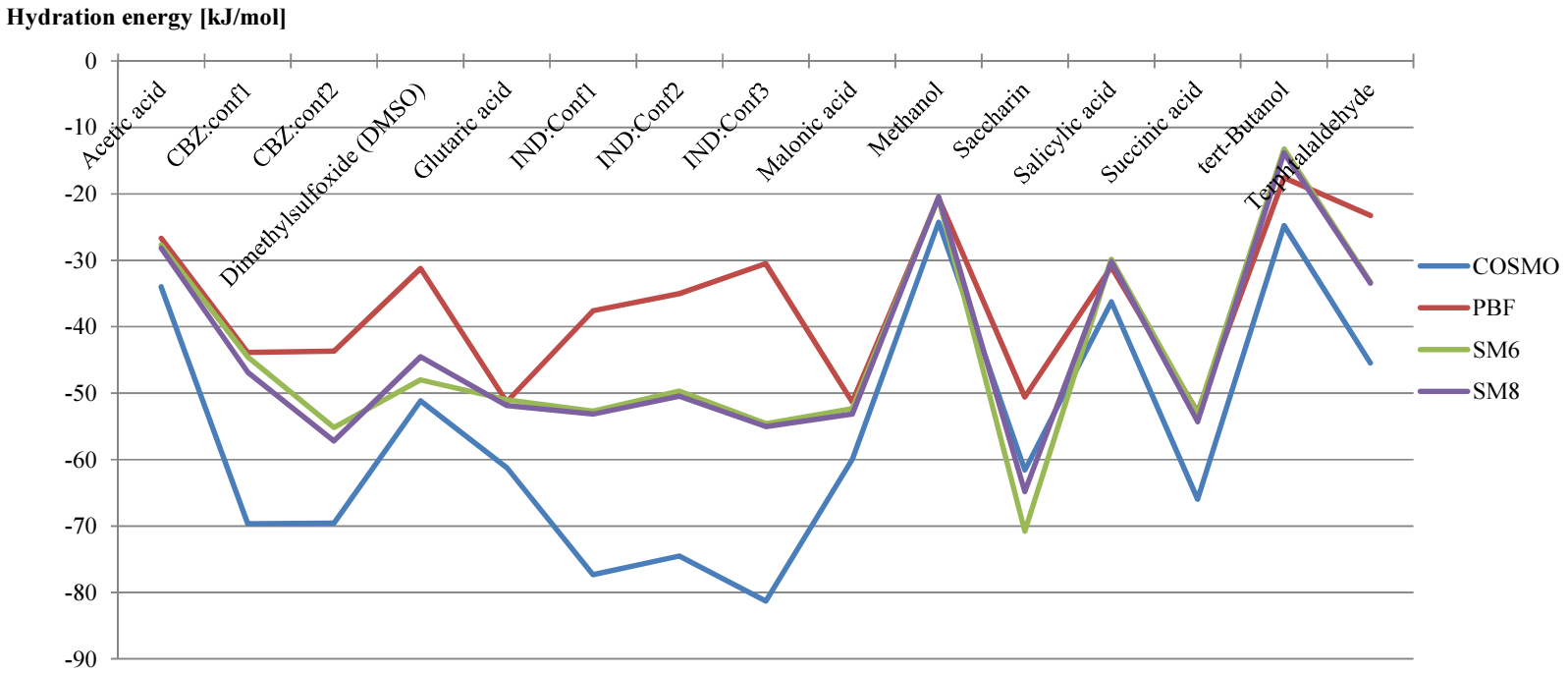

Figure 24. Comparison of different methods used for determining the hydration energies.

There is a difference of how the quality settings are defined between the methods. The settings *"Fine"* was used in DMOL3 and *"Accurate"* in Jaguar were deemed equal. From the comparison in Figure 24, it was decided to use PBF as solvation method.

Simulated lattice and hydration free energies and the calculated solubility are summarized in Table 9.

Table 9. Summary of lattice and solvation free energies as well as the corresponding solubility.

IND Crystal form	**Lattice free energy [kJ/mol]**	**Solvation free energy [kJ/mol]**	**Calc. solubility [mol/L]**
α-Polymorph	-10.79	-23.54	5.48E+08
γ-Polymorph	-11.28	-23.05	4.50E+08
tert-Butanol hemi-solvate	-67.17	-15.63	2.25E+07
Methanol hemi-solvate	-40.31	-24.01	6.60E+08
Saccharin co-crystal	-4.29	-41.67	8.26E+11
CBZ Crystal form	**Lattice free energy [kJ/mol]**	**Solvation free energy [kJ/mol]**	**Calc. solubility [mol/L]**
Polymorph I	-77.25	33.54	5.40E-02
Polymorph III	-89.23	45.52	4.30E-04
Acetic acid mono-solvate	-57.70	22.51	4.63E+00
Dimethylsulfoxide hemi-solvate	-71.03	33.57	5.35E-02
Glutaric acid co-crystal	-103.31	55.83	6.68E-06
Malonic acid co-crystal	-32.26	-14.03	1.18E+07
Saccharin co-crystal	-36.05	-11.07	3.57E+06
Salicylic acid co-crystal	-42.03	4.70	6.13E+03
Succinic acid co-crystal	-42.12	-4.68	2.70E+05
Terephtalaldehyde co-crystal	-85.38	48.44	1.32E-04

4.4. CORRELATIONS

In Figure 25 and Figure 26, the experimentally determined melting temperature was plotted against calculated sublimation free energy for indomethacin and carbamazepine crystals. Even though the trend was quite clear for the benchmark set, the correlation found is not applicable on the two APIs. They both show a trend where an increasing sublimation free energy corresponds to a decrease in the melting temperature. This is something that is hard to understand and the results here might be erronous. The errors might originate from the entropy predictions, since the entropy was not simulated, merely predicted using an extrapolated formula.

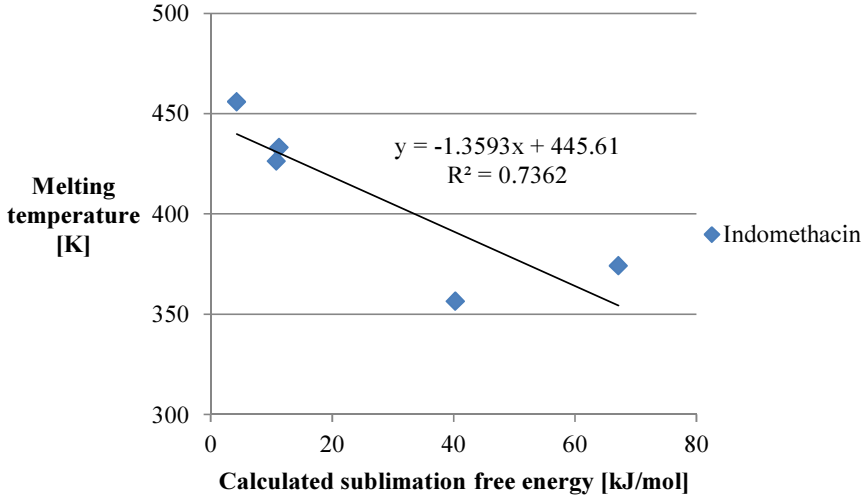

Figure 25. Plot of experimentally determined melting temperatures as a function of simulated sublimation energies for the crystals of indomethacin.

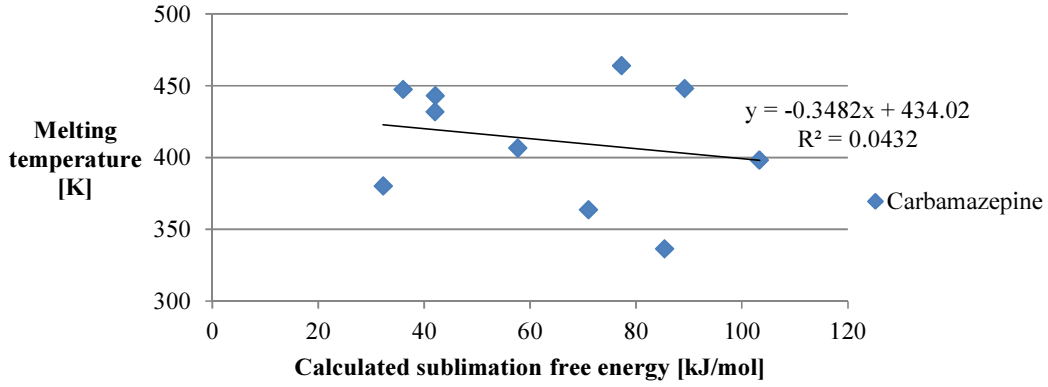

Figure 26. Plot of experimentally determined melting temperatures as a function of simulated sublimation energies for the crystals of carbamazepine.

The experimentally determined dissolution rates, adjusted for surface area, are compared to the simulated solubilities in Table 10. For indomethacin, the simulated ranking is fairly accurate even though the simulated absolute values of the solubility is obviously erroneous. The ranking of carbamazepine crystals is not so accurate. This might be due to the variation in solubility for carbamazepine which is even greater than for indomethacin.

Table 10. Comparison of experimentally determined data and simulated solubilites for indomethacin and carbamazepine.

	Experiments		Simulations	
IND Crystal form	**Dissolution rate [nmol/cm^2 min]**	**Ranking**	**Solubility [mol/L]**	**Ranking**
α-polymorph	10.0	3	5.5E+8	4
γ-polymorph	10.2	2	7.8E+10	2
tert-Butanol hemi-solvate	2.2	5	2.3E+7	5
Methanol hemi-solvate	5.2	4	6.6E+8	3
Saccharin co-crystal	21.3	1	8.3E+11	1
CBZ Crystal form	**Dissolution rate [nmol/cm^2 s]**	**Ranking**	**Solubility [mol/L]**	**Ranking**
Polymorph I	6.9	3	5.4E-2	6
Polymorph III	2.6	7	4.3E-4	7
Acetic acid mono-solvate	2.9	6	4.6E+0	5
Glutaric acid co-crystal	1.4	9	6.7E-6	9
Malonic acid co-crystal	3.1	5	1.2E+7	1
Saccharin co-crystal	4.6	4	3.6E+6	2
Salicylic acid co-crystal	1.6	8	6.1E+3	4
Succinic acid co-crystal	7.7	2	2.7E+5	3
Terephtalaldehyde co-crystal	22.7	1	1.3E-4	8

Having all the results from the calculations it is obvious that something in the method is not ideal. The first and most obvious source is probably the entropy contribution, which has been mentioned earlier. Another more complex error source might be that the thermodynamic cycle used in the project is not the correct one. It is possible that the co-former and the main molecule are closely bound together during either the sublimation and/or the hydration steps. These two alternatives differ substantially from the cycle used in this project and can be seen in Figure 27. The cycles represent different ways of modeling the problem, that greatly affect the outcome of the simulations. There is possibly not a correct definitive answer due to the fact that they both probably occur simultaneously (perhaps with different probabilities) during dissolution. This shows that solvation of a multi-component crystal is a complex process.

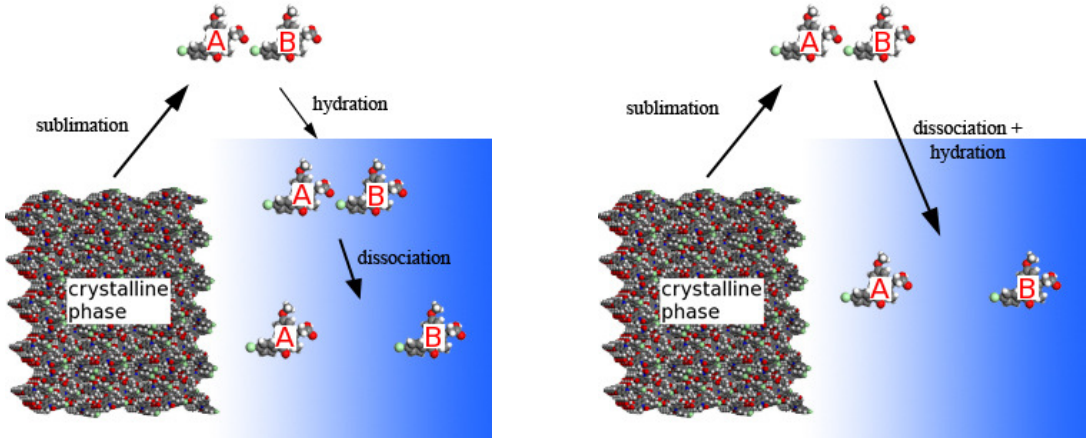

Figure 27. Alternative proposed thermodynamic cycles. In (a) the main molecule and co-former are treated as one unit. Dissociation occurs after solvation of the unit of the molecules. In (b) dissociation takes place so that the hydration processes are independent of each other. The blue background represents solvent.

5. CONCLUSIONS

Analyzing a sample with XRPD is a quick way to gain insights in the crystal structure of a material. If the XRPD data is combined with data from DSC and TGA, characterization of the crystal form of the material is in most cases possible.

Dissolution rates could be determined with high accuracy for different samples using µDISS. To be able to draw quantitative conclusions, the solvent available surface area is necessary. Using the experimentally determined particle size distribution data in combination with the surface morphology images visualized by SEM, quantitative interpretation of the dissolution data was possible.

Calculated lattice energies for the benchmark set $C19_{RT}$ is in good agreement with experimental data using the procedure established in the project. If Grimme D3 dispersion method was to be implemented into CASTEP it would probably lead to slightly more accurate geometries as well as crystal energies. The crystals with molecules allowing hydrogen bonding show a discrepancy in the results, indicating that first principles methods are less adequate to describe hydrogen bonding with a high degree of accuracy.

A correlation for the melting temperature as a function of sublimation energy was found using the benchmark set $C19_{RT}$, with good agreement of experimental data. The relationship between melting temperature and sublimation energy was investigated for the model system containing indomethacin and carbamazepine. No correlation between melting temperature and sublimation energy could be found. This could possibly be explained that the correlation was found for a system of fairly small crystal structures. It might be that the correlation is not transferable to such large unit cells as the indomethacin and carbamazepine systems have. It could also be that error propagation in the calculations is a significant factor.

An entropy correlation was found for the unit cell volume using the benchmark set $C19_{RT}$. In order to generalize the correlation, it would be beneficial to perform simulations on crystals with an even wider range of unit cell volumes.

The zero point energy correlation found can be used instead of using time consuming simulations. The exceptional agreement of the correlation coefficients indicates that the zero point energy is very strongly dependent on the unit cell volume. The dependence could be checked by performing additional energy and entropy calculations for isolated molecules using different sized vacuum cells.

Hydration free energies can be calculated using several different methods, each having advantages and drawbacks. Since hydration free energies are somewhat difficult to measure experimentally it is often hard to find accurate data. One way to evaluate the methods is to use a benchmark set where both sublimation free energies and solubilities are known, making it possible to evaluate the hydration free energy.

Solubility could be accurately calculated to the correct order of magnitude for a selection of the benchmark set $C19_{RT}$. No reasonable solubilities could be calculated for the indomethacin and carbamazepine systems. The entropy contribution is thought to be a major factor in the results. Another factor that contributes is the uncertainty of how the dissolution process takes place for co-crystals and solvates. It might be that the process is different than the thermodynamic cycle used in this project. Thorough investigation is needed in order to establish the primary dissolution cycle for co-crystals and solvates.

More investigations are also necessary to further evaluate the procedure used and to test the melting temperature correlation.

ACKNOWLEDGEMENTS

I would like to thank my supervisors S. Nilsson Lill, M. Lindkvist, A. Broo and C. Boissier for their outstanding guidance, support and patience and corrective reading of my report.

Special thanks go to E. Karlsson for practical guidance of how to handle the uDISS instrument, U. Tehler for her advice and constructive discussions concerning dissolution measurements and their analysis and R. Johansson for his practical guidance in determining particle size distributions.

I would also like to thank my examiner A. Matic for his support and help with the many administrative necessities regarding this Master's Thesis project.

REFERENCES

1. Ng, R., *Drugs: From Discovery to Approval*. 2011: Wiley.
2. McCurdy, C.R. and Scully, S.S., *Analgesic substances derived from natural products (natureceuticals).* Life Sciences, 2005. **78**(5): p. 476-484.
3. Serajuddin, A.T.M., *Salt formation to improve drug solubility.* Advanced Drug Delivery Reviews, 2007. **59**(7): p. 603-616.
4. Bennema, P., et al., *Solubility of molecular crystals: Polymorphism in the light of solubility theory.* International Journal of Pharmaceutics, 2008. **351**(1–2): p. 74-91.
5. Good, D.J. and Rodríguez-Hornedo, N., *Solubility Advantage of Pharmaceutical Cocrystals.* Crystal Growth & Design, 2009. **9**(5): p. 2252-2264.
6. Palmer, D.S., et al., *First-Principles Calculation of the Intrinsic Aqueous Solubility of Crystalline Druglike Molecules.* Journal of Chemical Theory and Computation, 2012. **8**(9): p. 3322-3337.
7. Westergren, J., et al., *In Silico Prediction of Drug Solubility: 1. Free Energy of Hydration.* The Journal of Physical Chemistry B, 2007. **111**(7): p. 1872-1882.
8. Hart, F.D. and Boardman, P.L., *Indomethacin: A New Non-steroid Anti-inflammatory Agent*. Vol. 2. 1963. 965-970.
9. Slavin, P.A., et al., *Morphological evaluation of the γ-polymorph of indomethacin.* Journal of Crystal Growth, 2002. **237–239, Part 1**(0): p. 300-305.
10. Callister, W.D. and Rethwisch, D.G., *Fundamentals of Materials Science and Engineering: An Integrated Approach*. 2012: Wiley.
11. Morissette, S.L., et al., *High-throughput crystallization: polymorphs, salts, co-crystals and solvates of pharmaceutical solids.* Advanced Drug Delivery Reviews, 2004. **56**(3): p. 275-300.
12. Amidon, G.L., et al., *A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability.* Pharm Res, 1995. **12**(3): p. 413-20.
13. Aulton, M.E. and Taylor, K.M.G., *Aulton's Pharmaceutics: The Design and Manufacture of Medicines*. 2013: Elsevier Health Sciences UK.
14. Warren, B.E., *X-ray Diffraction*. 1969: Dover Publications.
15. Coats, A.W. and Redfern, J.P., *Thermogravimetric analysis. A review.* Analyst, 1963. **88**(1053): p. 906-924.
16. Michael H Chiu, E.J.P., *Differential scanning calorimetry: An invaluable tool for a detailed thermodynamic characterization of macromolecules and their interactions.* Journal of Pharmacy & BioAllied Sciences, 2011. **3**(1): p. 39-59.
17. Skoog, D.A., *Fundamentals of Analytical Chemistry*. 2004: Thomson-Brooks/Cole.
18. McMullan, D., *Scanning electron microscopy 1928–1965.* Scanning, 1995. **17**(3): p. 175-185.
19. Whelan, C.T., et al., *Electron Scattering From Atoms, Molecules, Nuclei, and Bulk Matter*. 2005: Springer US.
20. Gouesbet, G. and Gréhan, G., *Optical particle sizing: theory and practice*. 1988: Plenum Press.
21. Hohenberg, P. and Kohn, W., *Inhomogeneous Electron Gas.* Physical Review, 1964. **136**(3B): p. B864-B871.
22. Jensen, F., *Introduction to Computational Chemistry*. 2007: Wiley.
23. Schwerdtfeger, P., *The Pseudopotential Approximation in Electronic Structure Theory.* ChemPhysChem, 2011. **12**(17): p. 3143-3155.

24. Burke, K., *Perspective on density functional theory.* The Journal of Chemical Physics, 2012. **136**(15): p. -.
25. Kohn, W. and Sham, L.J., *Self-Consistent Equations Including Exchange and Correlation Effects.* Physical Review, 1965. **140**(4A): p. A1133-A1138.
26. von Barth, U. and Hedin, L., *A local exchange-correlation potential for the spin polarized case. i.* Journal of Physics C: Solid State Physics, 1972. **5**(13): p. 1629.
27. Perdew, J.P., Burke, K., and Ernzerhof, M., *Generalized Gradient Approximation Made Simple.* Physical Review Letters, 1996. **77**(18): p. 3865-3868.
28. Lee, K., et al., *Higher-accuracy van der Waals density functional.* Physical Review B, 2010. **82**(8): p. 081101.
29. Bučko, T., et al., *Tkatchenko-Scheffler van der Waals correction method with and without self-consistent screening applied to solids.* Physical Review B, 2013. **87**(6): p. 064110.
30. Grimme, S., *Semiempirical GGA-type density functional constructed with a long-range dispersion correction.* Journal of Computational Chemistry, 2006. **27**(15): p. 1787-1799.
31. Grimme, S., et al., *A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu.* The Journal of Chemical Physics, 2010. **132**(15): p. 154104.
32. Grimme, S., et al., *DFT-D3 v.3.1.0.* 2014.
33. Kaneniwa, N., Otsuka, M., and Hayashi, T., *Physicochemical Characterization of Indomethacin Polymorphs and the Transformation Kinetics in Ethanol.* Chemical & pharmaceutical bulletin, 1985. **33**(8): p. 3447-3455.
34. Hamdi, N., et al., *Solvates of Indomethacin.* Journal of Thermal Analysis and Calorimetry, 2004. **76**(3): p. 985-1001.
35. Basavoju, S., Boström, D., and Velaga, S., *Indomethacin–Saccharin Cocrystal: Design, Synthesis and Preliminary Pharmaceutical Characterization.* Pharmaceutical Research, 2008. **25**(3): p. 530-541.
36. Grzesiak, A.L., et al., *Comparison of the four anhydrous polymorphs of carbamazepine and the crystal structure of form I.* Journal of Pharmaceutical Sciences, 2003. **92**(11): p. 2260-2271.
37. Fleischman, S.G., et al., *Crystal Engineering of the Composition of Pharmaceutical Phases: Multiple-Component Crystalline Solids Involving Carbamazepine.* Crystal Growth & Design, 2003. **3**(6): p. 909-919.
38. Johnston, A., Florence, A.J., and Kennedy, A.R., *Carbamazepine N,N-dimethylformamide solvate.* Acta Crystallographica Section E, 2005. **61**(5): p. 1509-1511.
39. Childs, S.L., et al., *Screening strategies based on solubility and solution composition generate pharmaceutically acceptable cocrystals of carbamazepine.* CrystEngComm, 2008. **10**(7): p. 856-864.
40. Lu, E., Rodriguez-Hornedo, N., and Suryanarayanan, R., *A rapid thermal method for cocrystal screening.* CrystEngComm, 2008. **10**(6): p. 665-668.
41. Clark, S.J., Segall, M.D., and Pickard, C.J., *First principles methods using CASTEP.* Zeitschrift für Kristallographie, 2005. **220**(5/6): p. 567-570.
42. Delley, B., *An all-electron numerical method for solving the local density functional for polyatomic molecules.* The Journal of Chemical Physics, 1990. **92**(1): p. 508-517.
43. Delley, B., *From molecules to solids with the DMol3 approach.* The Journal of Chemical Physics, 2000. **113**(18): p. 7756-7764.
44. Bochevarov, A.D., et al., *Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences.* International Journal of Quantum Chemistry, 2013. **113**(18): p. 2110-2142.

45. *Jaguar v.8.3*. 2014, Schrödinger LLC.
46. Delley, B., *The conductor-like screening model for polymers and surfaces.* Molecular Simulation, 2006. **32**(2): p. 117-123.
47. Marenich, A.V., et al., *Self-Consistent Reaction Field Model for Aqueous and Nonaqueous Solutions Based on Accurate Polarized Partial Charges.* Journal of Chemical Theory and Computation, 2007. **3**(6): p. 2011-2033.
48. Olson, R.M., et al., *Charge Model 4 and Intramolecular Charge Polarization.* Journal of Chemical Theory and Computation, 2007. **3**(6): p. 2046-2054.
49. Kelly, C.P., Cramer, C.J., and Truhlar, D.G., *SM6: A Density Functional Theory Continuum Solvation Model for Calculating Aqueous Solvation Free Energies of Neutrals, Ions, and Solute−Water Clusters.* Journal of Chemical Theory and Computation, 2005. **1**(6): p. 1133-1152.
50. Carter, D.J. and Rohl, A.L., *Benchmarking Calculated Lattice Parameters and Energies of Molecular Crystals Using van der Waals Density Functionals.* Journal of Chemical Theory and Computation, 2014. **10**(8): p. 3423-3437.
51. Allen, F., *The Cambridge Structural Database: a quarter of a million crystal structures and rising.* Acta Crystallographica Section B, 2002. **58**(3 Part 1): p. 380-388.
52. Bruno, I.J., et al., *New software for searching the Cambridge Structural Database and visualizing crystal structures.* Acta Crystallographica Section B, 2002. **58**(3 Part 1): p. 389-397.
53. Macrae, C.F., et al., *Mercury: visualization and analysis of crystal structures.* Journal of Applied Crystallography, 2006. **39**(3): p. 453-457.
54. Moellmann, J. and Grimme, S., *DFT-D3 Study of Some Molecular Crystals.* The Journal of Physical Chemistry C, 2014. **118**(14): p. 7615-7621.

APPENDIX A Benchmark

Table 11. Summary of the different quality settings, explaining the differences. *Geometry optimization convergence tolerance.

Quality setting	Energy* [eV/atom]	Force* [eV/Å]	Stress* [GPa]	Displacement* [Å]	Cutoff [eV]	SCF tolerance [eV/atom]
Medium	2E-5	0.05	0.1	2E-3	550	2E-6
Fine	1E-5	0.03	0.05	1E-3	550	1E-6
Ultra-fine	5E-6	0.01	0.02	5E-4	550	5.0E-7

Table 12. Percentage deviation of simulated lattice parameters, ρ and cell volume for different quality settings, compared to experimental data. N.B. the decreasing volume deviation as the quality is increased. *No convergence.

Quality setting	Compound	Dispersion	a	b	c	ρ	V_{cell}
Medium	β-oxalic acid	Grimme	-0.77	7.78	7.62	-13.26	15.29
		TS	-0.58	6.26	9.56	-13.78	15.98
	Urea	Grimme	1.21	1.21	0.73	-2.02	2.06
		TS	-0.85	-0.85	0.51	2.32	-2.27
	Pyrazine	Grimme	-*	-*	-*	-*	-*
		TS	-*	-*	-*	-*	-*
Fine	β-oxalic acid	Grimme	-1.08	3.74	0.64	-4.6	4.82
		TS	-0.99	-0.06	2.58	-1.7	1.73
	Urea	Grimme	-0.05	-0.05	0.45	-1.7	-0.73
		TS	-2.07	-2.07	-0.01	5.43	-5.15
	Pyrazine	Grimme	-1.57	-4.11	1.28	4.61	-4.41
		TS	-2.1	-5.7	2.14	6.05	-5.71
Ultra-fine	β-oxalic acid	Grimme	-1.21	2.56	-0.65	-1.39	1.41
		TS	-1.11	-0.42	0.67	1.07	-1.06
	Urea	Grimme	-0.06	-0.06	0.19	1.02	-1.01
		TS	-2.08	-2.08	-0.36	5.82	-5.5
	Pyrazine	Grimme	-0.26	-3.39	-2.9	6.87	-6.43
		TS	-0.04	-5.42	-3.5	9.62	-8.77

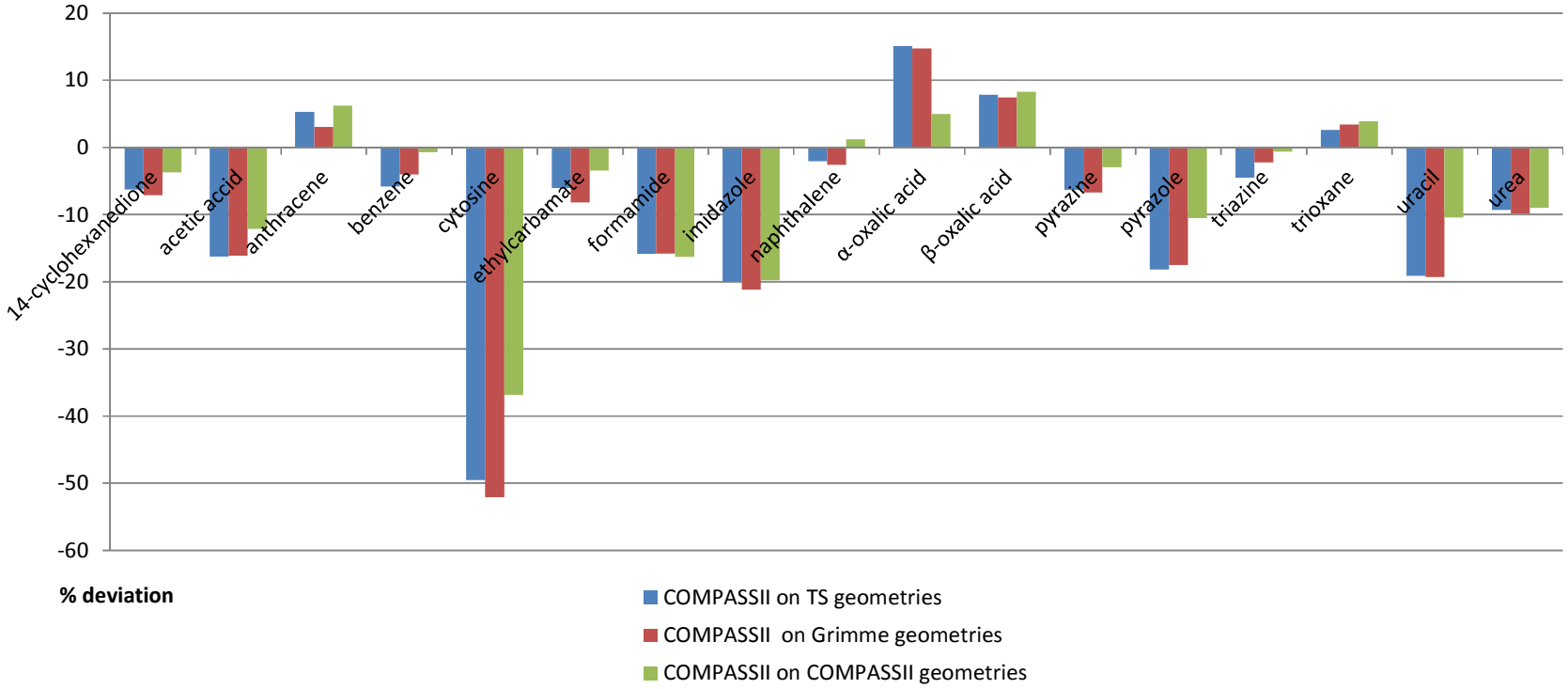

Figure 28. Summary of energy calculations performed with COMPASSII on geometries optimized using TS, Grimme and COMPASSII, showing the percentage deviations compared to the experimental values. It can be seen that geometries optimized by COMPASSII gives the best result for the lattice energy, which is rather intuitive.

APPENDIX B Indomethacin

Polymorphs

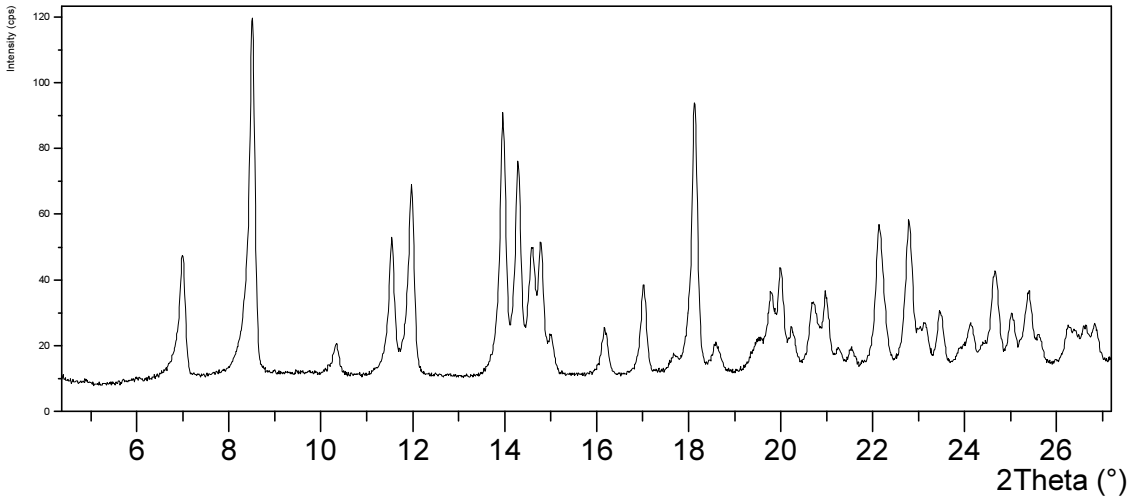

Figure 29. X-ray diffractogram of the α-polymorph of IND.

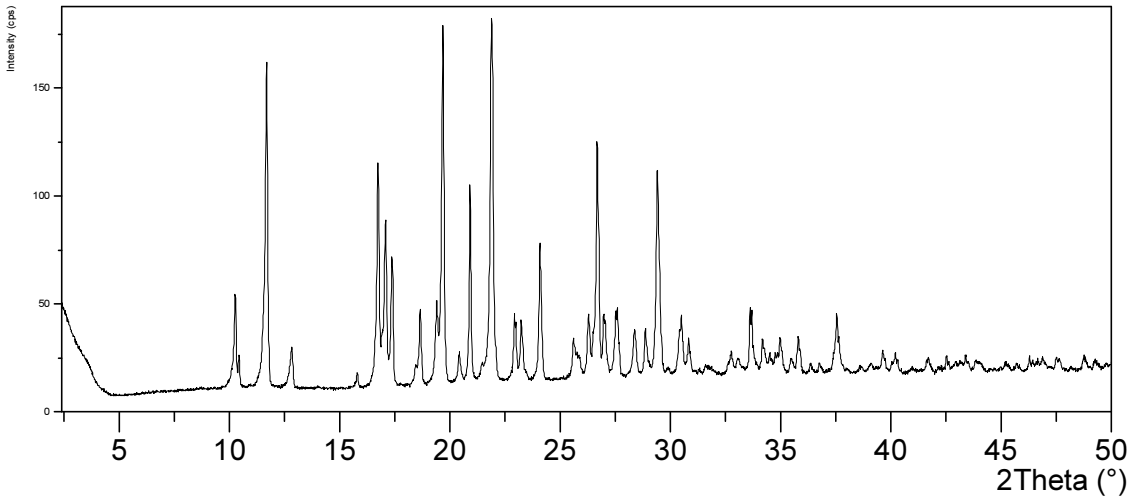

Figure 30. X-ray diffractogram of the γ-polymorph of IND

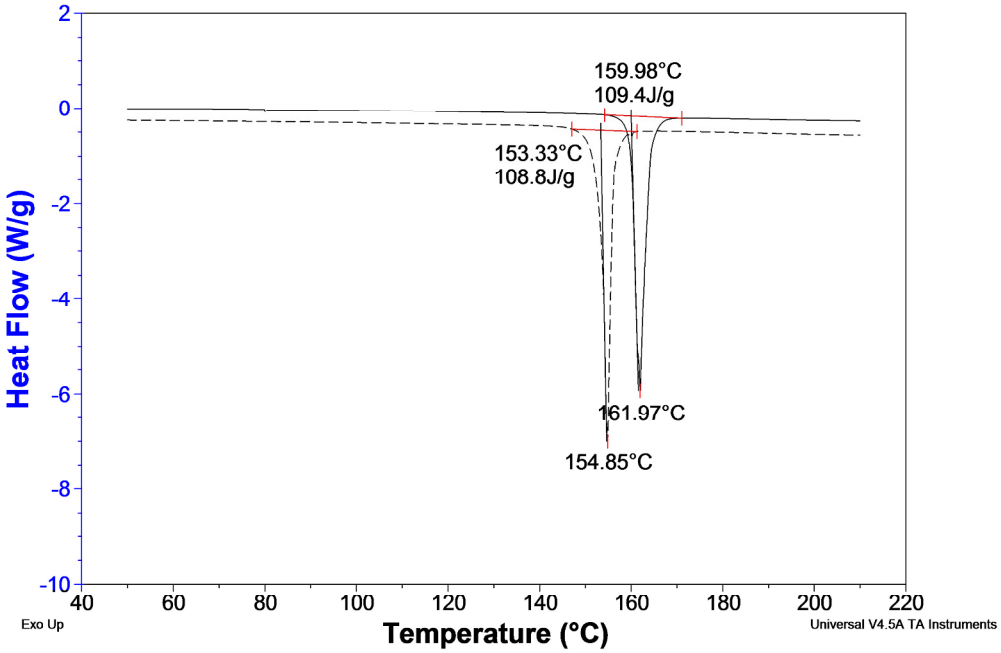

Figure 31. DSC thermogram of the γ-polymorph (a) and α-polymorph (b) of IND.

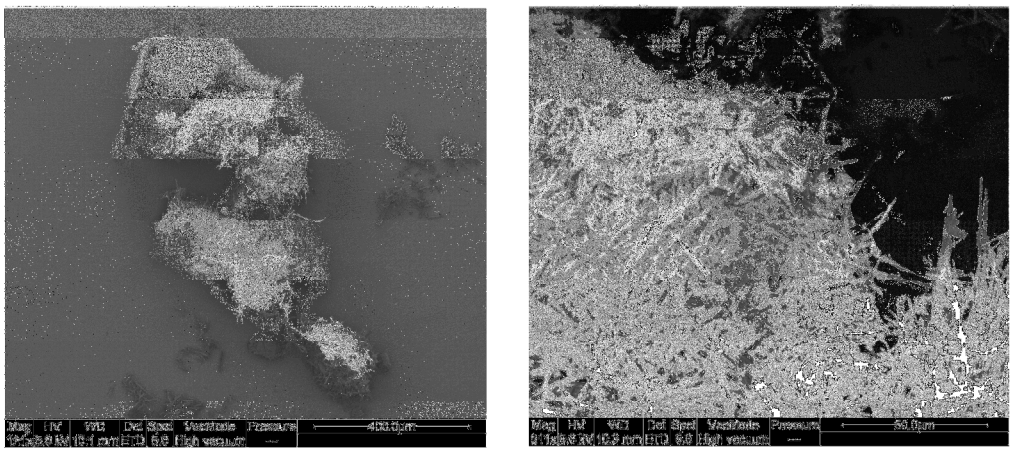

Figure 32. SEM micrographs for α–polymorph, overview (left) and close-up (right).

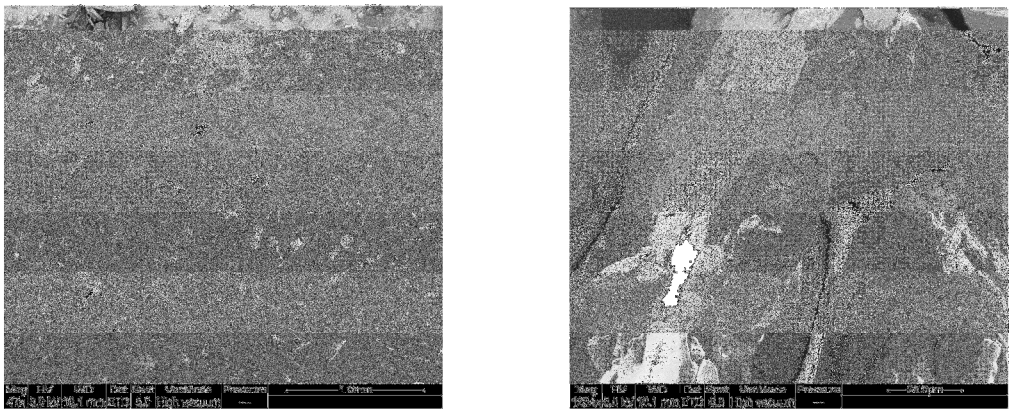

Figure 33. SEM micrographs for γ–polymorph, overview (left) and close-up (right).

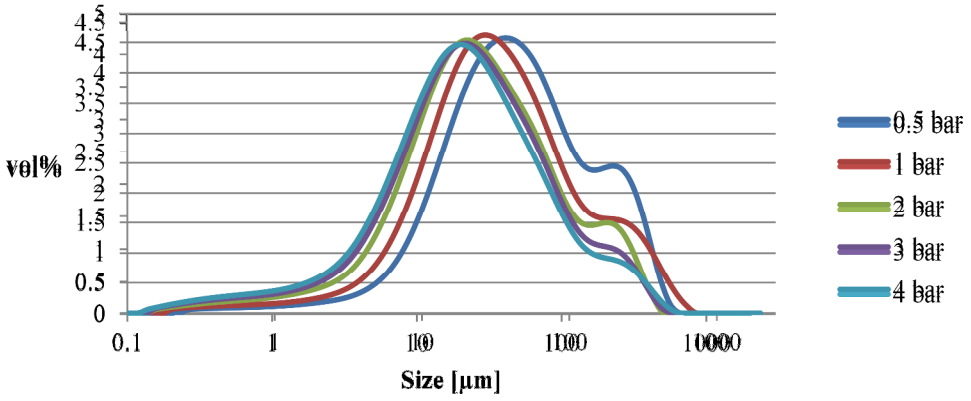

Figure 34. PSDs for IND γ-polymorph for varying pressures.

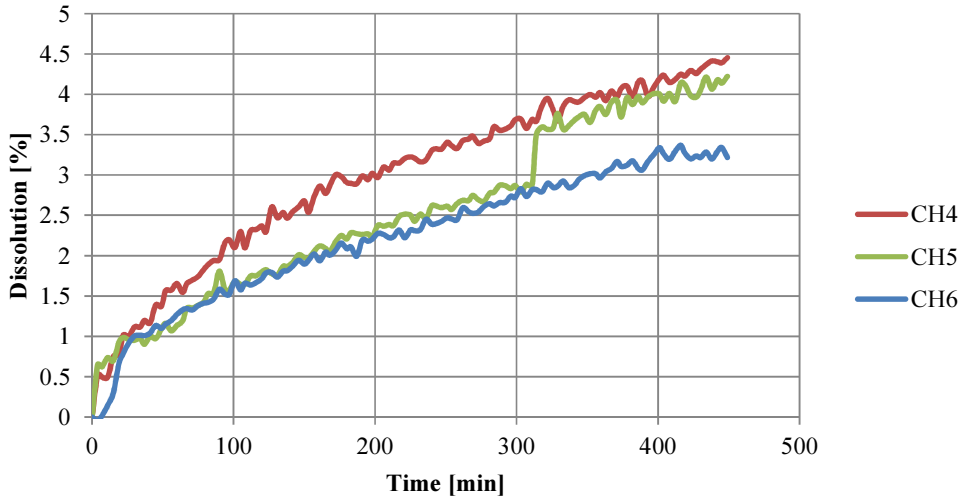

Figure 36. Dissolution rate measurements for IND α-polymorph. Sample weights were 25 µg, 31 µg and 26 µg for CH 4, 5 and 6 respectively.

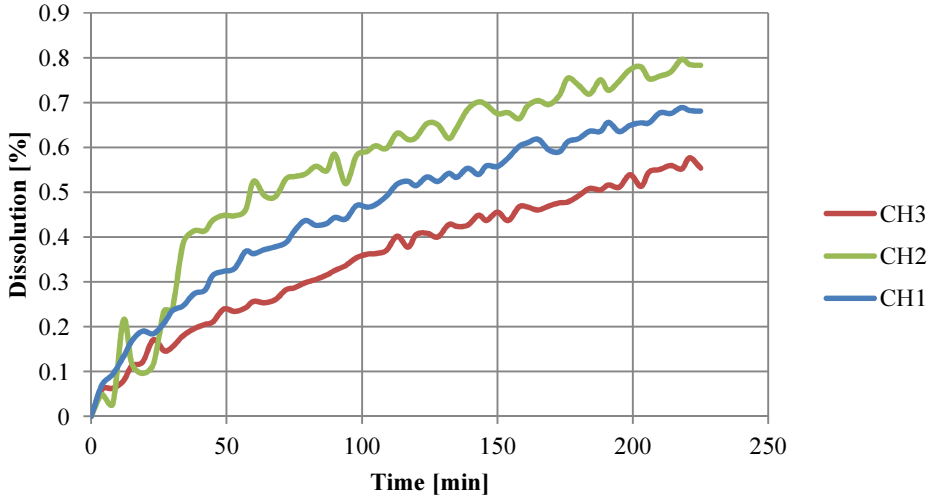

Figure 37. Dissolution rate measurements for IND γ-polymorph. Sample weights were 111 µg, 91 µg and 112 µg for CH 1, 2 and 3 respectively.

tert-Butanol hemi-solvate

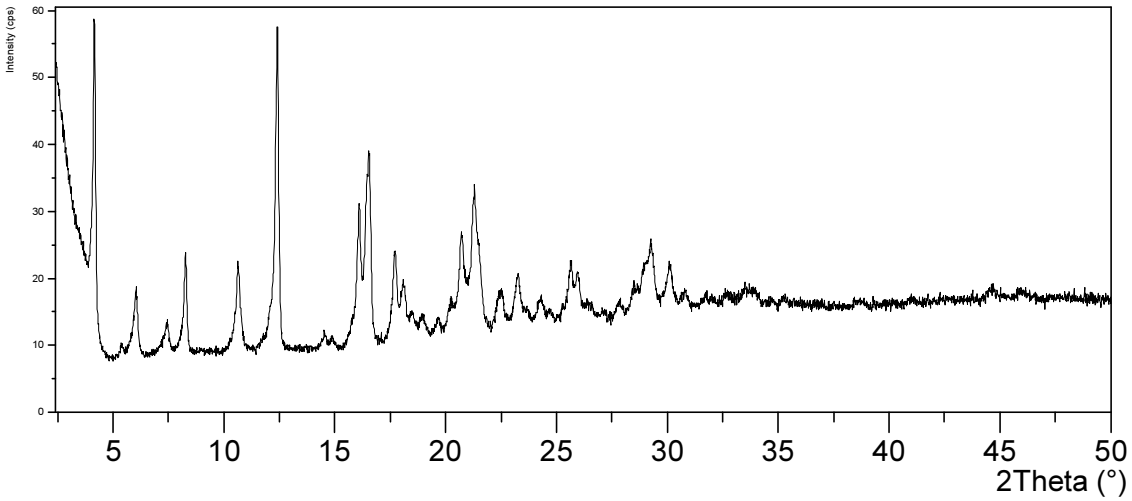

Figure 38. X-ray diffractogram of the IND:*tert*-Buthanol hemi-solvate.

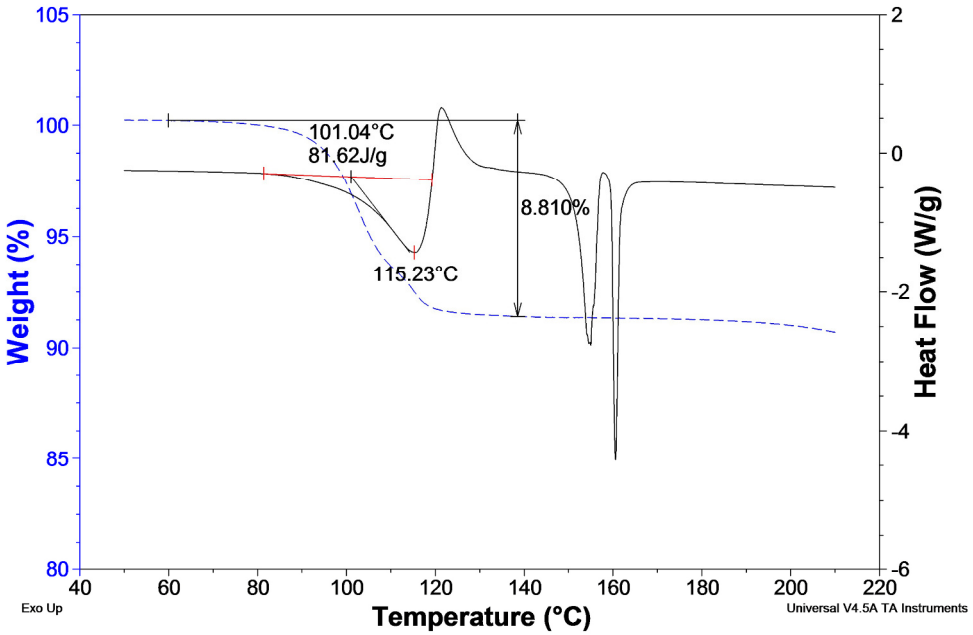

Figure 39. TGA and DSC thermograms of IND:*tert*-Butanol hemi-solvate. The weight loss from 80 -> 120 °C was measured to 8.8%. This corresponds to molar ratio of 1:0.5 (IND:*tert*-Butanol) which is a hemi-solvate.

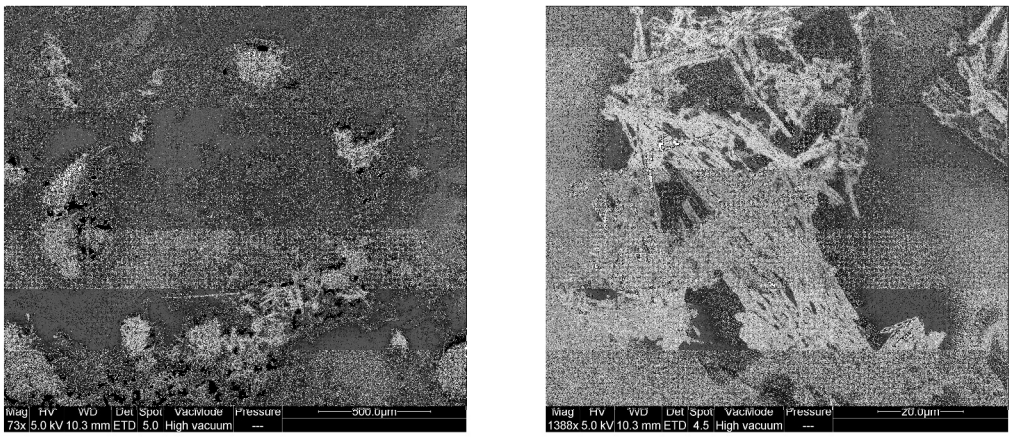

Figure 40. SEM micrographs for IND:*tert*-Buthanol hemi-solvate, overview (left) and close-up (right).

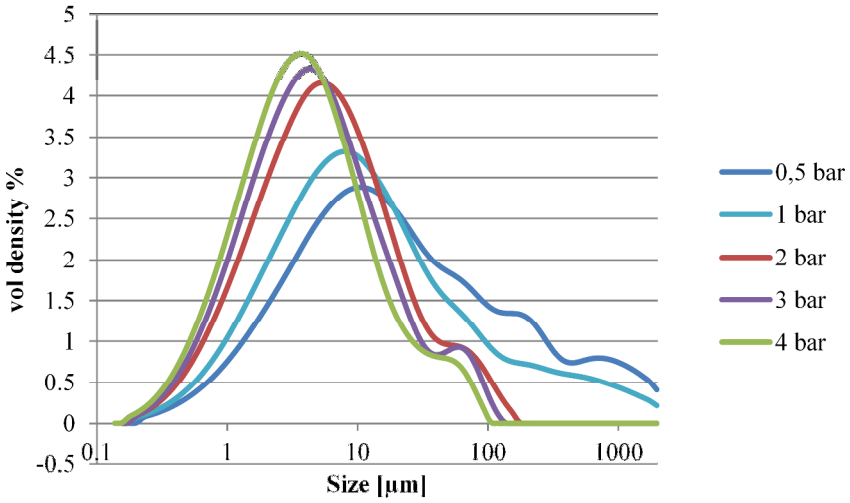

Figure 41. PSDs for IND:*tert*-Buthanol hemi-solvate for varying pressures.

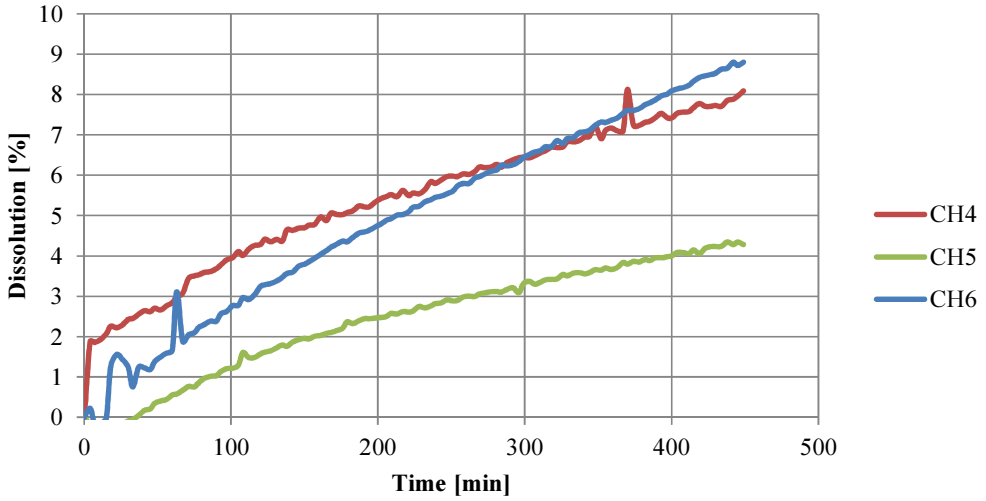

Figure 42. Dissolution rate measurements for IND:tert-Buthanol hemi-solvate. Sample weights were 73 µg, 57 µg and 63 µg for CH 4, 5 and 6 respectively.

Methanol hemi-solvate

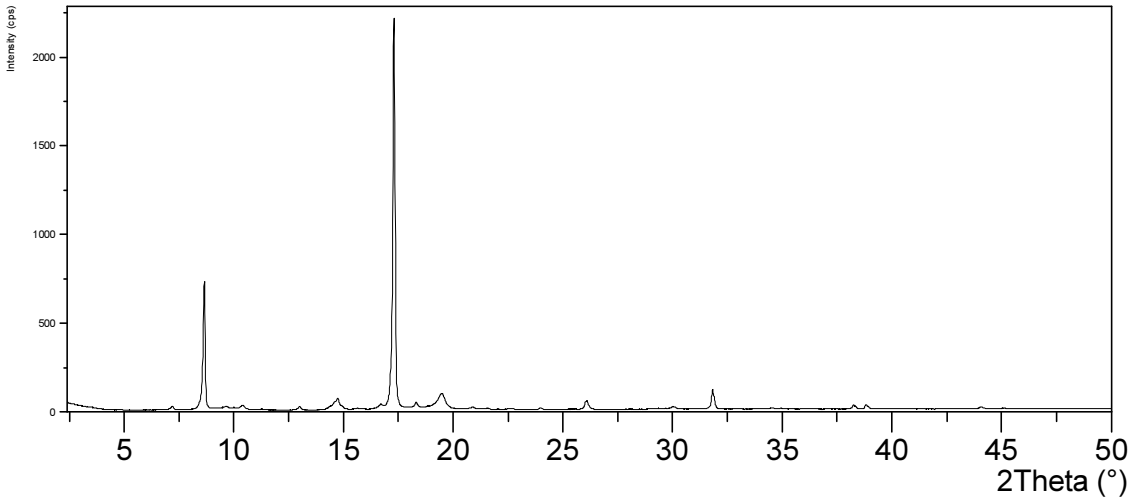

Figure 43. X-ray diffractogram of the IND:Methanol hemi-solvate.

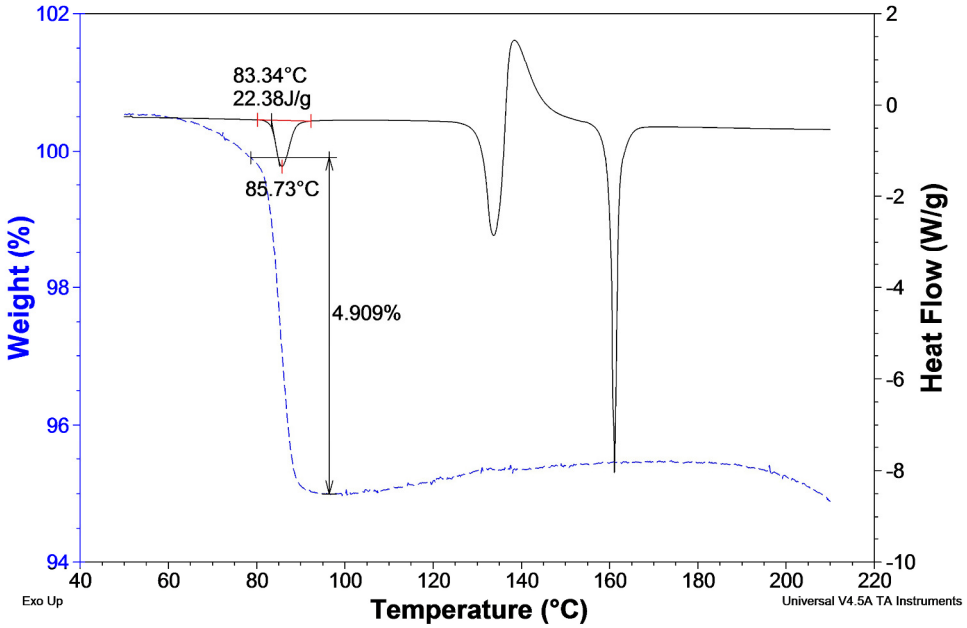

Figure 44. TGA and DSC thermograms for IND: Methanol hemi-solvate. The weight loss at around 80 °C was measured to roughly 5.1% corresponding to a hemi-solvate.

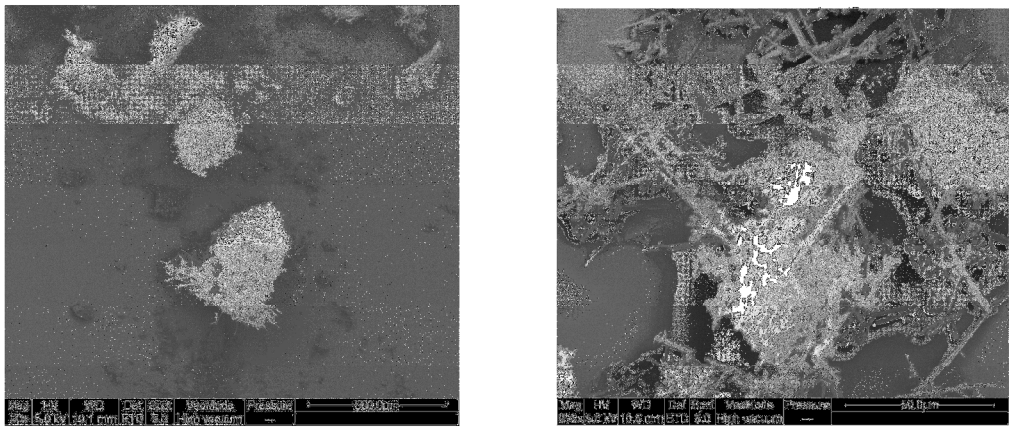

Figure 45. SEM micrographs for IND:Methanol hemi-solvate, overview (left) and close-up (right).

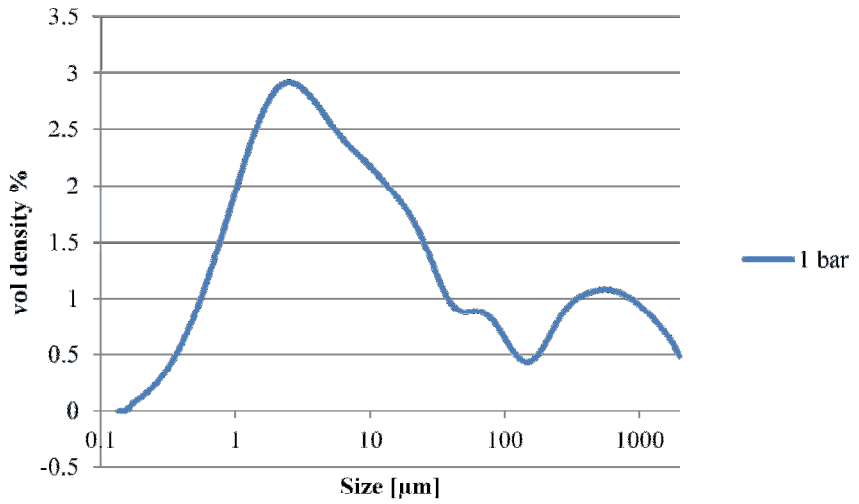

Figure 46. PSD for IND: Methanol hemi-solvate. Lack of material allowed only one pressure to be tested

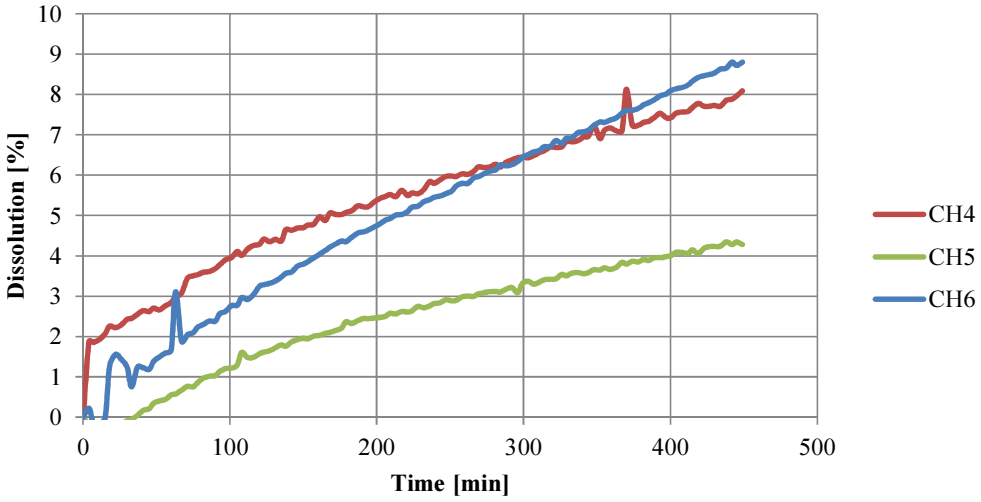

Figure 47. Dissolution rate measurements for IND:Methanol hemi-solvate. Sample weights were 28 µg, 36 µg and 38 µg for CH 4, 5 and 6 respectively.

Saccharin co-crystal

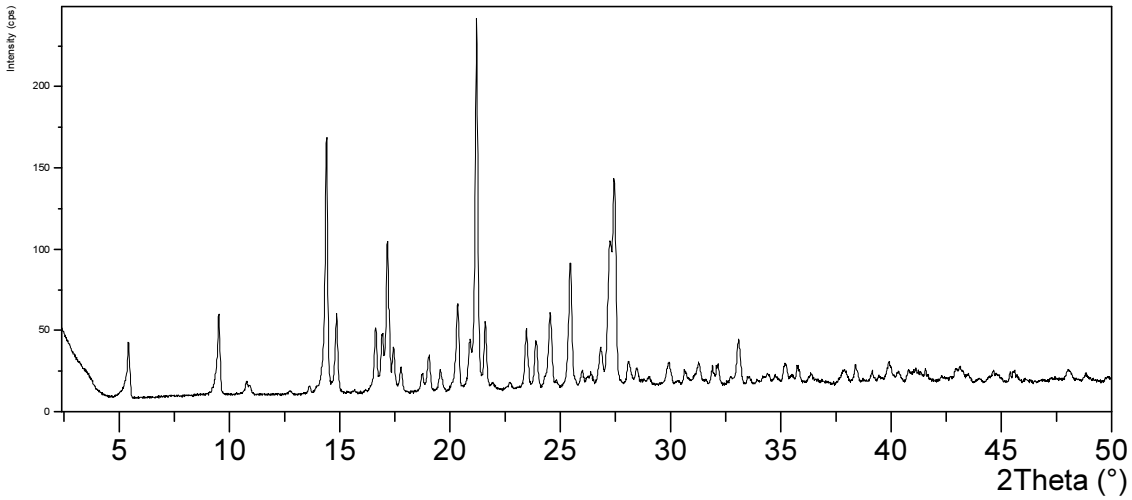

Figure 48. X-ray diffractogram of the IND:Saccharin co-crystal.

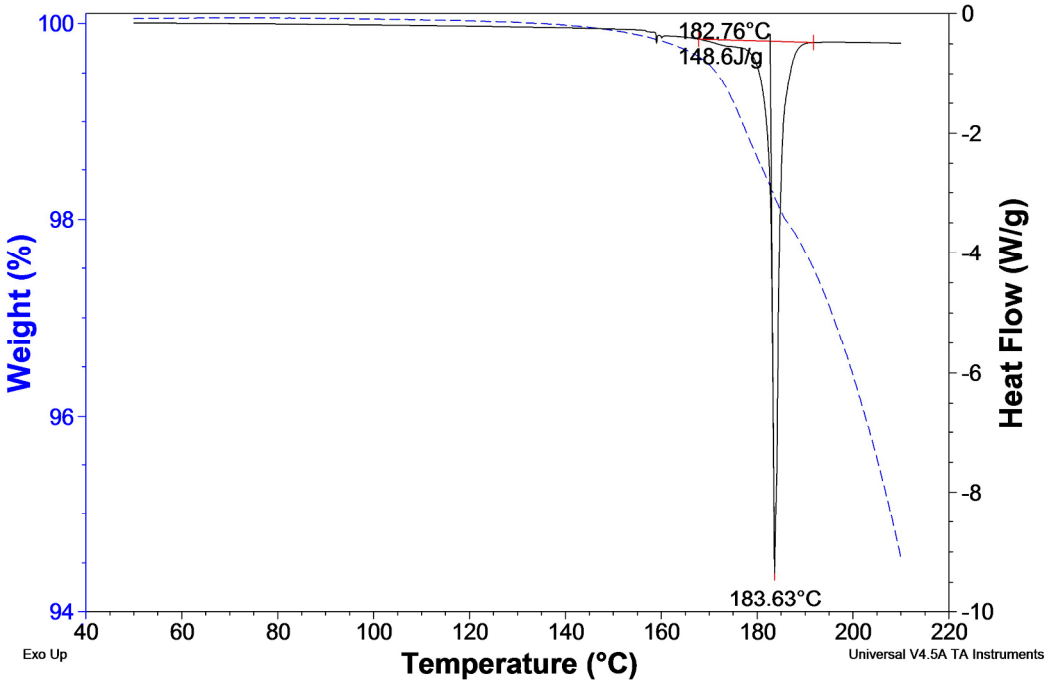

Figure 49. TGA and DSC thermograms IND:Saccharin co-crystal.

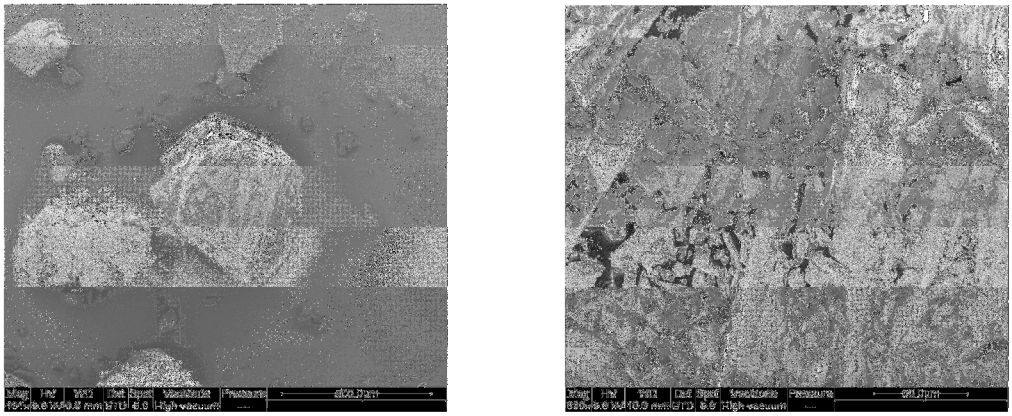

Figure 50. SEM micrographs for IND:Saccharin co-crystal, overview (left) and close-up (right).

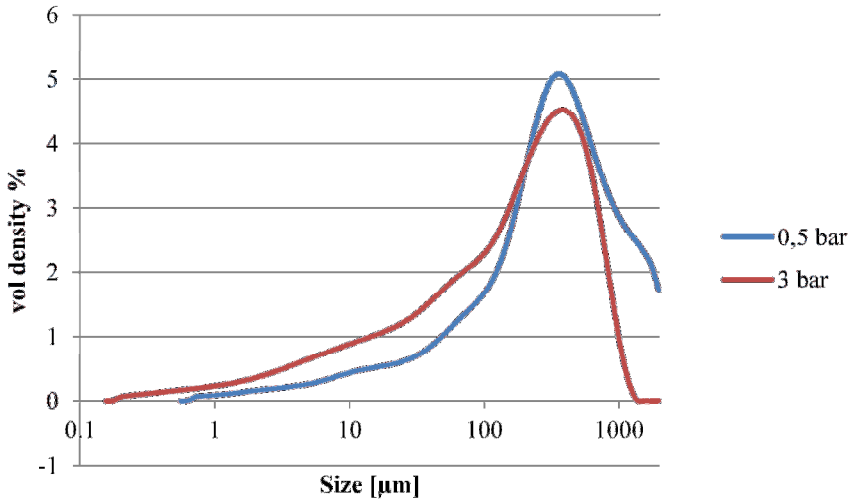

Figure 51. PSDs for IND:Saccharin co-crystal for varying pressures. Lack of material resulted in only two tested pressures.

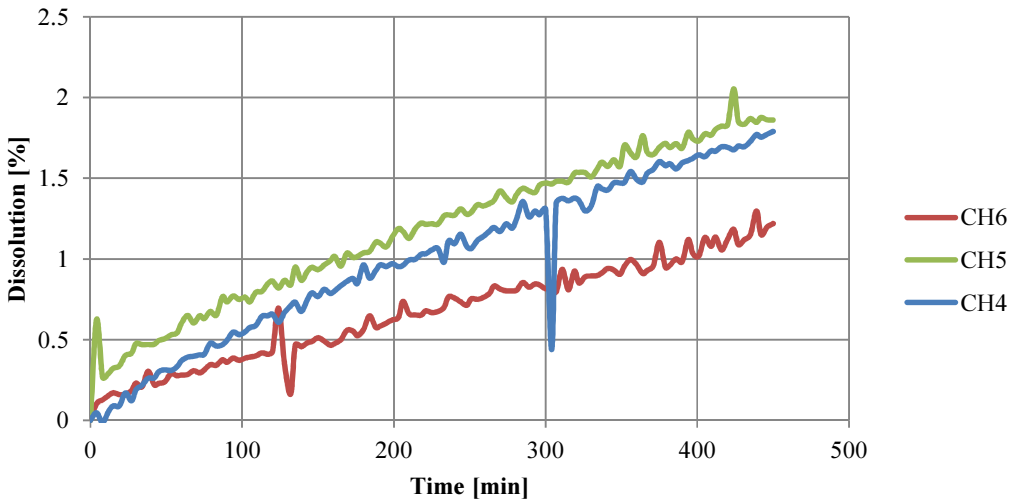

Figure 52. Dissolution rate measurements for IND:Saccharin co-crystal. Sample weights were 74 µg, 85 µg and 67 µg for CH 4, 5 and 6 respectively.

APPENDIX C Carbamazepine

Polymorph I

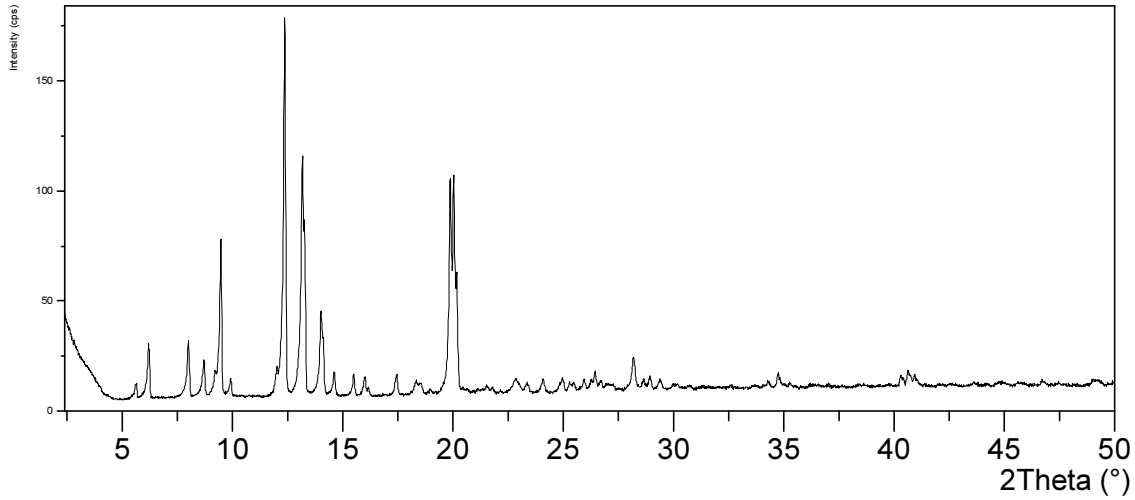

Figure 53. X-ray diffractogram of CBZ polymorph I.

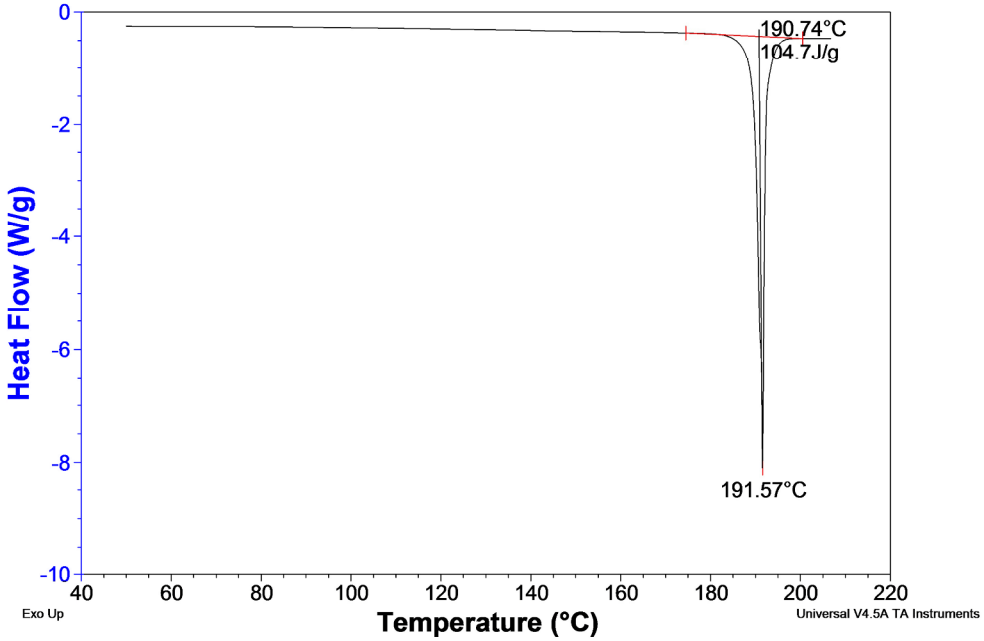

Figure 54. DSC of CBZ polymorph I.

TGA experiments were not performed since no solvent was used for preparation of form I and form III was the bulk form, supplied from the manufacturer.

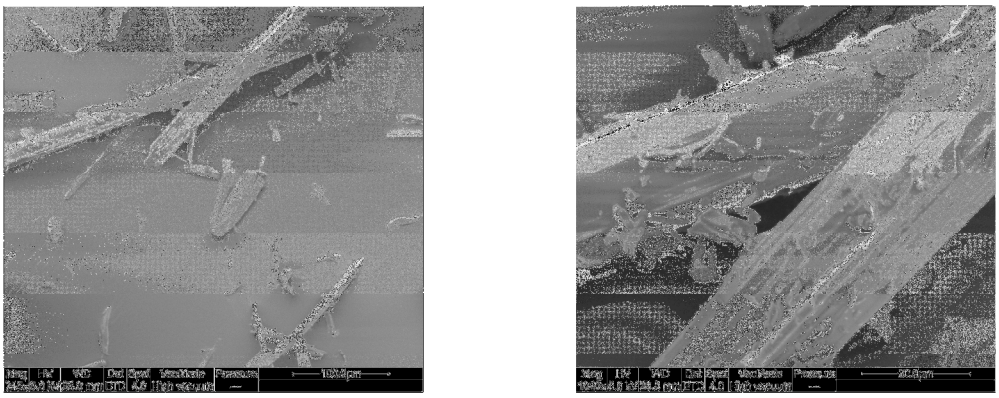

Figure 55. SEM micrographs for CBZ Polymorph I, overview (left) and close-up (right).

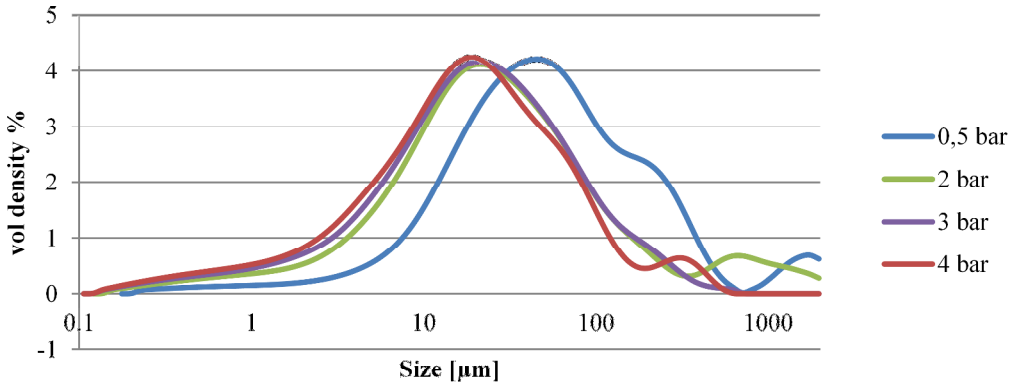

Figure 56. PSDs for CBZ polymorph I for varying pressures.

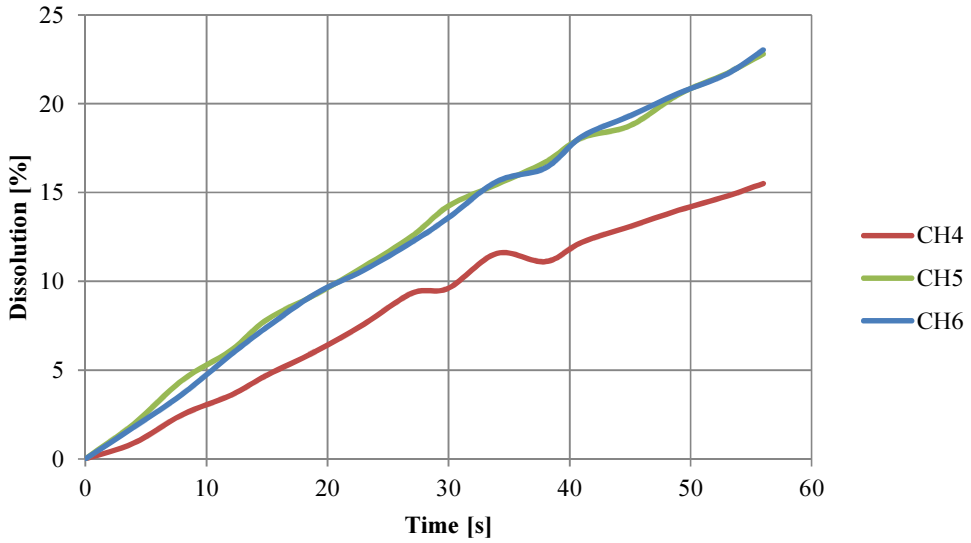

Figure 57. Dissolution rate measurements for CBZ polymorph I. Sample weights were 102 µg, 103 µg and 139 µg for CH 4, 5 and 6 respectively.

Polymorph III

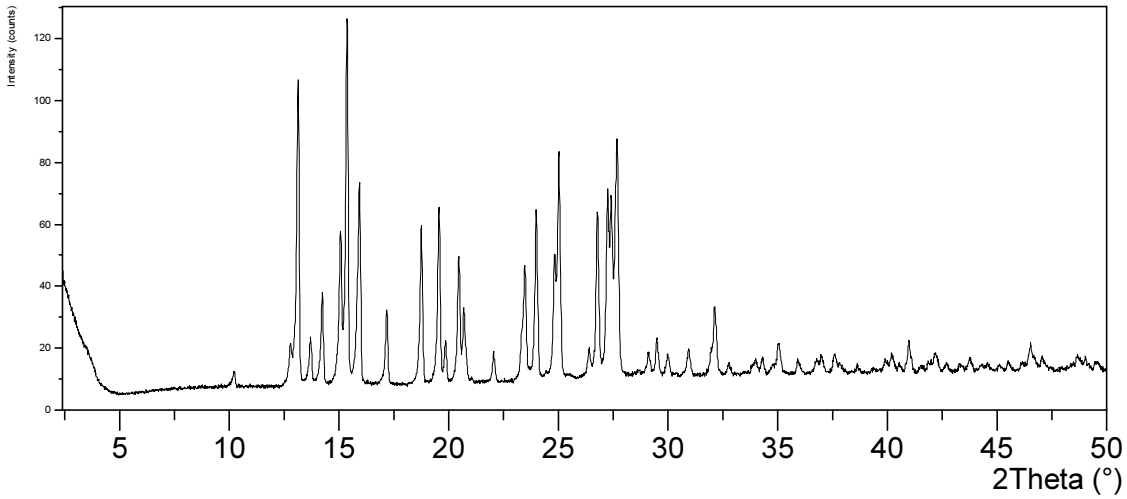

Figure 58. X-ray diffractogram of the CBZ polymorph III.

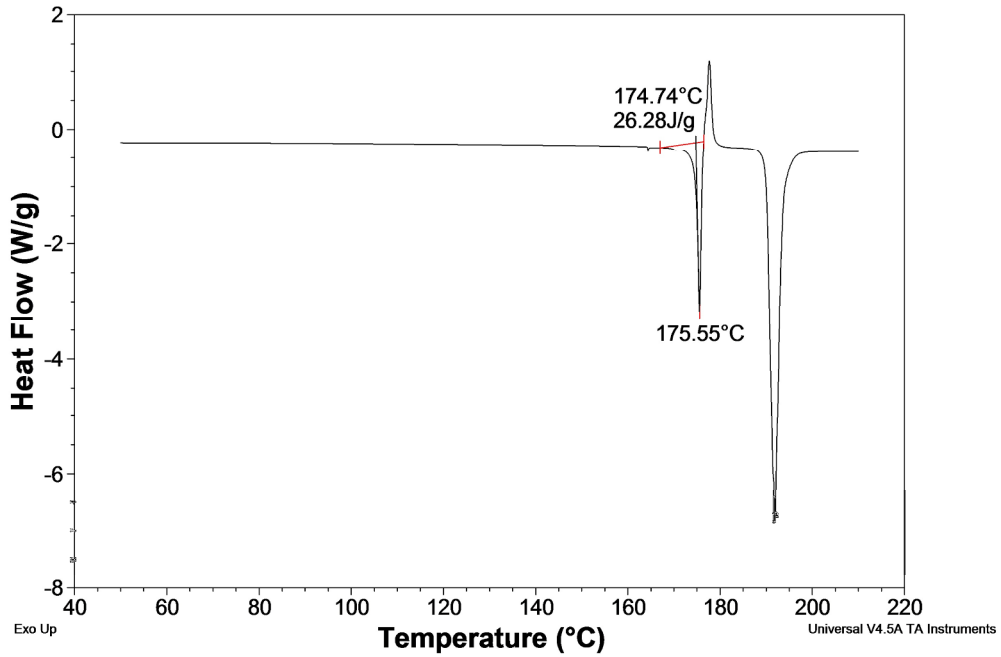

Figure 59. DSC of CBZ polymorph III. It can be seen that form III transforms into form I which is the more stable high temperature crystal form.

TGA experiments were not performed since the powder was supplied from the manufacturer.

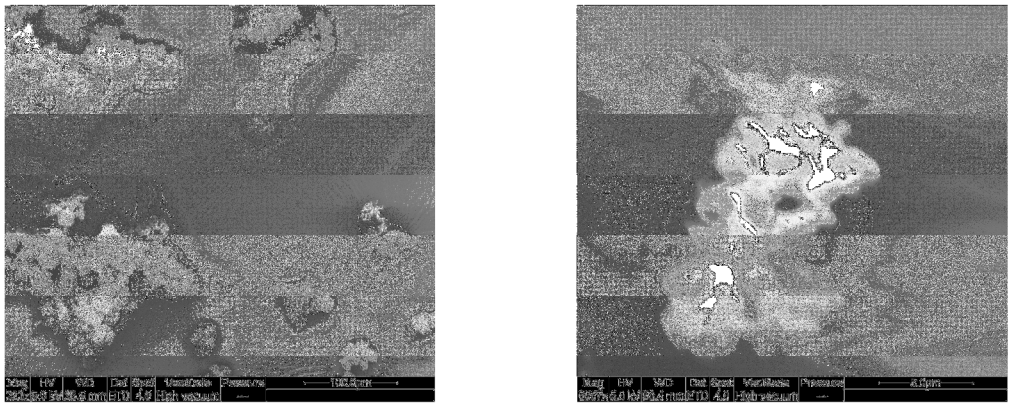

Figure 60. SEM micrographs for CBZ Polymorph III, overview (left) and close-up (right).

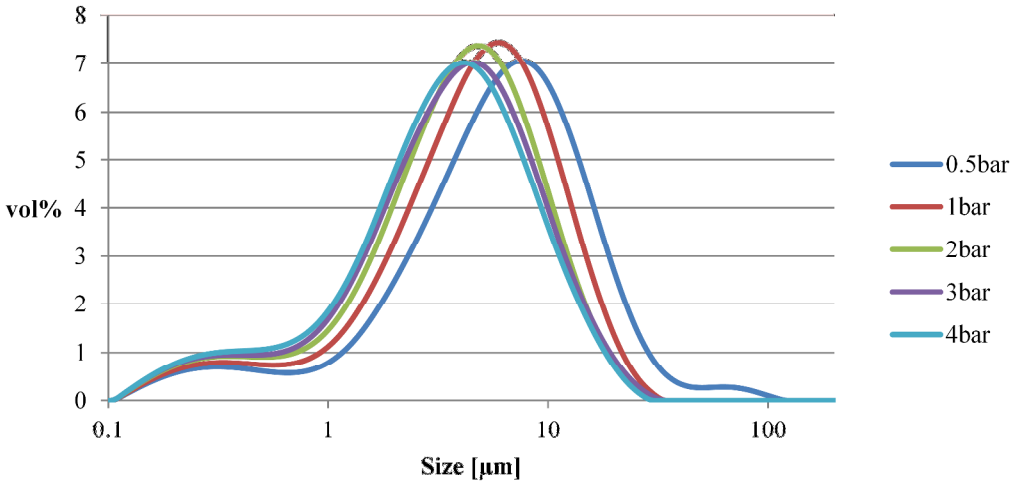

Figure 61. PSDs for CBZ polymorph III for varying pressures.

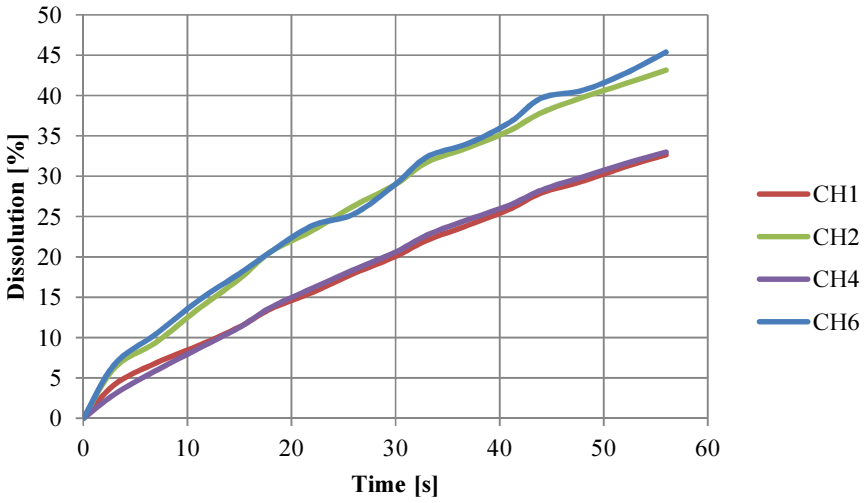

Figure 62. Dissolution rate measurements for CBZ polymorph III. Sample weights were 132 µg, 107 µg, 312 µg and 62 µg for CH 1, 2, 4 and 6 respectively.

Acetic acid mono-solvate

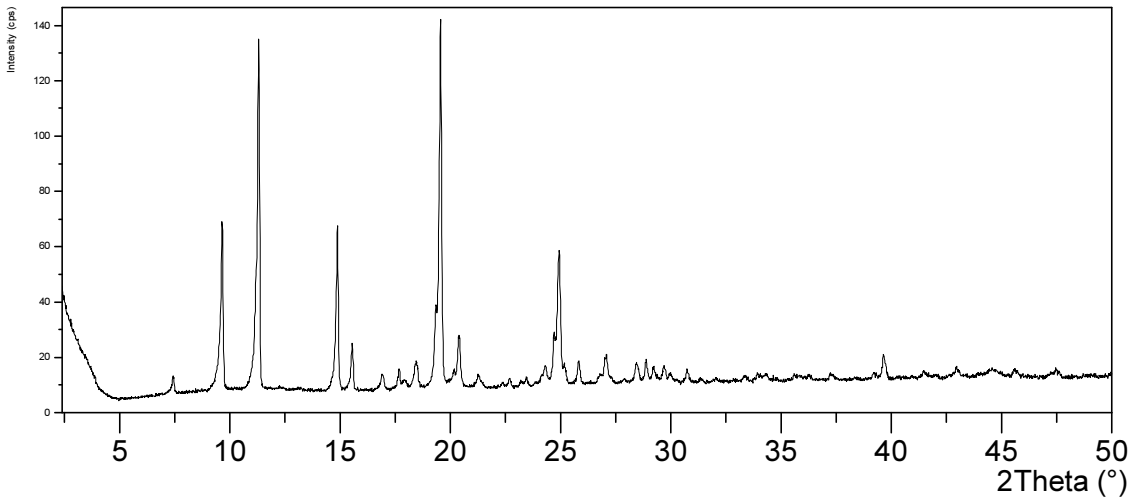

Figure 63. X-ray diffractogram of the acetic acid mono-solvate.

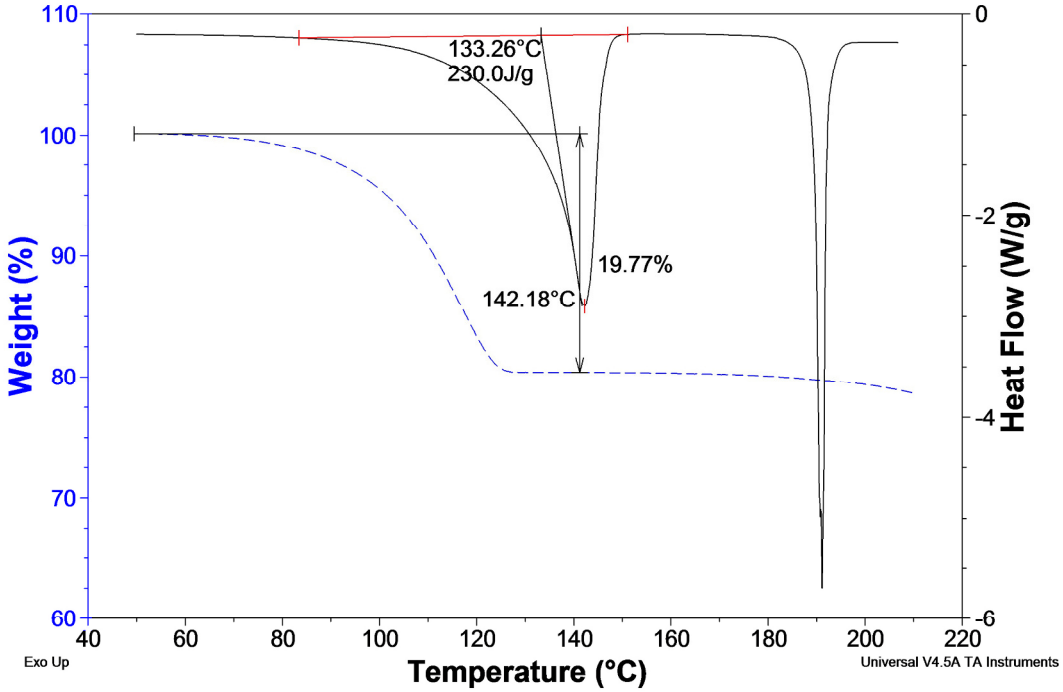

Figure 64. TGA (blue) and DSC (black) analysis of CBZ:Acetic acid mono-solvate.

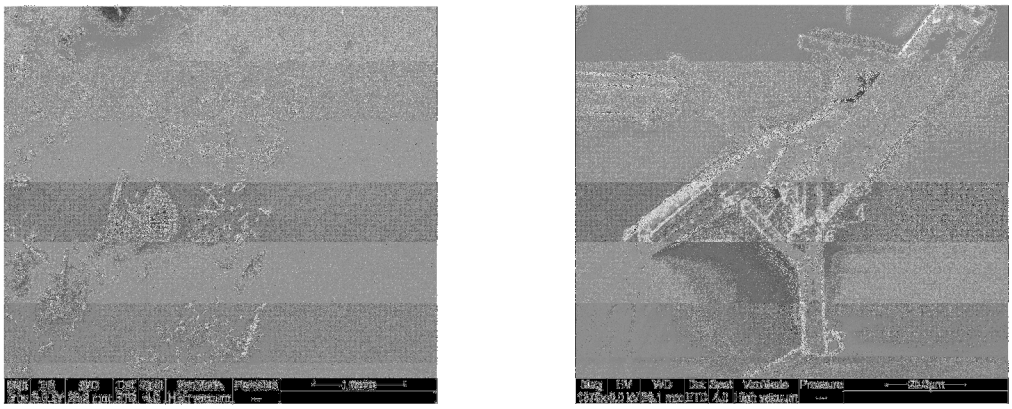

Figure 65. SEM micrographs for CBZ:Acetic acid mono-solvate, overview (left) and close-up (right).

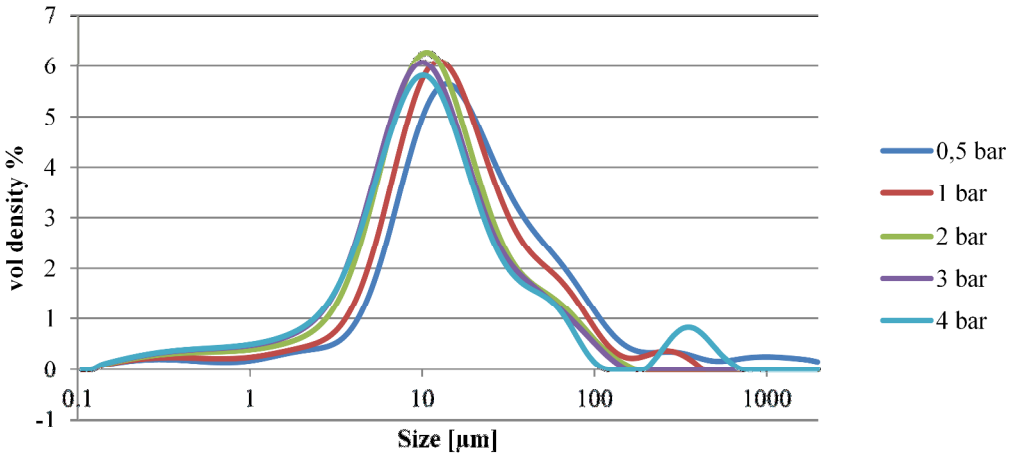

Figure 66. PSDs for CBZ:Acetic acid mono-solvate for varying pressures.

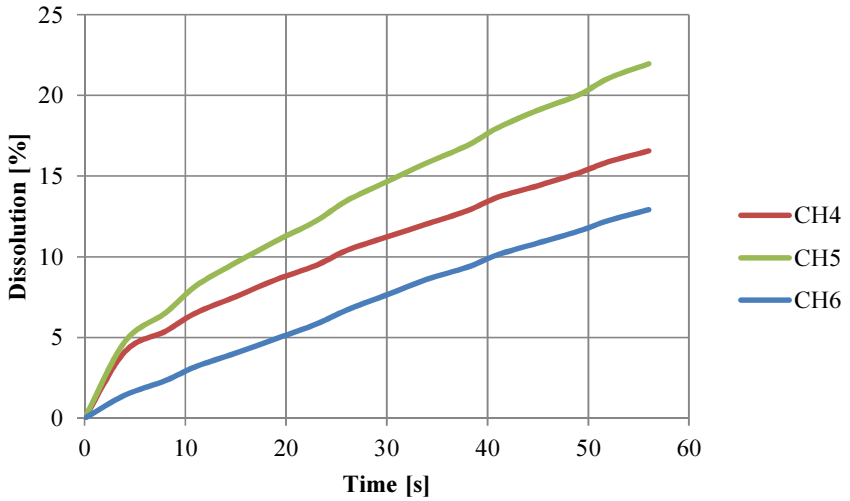

Figure 67. Dissolution rate measurements for CBZ:Acetic acid mono-solvate. Sample weights were 115 µg, 88 µg and 130 µg for CH 4, 5 and 6 respectively.

DMSO hemi-solvate

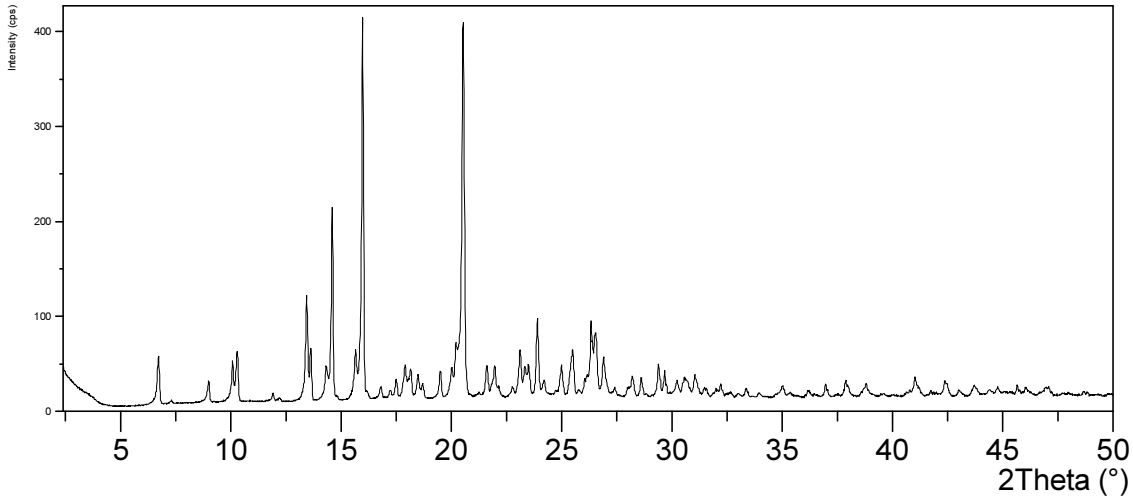

Figure 68. X-ray diffractogram of the CBZ:DMSO hemi-solvate.

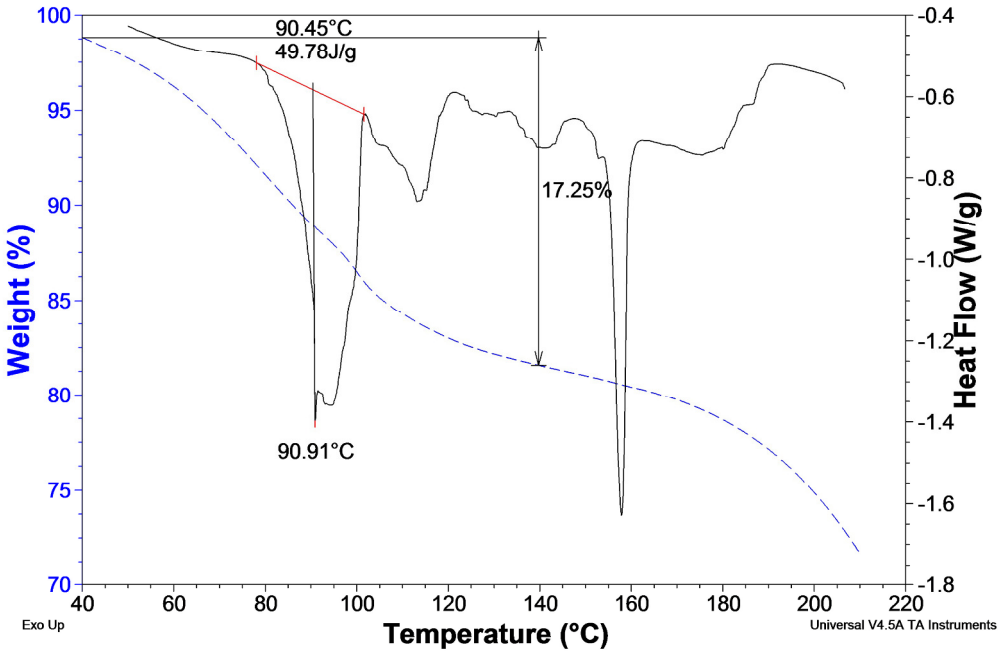

Figure 69. TGA (blue) and DSC (black) analysis of CBZ:DMSO hemi-solvate.

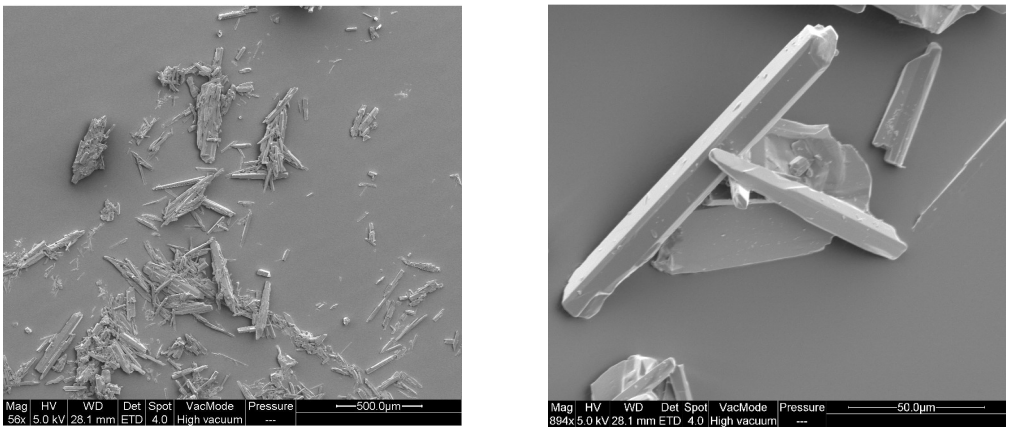

Figure 70. SEM micrographs for CBZ:DMSO hemi-solvate, overview (left) and close-up (right).

The PSD of CBZ:DMSO solvate could not be determined due to lack of material.

Glutaric acid co-crystal

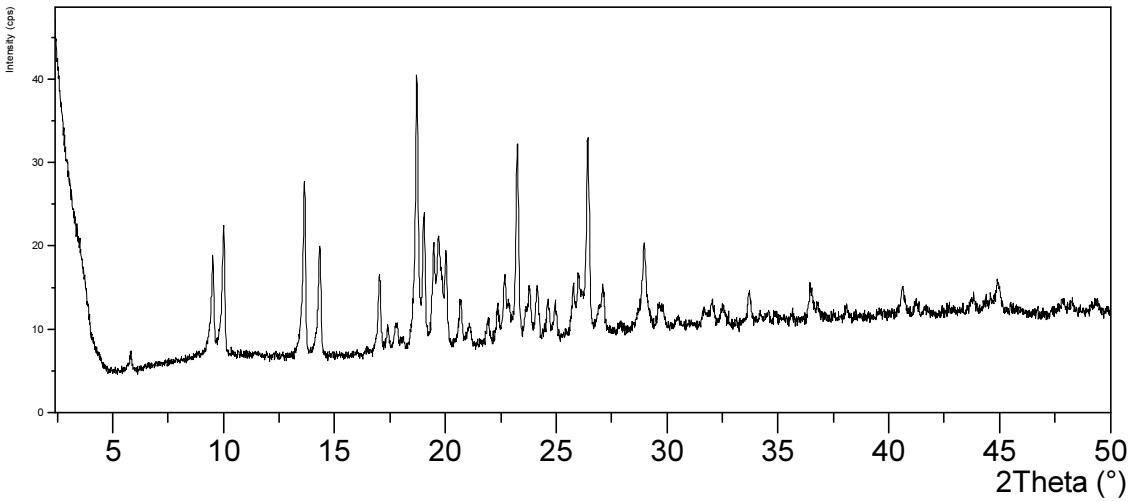

Figure 71. X-ray diffractogram of the CBZ:Glutaric acid co-crystal.

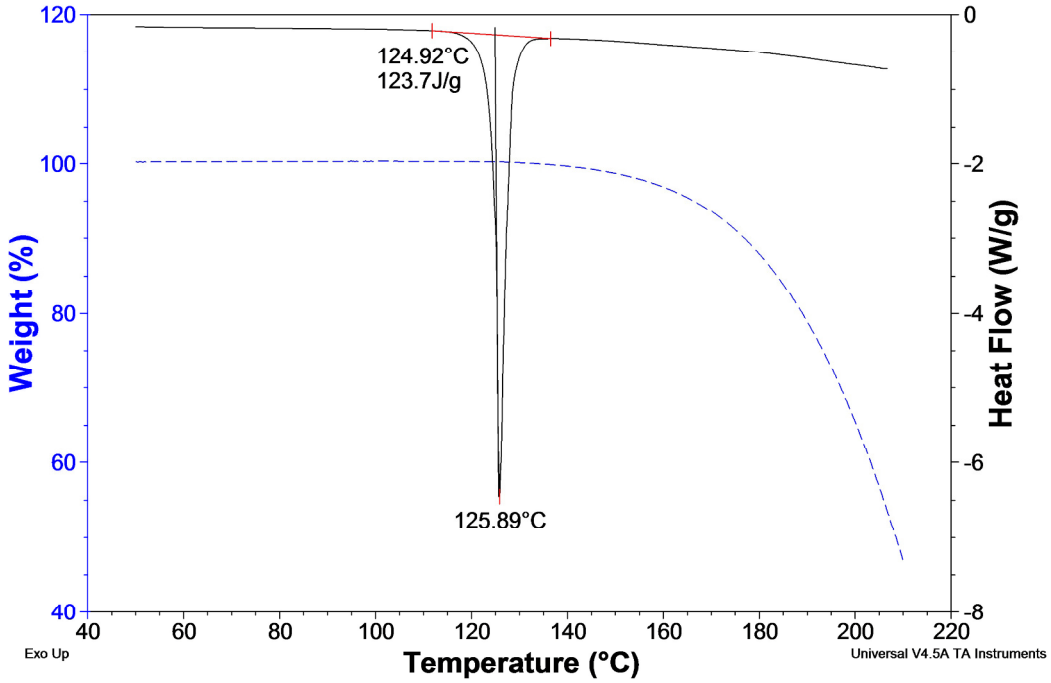

Figure 72. TGA (blue) and DSC (black) analysis of CBZ:Glutaric acid co-crystal.

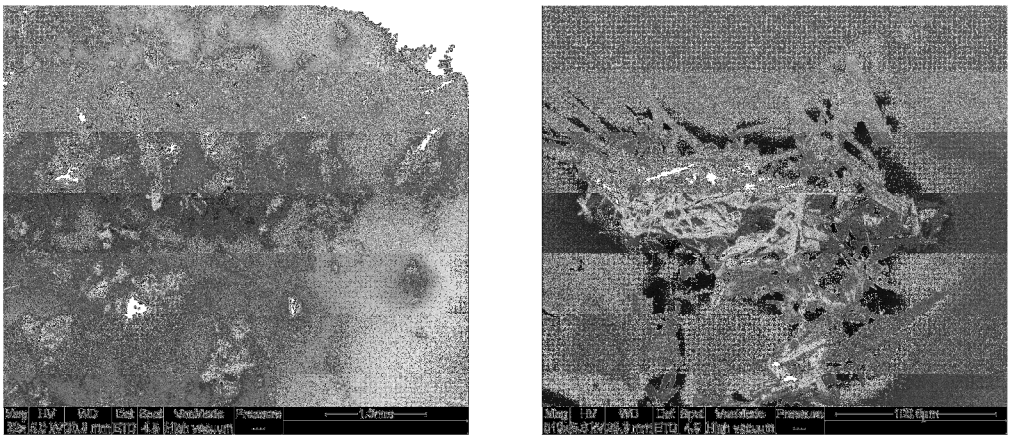

Figure 73. SEM micrographs for CBZ:Glutaric acid co-crystal, overview (left) and close-up (right).

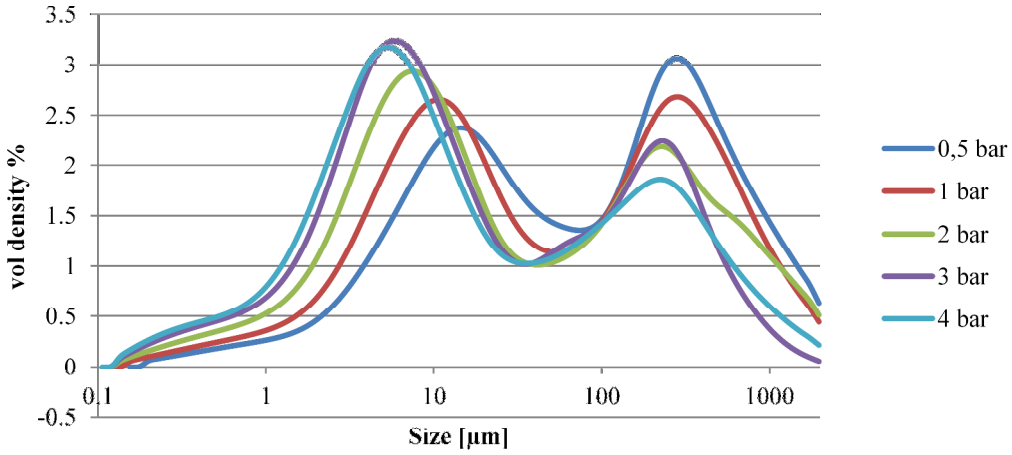

Figure 74. PSDs for CBZ:Glutaric acid co-crystal for varying pressures.

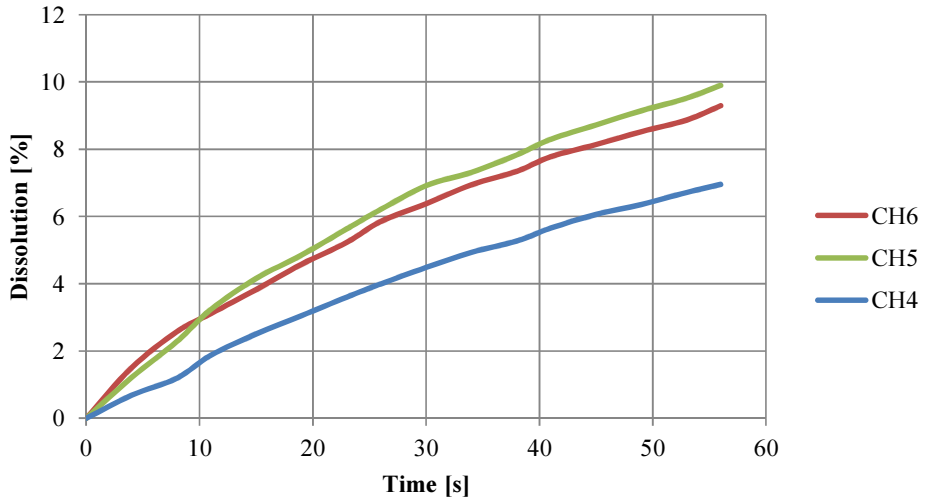

Figure 75. Dissolution rate measurements for CBZ:Glutaric acid co-crystal. Sample weights were 81 µg, 97 µg and 88 µg for CH 4, 5 and 6 respectively.

Malonic acid co-crystal

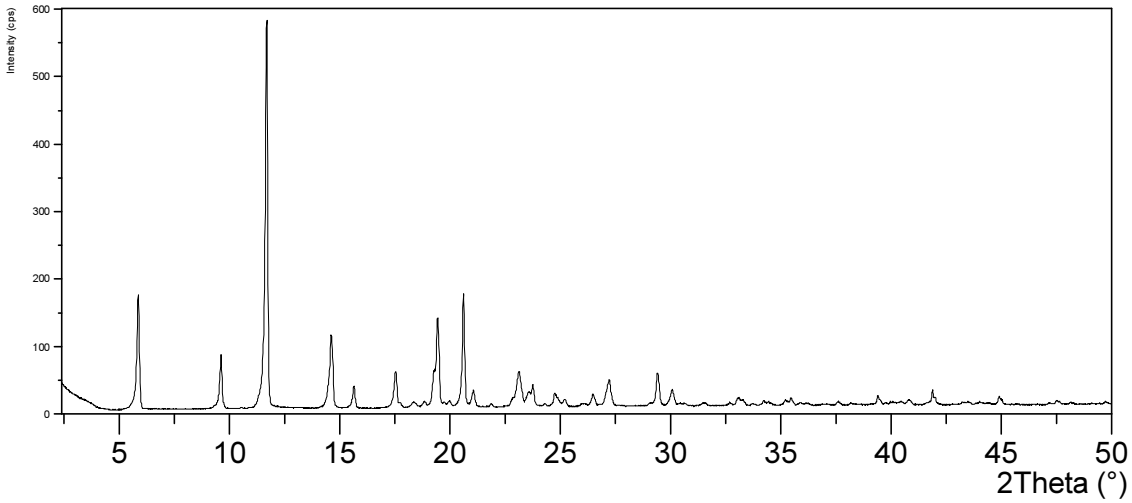

Figure 76. X-ray diffractogram of the CBZ:Malonic acid co-crystal.

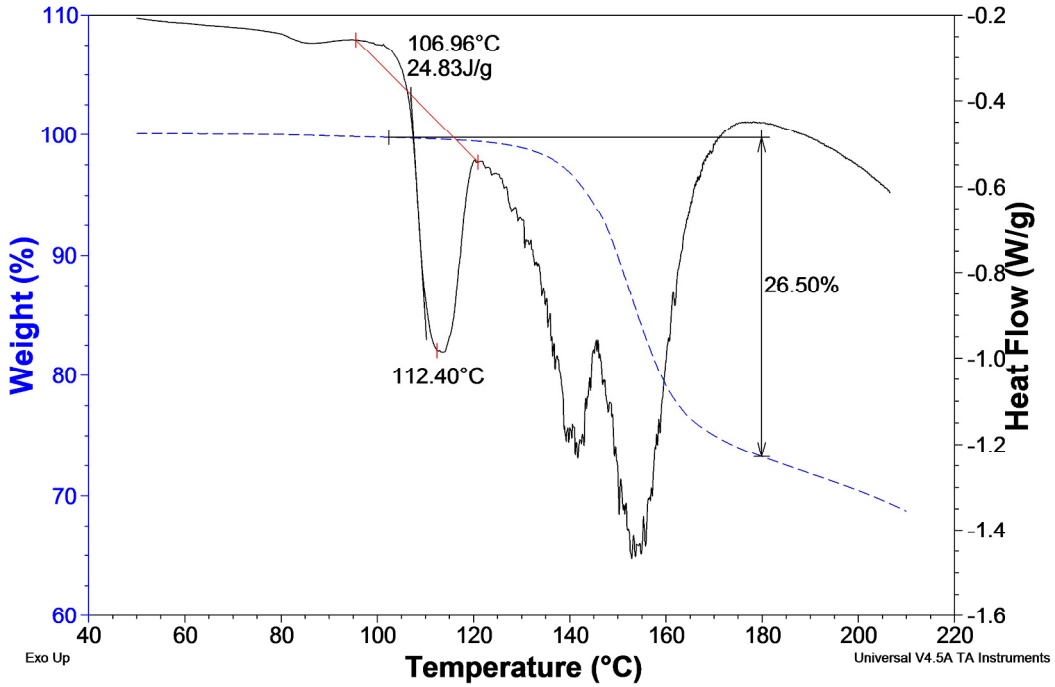

Figure 77. TGA (blue) and DSC (black) analysis of CBZ:Malonic acid co-crystal.

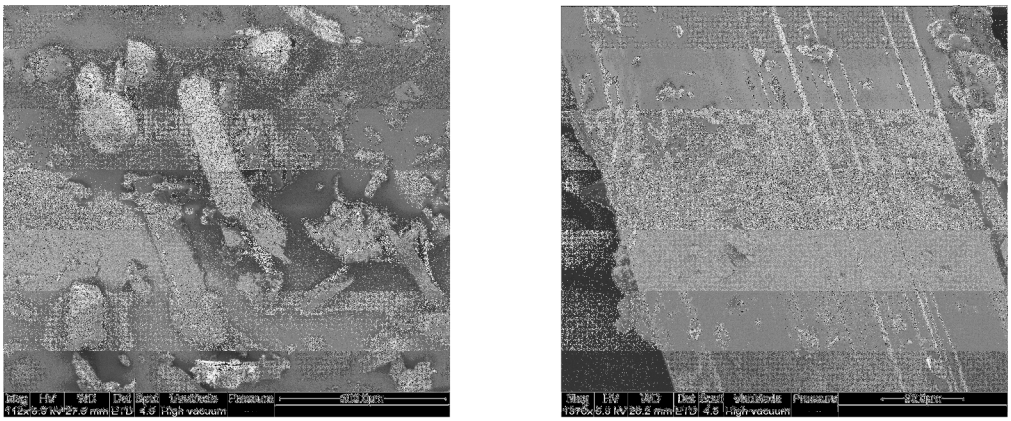

Figure 78. SEM micrographs for CBZ:Malonic acid co-crystal, overview (left) and close-up (right).

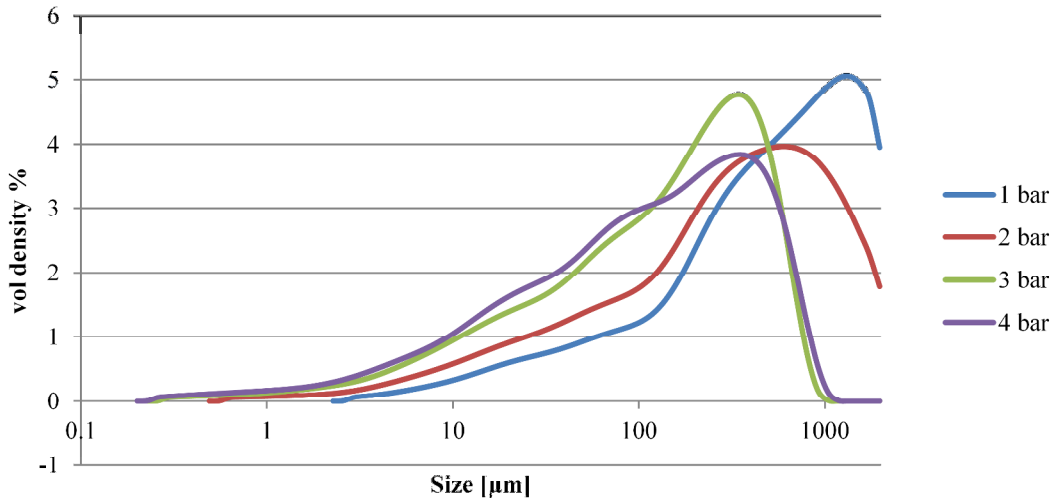

Figure 79. PSDs for CBZ:Acetic acid solvate for varying pressures.

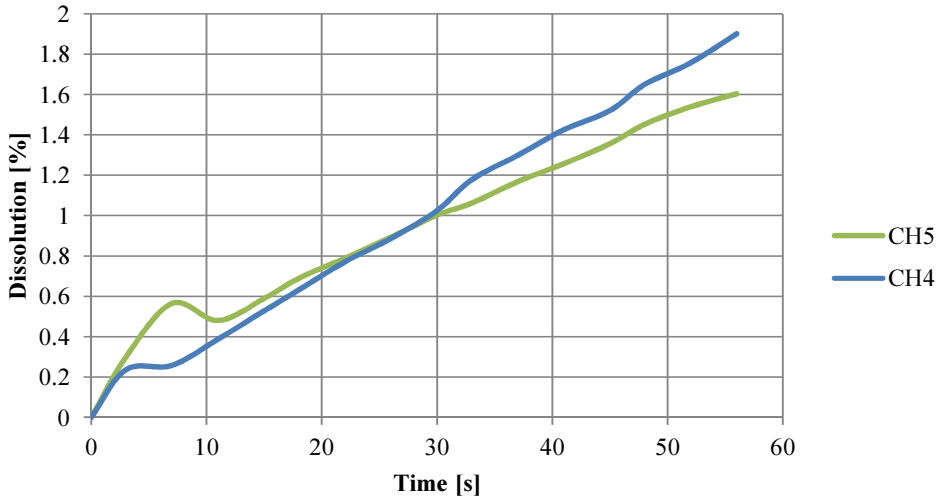

Figure 80. Dissolution rate measurements for CBZ::Malonic acid co-crystal. Sample weights were 122 µg and 138 µg for CH 4 and 5.

Saccharin co-crystal

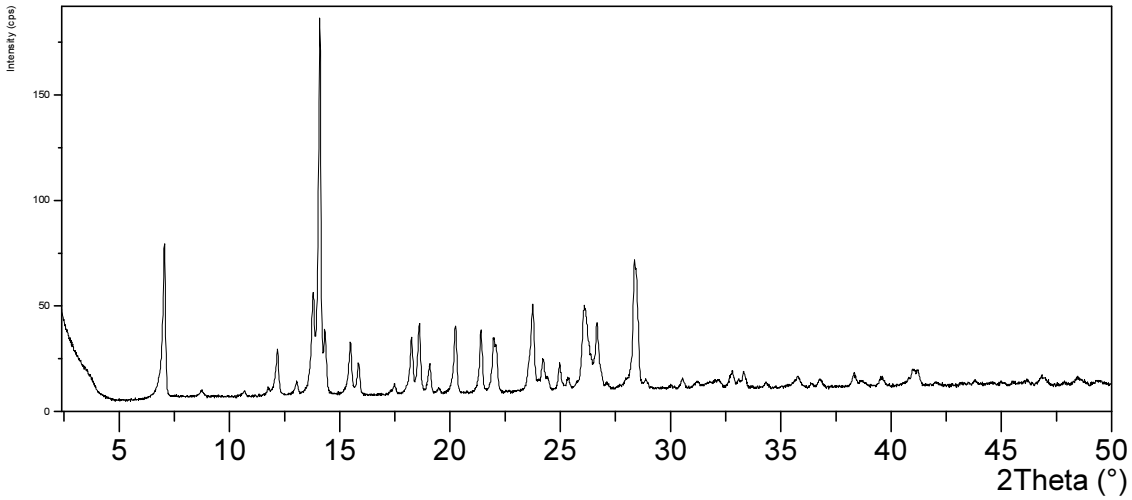

Figure 81. X-ray diffractogram of the CBZ:Saccharin co-crystal.

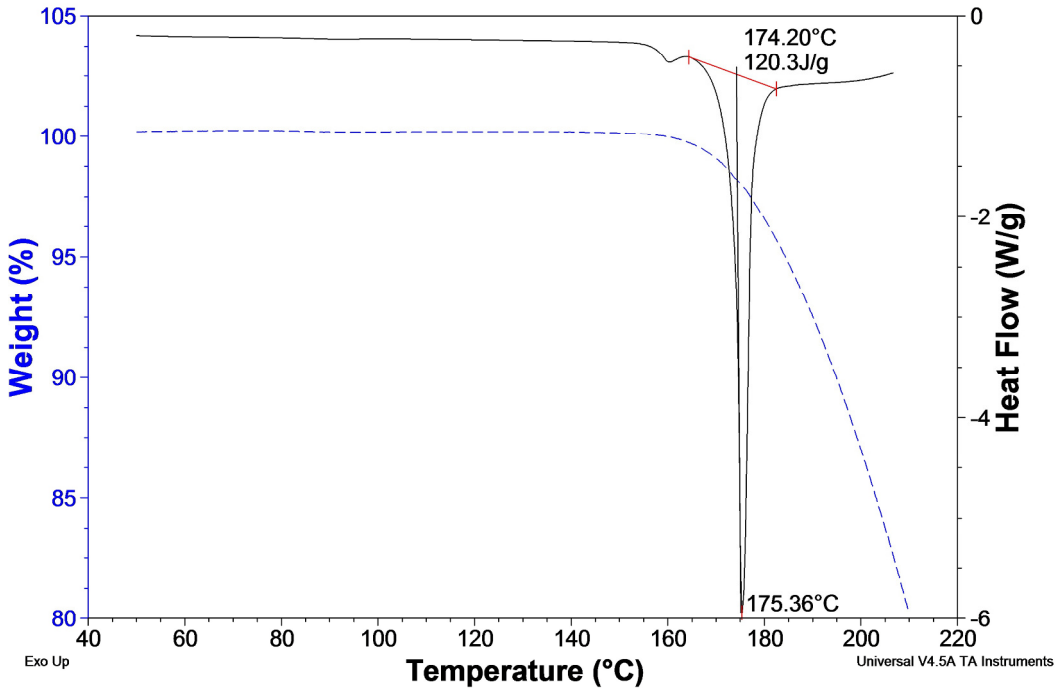

Figure 82. TGA (blue) and DSC (black) analysis of CBZ:Saccharin co-crystal.

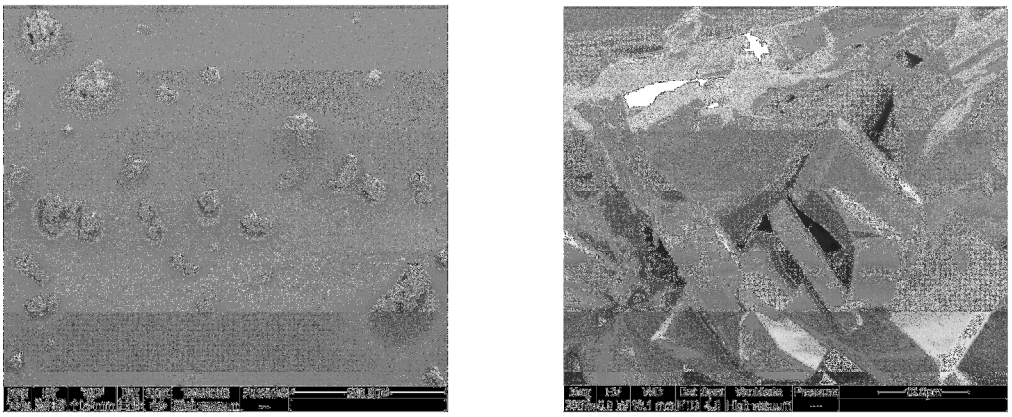

Figure 83. SEM micrographs for CBZ:Saccharin co-crystal, overview (left) and close-up (right).

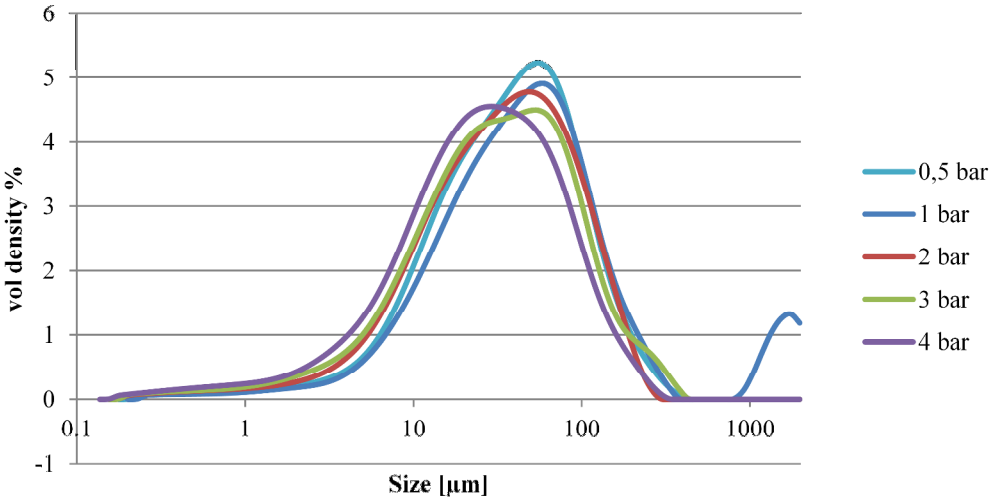

Figure 84. PSDs for CBZ:Saccharin co-crystal for varying pressures.

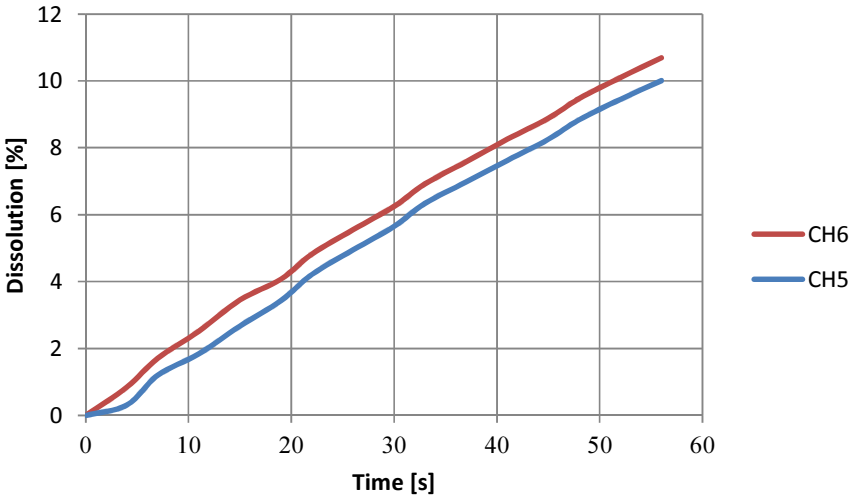

Figure 85. Dissolution rate measurements for CBZ:Saccharin acid co-crystal. Sample weights were 142 µg and 171 µg for CH 5 and 6 respectively. CH4 was omitted due to large variation probably arising from instrument error.

Salicylic acid co-crystal

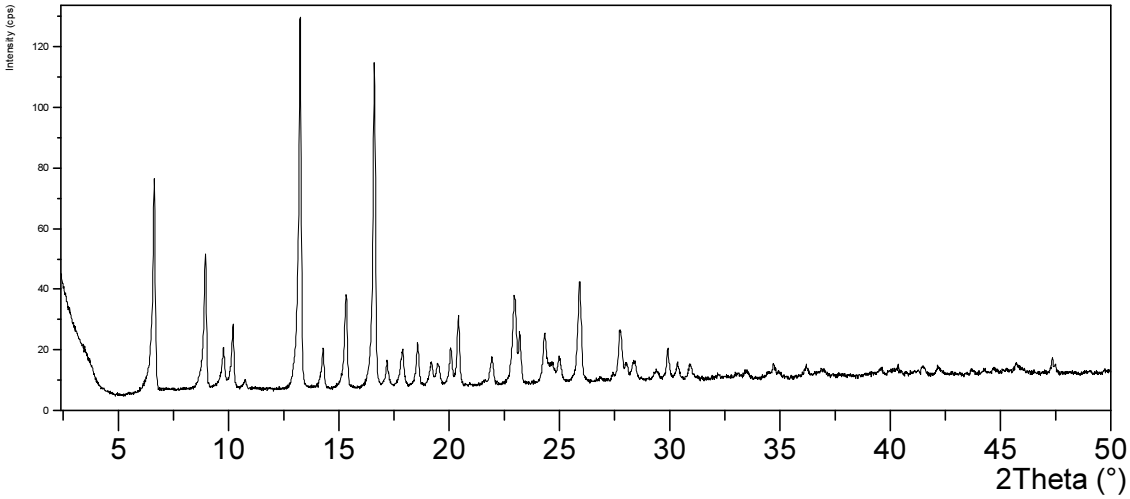

Figure 86. X-ray diffractogram of the CBZ:Salicylic acid co-crystal.

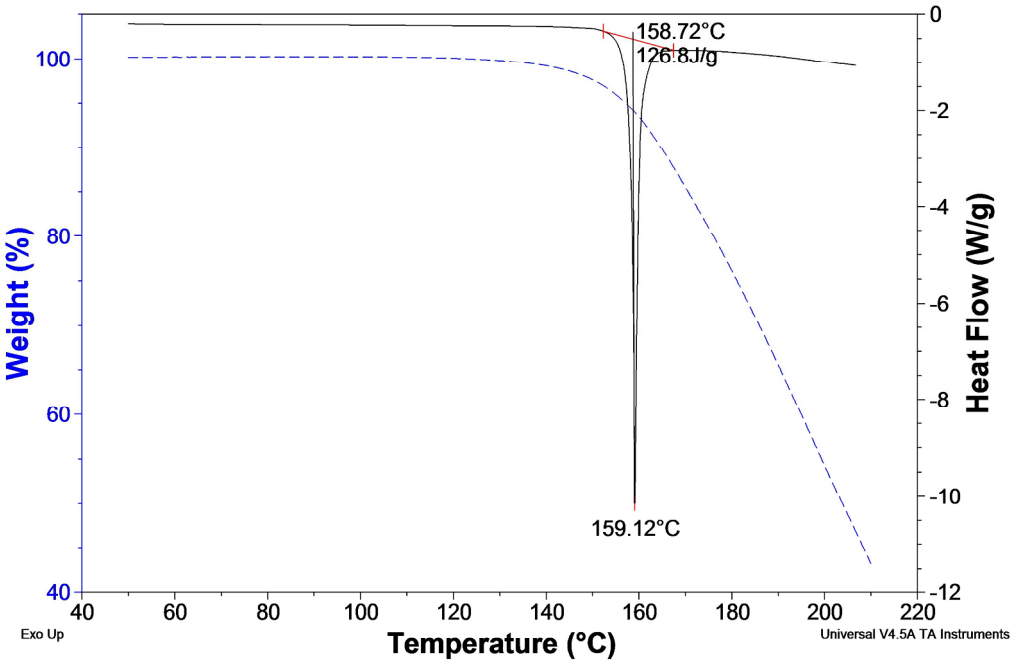

Figure 87. TGA (blue) and DSC (black) analysis of CBZ:Salicylic acid co-crystal.

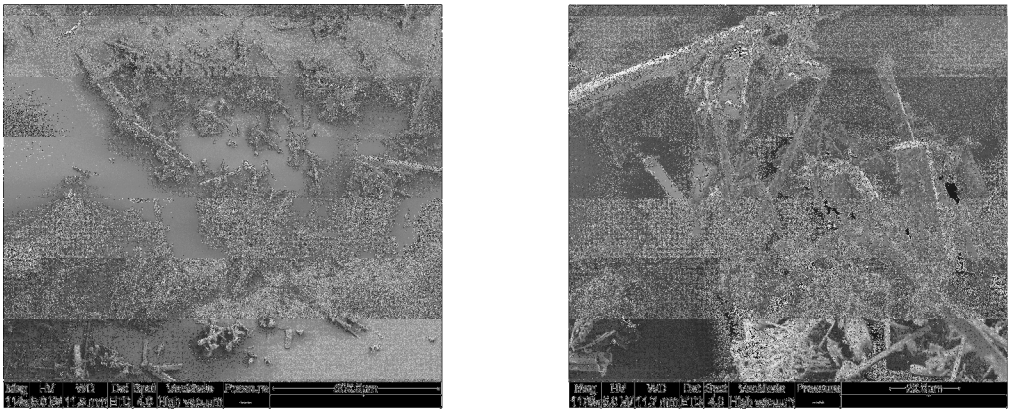

Figure 88. SEM micrographs for CBZ:Salicylic acid co-crystal, overview (left) and close-up (right).

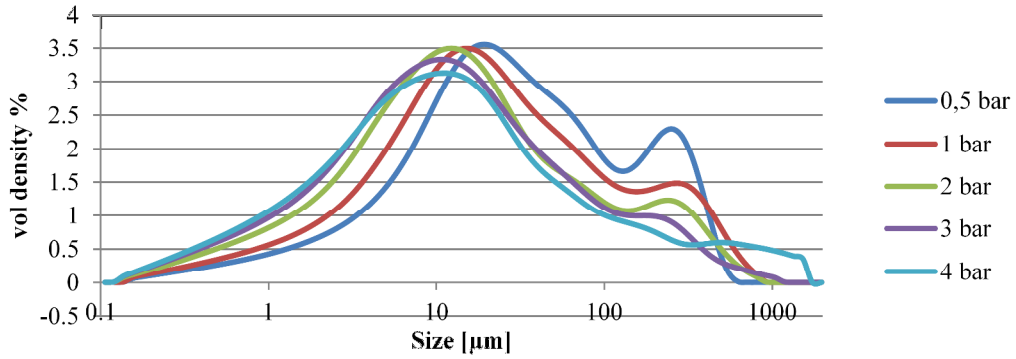

Figure 89. PSDs for CBZ:Salicylic acid co-crystal for varying pressures.

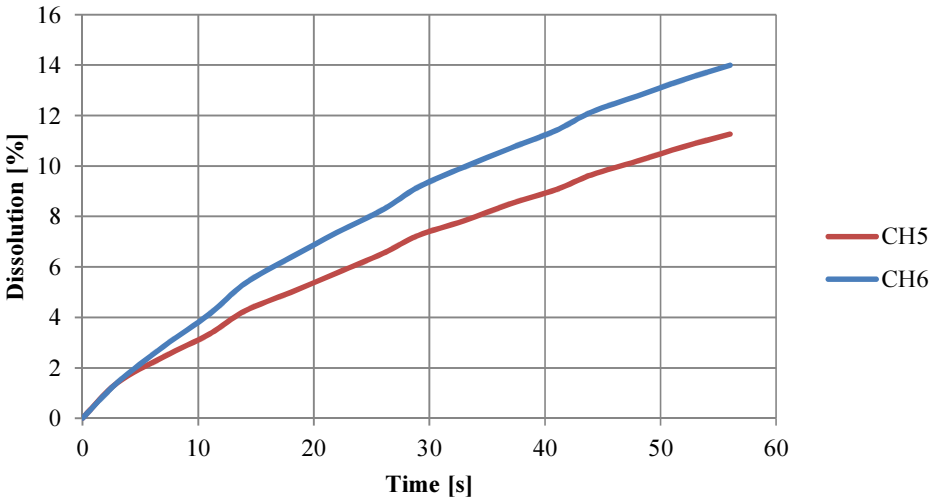

Figure 90. Dissolution rate measurements for CBZ:Salicylic acid co-crystal. Sample weights were 131 µg and 102 µg for CH 5 and 6 respectively. CH4 was omitted due to large variation probably arising from instrument error.

Succinic acid co-crystal

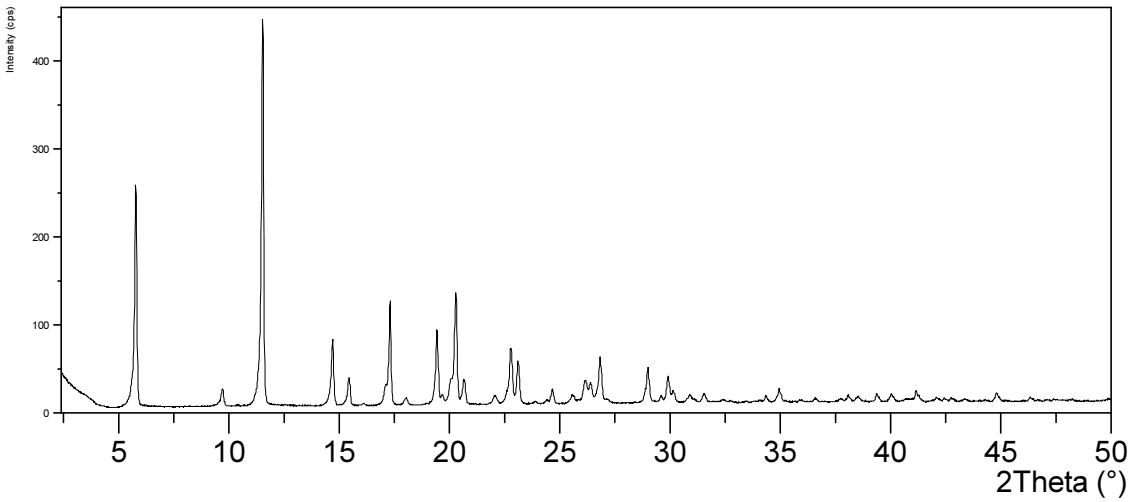

Figure 91. X-ray diffractogram of the CBZ:Succinic acid co-crystal.

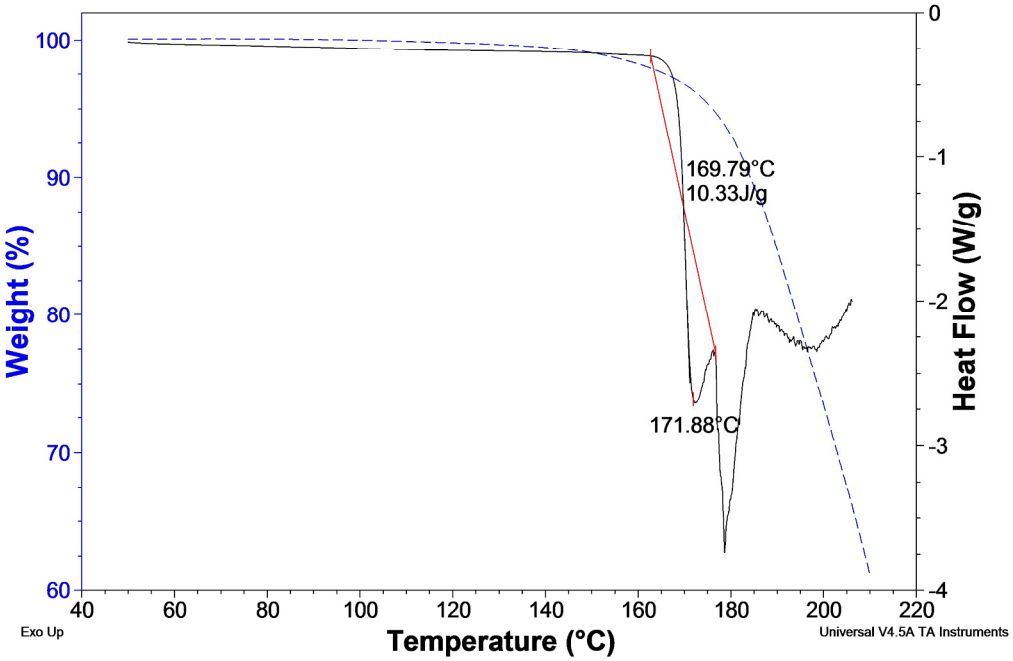

Figure 92. TGA (blue) and DSC (black) analysis of CBZ:Succinic acid co-crystal.

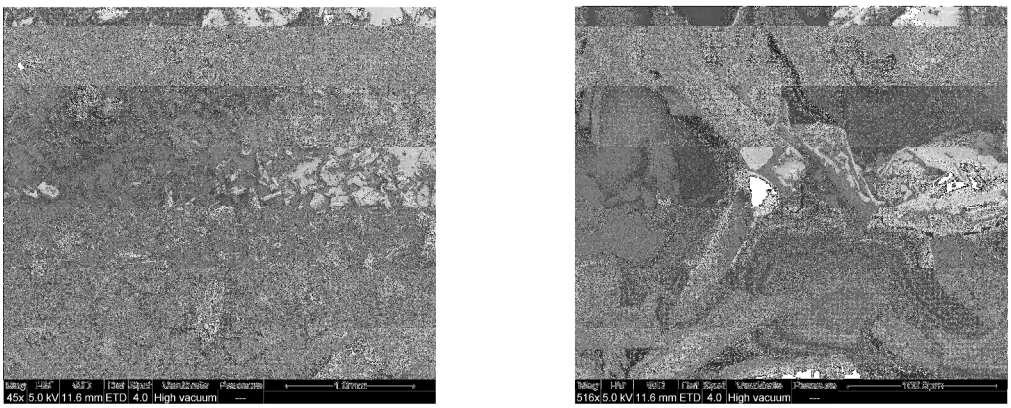

Figure 93. SEM micrographs for CBZ:Succinic acid co-crystal, overview (left) and close-up (right).

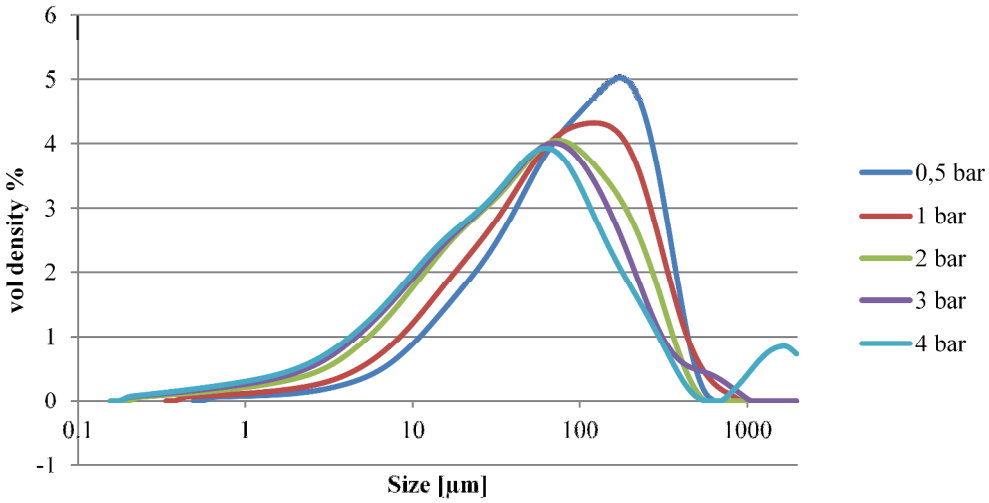

Figure 94. PSDs for CBZ:Succinic acid co-crystal for varying pressures.

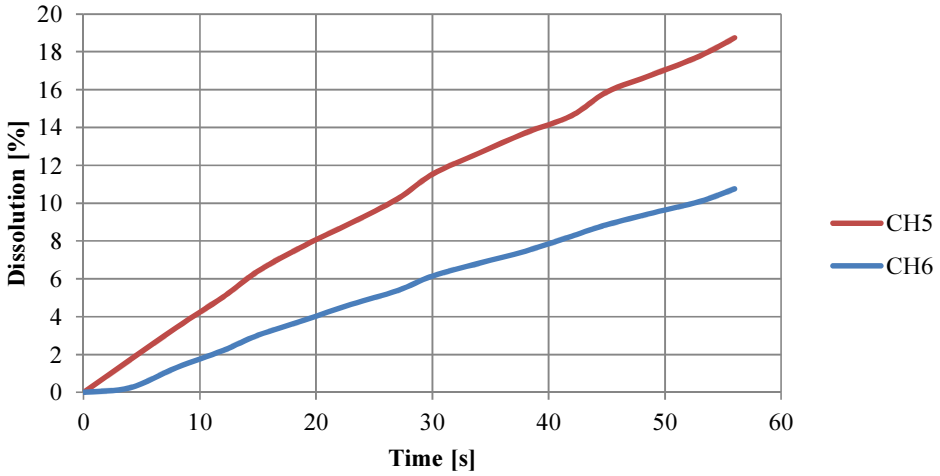

Figure 95. Dissolution rate measurements for CBZ:Succinic acid co-crystal. Sample weights were 108 µg and 83 µg for CH 5 and 6 respectively. CH4 was omitted due to large variation probably arising from instrument error.

Terephtalaldehyde co-crystal

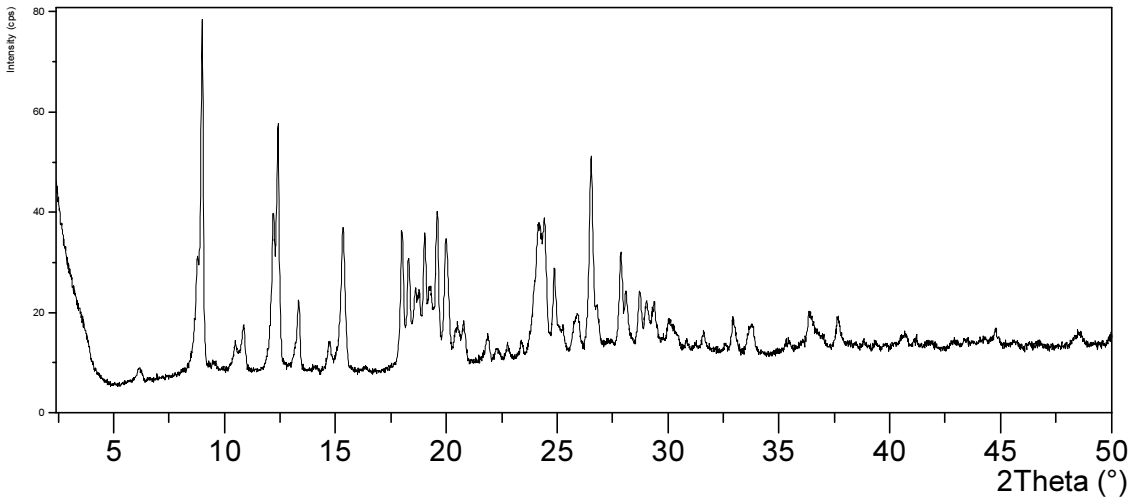

Figure 96. X-ray diffractogram of the CBZ:Terephtalaldehyde co-crystal.

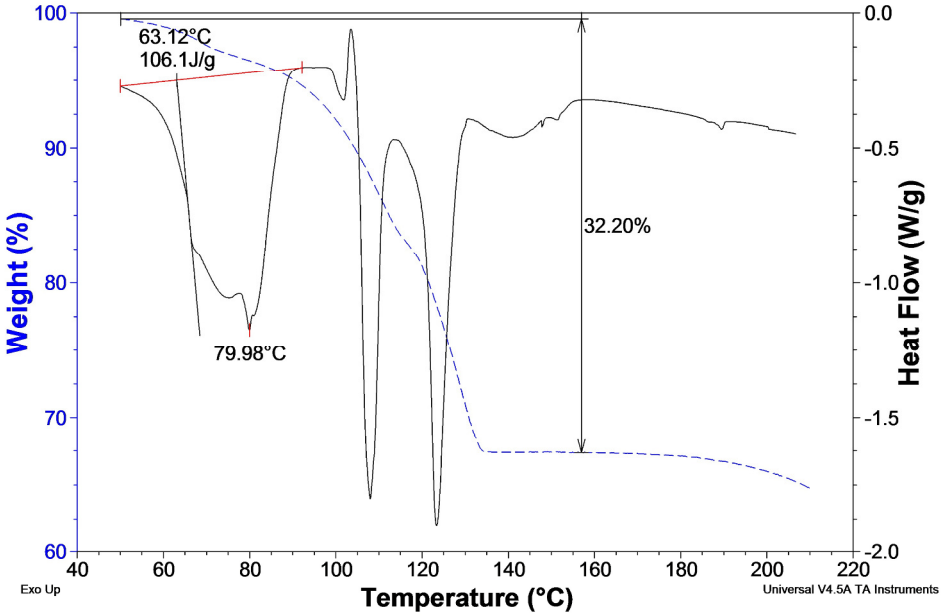

Figure 97. TGA (blue) and DSC (black) analysis of CBZ:Terephtalaldehyde co-crystal.

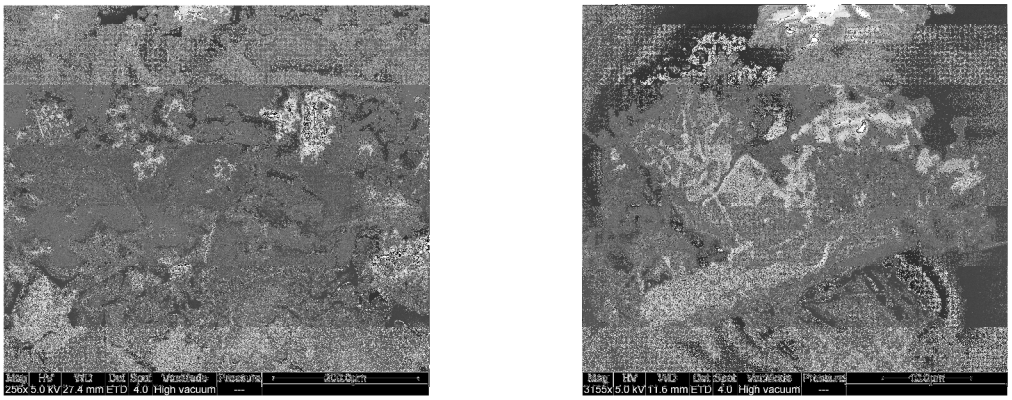

Figure 98. SEM micrographs for CBZ:Acetic acid solvate, overview (left) and close-up (right).

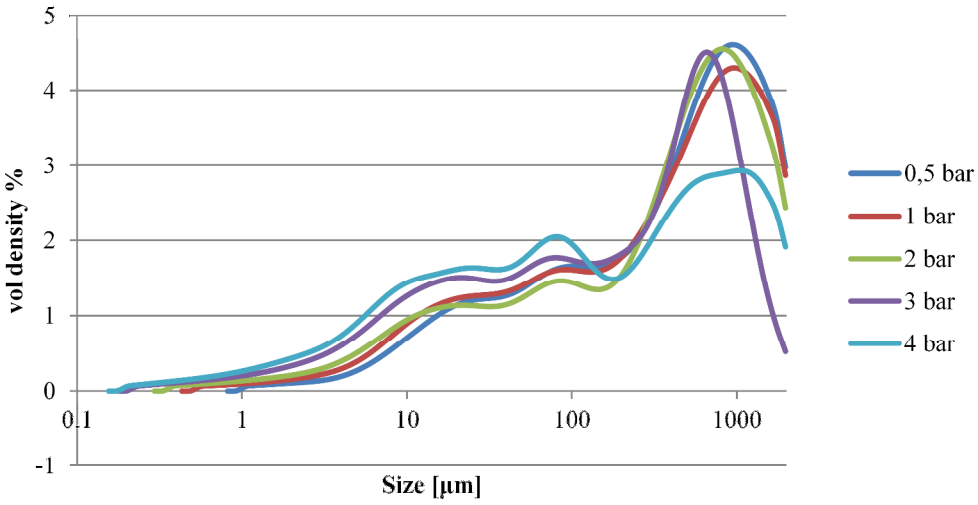

Figure 99. PSDs for CBZ:Terepthalaldehyde co-crystal for varying pressures.

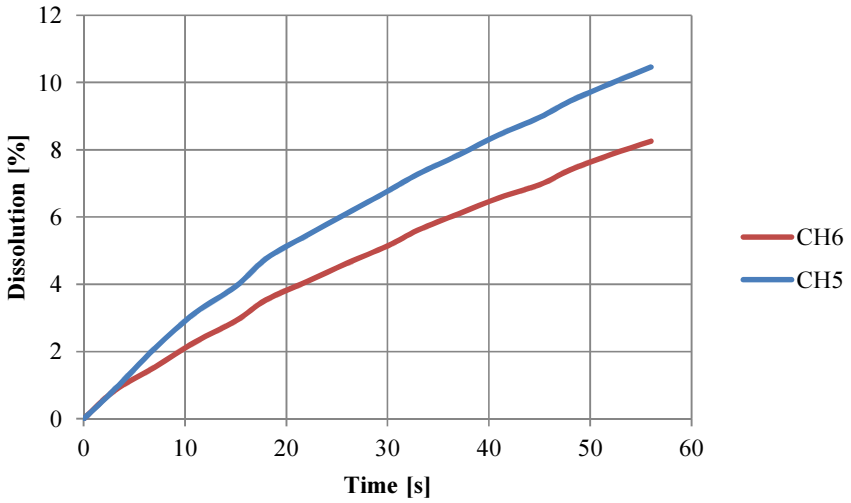

Figure 100. Dissolution rate measurements for CBZ:Terephtalaldehyde acid co-crystal. Sample weights were 106 µg and 93 µg for CH 5 and 6 respectively. CH4 was omitted due to large variation probably arising from instrument error.

APPENDIX D Experimental information

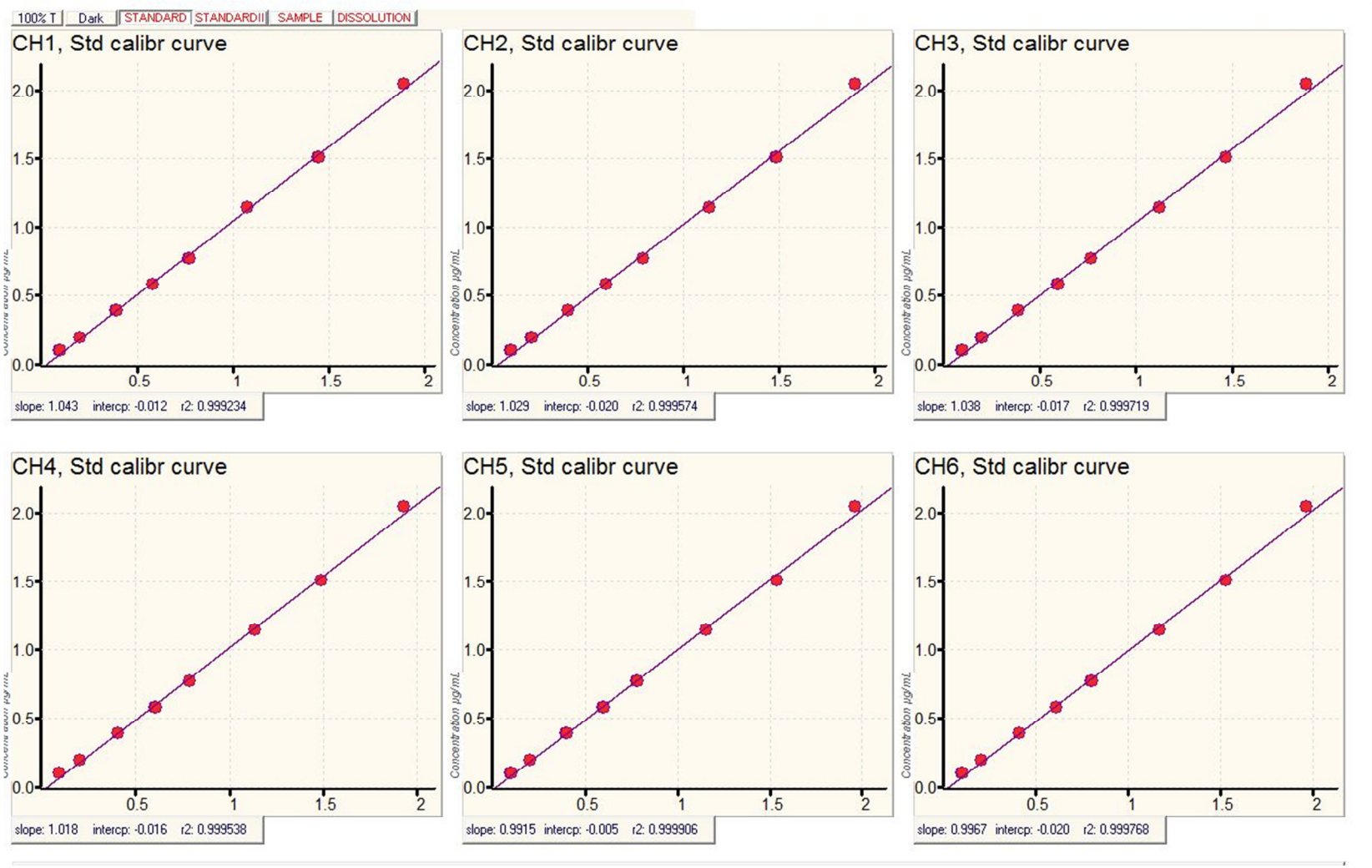

Figure 101. Calibration curve for carbamazepine as determined by uDISS. Notice the very low concentrations used, which is due to fact that the initial dissolution rate is the most interesting.

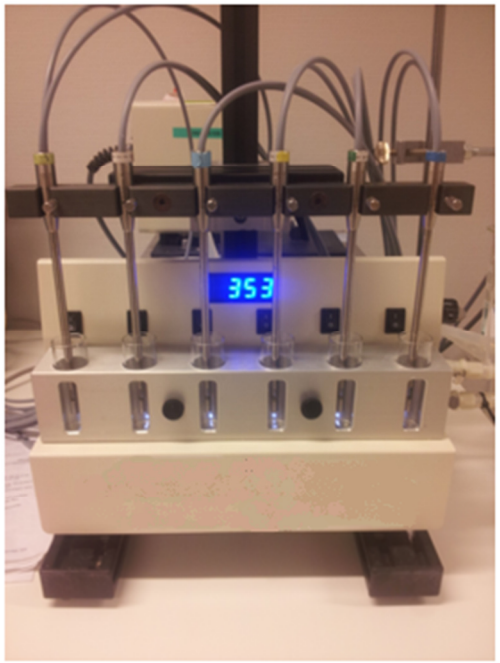

Figure 102. Experimental setup of the dissolution experiments. Here the six µDISS probes can be seen which collect data simultaneously. A cross-type magnetic stirrer was used to ensure good mixing. The blue number show the RPM of the magnetic stirring.