

Amide Pyramidalization in Carbamazepine: a Flexibility Problem in Crystal Structure Prediction?

**Aurora J. Cruz-Cabeza,^a Graeme M. Day,^a W. D. Sam Motherwell^b and William
Jones^a**

^aThe Pfizer Institute for Pharmaceutical Materials Science, Department of Chemistry,
University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK.

^bCambridge Crystallographic Data Centre, 12 Union Road, Cambridge, CB2 1EZ, UK.

Abstract

Carbamazepine is known to exist in various polymorphic forms. Here we report on crystal structure prediction calculations for carbamazepine in an attempt to examine the predictability and relative stability of the various polymorphs. Hypothetical crystal structures have been generated in ten of the most common space groups and these were compared to the four known polymorphs. Particular attention was given to the influence of amide pyramidalization on the relative energies of the predicted structures. While the actual generation of individual structures was not found to be dependent on the degree of deformation of the amide group, the final ranking of the predicted crystals was found to be greatly affected by pyramidalization of the amide nitrogen. This effect was examined in detail through systematic variation of the NH₂ geometry for each of the low energy crystal structures. A range of amide geometries were found to be favourable amongst the low energy structures, demonstrating that energetically feasible deformation of the amide

group could be adopted in order to optimize hydrogen bonding interactions in the crystals. We conclude that the neglect of pyramidalization will produce significant errors in crystal structure prediction for similar molecules.

Keywords: crystal structure prediction, carbamazepine, lattice energy minimization, amide pyramidalization.

1. Introduction

Polymorphism, the ability of a substance to crystallise in a variety of different crystal structures,¹ is of considerable interest. For a drug molecule polymorphism is recognised as an undesirable feature since the various polymorphs may display quite different physical properties. The ability to predict the relative thermodynamic stability and likely properties of known and latent polymorphs is, therefore, clearly of importance to various applications of molecular solids, including pharmaceutical applications. Such predictions, however, remain a formidable challenge with many difficulties to be overcome before successful prediction may be routinely made.²⁻⁶ One particular aspect is in the development of strategies and methodologies for dealing with conformationally flexible molecules.

Carbamazepine (CBZ, fig. 1), an anticonvulsant drug used in the treatment of epilepsy and trigeminal neuralgia, is known to crystallize in four polymorphic forms (Form I, II, III and IV, fig. 2),⁷⁻¹⁰ with the most recently observed form being characterized only two years ago (despite 30 years of study of this system). Concurrent with the work described here, automated parallel crystallization methods have been used in conjunction with crystal structure prediction to explore the polymorphism of carbamazepine under a wide range of crystallization conditions. Such experimental studies were accompanied by a parallel computational assessment of possible polymorphs.¹¹ No new polymorphs were observed after 594 individual crystallizations and, although calculations suggested that structures based on a chain hydrogen bond motif were found to be energetically stable,

the experimental data suggested that kinetic factors favoured the formation of the $R_2^2(8)$ carboxamide dimer - the motif observed in all four polymorphs.

Here we report a more detailed prediction study of the crystal structures of carbamazepine, with particular consideration given to the influence of amide pyramidalization on the energy ranking of the hypothetical structures. Although amides groups are generally planar, the planarity can be distorted through C-N bond rotation (twisted amides) or nitrogen pyramidalization (pyramidal amides). Nitrogen pyramidalization is energetically more feasible than rotation but, in general, both processes are very closely related and neither occurs without some contribution of the other. This is a well known problem in the computer simulation of peptides.¹² Although some force fields attempt to address this matter,^{12,13} the problem is complicated by the fact that a change in nitrogen hybridization will be accompanied by large changes in partial atomic charges. As a result, polarisation effects also need to be considered in order to correctly describe the molecular electrostatic potential. This problem has not previously been addressed in the prediction of crystal structures, due in part to the lack of accurate experimental data. Hydrogen atoms, often the most important atoms involved in intermolecular interactions in molecular solids, are the most poorly resolved by X-ray diffraction - the most commonly used technique for structure determination.

We report here, therefore, on our studies concerning the predictability of three (Forms II – IV) of the four polymorphs of carbamazepine. We exclude Form I, which has four molecules in the asymmetric unit, and limit our studies to crystal structures with $Z' = 1$.

We compare two descriptions of the electrostatic intermolecular interaction – (i) ESP derived atomic point charges^{14,15} and (ii) atomic multipoles obtained from a distributed multipole analysis.¹⁶ In addition, the influence of the inversion of the amide nitrogen on the final energies of both the known and hypothetical polymorphs is examined. Since adequate force fields with suitable terms for the description of the NH₂ pyramidalization, together with an accurate description of intermolecular interactions, are not available, we are limited to treating the molecules as rigid bodies in a series of calculations using systematically varied molecular geometries.

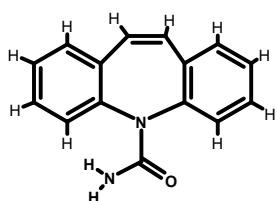


Figure 1. Molecular structure of carbamazepine.

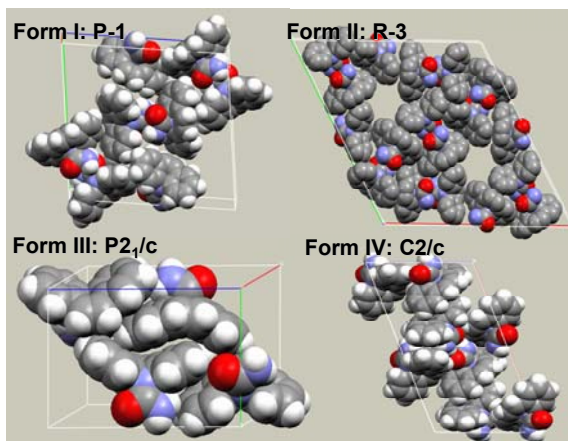


Figure 2. Unit cells of the experimentally known carbamazepine polymorphs.

2. Computational Methods

Molecular model. All molecular models were treated as rigid during the crystal structure search and lattice energy minimizations. The initial molecular structures were obtained using density functional theory (DFT) gas phase molecular geometry optimizations using the code *Dmol3* as implemented in the Accelrys package *Materials Studio*¹⁷ and the PW91 hybrid functional¹⁸ along with a double numerical polarised (DNP) basis set.¹⁹ In studying the effect of nitrogen pyramidalization, conformational energies were evaluated at the same level of theory by partially optimizing molecular structures with constrained HNCN torsion angles.

Potential model. The repulsion-dispersion contributions to the intermolecular potential were evaluated as:

$$U^{vdw} = \sum_{M,N}^{Nmol} \left(\sum_{i \in M < k \in N} U_{ik}^{vdw} \right) = \sum_{M,N}^{Nmol} \left(\sum_{i \in M < k \in N} \left(A_{ik} e^{-B_{ik} R_{ik}} - \frac{C_{ik}}{R_{ik}^6} \right) \right),$$

where atoms in molecules *M* and *N* are of type *i* and *k* respectively, and parameters A_{ik} , B_{ik} and C_{ik} are characteristic of the atom types. The Williams W99 empirically derived atom-atom potential parameters²⁰⁻²² were used and interactions were summed to a 15 Å cutoff. In this potential, the centre of interaction for all hydrogen atoms is shifted 0.1 Å along the X-H bond (X= C, N, O) towards the heavy atom; X-H distances were shortened by 0.1 Å after DFT optimization.

Electrostatic models. Electrostatic contributions to the intermolecular potential were calculated using two models for the molecular charge distribution: atom centred charges

and atomic multipoles. Atom centred point charges were derived to reproduce the electrostatic molecular potential (ESP charges)^{14,15} using the fitting procedure implemented in *Dmol3*. Atomic multipoles up to $l=4$ (charge, dipole, quadrupole, octupole and hexadecapole) were obtained by performing a Distributed Multipole Analysis¹⁶ of a B3P91/6-31G(d,p) electron density obtained using the *CADPAC* code.²³ All charge-charge, charge-dipole and dipole-dipole interactions were evaluated using an Ewald summation, while higher order terms (up to R^{-5}) were summed to a 15 Å cutoff based on whole molecules.

Crystal structure prediction. Crystal structures were generated in the nine most common space groups ($P2_1/c$, $P-1$, $P2_12_12_1$, $P2_1$, $C2/c$, $Pbca$, $Pnma$, $Pna2_1$ and $Pbcn$) using the simulated annealing algorithm of Karfunkel and Gdanitz,²⁴⁻²⁶ as implemented in the Accelrys *Polymorph Predictor* (PP) module of the Cerius² software suite.²⁷ In addition the $R-3$ space group was also considered because one of the observed polymorphs exists in this less common space group. Consecutive independent searches from different random starting points (up to 3 or 4) were performed within each space group until the generation of structures converged, i.e. no additional new structures were generated. These initial searches were carried out with only one molecule in the asymmetric unit ($Z'=1$) followed by lattice energy minimization (W99 + ESP charges). Hypothetical structures were clustered (using the algorithm implemented in the Cerius2 software²⁶) to remove identical structures. The adjustable parameters in the simulated annealing steps were taken from previous studies.²⁸ Lattice energy minimization of the

predicted structures with the atomic multipole electrostatic model and the same W99 model potential were performed using the program *DMAREL*.²⁹

Final clustering. Predicted structures were finally clustered using the *COMPACT*³⁰ algorithm with a tolerance of 0.15 Å on atomic separations. The same algorithm was used for comparison of predicted structures with the experimental structures. The *COMPACT* algorithm compares the structures on the basis of atomic separations in a coordination sphere of 14 molecules.

3. Results

3.1 Exploring the carbamazepine molecular energy surface and choosing a starting conformation

Prior to searching for crystal structures, it is necessary to identify possible low energy molecular conformations. Gas phase geometry optimizations of the isolated carbamazepine molecule were therefore performed and two molecular conformations identified as minima in the potential energy surface (fig. 3), located. The first was a global minimum (denoted as H_{in}). This was about 2 kJ/mol more stable than a second local minimum (designated H_{out}). The only significant difference between the two geometries is the direction of the nitrogen pyramidalization and, as a consequence, the orientation of the hydrogen atoms of the amide group. Using the OCN molecular plane as reference, the amide hydrogen atoms may be described as pointing towards the inner (H_{in}) or the outer (H_{out}) part of the molecule – see Figure 3.

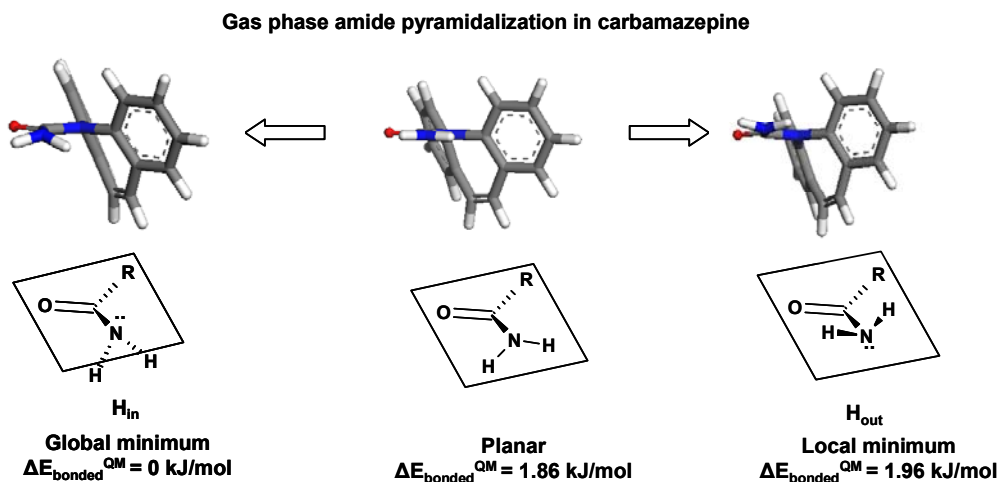


Figure 3. Gas phase carbamazepine molecular geometry minima and planar geometry and their relative energy ($\Delta E_{\text{bonded}}^{\text{QM}}$) referred to the most stable conformation (*Dmol3*, *pw91/dnp*).

The lowest energy interconversion pathway between these two minima was then studied using *Dmol3* and the relative conformational energies obtained as a function of selected the HNCN dihedral angle (fig. 4). (Calculations were performed by constraining the HNCN dihedral angle at different values from -60° to $+50^\circ$, allowing the remainder of the molecule to optimize at each point.) Each value of the HNCN dihedral angle corresponds to a different degree of amide pyramidalization, with negative values result in the amide hydrogens pointing below the plane (fig. 3) and positive values above the plane: 0° corresponding to the planar geometry. As expected, the low energy barrier between molecular minima is sufficiently small (2.06 kJ/mol) to allow interconversion in the gas phase, although in the crystal, hydrogen bonding may limit the mobility of the hydrogen atoms. Rotation about C-N bond was also monitored, but was found to be significantly more energetically expensive (the calculated rotation barrier was almost 46 kJ/mol).

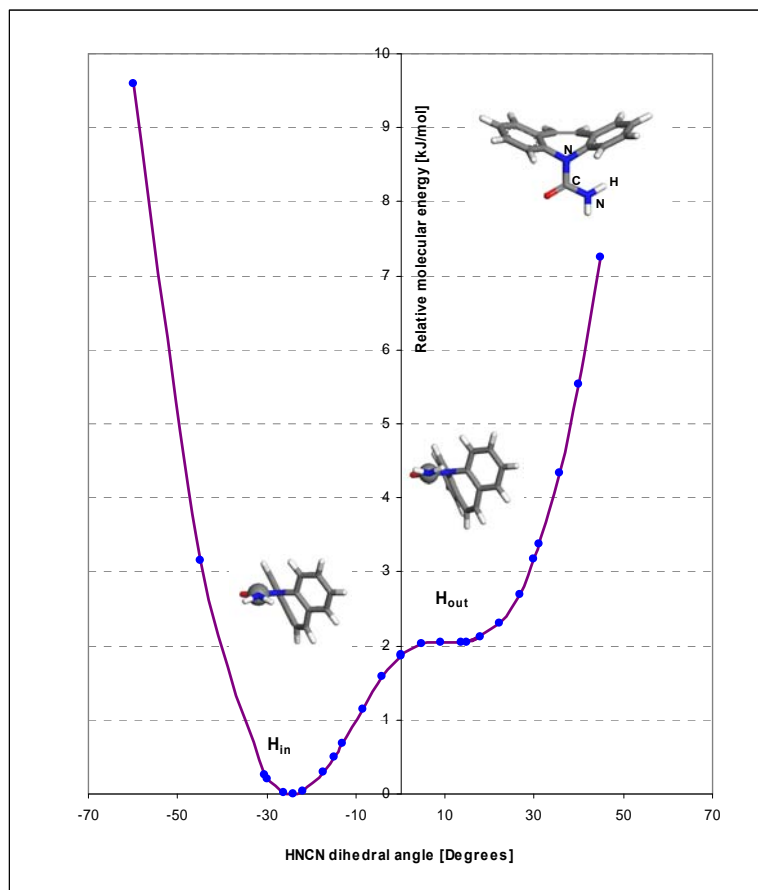


Figure 4. H_{in} - H_{out} interconversion pathway (Dmol3, pw91/dnp).

We also calculated non-bonded intramolecular interactions of the amide group with the rest of the molecule for the conformations involved in the interconversion pathway (using the W99 parameters for the repulsion-dispersion terms and the atomic point charge model for the electrostatic component). Subtracting this non-bonded contribution from the total energy obtained from DFT calculations revealed the natural geometry of the isolated amide group; the surface converged to a surface with one minimum corresponding to the completely planar amide group (the curve takes the shape of a parabola with minimum at

0°). These calculations suggest that the amide deformation from planarity in the gas phase is a consequence of the non-bonded intramolecular interactions of the hydrogen atoms with the rest of the molecule. These interactions are of the same type and magnitude as those between molecules in crystals, so it is to be expected that crystal packing forces will affect the degree of deformation of the amide group. As a result it is important to investigate how the choice of molecular model influences the various steps involved in crystal structure prediction: the generation of all closely packed structures and the evaluation of their relative stabilities. Although the conformation adopted by the molecule in the crystal does not necessarily correspond to an energy minimum for the isolated molecule, we first performed calculations with both molecular minima to see how the choice of conformation affected the generation of structures and their relative energies.

3.2 How does the choice of molecular conformation affect the crystal structure generation and ranking?

Structure generation was for both conformations performed with one molecule in the asymmetric unit. Values of relative lattice energy ($\Delta E_{non-bonded} = E_{non-bonded,i} - E_{non-bonded,1}$), referred to the most stable crystal structure (ranked $i = 1$), and the density of the structures were plotted after minimization with the atomic point charge and the atomic multipole models – see figure 5. As mentioned earlier, predictions for Form I were not attempted. Instead, given the experimental crystal structure, Form I was lattice energy minimized using the experimental coordinates after replacing the observed molecular geometries by each of our molecular models.

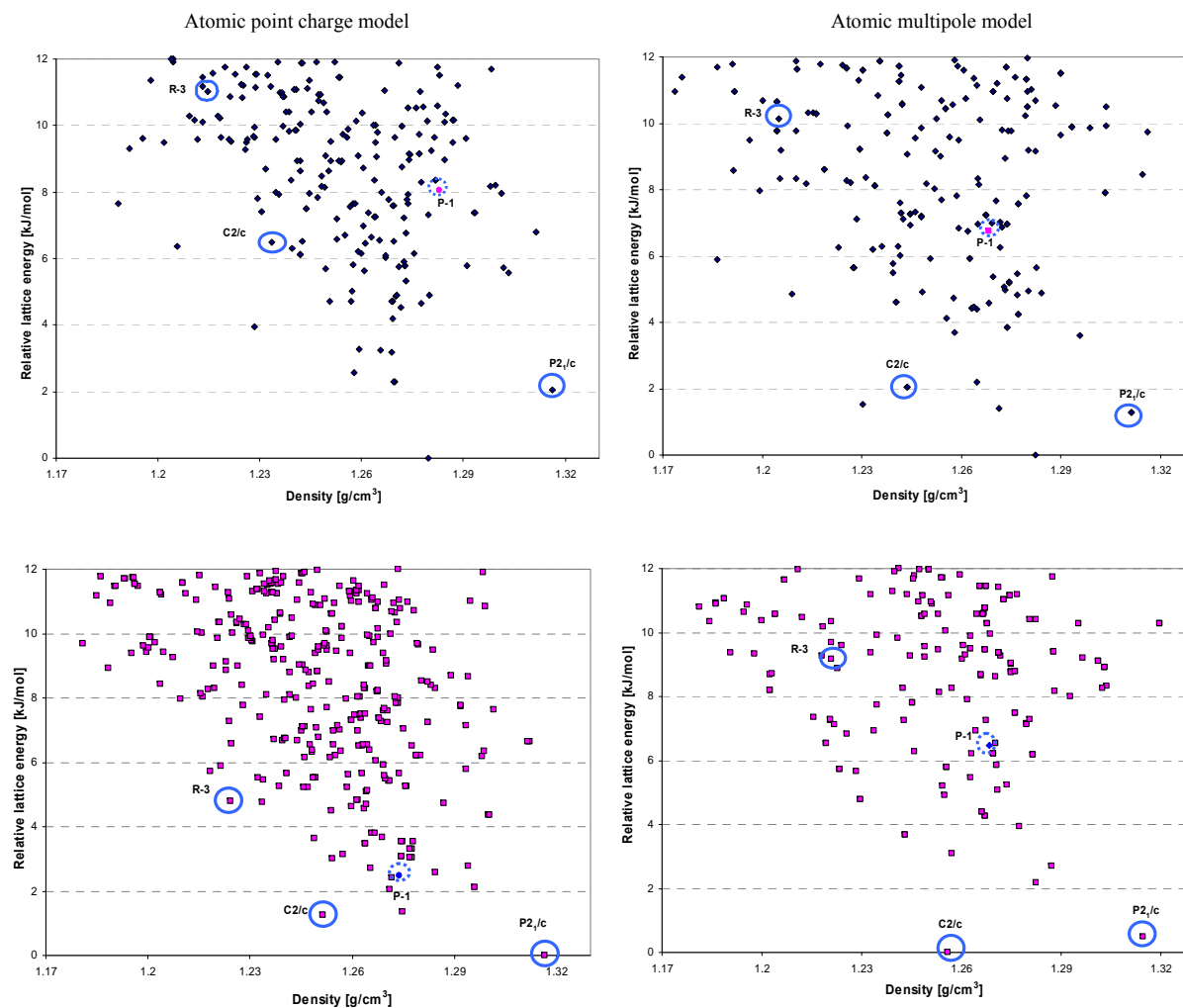


Figure 5. Relative energies and densities of predicted crystal structures of carbamazepine obtained with H_{in} (diamonds, above) and H_{out} (squares, below) molecular conformations, after minimization using the atomic point charge model (left) and the atomic multipole model (right). Circled structures correspond to those matching the experimentally observed polymorphs. Form I is lattice energy minimized from the experimental coordinates.

Our predictions showed that the same set of hypothetical crystal structures were generated with both molecular geometries in the low energy regions of packing space (where the sampling is complete), including the three $Z'=1$ experimental polymorphs (circled structures) – i.e. the conformational differences between the molecular models were not important enough to affect the generation of crystal structures. Nevertheless, the position of the NH_2 hydrogen atoms is clearly seen to affect the calculated lattice energies and, therefore, the ranking of the structures. This is to be expected since these atoms play a major role in the hydrogen bond interactions within the crystal. In the H_{out} set of predictions, the first and second ranked hypothetical crystals with both electrostatic model correspond to the experimental monoclinic structures (Forms III and IV) whereas with H_{in} an unobserved structure was found to be the most stable. Furthermore, we note that the order of stability of Forms III and IV for the two models is reversed depending on the conformation used.

Comparison of the total energy ($E_i^{total} = E_{non-bonded,i} + \Delta E_{bonded}^{QM}$) of identical predicted crystal structures, found within 3 kJ/mol of the plot energy minimum, is shown in figure 6 for H_{in} and H_{out} arrangements. All the structures in this energy window contain either hydrogen bonded dimers or chain motifs. The lines indicate matching crystals with H_{in} and H_{out} molecular structures - solid lines for crystals with hydrogen bond dimers and dashed lines indicating structures containing chains. The changes in relative total energy observed on going from the H_{in} to H_{out} conformation are greater in some structures than in others, and in some cases by values in excess of 8 kJ/mol. These shifts in relative energies seem to indicate that amide pyramidalization has a greater effect on the energy

of structures that involve chains than those based on dimers (see supplementary material for an example).

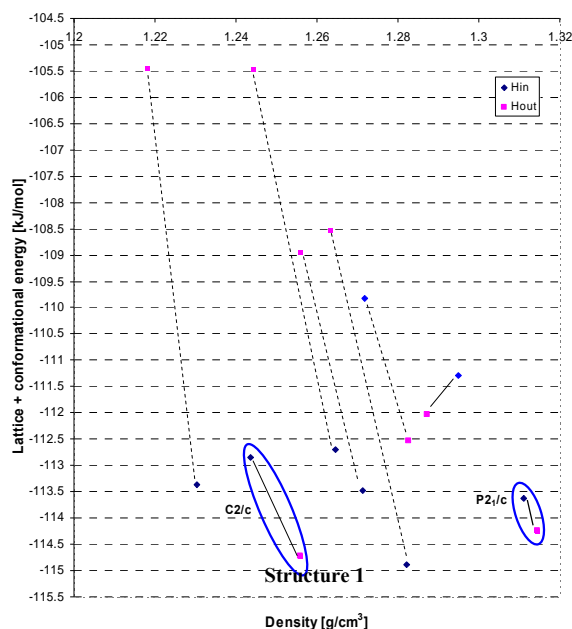


Figure 6. Total energy shift of the same crystal structures when used H_{in} and H_{out} conformations (multipole model). Solid lines match crystal structures containing dimers whereas structures joined by dashed lines contain chain motifs.

3.3. Influence of amide pyramidalization on the relative total energy and ranking of the predicted carbamazepine crystals - refining the calculations

Given that the lattice energy is clearly dependent on the NH_2 geometry, it is necessary to refine the calculations in order to consider the possible deformation of the amide group away from the two gas phase minima, at a range of geometries along the path of nitrogen inversion (fig. 4). We approached the problem by considering a selection of molecular conformations along the H_{in} - H_{out} interconversion pathway and re-minimizing the

previously generated crystal structures after incorporating successively each molecular conformation. Six molecular conformations (Table 1) were selected.

Table 1. Two structural angles that define the conformations used to re-minimize the previously generated crystal structures. *H* is the amide hydrogen that points towards the aromatic ring system whereas *H'* points away from it. *HNCN* is a proper torsion whereas *H'CNH* is improper (out of plane).

Conformation	HNCN	H'CNH
	$/\circ$	$/\circ$
1	-30.0	-136.9
2 (<i>H_{in}</i>)	-24.3	-141.8
3	-15.0	-150.7
4 (planar geometry)	0.0	180.0
5 (<i>H_{out}</i>)	14.6	153.4
6	30.0	138.9

For each previously generated crystal structure, the original molecular model (either *H_{in}* or *H_{out}*) was replaced by each of those listed in Table 1, matching all the atoms but the amide hydrogens. Each structure was then lattice energy minimized using the W99 + atomic point charge model. The relative total energy was calculated as a sum of the relative intermolecular and conformational energies (intermolecular energies calculated with the W99 + point charge model and conformational energies from DFT):

$$\Delta E^{total} = E_i^{total} - E_1^{total} = (E_{non-bonded,i} + \Delta E_{bonded,i}^{QM}) - (E_{non-bonded,1} + \Delta E_{bonded,1}^{QM})$$

and curves of the total energy for each crystal structure with respect to the HNCN dihedral angle were obtained – see Figure 7. In some cases (e.g. *R*-3 or *P*-1), when the lowest energy point was located at the limit of the chosen conformational range, an extra conformation was considered (HNCN = 45.0°, H'CNH = 128.4°).

Each of our chosen conformations was found to provide a minimum energy in one or more of the predicted structures with many of the low energy crystals favouring a molecular geometry away from the gas phase minima. The lattice energy of some structures were greatly affected by a change of conformation, as in the case of the lowest energy crystal (triangles, chain motif) whereas others, such as the *P*2₁/*c* (Form III), show a greater freedom of molecular deformation with low accompanying energetic cost. The importance of molecular flexibility in the ordering of the possible low energy structures is clear; no single molecular geometry gives the correct energetic ordering, so consideration of molecular deformations are necessary for obtaining the correct relative stabilities.

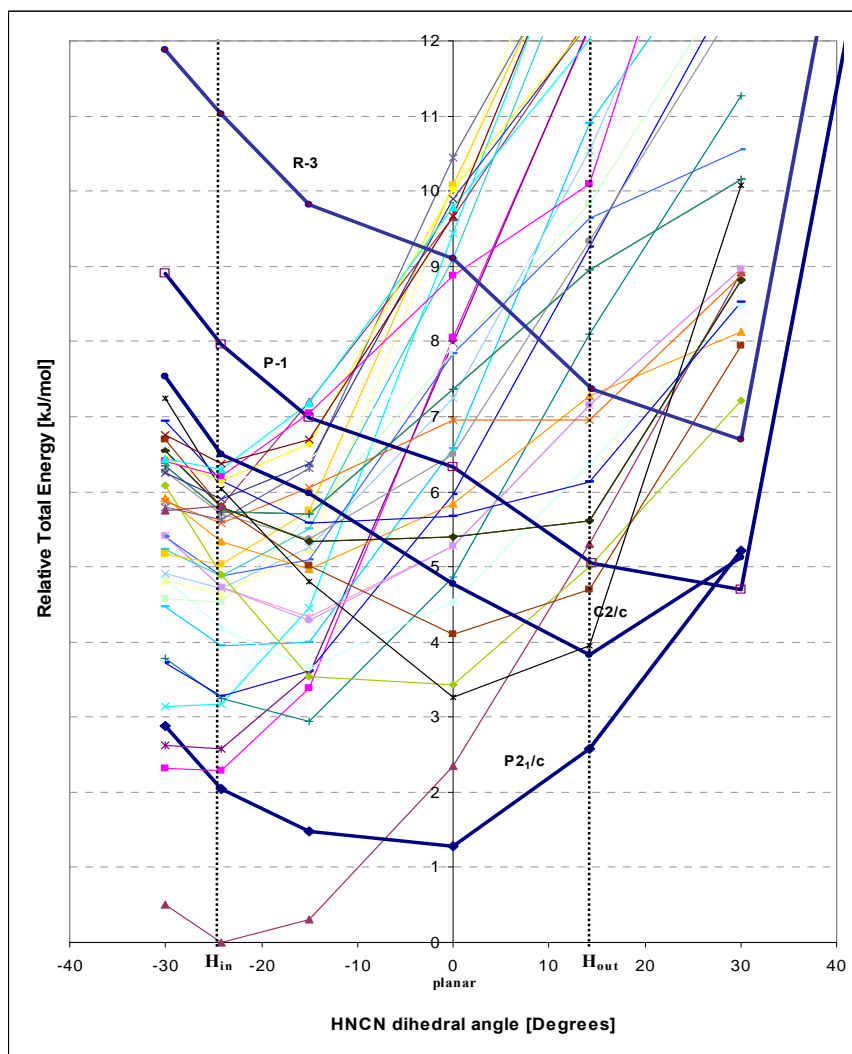


Figure 7. Influence of the molecular amide pyramidalization on the relative total energy of the predicted structures using the atomic point charge model. The thickest lines correspond to the structures matching the observed crystals.

Similar calculations were also performed with the atomic multipole model for a subset of these structures (see supplementary information). The results indicated that the favoured geometry for each structure did not depend on the electrostatic model - when multipoles were used, the crystal energy curves were shifted toward lower energies but with

identical shapes to those in Figure 7. Therefore, as the final step in our calculations, all crystal structures at various minima in Figure 7 were re-minimized with the multipole model (Figure 8).

The multipole description of the electrostatics improves the ranking over the point charge model for the two monoclinic structures (Forms III and IV), which are found as the most stable at $T = 0$ K, together with a third form which has not been experimentally observed (Figure 8). The triclinic and trigonal polymorphs (Forms I and II) were not so well ranked on total energy, being 8.3 and 6.2 kJ/mol above the global minimum; if they had not been observed experimentally the calculations described here would not have reliably predicted either.

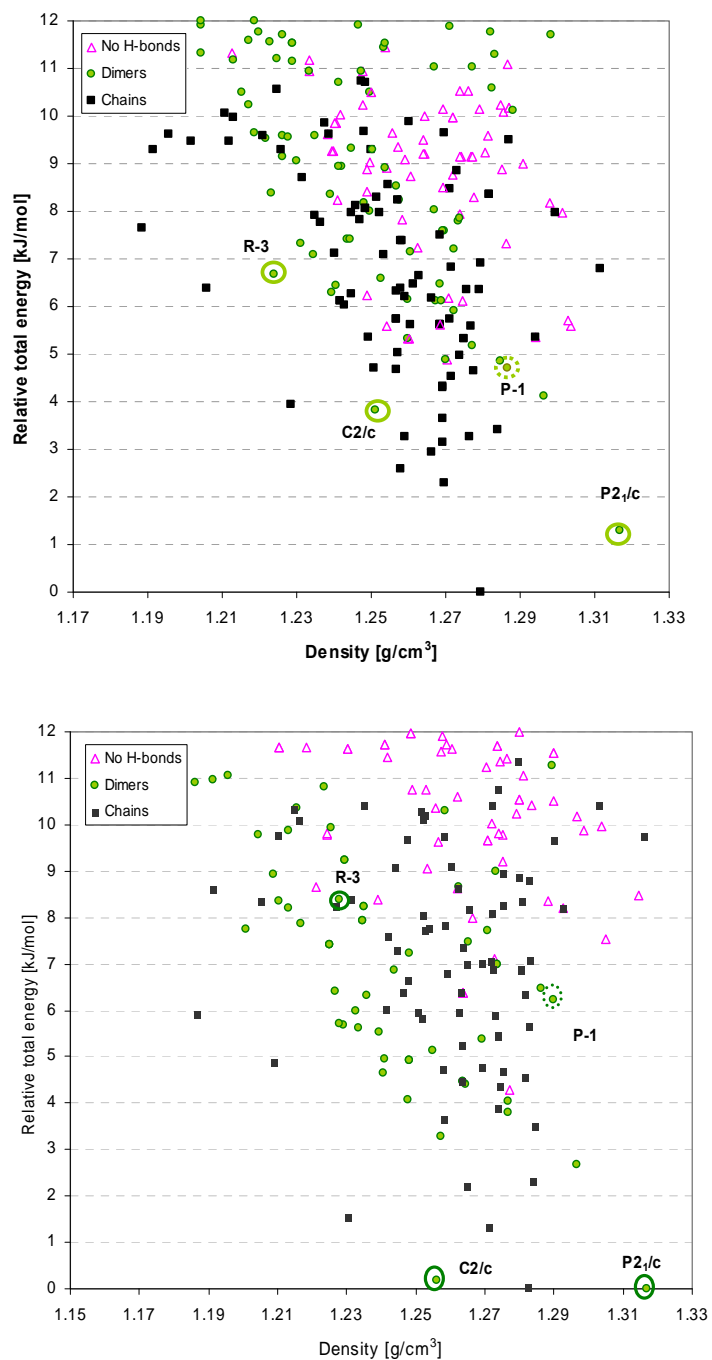


Figure 8. Final carbamazepine crystal structure predictions with atomic point charge (top) and atomic multipole (bottom) models. Each structure exhibits a different conformation of the amide group and corresponds to the energy minima on the total energy surface. Structures are labeled by their hydrogen bond motifs: dimers (circles), chains (squares) and structures with no conventional hydrogen-bonds (triangles).

Table 2. Structural information of the experimental and energy minimized structures, rms deviations and calculated energies.

Space group	Crystal Structure	Molecular conf. at the minimum	ρ /gcm ⁻³	a /Å	b /Å	c /Å	α /°	β /°	γ /°	T /K	Rms dev atom positions	F*	ΔE Conf /kJmol ⁻¹	E Lattice /kJmol ⁻¹	E Conf + Lattice /kJmol ⁻¹
P2 ₁ /c	Exp. CBMZPN01		1.35	7.53	11.15	15.47	90.0	116.17	90.0	298					
	H_{in}		1.31	7.88	11.00	15.74	90.0	118.73	90.0		0.257	40.50	0	-113.62	-113.62
	H_{out}		1.31	7.80	10.98	15.78	90.0	117.97	90.0		0.205	28.09	+1.96	-116.19	-114.23
	Min-conf	Planar HNCN = 0°	1.32	7.86	10.96	15.73	90.0	118.45	90.0		0.241	34.99	+1.86	-116.77	-114.91
C2/c	Exp. CBMZPN12		1.30	26.61	6.93	13.96	90.0	109.70	90.0	158					
	H_{in}		1.24	26.87	7.00	14.46	90.0	111.87	90.0		0.338	30.72	0	-112.84	-112.84
	H_{out}		1.24	25.73	6.77	14.58	90.0	100.03	90.0		0.233	42.07	+1.96	-116.68	-114.72
	Min-conf	H_{out} HNCN = 14.6°	1.24	25.73	6.77	14.58	90.0	100.03	90.0		0.233	42.07	+1.96	-116.68	-114.72
R-3	Exp. CBMZPN03		1.24	35.45	35.45	5.25	90.0	90.0	120.0	298					
	H_{in}		1.21	36.84	36.84	4.98	90.0	90.0	120.0		0.414	73.05	0	-104.74	-104.74
	H_{out}		1.22	36.54	36.54	5.00	90.0	90.0	120.0		0.391	55.02	+1.96	-107.53	-105.57
	Min-conf	HNCN = 30°	1.23	36.51	36.51	4.98	90.0	90.0	120.0		0.414	57.94	+3.18	-109.71	-106.53
P-1	Exp. CBMZPN11		1.34	20.57	5.17	22.24	88.01	84.12	85.19	158					
	H_{in}		1.27	21.46	5.01	23.13	90.4	84.9	86.7		-		0	-108.14	-108.14
	H_{out}		1.27	21.41	5.01	23.20	90.3	85.2	85.3		-		+1.96	-110.22	-108.26
	Min-conf	HNCN = 30°	1.29	21.33	5.00	23.03	87.4	84.9	86.7		-		+3.18	-111.87	-108.69

* “Structural drift factor”³¹ defined as $F = \left(\frac{\Delta\theta}{2}\right)^2 + (10\Delta x)^2 + \left(\frac{100\Delta\alpha}{a}\right)^2 + \left(\frac{100\Delta b}{b}\right)^2 + \left(\frac{100\Delta c}{c}\right)^2 + \Delta\alpha^2 + \Delta\beta^2 + \Delta\gamma^2$, where $\Delta\theta$ = total rigid body rotational displacement after

minimisation and Δx = total rigid-body translational displacement. Values were obtained from the DMAREL calculations.

4. Discussion

4.1 Predicted stability of polymorphs.

We have identified several factors that are important for the modelling and prediction of the crystal structures of carbamazepine. As demonstrated in earlier work³² the form of the electrostatic model can have a great impact on the lattice energy ranking of computer generated hypothetical crystal structures, with the theoretically better founded atomic multipole model typically yielding more reliable results, with the experimentally observed structures closer to the global minimum. Changes to the geometry of the amide group also have an important effect on the relative stabilities of the known and low energy hypothetical structures. Our final model finds two of the known polymorphs (III and IV) amongst the three lowest energy predictions, while the other two (I and II) are ranked above many competing hypothetical polymorphs and more than 6 kJ/mol above the global minimum. It is to be noted that there is a 10% variation in densities (1.19 to 1.32 g/cm³) amongst the set of low energy structures, and this could lead to important differences in thermal contributions to the relative free energies at temperatures above 0 K. Dynamical calculations might, therefore significantly alter the energy ranking of the real and hypothetical crystals.

Form I (P-1) is obtained from the melt of another carbamazepine crystal form and shows many similarities with Form II (R-3). It is known that Form II transforms into Form I upon heating. Our calculations place Form I at 1.73 kJ/mol lower energy than Form II, an energy difference which is within the range of the experimentally observed exothermic Form II \rightarrow I transition enthalpies (-1.46 and -2.34 kJ/mol).^{10,33} Nevertheless, we calculate

Form I to be less stable than Form III (P2₁/c) by 6.2 kJ/mol. This value is much higher than the experimentally observed transition enthalpy (1.30 to 3.60 kJ/mol).^{10,34-38} The Form II → I transition enthalpy can be measured directly with DSC, whereas ΔH for Form III → I is always obtained indirectly, with little agreement in the literature reported values. In comparing the experimental and theoretical data, recognition must be given to possible errors that affect both the measurements and calculations. Experimental accuracy might be poor if multiple processes are involved during the DSC measurements, while the accuracy of the calculations is limited by the use of empirical atom-atom repulsion-dispersion potentials, the static energy minimization approach and the uncertainties in the molecular geometry. Nevertheless, a qualitative agreement in the relative stability of the solid modifications should be expected.³⁹ Our final calculations give the following order of stability at 0K: Form III (P2₁/c) = Form IV (C2/c) > Form I (P-1) > Form II (R-3), in agreement with experimental studies (Form III > Form I > Form II). The only polymorph in disagreement is Form IV which is reported¹⁰ to lie between Forms I and II. Only one experimental investigation of the stability of this polymorph, however, has been reported, with relative stabilities deduced indirectly.

4.3 Molecular conformations in the crystals.

Different degrees of amide pyramidalization were found in the final energy minimized for Forms II, III and IV (Table 2). The final minimum for Form I has the same molecular geometry as Form II, though independent variation of the geometry of the four independent molecules was not allowed. Due to the limitations of positioning H atoms

accurately from X-ray data, our optimized conformations cannot be compared with the existing experimental structures.

Although all four polymorphs contain hydrogen bonding dimers, they exhibit small structural differences in the orientation of the two molecules (Figure 9) - the NCO planes in each molecule are not coplanar. In Form III, the NCO planes are aligned, leading to optimal hydrogen bonding with the planar NH_2 geometry. However, a planar amide group in Form IV gives a non-linear $\text{N-H}\cdots\text{O}=\text{C}$ geometry, so a more effective linear hydrogen bond is obtained by pyramidalization of the amide away from the aromatic rings, towards the oxygen of the second molecule. Such pyramidalization is observed in the benzamide⁴⁰ (BZAMID02) and nicotinamide⁴¹ (NICOAM03) crystal structures, where neutron diffraction data was used to accurately determine the hydrogen atom positions.

Unlike the results for Forms III and IV, the NH_2 geometry optimizations for Forms I and II do not favour linear $\text{N-H}\cdots\text{O}$ bonds. Instead, a lower total energy is obtained with amide pyramidalization away from the carbonyl oxygen of the second molecule; a negative HNCN angle is needed for a linear hydrogen bond, contrary to predicted. Analysis of the energy minimized structure suggests that this counterintuitive NH_2 distortion is a result of stabilizing $\text{CH}\cdots\text{N}$ contacts with the aromatic ring of a third molecule. Indeed, such contacts ($d_{\text{H}\cdots\text{N}}=2.57 \text{ \AA}$) are increasingly considered to be weak hydrogen bonds,⁴² with structure-directing capabilities.

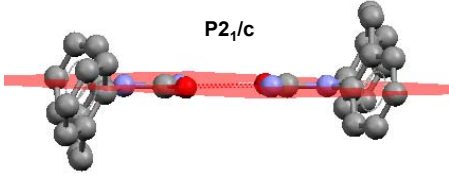
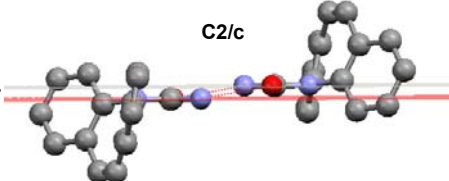
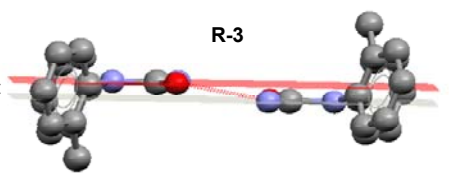
	Experimental distance between NCO planes d/Å	Predicted Conformation HNCN /degrees
 <p>P2₁/c</p>	0	0 - planar
 <p>C2/c</p> <p>d ⇄</p>	0.45	14.6 - H _{out}
 <p>R-3</p> <p>d ⇄</p>	0.58	+30

Figure 9. Geometries of the dimer motifs found in the different crystals, distances between the NCO planes (*d*) and predicted conformations.

4.3 Analysis of the hydrogen bonding and packing of the structures

Although all experimental polymorphs exhibit the same hydrogen bonding arrangement of anti-carboxamide dimers, other stable patterns are present in the predicted structures. The low energy hypothetical crystal structures were analysed and grouped by type of hydrogen bonding motif (dimers, chains and structures lacking typical hydrogen bonds) and the results shown in Figure 8. Non-hydrogen bonded hypothetical crystals are the least stable, though they represent some of the most dense structures, clearly optimisation of hydrogen bonding is likely to lead to less efficient packing. It is interesting to note that hydrogen bond chains are as abundant and stable as dimers, although no structures with chain motifs have been experimentally obtained. Only one of the three most stable

structures exhibits chains, but there are many others more stable than both Forms I and II. The lack of observed chain structures, even after an extensive polymorph screen,¹¹ suggests the importance of molecular recognition effects during crystallization. Dimers pack as a unit and, if they are present in the solution or melt, chain formation would require disruption of this supramolecular synthon during crystal nucleation and growth. While such kinetic factors may be responsible for the lack of observed hydrogen bond chains, we note that the structurally similar 10,11-dihydrocarbamazepine⁴³ (CSD refcode VACTAU01) does crystallise with the amide groups forming H-bonded chains. That such a small perturbation of the molecular structure switches the hydrogen bond motif shows that the dimer motif may not be so robust and further investigation of crystal growth factors, as well as entropy contributions to the crystal energies, are needed.

It is the packing of the dimer unit that differentiates between the known polymorphs. It is well known that aromatic rings in crystals tend to organise themselves in two fundamentally different ways: parallel stacked and herringbone arrangements. In the stacking arrangement, the aromatic rings are aligned parallel ($d_z = 3.1$ to 3.4 Å) and shifted relative to each other ($d_{xy} = 3.6$ to 4.3 Å). In the herringbone interaction, the preferred geometry has angles between the rings in the range of 40 to 90° , with the closest contacts found between carbon and hydrogen atoms.⁴⁴ We therefore searched the hypothetical structures for common patterns of aromatic interactions. The two most abundant are π - π stacking (blue arrows) and edge to face (red arrows) interactions that were classified as offset π stacking and sandwich-herringbone (figure 10);^{45,46} the three

most stable crystals showed sandwich-herringbone packing, while both Forms I and II, as well as most of the other low energy predictions, showed offset π stacking.

Crystal structures are a balance of many types of interactions. Clearly, hydrogen-bonding is very important, but aromatic interactions also appear to be important in this case and differentiate between the few most stable from the metastable crystals. The majority of the molecular surface is non-polar and interactions between these hydrophobic surfaces must contribute importantly to the packing energy, with perhaps as much structure-directing influence as the hydrogen bonds.

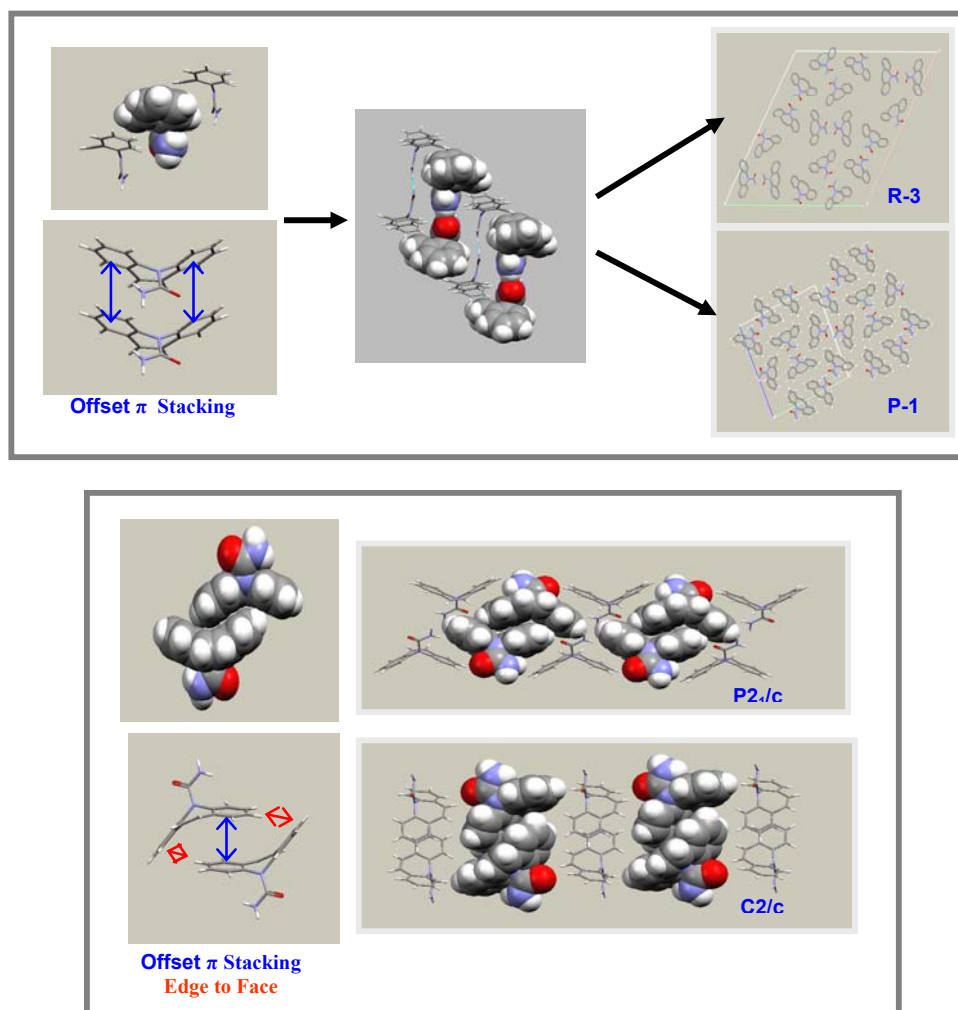


Figure 10. Common aromatic patterns found in the observed crystals: a) offset π stacking packing in the trigonal and triclinic polymorphs and b) sandwich-herringbone packing in both monoclinic crystals as well as the most stable but yet unobserved structure.

5. Conclusions

The low molecular weight pharmaceutical compound carbamazepine exhibits a certain degree of conformational flexibility due to the pyramidalization of the amide group

nitrogen. It has been demonstrated that the influence of this apparently small amount of flexibility on lattice energy calculations plays a significant role in the modelling and crystal structure prediction of the polymorphs of carbamazepine. Many of the low energy predicted crystal structures gave different optimal amide geometries and the relaxation of this molecular degree of freedom has an important effect on the energy ranking of the predicted structures.

The calculated relative stabilities of the polymorphs show qualitative agreement with the available experimental data - only the stability of Form IV was in disagreement with the one experimental study available for this polymorph. The final calculations found the two known monoclinic polymorphs (Forms III and IV) as the most stable possible structures, together with a third yet unobserved crystal that contains the alternate chain hydrogen-bonding motif. The remaining polymorphs (Forms I and II) were found at 6 and 8 kJ/mol above the global minimum, higher in energy than many alternative computer-generated crystal structures; these polymorphs would not have been suggested as likely observed structures by lattice energy calculations. Entropic or crystal growth considerations are suggested therefore as important factors to be considered in any further analysis.

Acknowledgements

We would like to thank Dr Neil Feeder and Dr. Pete Marshall for valuable discussions in this project. We would also thank Prof. Matzger for providing us with the structure of carbamazepine Form IV and Dr. James Chisholm for the COMPACK algorithm

implemented in COSET. Finally, we thank the Pfizer Institute for Pharmaceutical Materials Science for funding.

References

- (1) Bernstein, J. *Polymorphism in Molecular Crystals*; Oxford University Press, 2002.
- (2) Price, S. L. *Adv. Drug Deliv. Rev.* **2004**, *56*, 301-319.
- (3) Chemburkar, S. R.; Bauer, J.; Deming, K.; Spiwek, H.; Patel, K.; Morris, J.; Henry, R.; Spanton, S.; Dziki, W.; Porter, W.; Quick, J.; Bauer, P.; Donaubaue, J.; Narayanan, B. A.; Soldani, M.; Riley, D.; McFarland, K. *Org. Process Res. Dev.* **2000**, *4*, 413-417.
- (4) Lommerse, J. P. M.; Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Mooij, W. T. M.; Price, S. L.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E. *Acta Crystallogr. Sect. B-Struct. Sci.* **2000**, *56*, 697-714.
- (5) Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Dzyabchenko, A.; Erk, P.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Lommerse, J. P. M.; Mooij, W. T. M.; Price, S. L.; Scheraga, H.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E. *Acta Crystallogr. Sect. B-Struct. Sci.* **2002**, *58*, 647-661.
- (6) Day, G. M.; Motherwell, S.; Ammon, H. L.; Boerrigter, S. X. M.; Della Valle, R. G.; Venuti, E.; Dzyabchenko, A.; Dunitz, J. D.; Schweizer, B.; van Eijck, B. P.; Facelli, J. C.; Bazterra, V. E.; Ferraro, M. B.; Hofmann, D. W. M.; Leusen, F. J. J.; Liang, C.; Pantelides, C. C.; Karamertzanis, P. G.; Price, S. L.; Lewis, T. C.; Nowell, H.; Torrisi, A.; Scheraga, H.; Arnautova, Y. A.; Schmidt, M. U.; Verwer, P. **2005**.
- (7) Reboul, J. P.; Cristau, B.; Soyfer, J. C.; Astier, J. P. *Acta Crystallogr. Sect. B-Struct. Sci.* **1981**, *37*, 1844-1848.
- (8) Himes, V. L.; Mighell, A. D.; De Camp, W. H. *Acta Crystallogr. Sect. B-Struct. Sci.* **1981**, *37*, 2242-2245.
- (9) Lowes, M. M. J.; Caira, M. R.; Lötter, A. P.; Van Der Watt, J. G. J. *Pharm. Sci.* **1987**, *76*, 744-752.

- (10) Grzesiak, A. L.; Lang, M. D.; Kim, K.; Matzger, A. J. *J. Pharm. Sci.* **2003**, *92*, 2260-2271.
- (11) Florence, A. J.; Johnston, A.; Price, S. L.; Nowell, H.; Kennedy, A. R.; Shankland, N., 2005.
- (12) Mannfors, B. E.; Mirkin, N. G.; Palmo, K.; Krimm, S. *J. Phys. Chem. A* **2003**, *107*, 1825-1832.
- (13) Mannfors, B.; Palmo, K.; Krimm, S. *J. Mol. Struct.* **2000**, *556*, 1-21.
- (14) Momany, F. A. *Journal of Physical Chemistry* **1978**, *82*, 592.
- (15) Cox, S. R.; Williams, D. E. *J. Comput. Chem.* **1981**, *2*, 304.
- (16) Stone, A. J.; Alderton, M. *Mol. Phys.* **1985**, *56*, 1047-1064.
- (17) MS Modelling, Release 3.0.1, Accelrys Inc., A.: San Diego, 2004.
- (18) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13288-13249.
- (19) Delley, B. *J. Chem. Phys.* **1990**, *92*, 508-517.
- (20) Williams, D. E. *J. Mol. Struct.* **1999**, *486*, 321-347.
- (21) Williams, D. E. *J. Comput. Chem.* **2001**, *22*, 1154-1166.
- (22) Williams, D. E. *J. Comput. Chem.* **2001**, *22*, 1-20.
- (23) Amos, R. D.; with contributions from Alberts, I. L.; Andrews, J. S.; Colwell, S. M.; Handy, N. C.; Jayatilaka, D.; Knowlles, P. J.; Kobayashi, R.; Koga, N.; Laidig, K. E.; Maslen, P. E.; Murray, C. W.; Rice, J. E.; Sanz, J.; Simandiras, E. D.; Stone, A. J.; Su, M. D. CADPAC, version 6.0: 1995.
- (24) Karfunkel, H. R.; Gdanitz, R. J. *J. Comput. Chem.* **1992**, *13*, 1171.
- (25) Karfunkel, H. R.; Leusen, F. J. J. *Speedup* **1992**, *6*, 43.
- (26) Verwer, P.; Leusen, F. J. J. *Rev. Comput. Chem.* **1998**, *12*, 327-365.
- (27) Cerius2, version 4.6, Accelrys Inc., A.: San Diego, 1997.
- (28) Day, G. M.; Chisholm, J.; Shan, N.; Motherwell, W. D. S.; Jones, W. *Cryst. Growth Des.* **2004**, *4*, 1327-1340.
- (29) Price, S. L.; Willock, D. J.; Leslie, M.; Day, G. M. DMAREL, version 3.1: 2001.
- (30) Chisholm, J. A.; Motherwell, S. *J. Appl. Crystallogr.* **2005**, *38*, 228-231.
- (31) Filippini, G.; Gavezzotti, A. *Acta Crystallogr. Sect. B-Struct. Sci.* **1993**, *49*, 868-880.

- (32) Day, G. M.; Motherwell, W. D. S.; Jones, W. *Cryst. Growth Des.* **2005**, *5*, 1023-1033.
- (33) Edwards, A. D.; Shekunov, B. Y.; Forbes, R. T.; Grossmann, J. G.; York, P. *J. Pharm. Sci.* **2001**, *90*, 1106-1114.
- (34) Behme, R. J.; Brooke, D. *J. Pharm. Sci.* **1991**, *80*, 986-990.
- (35) Urakami, K.; Shono, Y.; Higashi, A.; Umemoto, K.; Godo, M. *Chem. Pharm. Bull.* **2002**, *50*, 263-267.
- (36) Park, K.; Evans, J. M. B.; Myerson, A. S. *Cryst. Growth Des.* **2003**, *3*, 991-995.
- (37) Gu, C. H.; Grant, D. J. W. *J. Pharm. Sci.* **2001**, *90*, 1277-1287.
- (38) Ceolin R., T. S., Gardette M.F., Agafonov V.N., Dzyabchenko A.V., Bachet B. *J. Pharm. Sci.* **1997**, *86*, 1062-1065.
- (39) Gavezzotti, A.; Filippini, G. *J. Am. Chem. Soc.* **1995**, *117*, 12299-12305.
- (40) Gao, Q.; Jeffrey, G. A.; Ruble, J. R.; McMullan, R. K. *Acta Crystallogr. Sect. B-Struct. Sci.* **1991**, *47*, 742.
- (41) Miwa, Y.; Mizuno, T.; Tsuchida, K.; Taga, T.; Iwata, Y. *Acta Crystallogr. Sect. B-Struct. Sci.* **1999**, *55*, 78.
- (42) Allen, F. H.; Bird, C. M.; Rowland, R. S. *Acta Crystallogr. Sect. B-Struct. Sci.* **1995**, *B51*, 1068-1081.
- (43) Bandoli, G.; Nicolini, M.; Ongaro, A.; Volpe, G.; Rubello, A. *Journal of Crystallographic Spectroscopic Res.* **1992**, 177.
- (44) Hunter, C. A.; Sanders, J. K. M. *J. Am. Chem. Soc.* **1990**, *112*, 5525-5534.
- (45) Gavezzotti, A. *Acta Crystallogr. Sect. B-Struct. Sci.* **1988**, *44*, 427-434.
- (46) Desiraju, G. R.; Gavezzotti, A. *Acta Crystallogr. Sect. B-Struct. Sci.* **1989**, *45*, 473-482.