Supplementary Material

A) Discussion of results and methodology; each section prepared by the named author(s).

H. L. Ammon

Test VIII…

Determined with PMC (Dzyabchenko, Agafonov, Davydo, 1999) in space group Cc with a model based on a Ci-dimer. Lattice E refinement and structure optimization with WMIN (Busing, 1981).

Test IX…

Problems with finding a 631g* basis set for iodine, required a reparameterization with a 321g** basis set. Attempted unsuccessful solution with MOLPAK (Holden, Du, Ammon, 1993) and WMIN.

Test X…

Solution achieved with ROTPAK (unpublished) and WMIN. Semi-empirical PM3 calculations were used for the intramolecular part of total energy. Initially, the search angle range for the amide C-N was inadequate based on a b3lyp/631g* model of 0 degs. Later, a range of +/- 45 degs gave the correct solution. Following table shows the exptl and ROTPAK/WMIN values.

Angle	Exptl (^{o})	ROTPAK/WMIN (^{o})	Difference (^{o})**
ϕ6-1-N-C	-41.4	-45.4	4.0
ω, 1-N-C-C	173.1	-177.4*	N/A
τ_1, 1-2-C-H	-159.0	-179.5*	N/A
τ_2, 3-4-N-O (cis)	25.8	31.9	-6.1
τ_3, 4-5-N-O (cis)	52.0	48.6	3.4

* Angles were not varied. Values are from original b3lyp/631g* optimized model.
** Difference = Exptl – ROTPAK/WMIN

Test XI…

Solution attempted with MOLPAK/WMIN and a single molecule. Did not attempt anything with Z’ > 1. When the exptl Z’ = 2 became known, a H-bonded dimer used originally to access the N…N intermolecular distance was used to solve the structure.

Busing, W. R. (1981). *WMIN*. Report ORNL-5747. Oak Ridge National Laboratory, Oak Ridge, TN, USA.
Dzyabchenko, A. V., Agafonov, V. & Davydov, V. A. (1999). *J. Phys. Chem. A*, **103**, 2812-2820.
Holden, J. R., Du, Z. & Ammon, H. L. (1993). *J. Comput. Chem.* **14**, 422-437.

Della Valle and Venuti (Xfind/Gaussian98/Wmin/Ionic/Platon)

Raffaele Guido Della Valle & Elisabetta Venuti

Dipartimento di Chimica Fisica e Inorganica and INSTM-UdR Bologna, Universit`
a di Bologna, Viale Risorgimento 4, I-40136 Bologna, Italy

Since our Xfind procedure for crystal structure prediction has been presented elsewhere (Della Valle et al., 2003), only a brief discussion is given here. The crystals are described in terms of rigid molecules interacting through a model potential. We sample several thousands of different structures for each compound under study. Since random sampling often exhibit gaps and clusters of points, which may be wasteful of computer time, we prefer to follow a low-discrepancy Sobol' sequence (Press et al., 1992), which ensures a more uniform, and thus more efficient, sampling (Della Valle et al., 2003). Starting from each structure, we minimize the lattice potential energy by adjusting the cell axes and angles, together with the positions and orientations of the molecules. Finally, we analyze the structures to identify their space group and to recognize minima encountered more than once.

We used this procedure to predict possible stable crystal structures for compounds VIII ($C_3H_4N_2O_2$), IX ($C_{22}H_8I_2O_2$) and XI (C_3H_7N). Compound X cannot be considered as a rigid molecule and was ignored. The lattice potential energy constituted our primary structure selection criterion. For the deepest potential minima, we employed also the Gibbs energy, as a secondary structure confirmation criterion, as discussed below. No use was made of the information on compound VIII made public (Yu & Schwalbe, 2002) before the end of the blind test.

Molecular geometry. The geometry was determined for each isolated molecule with Gaussian98 (Frisch et al., 1998), using the exchange-correlation functional B3LYP (Lee et al., 1988) in combination with a suitable basis set (from STO-3G to 6-31G(d,p), depending on molecule and available computational resources). Compounds IX and XI exhibit point group symmetries C_{2h} (planar) and C_s (bent, with the amminic hydrogen in equatorial conformation), respectively. In the post analysis phase we have found that both computed geometries agree reasonably well with those of the experimental crystal structures.

For compound VIII, with the basis B3LYP/6-31(d,p), we have found two stable conformers with C_1 symmetry, both nearly planar, with the amminic hydrogens in syn or anti positions with respect to the ring. The anti conformer is stabler that the syn one, by $\approx 1.4\times10^{-4}$ kJ/mole, and was selected for our calculations. The two conformers are separated by a planar (C_s symmetry) energy saddle $\approx 1.7\times10^{-2}$ kJ/mole high. The energetic and geometric differences are anyway very small.

The crystal field probably modifies the various relative energies, since the experimental conformer is the syn one.

Potential model. The intermolecular potential was represented by an atom-atom Buckingham model (Born & Huang, 1954, Pertsin & Kitaigorodsky, 1987). Williams parameters for C and H atoms (Williams, 1967) were combined, using standard mixing rules (Pertsin & Kitaigorodsky, 1987), with parameters for nitrogen (Williams & Cox, 1984) (compounds VIII and XI), oxygen (Cox et al., 1981) (compounds VIII and IX) and iodine (Brillante et al., 1997) (compound IX). Coulombic interactions were added through atomic charges derived from the ab-initio electrostatic potential (ESP charges) and evaluated using the Ewald's method (Born & Huang, 1954; Pertsin & Kitaigorodsky, 1987). Due to the notorious problems with the iodine in Gaussian98 (Frisch et al., 1998), the charges for compound IX were transferred from those of the analogous $C_{22}H_8Cl_2O_2$ compound.

Calculations. For each of the three compounds (VIII, IX and XI), we have generated 3000 different triclinic structures with two independent molecules in the primitive unit cell, and then minimized their potential energy with Wmin (Busing & Matsui, 1984). All structures that failed to converge to stable bound states (i.e. to compact structures with negative potential energy) were discarded. Since the triclinic lattice has the lowest possible symmetry, and since no constraints were imposed on the cell parameters, we allowed directly for all possible space groups with two molecules for primitive unit cell (Z = 2). Primitive cells with Z = 1 were also found at this stage, by halving the unit cell whenever the two molecules happened to be translationally equivalent. Primitive cells with Z = 4 were meant to be identified indirectly. In fact, we detected instabilities leading to cell doubling by computing with Ionic (Signorini et al., 1991) the lattice vibrational frequencies the at the boundaries of the Brillouin zone. The frequencies ν_k must be real and positive for all wave-vectors k, as a necessary mathematical condition for local stability (Born & Huang, 1954). Non positive frequencies indicate that the system is at a saddle point, rather than at an energy minimum. In most cases we could still reach a nearby minimum by repeating the energy minimization after properly doubling the unit cell.

Once found the potential minima, we analyzed all structures with Platon (Spek) to identify their space group, a step which occasionally involved cell doubling to obtain a conventional reduced Niggli cell. Then, we searched all structures encountered more than once, by comparing energies, densities, phonon frequencies, axes and angles of the reduced cells. For all compounds, we have probably identified a large part of the minima accessible with our method. While increasing the coverage of the search space, in fact, we approached a saturation plateau where to find new minima was becoming progressively more difficult, because new configurations tended to fall more and more frequently onto previously encountered minima.

Finally, the minima were sorted by potential energy Φ, which was our primary selection criterion. The dependence of the energy rank on the potential model is especially critical for the deepest minima, which often tend to be clustered in a narrow energy range. For this reason, we explored the possibility of using the Gibbs energy G at 0 K as an alternative selection criterion. Starting from the deepest 20 energy minima of each compound, we determined the structures of minimum G at 0 K, by repeating the minimization after adding the zero point lattice vibrational energy $\Sigma h\nu_k /2$ (van Eijck, 2001; Della Valle et al., 1996). We thus found that Φ and G yield matching ranks, with very few exceptions. Nothing can be said about the relative stabilities at non zero temperatures T since we decided to avoid the time consuming procedure necessary to minimize G at $T > 0$.

Search results. Among the 3000 different initial configurations for compound VIII, ≈ 70% failed to converge to stable bound states. After cell reduction and identification of structures found more than once, we were left with 192 distinct minima in an energy range of about 20 kJ/mole, in the following space groups: 159 P 1, 30 P 2_1, 2 C2 and 1 P 1. About 25% of the distinct minima presented non positive frequencies at the boundaries of the Brillouin zone, requiring cell doubling in one or more directions (Z = 4).

For compound IX about 70% of the generated structures converged successfully, leading to a very high number of distinct minima (494) distributed over a range of ≈ 90 kJ/mole in the following space groups: 423 P 1, 43 P 2_1/c, 14 P 1, 9 C2/c, 4 P 2/c and 1 Cc. About 3% of these minima required cell halving (Z = 1). Non positive frequencies and cell doubling occurred for 5% of the distinct minima.

For compound XI about 80% of the initial structures reached a minimum. After cell reduction, we found 321 distinct minima spanning about 10 kJ/mole, in the following space groups: 253 P 1, 18 P 2_1, 17 P 1, 13 Cc, 12 P c, 5 P m, 1 P 2_1/m, 1 Cmc2_1 and 1 C2. Nearly 1% of the distinct minima were subject to cell halving, and about 15% to cell doubling.

Post analysis. Once the actual crystallographic data were disclosed, we analyzed the experimental structures and most of the three best predictions selected by the other groups. These structures were used as the starting points of further lattice energy minimizations, with the same potential model and molecular geometries used in our calculations.

The potential model appears to be at least locally reliable, since starting from the experimental structures for compounds VIII, IX and XI we reached energy minima VIII$_0$, IX$_0$ and XI$_0$ closely resembling the experimental crystals. In fact, the experimental (Prediction2004) cell parameters (a,b,c,β) are well matched by the computed ones, which are (9.381, 12.194, 7.864, 98.52°) for VIII$_0$, (4.085, 21.645, 9.123, 96.73°) for IX$_0$, and (9.210, 9.143, 9.869, 108.36°) for XI$_0$ (lengths in Å). However, none of the these minima correspond to one of our selected predictions which, therefore, are not correct. The potential model is certainly defective on a global scale, since the three post-analysis minima are all above the corresponding deepest minima, by +0.95, +13.60 and +4.61 kJ/mole for compounds VIII, IX and XI, respectively.

For compound VIII the best prediction (van Eijck 1) also converged to minimum VIII$_0$, as expected. The other correct predictions should also converge to VIII$_0$, although we did not check this. The minimum VIII$_0$, never encountered during our search process, is very stable, with an energy between our two deepest minima (energy rank 1 and 2). The experimental structure belongs to space group C2/c, with Z = 8 in the conventional unit cell, i.e. Z = 4 in the primitive cell. Since we did encounter many Z = 4 minima, we should have found the correct minimum. Ruling out the case of pure bad luck, we must have a problem with the search process.

For compound IX the only correct prediction (Day 2) converges to the same minimum IX$_0$ as the experimental structure. In this case we had already encountered IX$_0$, at energy rank 35, although, due to too strict tolerance thresholds in the analysis, we had not recognized the proper space group, P 2$_1$/c with Z = 4.

Finally, for compound XI, the minimum XI$_0$ reached starting from the experimental structure is very shallow, with energy rank above 193. Our calculations had no chance to identify the correct structure, P 2$_1$/c with Z = 8, since no Z = 8 minima were sampled for this compound. Overall, the global properties of the model appear to be acceptable for compound VIII, quite inaccurate for compound IX, and totally wrong for compound XI.

The analysis of the minima reached starting from a subset of the structures selected by the other groups provides a test on the completeness of the search process. For all compounds we find that all tested structures with Z =1 and most of the structures with Z = 2 (especially the deepest ones) converge to minima already identified by our search. Thus, we do have an excellent coverage of the Z = 1 and Z = 2 search space. Instead, for Z = 4 and Z = 8 our coverage is indeed inadequate, as none of the proposed minima converges to any of ours. Clearly, the strategy of sampling Z = 4 and Z = 8 spaces only on the basis of instabilities detected in the Z = 2 space introduces a bias, allowing us to reach only a reduced part of the search space of interest.

References.

Born, M.; Huang, K. (1954). *Dynamical Theory of Crystal Lattices* (Oxford University Press, New York).

Brillante, A.; Della Valle, R. G.; Ulrich, C.; Syassen, K. (1997). J. Chem. Phys. 107, 4628.

Busing, W. R.; Matsui, M. (1984). Acta Cryst. A40, 532.

Cox, S. R.; Hsu, L. Y.; Williams, D. E. (1981). Acta Cryst. A37, 293.

Della Valle, R. G.; Venuti, E.; Brillante, A. (1996). Chem. Phys. 202, 231.

Della Valle, R. G.; Venuti, E.; Brillante, A.; Girlando, A. (2003). J. Chem. Phys. 118, 807-815.

Eijck, B. P. van (2001). J. Comput. Chem. 22, 816.

Frisch, M. J., et al. (1998). *Gaussian 98, Revision A. 7* (Gaussian Inc., Pittsburgh PA).

Lee, C.; Yang, W.; Parr, R. G. (1988). Phys. Rev. B37, 785.

Pertsin, A. J.; Kitaigorodsky, A. I. (1987). *The Atom-Atom Potential Method* (Springer-Verlag, Berlin).

Crystal Structure Prediction (CSP); Acta Cryst. (2004).

Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. (1992). *Numerical Recipes in Fortran* (Cambridge University Press, Cambridge).

Signorini, G. F.; Righini, R.; Schettino, V. (1991). Chem. Phys. 154, 245.

Spek, A. L., PLATON, Utrecht University, The Netherlands; http://www.cryst.chem.uu.nl/platon/.

Williams, D. E. (1967). J. Chem. Phys. 47, 4680.

Williams, D. E.; Cox, S. R. (1984). Acta Cryst. B40, 404.

Yu, F.; Schwalbe, C. H. (2002). Poster presented at the British Crystallographic Association Spring Meeting, Nottingham, March 2002.

G. M. Day

University Chemical Laboratory, Cambridge University

Overview of Methodology

The simulated annealing method implemented in the Cerius2 [1] software (Polymorph Predictor module) has been a successful method for identifying candidate structures in crystal structure prediction. However, an accurate model potential is necessary to calculate the relative energies of possible structures to the accuracy required for crystal structure prediction by energy minimisation. Experience shows [2,3] that the reliability of predictions can be improved significantly by using a better description of electrostatic interactions than the commonly used atomic charge model. Therefore, the candidate structures generated by the Polymorph Predictor package were only used as a starting point and further calculations were performed with the program DMAREL [4], which allows the use of anisotropic atom-atom models, such as multipole expansions for electrostatics and anisotropic repulsion models. DMAREL was also used for elastic constant and rigid molecule phonon calculations which were used to ensure the stability of minimised structures and, for molecule XI, to perform constrained free energy minimisations on the lowest lattice energy structures.

The methodology used for each molecule depended on the difficulties of the molecule and time available for the calculations. Details are given below for each molecule.

Molecule VIII

Many molecules of this size and with about the same rigidity have been considered in a recent evaluation of crystal structure prediction by lattice energy minimisation [2,3]. Therefore, exactly the same approach was used as described in these references, except that the number of space groups considered was increased.

The molecular structure was optimised in the gas phase using density functional theory, as implemented in the DMol3 program. The VWN-BP functional was used, along with the double numerical polarized basis set (DNP), inside a dielectric of dielectric constant 3.5 (which seems typical for an organic molecular crystal). Molecular frequencies were calculated; all were positive, confirming that the structure is a real minimum in energy. This molecular structure was then kept rigid for all further calculations. A simple model potential was used for the crystal structure search to allow time to search the 16 most common space groups with one molecule in the asymmetric unit ($P2_1/c$, P-1, $P2_1$, $P2_12_12_1$, $C2/c$, $Pbca$, $Pbcn$, $Pna2_1$, $Pnma$, $Pca2_1$,$P2_12_12$, Cc, $C2$, $C2/m$, $P2_1/m$, $P2/c$) and the 10 most common space groups with Z'=2 ($P1$, $P2_1/c$, P-1, $P2_1$, $P2_12_12_1$, $C2/c$, $Pbca$, $Pbcn$, $Pna2_1$, $Pnma$). The simulated annealing was repeated in each space group until the list of low energy structures had converged (at least 4 repeats of the search were performed in each space group/Z' combination). The W99 *exp-6* model was chosen for repulsion-dispersion interactions and atomic point charges were fitted to reproduce the PW91/DNP molecular electrostatic potential on a grid of points around the molecule. All structures within 20 kJ/mol of the global minimum were then re-energy minimised in DMAREL, with the atomic charges replaced by a distributed multipole analysis [5] of the calculated wavefunction (this time calculated in CADPAC [6], using the B3P91 functional and 6-31G(d,p) basis set). Multipoles up to hexadecapole were included on each atomic site. The elastic constants and (**k**=0) phonons were calculated for each minimum using DMAREL - if the elastic constant matrix is positive definite and all phonon frequencies positive, then a true minimum has been found. Otherwise, the space group symmetry of the crystal was relaxed, the structure was distorted along the path to lower energy and energy minimisation was continued until a true minimum was found. The structures were then ranked on lattice energy - 69 structures were found within 5 kJ/mol of the global minimum. I intended to try free energy minimisation of these structures, but stopped when the abstract describing the known crystal was discovered. Therefore, the three lowest lattice energy crystal structures were submitted, with no further analysis or re-ranking on other criteria. The lowest lattice energy structure was found to be a very good reproduction of the observed crystal structure.

Table 1. Predicted and experimental crystal structures of molecule VIII.

	space group	a (Å)	b (Å)	c (Å)	β (deg)	density (g/cm^3)	lattice energy (kJ/mol)
1st submitted structure	*C*2/*c* (Z'=1)	8.962	12.287	7.857	102.97	1.577	-89.34
2nd submitted structure	*C*2/*c* (Z'=2)	13.104	12.266	10.514	94.05	1.577	-89.09
3rd submitted structure	$P2_1/c$ (Z'=1)	4.203	12.255	8.382	102.58	1.578	-89.00
observed crystal	*C*2/*c* (Z'=1)	9.354	12.176	7.229	104.59	1.669	-

Molecule XI

The second small rigid molecule was treated in much the same way as molecule VIII - a rigid body search of the same space groups, in Z'=1 and Z'=2, from a DFT optimisation and using the W99 model potential and atomic charges fitted the molecular electrostatic potential. The structures within 20 kJ/mol of the global minimum were re-energy minimised in DMAREL with atomic charges replaced by multipoles, then phonons and elastic constants were used to check stability. However, instead of lattice energy, the calculated free energy was used to rank the stability of the structures. 19 structures were found within 1 kJ/mol of the global minimum in lattice energy. Only **k**=0 phonon calculations are implemented in DMAREL, so each crystal was expanded into a supercell of 16 molecules to include extra k-points (between 2 and 8) in the calculated density of states. The sampling is still crude and more k-point are probably necessary to converge the free energy calculation, but computational time was only available for this size of calculation. The structure of the crystals was minimised within the quasi-harmonic approximation, i.e. the unit cell was allowed to vary to minimise the calculated free energy and for a given set of cell constants, the phonons were calculated in the harmonic approximation. The free energy was calculated from the phonons using the Einstein approximation to extend the phonons across the Brillouin Zone and the Debye approximation for the missing low frequency modes [7]. As a further computational constrain, the unit cell dimensions were not allowed to vary independently, but the ratio of expansion along **a**, **b** and **c** was fixed proportionally to S_{11}, S_{22} and S_{33} of the elastic compliance matrix. In this way, the free energy only had to be calculated as a function of one expansion parameter. All free energise were calculated at 170K, the temperature at which the crystals were grown. This free energy calculation expanded crystals by 6.6% to 7.4% from the lattice energy minimum to T=170K and there was some re-ranking of the order of stability of the candidate structures. While there was significant re-ranking of structures, the lowest 2 in lattice energy remained the lowest in free energy. The third lowest in calculated free energy was originally the 6th lowest in lattice energy.

Table 2. Predicted and experimental crystal structures of molecule XI.

	space group	a (Å)	b (Å)	c (Å)	β (deg)	density (g/cm^3)	lattice energy [a] (kJ/mol)	Helmholtz free energy [b] (kJ/mol)
1st submitted structure	$P2_1/c$ (Z'=1)	5.538	4.729	16.090	93.28	0.901	-40.02	-43.26
2nd submitted structure	$P2_1$ (Z'=1)	5.515	4.784	8.013	93.68	0.899	-39.90	-43.25
3rd submitted structure	$C2/c$ (Z'=2)	20.348	4.458	20.348	117.38	0.926	-39.76	-43.22
Match to experimental structure	$P2_1/c$ (Z'=2)	9.832	8.983	10.054	119.22	0.979	-39.79	- [c]
expt structure lattice energy minimised with symmetry relaxed	*P-1* (Z'=4)	9.830	8.990	10.059	α = 92.41 β = 119.24 γ = 87.52	0.979	-40.17	
free energy minimised expt structure	*P-1* (Z'=4)	10.023	9.141	10.298	α = 89.94 β = 119.31 γ = 92.80	0.923		-43.33
observed crystal	$P2_1/c$ (Z'=2)	9.507	9.122	9.790	117.47	1.007	-	-

[a] Lattice energy at the lattice energy minimised structure.

[b] T = 170K Helmholtz free energy at the free energy minimised structure.

[c] The vibrational contribution to the free energy could not be calculated due to the unstable mode.

The experimentally observed structure was not present amongst the 3 submitted predictions. Once the experimental structure was made available, it was located in the list of structures generated from the search. However, the calculated structure had not retained the observed $P2_1/c$ Z'=2 symmetry - an unstable phonon mode was found and the fully relaxed structure had *P*-1 Z'=4 symmetry (the overlay of the P-1 and $P2_1/c$ structures is shown in Figure 6 of the main text). The unstable $P2_1/c$ Z'=2 structure had been the 4th lowest lattice energy structure, 0.23 kJ/mol above the 1st submitted structure. When this structure was fully relaxed with no symmetry restrictions, it was actually lower in energy than any other structure (0.15 kJ/mol lower than the 1st submitted structure on lattice energy, 0.07 kJ/mol lower on free energy). However, because the energy differences between structures were so small and there was a restriction to Z' ≤ 2 in the blind test, this structure had not been considered for the three submitted guesses. Had there not been a restriction on Z', this structure might have been submitted as the first prediction - the global minimum on lattice energy and T=170K free energy. In retrospect, the *P*-1 Z'=4 structure stayed close enough to the $P2_1/c$ Z'=2 symmetry that it should have been kept as a potential structure. It is interesting to note that several groups observed the same instability in the calculated Z'=2 structure - perhaps this is a real feature of the potential energy surface and a thermal averaging over Z'=4 potential energy minima results in the observed higher symmetry.

Molecule IX

The molecular structure and electrostatic model for molecule IX were generated in the same way as for molecules VIII and XI, except that a smaller basis set had to be used for the CADPAC calculations for the distributed multipoles because of the presence of iodine and size of the molecule (a 3-21G basis set with polarisation functions on carbon and hydrogen atoms was used). I didn't have confidence in any published model potential parameters for iodine and had the feeling that the polar falttening that is observed in the repulsive wall around halogen atoms [8] would be important in describing the packing of this molecule. Thus, I adapted a model potential that was developed for modelling the crystals of chlorinated benzene molecules [9]. This model potential includes anisotropy in the repulsive wall around all atoms through an orientation-dependent term in the exponential of the *exp-6* model:

$$U = \sum_{i \in M, k \in N} \left[A^{\iota\kappa} \exp\left(-\alpha^{\iota\kappa}\left[R_{ik} - \rho\left(\Omega_{ik}\right)\right]\right) - C_6^{\iota\kappa} \big/ R_{ik}^6 + U_{elec}\left(DMA, \Omega_{ik}, R_{ik}^{-n}\; n \le 5\right) \right] \tag{1}$$

where

$$\rho^{\iota\kappa}(\Omega) = \rho_1^{\iota}\left(\mathbf{z}_i \cdot \mathbf{R}_{ik}\right) + \rho_1^{\kappa}\left(-\mathbf{z}_k \cdot \mathbf{R}_{ik}\right) + \rho_2^{\iota}\left(3\left[\mathbf{z}_i \cdot \mathbf{R}_{ik}\right]^2 - 1\right)\big/2 + \rho_2^{\kappa}\left(3\left[\mathbf{z}_k \cdot \mathbf{R}_{ik}\right]^2 - 1\right)\big/2 \tag{2}$$

R_{ik} is the interatomic axis between atoms i and k and the $\mathbf{z}_i$ are atomic z-axes, defined to point along bonds directed outward from the molecule. All parameters for carbon and hydrogen atoms were taken from the chlorobenzenes model and parameters describing oxygen and iodine were empirically derived. Oxygen was treated as spherical. The ρ_1 orientation-dependent terms in (2) were set to zero for iodine interactions and ρ_2 quadrupolar term was chosen to reproduce the ratio of major to minor axes observed in close contacts between iodine atoms in known crystal structures [8] (axial ratio = 0.83). The parameters $A_{\text{O-O}}$, $A_{\text{I-I}}$, $\alpha_{\text{O-O}}$, and $\alpha_{\text{I-I}}$, $C_{6,\text{O-O}}$ and $C_{6,\text{I-I}}$ were adjusted to reproduce the crystal structures of a set of similar molecules (1-iodo-anthraquinone, 1,8-di-iodo-anthroquinone and 1,5-di-iodoanthraquinone, CSD refcodes IONTRQ, IOANTQ and IANTRQ), while geometric combining rules were used for all I-X and O-X interactions. The parameters for this model potential are given in Table 3.

Table 3 The molecule IX anisotropic atom-atom model potential.

Atom types		$A^{\iota\kappa}$ kJ/mol	$B^{\iota\kappa}$ Å^{-1}	$C_6^{\iota\kappa}$ Å^6 kJ/mol	ρ_1^{ι} Å	ρ_1^{κ} Å	ρ_2^{ι} Å	ρ_2^{κ} Å
ι	κ							
I	I	1261494	3.1422	16632.3	0	0	–0.1216	–0.1216
C_I	C_I	28957	3.2131	2146.4	–0.2054	–0.2054	–0.3109	–0.3109
C_H	C_H	107333	3.1936	2146.4	–0.0026	–0.0026	+0.0419	+0.0419
O	O	205074	3.9600	833.5	0	0	0	0

H	H	2220	3.2575	200.0	–0.0449	–0.0449	+0.0036	+0.0036
I	C_I	191126	3.1776	5974.9	0	–0.2054	–0.1216	–0.3109
I	C_H	367967	3.1679	5974.9	0	–0.0026	–0.1216	+0.0419
I	H	52920	3.1998	1823.9	0	–0.0449	–0.1216	+0.0036
I	O	508625	3.5511	3723.3	0	0	–0.1216	0
C_I	C_H	61374	3.2443	2146.4	–0.2054	–0.0026	–0.3109	+0.0419
C_I	H	11254	3.3709	653.7	–0.2054	–0.0449	–0.3109	+0.0036
C_I	O	77061	3.5866	1337.5	–0.2054	0	–0.3109	0
C_H	H	16950	3.2654	653.7	–0.0026	–0.0449	+0.0419	+0.0036
C_H	O	148362	3.5768	1337.5	–0.0026	0	+0.0419	0
O	H	21337	3.6088	408.3	0	–0.0449	0	+0.0036

C_I is a carbon bonded to chlorine, C_H is a carbon bonded to hydrogen. Atomic **z**-axes are defined along the bonds pointing out from the aromatic ring.

Z'=1 or Z'=½ were considered most likely, so no search of Z'=2 was performed. Otherwise the same space groups were searched as for molecules VIII and XI. Z'=½ structures should be found in these Z'=1 searches. The anisotropic model potential could not be used in the initial search, so the intermolecular parameters in the *exp-6* form of the Dreiding force field were used for the initial search (keeping the molecule rigid). The 5000 lowest energy structures from this search were re-minimised in DMAREL using the specifically developed model potential, with DMA electrostatics. There was no time for free-energy minimisation for the predicted crystal structures of this molecule. On lattice energy, there were two low energy structures separated from the rest of the predictions by 1.7 kJ/mol (Figure 1). These were submitted as the first two predictions (Table 4). The next three in lattice energy were almost equi-energetic and the densest of these was submitted as the third guess. When the experimentally determined crystal structure was made available, it was easily matched with the first prediction (the global minimum in lattice energy), with only tiny deviations in the lattice parameters. An overlay of the observed and predicted structures is shown in the main text.

Figure 1. Predictions for molecule IX.

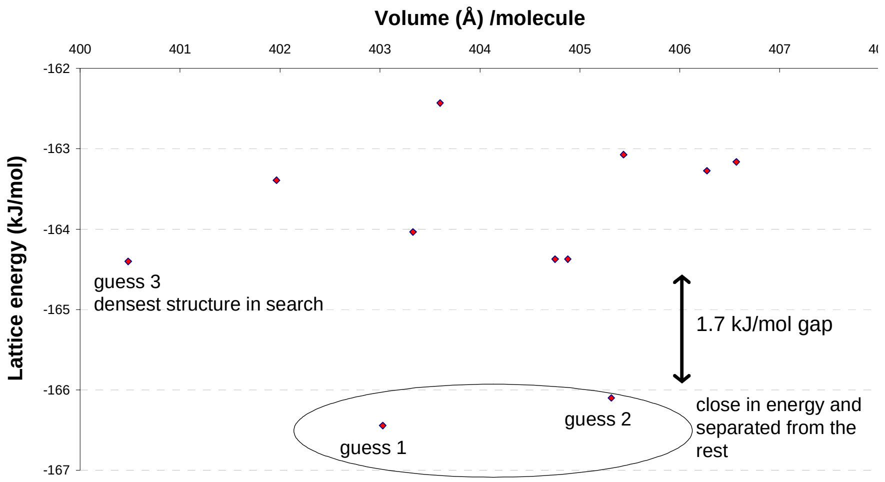


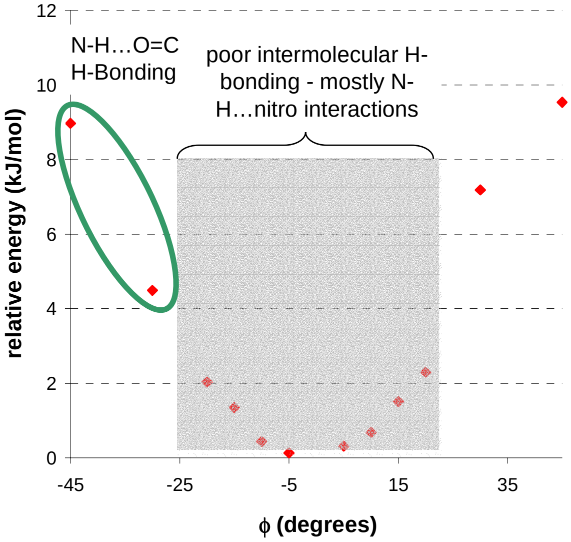
Table 4. Predicted and experimental crystal structures of molecule IX.

	space group	a (Å)	b (Å)	c (Å)	β (deg)	density (g/cm^3)	lattice energy (kJ/mol)
1st submitted structure	$P2_1/c$ (Z'=½)	4.224	20.925	9.306	101.05	2.296	-166.44
2nd submitted structure	$\mathrm{P2_12_12_1}$ (Z'=1)	35.321	4.141	11.100	90	2.284	-166.10
3rd submitted structure	P-1 (Z'=½)	4.204	9.337	10.818	α 77.06 $\beta = 83.96$ $\gamma = 77.49$	2.298	-164.40
observed crystal	$P2_1/c$ (Z'=½)	4.202	20.956	9.276	100.62	2.299	-

Molecule X

For this conformationally flexible molecule, the crystal structure search was performed with a series of conformations, each treated as rigid. Calculations of the conformational energies with varying sets of torsion angles were performed in DMol3. The trans conformation of the amide is much lower in energy than the cis conformation, so only the trans conformation was considered. Two sets of torsions around the nitro substituents seem possible and are almost equi-energetic, so searches were performed with each: (τ_2, τ_3) = (+15°, +75°) and (-60°, -30°). The energy profile around ϕ was scanned in DMol3 (Figure 2) within a dielectric medium of $\varepsilon = 3.5$, a typical value for organic crystals. The minimum, energy conformation is found with ϕ near 0 degrees. In this conformation, the polar hydrogen atom is buried in the neighbouring methyl group, so several conformations were considered with the amide substituent rotated out of the plane of the ring, by up to 45 degrees (searches were performed with ϕ = 0, ±5, ±10, ±15, ±20, ±30, ±45 degrees) to allow more efficient N-H...O=C hydrogen bonding. There is a local minimum conformation with ϕ = 122 degrees. While the energy of this conformation is calculated to be 12.3 kJ/mol higher than the best conformation, a search was performed with that molecular geometry because it seemed more likely to produce good intermolecular hydrogen bonding than with the N-H pointing into the methyl substituent. Due to the expense of searching all of these conformations, only the 9 most likely space groups (*P*2_1/*c*, *P*-1, *P*2_1, *P*$2_12_12_1$, *C*2/*c*, *Pbca*, *Pbcn*, *Pna*2_1, *Pnma*) could be searched, with Z'=1. The W99 *exp-6* parameters were used, with atomic charges fitted to the molecular electrostatic potential. There was not time to re-minimise the predicted crystal structures with a distributed multipole electrostatic model.

Figure 2. DMol calculated energy profile around ϕ for molecule X.



As expected, the crystal structures predicted from the lowest energy conformations showed poor intermolecular hydrogen bonding, with the N-H shielded from neighbouring molecules by the methyl substituent on the benzene ring. The carbonyl group could not get close enough to the polar hydrogen to hydrogen bond and all of the best lattice energy structures with conformations in the range ϕ = -20 to +20 degrees had N-H...nitro close contacts. A search for similar molecules in the Cambridge Structural Database shows that such interactions are not commonly observed. Therefore, all of the predicted structures with ϕ between -20 and +20 degrees were discarded. The predicted crystal structures with the amide twisted further out of the plane of the ring did have the expected N-H...O=C hydrogen bond patterns - the best were found with ϕ = -30 and -45 degrees. The two lowest lattice energy structures were essentially the same with both of these conformations and these were submitted as guesses 1 and 2. Guess 1 has a slightly better packing energy with the lower of these torsion angles, so this structure

was submitted. Guess 2 has a better packing energy with the larger torsion angle, so this structure was submitted; both form N-H ... O=C chains. As a third guess, I submitted the lowest energy structure from the search with the alternate conformation (ϕ = 122 degrees), which also forms N-H...O=C chains. The better hydrogen bonding in these structures does not compensate on total calculated energy (conformational + lattice), but this is viewed as a failure of the balance between inter- and intramolecular energy models. Accurately modelling the balance between conformational and packing energies for flexible molecules seems to be one of the greatest difficulties facing crystal structure prediction.

Table 5. Predicted and experimental crystal structures of molecule X.

	space group	a (Å)	b (Å)	c (Å)	β (deg)	density (g/cm^3)
1st submitted structure	$P2_1/c$	9.035	16.218	9.259	126.95	1.465
2nd submitted structure	$P2_1/c$	4.686	7.718	30.070	91.33	1.461
3rd submitted structure	$P2_1/c$	10.670	10.780	9.809	73.11	1.472
match to observed structure	$P2_1/c$	12.423	5.005	17.687	97.19	1.456
observed crystal	$P2_1/c$	12.569	4.853	17.266	99.16	1.528

The experimentally observed crystal structure is not present in the list of 3 guesses. However, it was located in the extended list of predicted structures - 3rd lowest in energy in the predictions with the $\phi = -45$ degree molecular conformation (the 1st and 2nd lowest in energy with were submitted). The lattice energy of this structure is 2.36 kJ/mol higher than the best calculated structure with this conformation. The chain structure is identical to that found in the 2nd submitted structure, but the chains are packed differently.

References

[1] Cerius2 version 4.6, Accelrys Inc (San Diego) USA.

[2] G.M. Day, J. Chisholm, N. Shan, W. D. S. Motherwell, W. Jones (2004), *Crystal Growth & Design*, 4, p.1327.

[3] G.M. Day, W. D. S. Motherwell, W. Jones (2004), *Crystal Growth & Design*, in press.

[4] Price, S. L., D. J. Willock, M. Leslie and G. M. Day (2001). DMAREL: http://www.ucl.ac.uk/~ucca17p/dmarelmanual/dmarel.html

[5] A. J. Stone and M. Alderton (1985), *Molecular Physics*, 56, p. 1047.

[6] Amos, R. D. with contributions from I. L. Alberts, S. M. Colwell, N. C. Handy, D. Jayatilaka, P. J. Knowles, R. Kobayashi, N. Koga, K. E. Laidig, P. E. Maslen, C. W. Murray, J. E. Rice, J. Sanz, E. D. Simandiras, A. J. Stone, and M.-D. Su (1995). CADPAC6.0: The Cambridge Analytical Derivatives Package, Cambridge.

[7] G.M. Day, S.L. Price, M.L. Leslie (2003), *J. Phys. Chem.*, 107, p.10919.

[8] S.C. Nyburg and C.H. Faerman (1985), *Acta Crys.*, B41, p.274.

[9] G.M. Day and S.L. Price (2003), *J. Amer. Chem. Soc.*, 125, p.16434.

A. Dzyabchenko (PMC)

Molecular structures were obtained by geometry optimization at the 6-31G level for VIII, X, and XI, and by AM1 method for IX. Lattice energy calculations were performed using the PMC program (Dzyabchenko, 2004). The systematic search method as described by Dzyabchenko & Scheraga (2004) was used for the global lattice energy minimization. Molecule X was treated as a flexible system consisted of five rigid fragments (phenyl cycle, peptide unit, two nitro and two methyl groups) connected by virtue of the *link* (bond variation) potentials scaled to preserve the standard molecular geometry. Its starting conformation was taken planar for the basic aromatic part, while both nitro groups were rotated 90° from the aromatic plane. The space groups tried were $P2_1/c$, $Pbca$, P-1, $C2/c$, $P2_12_12_1$, $P2_1$, $Pna2_1$, $Pca2_1$, Cc, $C2$, $Pbcn$ and Pc , Z corresponding to one independent molecule. These structural classes occur with some 75% of the total number of the homomolecular organic crystal structures (Belsky, Zorkaya & Zorky, 1995). The starting Euler angles were changed within the symmetry adapted ranges (Hirshfeld, 1968; Dzyabchenko, 1984) in steps of 30° to give 505, 294 and 168 unique orientations for the triclinic, monoclinic, and orthorhombic space groups, respectively, where the point group C_s was taken to account for equivalent rotations of molecules VIII, X, and XI, and C_{2h} for IX. Each starting center-of-mass coordinate was assigned 0 or ¼ to give the eight points in the unit cell; 0,0,0 was excluded in case of centrosymmetric group. The starting lattice dimensions were confined to a few cell shapes: ‘cubic’, ‘flat’, ‘tall’, ‘wide’, and ‘long’ (Dzyabchenko & Scheraga, 2004) and taken enough large to provide the density which amounts 50-75% of the expected minimized-structure density. Starting cell angles were not different from 90°. Rather, to account for the structures with β exceeding 120° (up to 135°) admitted by the monoclinic groups containing a *c*-glide or a *C*-centering, additional settings $P2_1/n$, Cn, $C2/n$, $I2/c$, and $I2$ have been considered. Further redundances in the starting points associated with the affine normalizer properties of particular space groups (Dzyabchenko, 1983) were eliminated with the program CRYCOM (Dzyabchenko, 1994).

The force field involved the atom-atom potentials of the Lennard-Jones (6-12) type for van der Waals interactions of C, H, O, and N and the hydrogen bonding N-H…O (Momany, et. al, 1974). The anisotropic interactions for iodine were modeled with the 6-12 potential as well, but the iodine interaction centre was placed closer to carbon by 0.4 Å as compared to the normal C-I distance 2.02 Å, while the equilibrium distance of the interaction potential was increased to maintain the experimentally observed crystal structure of hexaiodobenzene (HIBENZ01) as the energy minimum. The electrostatic energy for molecules VIII and IX was calculated with atomic net charges from AM1, whereas for X and XI the bond dipole scheme was introduced as more accurately reproducing the electrostatic potential of a free molecule from initio 6-31G calculation. The torsion potentials of the form $V(\tau) = t_1\sin^2\tau + t_2\sin^2 2\tau$ for the aromatic π-energy of molecule X have been derived from the lattice energy minimizations in 1,2,4-trinitrobenzene (BEYDOY) and 4-hydroxyacetanilide (Boldyreva, *et al*., 2000), in which various sets of constants t_1 and t_2 were tried to obtain the best agreement of the observed and calculated structures.

The minimized structures were sorted by energy and lists of unique structures were prepared using CRYCOM. The structures ranked 1-3 by their lattice energies were selected as crystal structure predictions.

Results

Molecule VIII. Structure predictions for this molecule were not submitted because the experimental structure has been disclosed. An energy minimum relevant to the experimental structure was available in the list of generated structures, however, it was of a high rank (83) being 6 kJ/mol above the global minimum. Its structural parameters showed deviations up to 7% in the lattice constants, the molecule was shifted by 0.1 Å and rotated 7° from experimental position/orientation to give the RSMSD 0.15Å.

Molecule IX. The energy minimum relevant for the experimental structure was 42 kJ/mol (!) above the global minimum and had energy rank 62. The cell parameter deviations were 5, 1, and 7% for *a*, *b*, and *c*, respectively, and 5° for β. The molecule is rotated 7° from the experimental orientation, RMSD 0.28Å. Remarkably, omitting the electrostatic potential with the van der Waals force field unchanged improves the ranking radically: the minimized experimental structure turns out of rank 5 (rank 1 in the $P2_1/c$ list), only 2.0 kJ/mol higher than the global minimum, while no significant change in the the parameters deviations.

Molecule X. Two structures, ranked 35 and 42 (ΔE = 6.6 and 7.3 kJ/mol with respect to the global minimum) in the extended submission list, have been recognized similar to the experimental structure. Their cell parameters deviate from the experimental structure up to 7.4 and 7.0%, the net molecular translations and rotation angles calculated by least squares fitting to the experimental structure are 0.38 and 0.17 Å and 8.6 and 4.3°, the RMSD's of the atoms apart hydrogens are 0.38 and 0.27 Å for the guesses 35 and 36, respectively. The molecular conformations are pretty close to the experimental with regard to the central torsion angle φ: deviations –3 and –7°; the nitro groups are rotated away from the experimental positions –14 and –1° for τ_2 and 23 and 8° for τ_3 while for both methyl groups the predicted orientations are incorrect: the respective deviations are –47 and 49° in τ_1 and –57 and 41° in ψ.

Molecule XI. The experimental structure belongs to a structural class with two independent molecules, which was not considered in this study. Anyway, it had no chance to be predicted with our force field because the minimized experimental structure was 14.6 kJ/mol above the lowest energy structure found, and far beyond of our list of one hundred lowest energy structures. The cell parameter deviations were pretty small (within 1% and 3°), however, both independent molecules of the minimized structure were markedly shifted (0.41 and 0.23Å) and rotated (16.7 and 14.8°), with the overall RMSD's 0.482 and 0.352Å, respectively.

Once again, we come to conclusion, that the force fields at our disposal yet far from being accurate enough for safe selection of the correct polymorph out of many alternatives. In the course of this work we made attempts to improve the electrostatic part of the potential, which severely influence the energy ranking rather than geometrical structure details of a particular polymorph. Thus, it was found that the electrostatic potentials of isolated molecules calculated by the ab initio method can be approximated much better by the bond dipoles rather than by net atomic charges of the molecule. Further work should be done, however, on balancing the van der Waals and electrostatic components of the total energy by modeling geometry and physical properties of the known structures.

Acknowledgements

We thank Prof. H. Scheraga for computer resources of his laboratory.

References

Belsky, V. K., Zorkaya, O. N. & Zorky, P. M. (1995). *Acta Cryst.* **A51**, 473-481.

Boldyreva, E. V., Shakhtshneider, T. P., Vasilchenko, M. A., Ahsbahs, H. & Uchtmann, H. (2000). *Acta Cryst.*, **B56**, 299-309.

Dzyabchenko, A.V. (1983). *Acta Cryst.* **A39**, 941-946.

Dzyabchenko, A.V. (1984). *J. Struct. Chem.* **25**, 416-420.

Dzyabchenko, A.V. (1994). *Acta Cryst.*, **B50**, 414-425.

Dzyabchenko, A.V. (2004). *PMC*, Version 2004. Karpov Institute of Physical Chemistry, Moscow.

Dzyabchenko A. & Scheraga H.A. (2004). *Acta Cryst.* B**60**, 228-237.

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R. Zakrzewski, V. G., Montgomery, J. A., Jr., Stratmann, R. E., Burant, J. C., Dapprich, S., Millam, J. M., Daniels, A. D. , Kudin, K. N., Strain, M. C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson, G. A., Ayala, P. Y., Cui, Q., Morokuma, K., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Cioslowski, J., Ortiz, J. V., Baboul, A. G., Stefanov, B. B., Liu, G. , Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong, M. W., Andres, J. L., Gonzalez, C., Head-Gordon, M., Replogle, E. S. & Pople, J. A. (1998). *Gaussian 98*, Revision A.7. Gaussian, Inc., Pittsburgh PA, 1998.

Hirshfeld, F. L. (1968). *Acta Cryst.* **A24**, 301-311.

Momany F.A., Carruthers L.M., McGuire R.F. & Scheraga H.A. (1974). *J. Phys. Chem.* **78**, 1595-1620.

Van Eijck (UPACK, XTINKER)

Force fields

Models of the free molecules were constructed by geometry optimization of suitable molecular fragments at the 6-31G* level, combined with average geometrical parameters from the CSD database (Allen and Kennard, 1993). Force field parameters for bond distances and bond angles were adjusted to reproduce these geometries, and reasonable values were guessed for the corresponding force constants. Nonbonded parameters were taken from two different force fields, using Lennard-Jones and Buckingham parameters respectively (van Eijck, 2002a). Charges for compounds VIII, X and XI were obtained from ESP fittings on 6-31G** wave functions for the complete molecules. For compound IX the charges were taken as in the standard force fields, augmented by values of ± 0.45 e for the atoms of the CO group.

All torsional angles involving sp^2 C atoms were restrained to planarity with aid of a harmonic potential. The same was done for the torsional angles in the five-membered ring in compound VIII, as both a database search and a 6-31G* ab-initio geometry optimization suggested that this ring should be essentially planar. In principle, compound X has six torsional degrees of freedom, which were modeled by twofold torsional potentials. A database search showed that ω is always in the trans conformation; this was confirmed by an energy difference of 14 kJ/mol between two 6-31G* geometry optimizations. Furthermore, $V(\tau_1)$ and $V(\psi)$ were set to zero, which leaves two dihedral force constants to be determined. Energy minimization of related crystal structures (see below) gave best results for $V(\tau_2) = V(\tau_3) = -2.4$ kcal/mol and $V(\varphi) = -4.0$ kcal/mol.

Methods

The random search technique implemented in the UPACK program package (van Eijck and Kroon, 2000) was used. The three essential torsional degrees of freedom in compound X were set to random values in each starting structure. For each compound 5000 structures with one independent molecule in the asymmetric unit (G = 1) were created in the thirteen space groups $P2_1/c$, P-1, $P2_12_12_1$, $P2_1$, Pbca, C2/c, $Pna2_1$, Cc, $Pca2_1$, C2, P1, Pbcn and Pc. Additionally, for compounds VIII, IX and XI 10000 structures with two independent molecules in the asymmetric unit (G = 2) were generated in the five space groups $P2_1/c$, P-1, $P2_1$, $P2_12_12_1$ and P1. (In our notation G is the number of independent molecules allowed in the structure generation, and Z" is the number of independent molecules that actually emerges - so Z" is often less than G (van Eijck, 2002b).) The generated structures were subjected to a preliminary energy minimization with fully flexible molecules. Equivalent structures were removed by clustering, after which the energy minimization was continued for structures within an energy window of 30 kJ/mol. A second clustering delivered the final lists of possible structures. Some more structures became equivalent after expansion to P1.

All these calculations were done with the Lennard-Jones potential to ensure that short distances remained always repulsive. (For compound IX the initial structure generation was done with a united-atom potential for CH groups to save computer time - even then the calculations for G = 2 were excessively tedious.)

The resulting lists of structures were used for comparison with other force fields, notably the Buckingham potential described above. For compounds IX and X that was the only further calculation. For the small molecules VIII and XI it was feasible to replace the point charge model by an anisotropic charge model, using ideas developed by Price and Stone. The atomic multipoles were obtained by fitting to the electrostatic potential of an SCF/DZ($2d^O$) wave function (Mooij et al, 1999a) , calculated once for a free molecule with optimized 6-31G* geometry. The intramolecular force field was still the Lennard-Jones potential described above, but intermolecular dispersion and attraction was modeled by a Buckingham potential taken from the work of Coombes and Price (1996) and Beyer and Price (2000). This approach will be referred to as MP/FF (multipole/force field), as implemented by Mooij et al (1999b) into the XTINKER package.

These two compounds are fairly rigid, so the intramolecular geometry should not be very important. Nevertheless, better crystal packings may be enabled by allowing intramolecular deformations which require only a small investment in energy. Therefore, a computationally more expensive approach was also undertaken for the most promising compounds. Here the intramolecular geometry was optimized in the field of the crystal environment at the 6-31G* level, and the atomic multipoles were recalculated for every individual structure (van Eijck et al, 2001a): the multipole/ab initio (MP/AI) approach.

Energies thus obtained should be accurate enough to make the correction to free energies worth while (van Eijck, 2001; van Eijck et al, 2001b). However, for nearly planar molecules like compounds VIII and XI we found that the results are sensitive to details of the intramolecular force constants. Therefore, such corrections were not attempted in this work.

Preliminary studies

In order to test and improve the force field, hypothetical structures were generated for related molecules found in the database. For compound VIII, BEPNIT and ADUQOF have the same ring fragment. Possible structures for these compounds were generated in the experimental space group ($P2_12_12_1$) as well as in $P2_1/c$. The Lennard-Jones force field gave values of over 6 kJ/mol for ΔE, which is defined as the energy difference between the energy-minimized experimental structure and the global energy minimum. The Buckingham force field performed already better, and excellent results were obtained when the point charge model was replaced by the MP/FF method. Here the experimental structure ranked first for both compounds, assuming the experimentally found CCCC trans structure for ADUQOF. Including gauche structures, with their relative energy estimated from a 6-31G* calculation, the experimental structure ranked second at ΔE = 2.2 kJ/mol. For compound IX no useful analoga were found, and compound XI was small enough to be studied exclusively by ab initio methods. For compound X the uncertainty in the force field is the largest, and here eleven analoga were studied: BEYDOY, FULYEQ, HNOBEN, ZZZFYW for the nitro groups, ACANIL, ACTOLD, DIXDIY, UGUJIP, UGUHEJ, WIVYIJ and LABJED for the acetanilide group. For each compound, space group $P2_1/c$ with G = 1 was investigated, augmented with the experimental space group if different. Charges were taken from 6-31G** calculations, and averaged for atoms that were equivalent, or could become so after internal rotation. The largest energy differences ΔE with respect to the global minimum were 9.8 kJ/mol for the Lennard-Jones potential and 4.8 kJ/mol for the Buckingham one, so the latter was preferred in the actual work. Note that the two torsional constants mentioned above were adjusted for optimum correspondence between observed and calculated cell parameters and molecular positions, as well as for the lowest possible average ΔE. Yet the results are not encouraging: although the ΔE values are no worse than may be expected, there are too many structures to obtain anything like a reliable structure prediction.

Structure predictions and post-analysis

First a general note regarding our lists of hypothetical structures: the space group reported in the CIF files is the space group in which the structure was generated. Especially for compound IX, which tends to lie on an inversion centre, the actual space group may be of higher symmetry. We did try, however, to report the correct number of independent molecules.

Compound VIII. Structures within a Lennard-Jones energy window of 10 kJ/mol were recalculated using the MP/FF method. Unfortunately, while this was in progress, Desiraju communicated that the structure of compound VIII had already been published by Yu and Schwalbe (2002). After completion of the calculation it was found that this experimental C2/c structure ranked first in our list. A limited MP/AI calculation confirmed this result. Although this satisfactory result cannot be properly considered as found in a blind test, we submitted the MP/FF list for comparison with other work. The RMS deviation with respect to the Lennard-Jones energies was 4.0 kJ/mol (maximum 9.5 kJ/mol), so we are confident that all relevant structures are present. The first structure with Z"= 2 has a relative energy of 2.1 kJ/mol.

Compound IX. Here the structures were recalculated with the Buckingham potential, and these results were considered as final. The correspondence between the two force fields is not too good, the RMS deviation being 3.6 kJ/mol (maximum 9.0). Fortunately, however, the two global energy minima correspond to the same structure. This structure has a molecule on an inversion center in space group P-1. So it is necessarily a layer structure, in contrast to the next low-energy structures. However, many of these also have the molecule on a center of symmetry, including the $P2_1/c$ structures submitted as second and third possibilities. The first structure with Z"= 2 has a relative energy of 3.2 kJ/mol.

For this compound the results are really disappointing: not only is $\Delta E = 10.0$ kJ/mol unexpectedly high, but, even worse, the experimental structure was not even present in the list of candidate structures. Such a failure is a rare event for UPACK. It could be traced to the use of a united-atom force field in the first phase of the structure generation. The experimental structure has $\Delta E = 21$ kJ/mol in that force field, so it had been discarded in an early stage. With this hindsight, we repeated the structure generation in $P2_1$ using the all-atom force field and found the experimental structure easily.

Compound X. In view of the three additional degrees of freedom in this compound, it was thought that a structure with Z" > 1 would not be predictable anyhow, so the calculations were performed only with G = 1. Again, the final submission was selected on the basis of the Buckingham force field, which had been seen to be more successful for the eleven analoga discussed above. The low-energy structures differ completely in their three relevant torsional angles.

Again the results are disappointing: $\Delta E = 7.1$ kJ/mol at rank 135 is much worse than for all the analoga which were used to optimize the force field.

Compound XI. When compound VIII was found to be already experimentally known, compound XI was presented as a substitute. A 6-31G* energy minimization indicated a folded structure. Since it was not immediately obvious whether the nitrogen-bonded hydrogen atom would point towards the CCC plane (trans) or away from it (cis), both forms were allowed in the structure generation. The energy differences are markedly smaller than for the first three compounds. There were 177 Z" = 1 structures within 8 kJ/mole and many more with Z" = 2. The molecule was small enough to allow MP/AI calculations for the former ones. It was immediately found that all cis structures were relatively high in energy. Some other structures became equivalent after this cycle of energy minimization, leaving finally 49 trans Z" = 1 structures. For the Z" = 2 structures, a range of 4 kJ/mol contained already 219 trans structures, which was more or less a practical limit for the computations. After clustering, 174 were left for a total of 223 structures. The RMS deviations with both empirical force fields was 1.4 kJ/mol (maximum 6.7). Assuming (erroneously!) that observable polymorphs tend to have one independent molecule in the unit cell, we discarded the Z"= 2 structures from the selection of the three structures to be proposed.

The experimental structure, with two independent molecules, was found at rank 30 with ΔE = 1.0 kJ/mol. Although not good enough to score as a success in the blind test, this is about all that can be expected from the MP/AI potential. We noted that the preceding 29 structures all have one short axis, in sharp contrast with the observed unit cell. Layer-type structures are evidently unrealistically favoured.

References

Allen, F. H. & Kennard, O. (1993). Chemical Design Automation News, 8, 31-37.

Beyer, T. & Price, S. L. (2000). J. Phys. Chem., B104, 2647-2655.

Coombes, D. S., Price, S. L., Willock, D. J., & Leslie, M. (1996). J. Phys. Chem., 100, 7352-7360.

Mooij, W. T. M., van Duijneveldt, F. B., van Duijneveldt - van de Rijdt, J. G. C. M., & van Eijck, B. P. (1999a). J. Phys. Chem., A103, 9872-9882.

Mooij, W. T. M., van Eijck, B. P., & Kroon, J. (1999b). J. Phys. Chem., A103, 9883-9890.

van Eijck, B. P. (2001). J. Comput. Chem., 22, 816-826.

van Eijck, B. P. (2002a). Phys. Chem. Chem. Phys., 4, 4789-4794.

van Eijck, B. P. (2002b). J. Comput. Chem., 23, 456-462.

van Eijck, B. P. & Kroon, J. (2000). Acta Cryst., B56, 535-542.

van Eijck, B. P., Mooij, W. T. M., & Kroon, J. (2001a). J. Comput. Chem., 22, 805-815.

van Eijck, B. P., Mooij, W. T. M., & Kroon, J. (2001b). J. Phys. Chem., B105, 10573-10578.

Yu, F. & Schwalbe, C. H. (2002). BCA spring meeting, Nottingham, England.

Julio C. Facelli, Victor E. Bazterra[1] and Marta B. Ferraro[a]

MGAC (Modified Genetic Algorithm for Crystals and Clusters)
Center for High Performance Computing
University of Utah, Salt Lake City, UT 84112, US

Genetic Algorithms (GAs) are a family of search techniques rooted on the ideas of Darwinian evolution.[1, 2] Operators analogue to crossover, mutation and natural selection are employed to perform a search able to explore and learn the multidimensional parameter space and to determine which regions of the space provide good solutions to a problem. To improve their convergence GAs are commonly coupled with local optimisations at each generation. This combined approach has been implemented in our program MGAC (Modified Genetic Algorithm for Crystal and Clusters).[3-6]

Crystal structure coding

When using GAs for the prediction of crystal structures, the structures have to be encoded in a “genome” that can be manipulated by the genetic operators as well as used to calculate the energy of the crystal. For organic crystals the molecular geometries are highly constrained by strong covalent bonds, this allows a reduction of the number of parameters for which a global search is necessary. It is possible to consider that all internal bond lengths and bond angles are fixed during the GA search, performing only local optimisations of these parameters. Because the rotational barriers around single covalent bonds are comparable to the intermolecular energies, these dihedral angles should be included both in the GA driven global and local optimisations. Therefore, for the n-molecules in the asymmetric cell the MGAC uses their molecule center of mass, $\{\mathbf{R}_1,...,\mathbf{R}_n\}$, its orientation $\{\mathbf{\Theta}_1,...,\mathbf{\Theta}_n\}$ and the values of the relevant dihedral angles $\{\mathbf{\Phi}_1,...,\mathbf{\Phi}_n\}$ as the parameters, describing the molecular coordinates, included in the GA driven global optimisations. In addition it is necessary to specify the space group and lattice parameters $\{a,b,c,\alpha,\beta,\gamma\}$ to define the crystal. In the MGAC program the space group is fixed and only the lattice angles are considered as independent parameters in the GA optimisation. The rest of the lattice parameters are determined as follows: given the molecular coordinates and the lattice angles that define the crystal system, the molecules are translated using the space group operations and the values of a,b,c chosen to define the smaller asymmetric cell that encloses the set n-molecules, the “Fit cell”. To reduce the possibility of producing crystals with very short inter-molecular distances, in this work it was chosen an asymmetric cell 8% larger than smallest possible cell. This procedure is exemplified in Figure 1. Once a crystal has been built using this procedure, its structure undergoes a full local minimization varying all the inter- and intra- molecular and crystal parameters. This last procedure allows that the effects of molecular interactions be included in the local refinement of the entire crystal structure.

---

[1] Also Departamento de Fisica, FCEyN, Universidad de Buenos Aires and CONICET, Argentina.

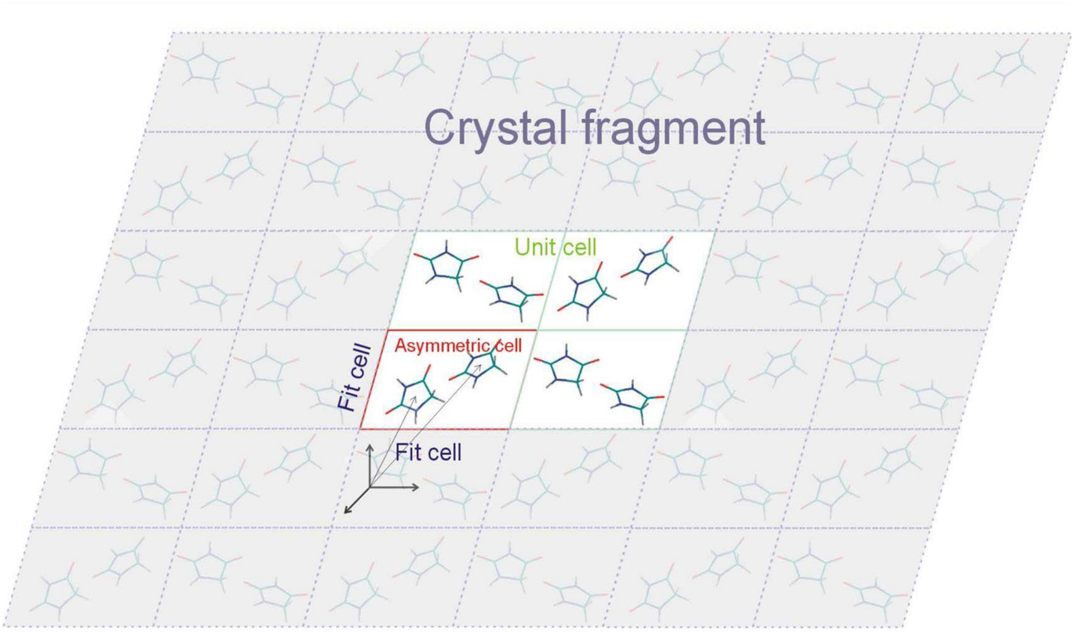


Figure 1: Use of the “Fit Cell” to derive the lattice parameters *a,b* and *c*.

Genetic operators and evolution

The crystal structure encoding described in the last section, except for the addition of the lattice parameters α, β, γ, resembles the one used for a molecular cluster. Note, that the other lattice parameters, *a, b, c,* are derived as described above and they are not included in the global minimization driven by the GA. The MGAC scheme for molecular cluster optimisation previously proposed by Niesse, Mayne, and White[7, 8] was applied in this work. This scheme uses one-point-crossover, two-point-crossover, N-point-crossover, uniform-crossover, arithmetic-crossover, inversion-crossover, geometric-crossover and Gaussian mutation. These operators were implemented in MGAC under the condition that they only act on the molecular coordinates. For α, β, γ only the arithmetic-crossover, geometric-crossover and Gaussian mutation operators are used, because they are the only ones that preserve the crystal system. During the evolution, when these operators are used to construct new candidates for crystal structures, a new set of a, b, c are chosen to fit the new structures in their corresponding asymmetric cells. This is followed by the verification that α, β, γ effectively define a set of linear independent lattice vectors and that the resulting crystal energies are under a certain energy cut-off limit to avoid structures with unphysical intermolecular distances. Those structures that do not fulfil these conditions are eliminated and replaced by entire new randomly chosen ones. After each GA evolution all the new structures are relaxed by a local optimisation within the desired space group, creating a new set of candidate solutions that compete with the existing solutions in the population. This competition is implemented by combining the new solutions with the old solutions in a larger population from which the worst ones are eliminated until the number of candidates equals the desired population size. This procedure is repeated and ends when the diversity of the population reaches the desired threshold or a predetermined number of generations have been completed.

For the work presented in this paper the MGAC calculations were done using 30 individuals, evolving for 60 generations with a mutation probability of 0.1 and a replacement factor of 0.5. For molecules VIII, X and XI calculations were performed twice for each of the most common fourteen space groups: P1, P-1, $P2_1$, C2, Pc, Cc, P21/c, C2/c, $P2_12_12_1$, $Pca2_1$, $Pna2_1$, Pbcn, Pbca and Pnma with one and two molecules per asymmetric unit cell, respectively. For molecule X the dihedral angles were not allowed to vary due to a typo in the input file ω was incorrectly set to 0°. No calculations were attempted for molecule IX.

Energy evaluation and local optimisation

MGAC uses the CHARMM program [9, 10] for the energy evaluation and local optimisation of the crystal structures. Within the CHARMM program the GAFF (Generic Amber Force Field)[11] was used in all the calculations with atomic charges determined with the RESP (Restrained Electrostatic Potential)[12] method using the Gaussian program[13] with the B3LYP exchange correlation functional and 6-31G* basis sets.

References

[1] D. E. Goldberg, *Genetic Algorithms in Search, Optimisation and Machine Learning* (Addison-Wesley, New York, 1989).

[2] R. L. Johnson and C. Robers, in *Soft Computing Approaches in Chemistry*, edited by H. M. Cartwright and L. M. Sztandera (Springer-Verlag, Heidelberg, 2003), Vol. 120, p. 161.

[3] V. E. Bazterra, M. B. Ferraro, and J. C. Facelli, J. Chem. Phys. 116, 5984 (2002).

[4] V. E. Bazterra, M. B. Ferraro, and J. C. Facelli, J. Chem. Phys. 116, 5992 (2002).

[5] V. E. Bazterra, O. Oña, M. C. Caputo, et al., Phys. Rev. A 69, 053202 (2004).

[6] V. E. Bazterra, M. B. Ferraro, and J. C. Facelli, Int. J. Quantum Chem. 96, 312 (2004).

[7] J. A. Niesse and H. R. Mayne, Journal of Computational Chemistry 18, 1233 (1996).

[8] R. P. White, J. A. Niesse, and H. R. Mayne, Journal of Chemical Physics 108, 2208 (1998).

[9] A. D. MacKerell, J. Brooks, B. , C. L. Brooks III, et al., in *The Encyclopedia of Computational Chemistry*, edited by P. v. R. S. e. al. (John Wiley & Sons, Chichester, 1998), p. 271.

[10] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, et al., J. Comp. Chem. 4, 187 (1983).

[11] J. Wang, R. M. Wolf, J. W. Caldwell, et al., Journal of Computational Chemistry 25, 1157 (2004).

[12] C. I. Bayly, P. Cieplak, W. D. Cornell, et al., Journal of Physical Chemistry 97, 10269 (1993).

[13] M. J. Frisch, G. W. Trucks, H. B. Schlegel, et al., (Gaussian, Inc, Pittsburgh, 2003).

F. J. J. Leusen

The search for crystal packing alternatives was performed using the Polymorph module [1, 2] in the Materials Studio modelling environment [3], running on a high performance Beowulf cluster with 80 Xeon CPUs and a number of graphic workstations. The standard Consistent Valence Force Field [4] was used throughout, with Ewald summation for van der Waals and Coulomb terms in the lattice energy calculations. In CVFF, atomic charges are assigned from bond increments which were fitted together with all other parameters in the original force field fitting procedure. All parameters in CVFF are therefore fully transferable. Note that these atomic charges are independent of molecular conformation, which enables a direct comparison of lattice energies calculated for conformational polymorphs (at the expense of accuracy in comparison to specifically derived, non-transferable, electrostatic models). CVFF uses a

Lennard-Jones 12-6 potential for van der Waals interactions and does not have a dedicated hydrogen bond term.

All simulations were started from the force field optimised molecular structures, with the flexible torsion angles of molecule X set at phi=0°, omega=180°, psi=180°, tau1=180°, tau2=90° and tau3=90°. This "saddle point" conformation of molecule X was chosen because any relevant conformation can be reached from this starting point during lattice energy minimisation with flexible molecules.

For each of the four molecules, the Monte Carlo simulated annealing search was carried out in the top 12 space groups ($P2_1/c$, P-1, $P2_12_12_1$, C2/c, $P2_1$, Pbca, $Pna2_1$, Cc, Pbcn, C2, P1 and $Pca2_1$) with one molecule in the asymmetric unit, and in the top 5 space groups ($P2_1/c$, P-1, $P2_12_12_1$, C2/c and $P2_1$) with two molecules in the asymmetric unit. Each run was performed five times in the "fine" setting to ensure adequate sampling. Molecules were kept rigid during the Monte Carlo searches, but were treated as fully flexible during the subsequent lattice energy minimisation stages. Note that there was no clustering step between the search and the minimisation stages. Clustering was performed after the lattice energy minimisation to remove duplicate structures. According to the rules of the blind test three predictions were allowed for each of the four molecules. The selection of crystal structures from the lists of predicted structures was based solely on calculated lattice energies, i.e., the three lowest-energy structures for each molecule were submitted.

The simulations on molecule VIII produced many distinct low-energy crystal structures: the ten lowest-energy structures fell in an energy window of only 0.9 kJ/mol, and the third structure (which corresponded to the experimentally observed crystal structure) was less than 0.2 kJ/mol above the global minimum. The simulations with two molecules per asymmetric unit did not lead to any unique low-energy structures, instead producing the same low-energy structures as the searches with one molecule in the asymmetric unit. The third predicted structure corresponded to the experimental structure, but the force field predicted a significant rotation of the molecule resulting in a severe deformation of the unit cell (deformation in unit cell dimensions relative to the experimental structure: a –12.2%, b –0.8%, c +18.4% and β -0.8%). Although the overall crystal packing, including the hydrogen bonding network, was predicted correctly, this result cannot be regarded as a success. If the goal of the exercise were to determine the crystal structure from experimental low-quality powder diffraction data by comparison to powder diffraction patterns simulated for the predicted structures, the unit cell deformation would be too large to pick this structure from a long list of predicted structures. Furthermore, the deformation is too large to allow for a meaningful prediction of solid state properties such as mechanical properties, colour, morphology, etc.

For molecule IX a problem with the force field implementation in Materials Studio resulted in the majority of predicted structures showing unreasonable molecular and lattice geometries. Of the few reasonable structures, the lowest-energy structure was in $Pna2_1$. Minimisation of the experimental structure in $P2_1/c$ lead to a large deformation in the unit cell and showed that the predicted global minimum structure was more than 200 kJ/mol above the force field minimised experimental structure.

A few of the predicted crystal structures of molecule X also showed unreasonable molecular and lattice geometries, and they were removed from the list of predicted structures. Minimisation of the experimental structure in $P2_1/n$ lead to an acceptably small deviation in molecular conformation and in unit cell dimensions. The energy difference between the predicted global minimum in $P2_1/c$ and the minimised experimental structure was 35 kJ/mol (in favour of the predicted global minimum). This large energetic error is probably due to the nitro groups, which are notoriously difficult to handle in force field calculations. The experimental structure had been sampled in the prediction sequence, but due to software limitations (lack of clustering functionality for combined output of multiple simulation runs) and considering the large energetic error, the rank of the experimental structure in the list of predicted structures could not be determined.

The force field predicted a flat conformation of molecule XI, which resulted in a large deformation of the unit cell when the experimental crystal structure was minimised in $P2_1/c$. The minimised experimental structure was nearly 10 kJ/mol above the predicted global minimum crystal structure in $Pna2_1$. Considering the small size of this molecule and the lack of ‘difficult’ functional groups, this energetic error is unacceptably large. Although the experimental structure had been sampled, it was not possible to determine its rank in the list of predicted structures because of limitations in the clustering functionality.

In conclusion, the CVFF force field has been shown to be inadequate to predict the crystal structures of these four molecules, both in terms of energy and in terms of geometry.

P. Verwer & F.J.J. Leusen, *Computer simulation to predict possible crystal polymorphs*, in “Reviews in Computational Chemistry”, Volume 12 (Eds.: K.B. Lipkowitz & D.B. Boyd), Wiley-VCH, John Wiley and Sons, Inc., New York, pp. 327 – 365 (1998).

F.J.J. Leusen, S. Wilke, P. Verwer & G.E. Engel, *Computational approaches to crystal structure and polymorph prediction*, in “Implications of Molecular and Materials Structure for

New Technologies" (Eds.: J.A.K. Howard, F.H. Allen & G.P. Shields), NATO Science Series E, Volume 360, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 303 – 314 (1999).

Materials Studio release 3.0.1, Accelrys Inc., San Diego, USA, 2004.

P. Dauber-Osguthorpe, V.A. Roberts, D.J. Osguthorpe, J. Wolff, M. Genest & A.T. Hagler, *Structure and energetics of ligand binding to proteins: E. coli dihydrofolate reductase-trimethoprim, a drug-receptor system*, Proteins: Structure, Function and Genetics, **4**, 31 – 47 (1988).

W.D.S. Motherwell

As in the previous blind test (Motherwell et al, 2002) the search methodology was based on a genetic algorithm search for an energy minimum, as implemented in the program RANCEL (Motherwell, 2001). A molecular model was built by a suitable molecular mechanics program, and treated as a rigid body thereafter. The force field parameters were those of Gavezzotti, 1999, without atomic charges. Each of the commonly occurring space groups was searched from random starting points, using the differential evolution technique on a population of 100. The simple parameters for the force field have been shown to reproduce the energy minimum reasonably well for molecules similar to VIII, IX and XI. The emphasis in this approach was post-analysis of the low energy structures, not paying great attention to the accuracy of the energy, but hopefully finding criteria for selection of the most likely polymorphs by consideration of the packing of similar molecules in the Cambridge Structural Database (CSD).

Attempted prediction of the structures for VIII, IX and XI with one molecule in the asymmetric unit failed to locate the correct minimum in the first 100 energy ranked

Structures. This is judged to be a failure of the search algorithm. Post-analysis showed that the experimental structures gave a minimum with reasonable energy which should have been located for VIII and IX.

As a matter of record, the following CSD based criteria were examined for effectiveness on a variety of similar molecules to the target structures, and also on the experimental structure coordinates without minimisation.

(a) Free space in the region of each atom.

This was probed by random exploration of a point around each atom using new code written for the Pluto program. The hope was that some structures would be revealed as either having some atoms too tightly packed, or even too loosely in certain regions. It was always found that there were no significant ‘void spaces’ in the low energy structures, and no evidence was found of any regions too tightly packed.

(b) Unusual H-bonding schemes

The idea was to eliminate any structures with H-bonding schemes (motifs) not

found, or rarely found, in the CSD for similar molecules. In the case of molecule VIII this did not help significantly, and all structures of XI showed a planar chain mediated by a screw axis, which is frequently found in the CSD.

(c) Hydrophobic interactions

Some testing was devised on criteria for hydrophobic atoms, i.e. CH, CH2 groups, having the maximum amount of contact, without success.

(d) Position of molecular centres

This was an attempt to use the observation that in certain space groups the position of the molecular centroid lies in preferred ranges, e.g. $P2_1/c$ shows a preference for the x coordinate to be at 0.25. This did not give any significant preference to the observed experimental structures over the other low energy candidates.

(e) Strongest binding of top 2, 4, 6 interactions

The idea was tested that perhaps it is important in crystal growth to form the strongest possible set of interactions for the first few molecules in the molecular coordination shell of 12 - 14 molecules, ranked by energy strength. Although some similar CSD molecules seemed to indicate that the summed energy of these top 2, 4, or 6 interactions might be an indicator, no success was evident for the test molecules.

Panagiotis G. Karamertzanis[2] and Constantinos C. Pantelides[3]

Centre for Process Systems Engineering
Department of Chemical Engineering and Chemical Technology
Imperial College London, South Kensington Campus, London SW7 2AZ, U.K.

Overview of methodology

The submitted crystal structures were determined with a new algorithm[1] for the prediction of molecular crystal structures using only the atomic connectivity of the molecule under consideration. The method is based on the global minimisation of the molar enthalpy of the crystal and has the following two key aspects:

---

[2] Current address: Department of Chemistry, University College London, 20 Gordon Street, London, WC1H 0AJ, United Kingdom. Email: p.karamertzanis@ucl.ac.uk
[3] Author to whom all correspondence should be addressed. Email: c.pantelides@imperial.ac.uk

(a) *An accurate model of the electrostatic interactions which is, however, efficient enough to allow the examination of hundreds of thousands of candidate structures*. The model comprises a set of distributed charges at both atomic and other (“satellite”) positions. The magnitudes of all charges and the positions of the satellite charges are determined via a least-squares estimation that seeks to fit the electrostatic field and its moments computed by quantum mechanical calculations, normally at the SCF/6-31G** level. An *ab initio* optimised molecular conformation is also obtained as part of the same calculation. The electrostatic potential was sampled on a dense geodesic grid starting at a distance of 1.4 times the van der Waals radii and involving 16-32 layers separated by distances of 0.025-0.05 times the van der Waals radii. The estimation was subject to molecular symmetry constraints and specified upper bounds on the errors associated with the charge estimates, and was solved to global optimality. A formal and fully automated methodology[2] has been developed which, for molecules of the type considered here, can produce an optimal distributed charge model within a few hours of computation.

(b) *An extensive search for local minima of the lattice enthalpy surface which uses deterministic low-discrepancy sequences*[3]. *The* latter are designed to produce an optimal coverage of the search space. Candidate structures are generated in the 59 space groups that appear with a frequency of more than 0.05% in the Cambridge Structure Database. Candidate structures with one or two chemically independent molecules in the asymmetric unit are generated, and these are used as starting points of local optimisation calculations with a sequential quadratic programming algorithm making use of analytical partial derivatives. A parallelised implementation of the algorithm allows minimisations from several hundreds of thousands of initial guesses to be carried out in reasonable time.

Van der Waals interactions were modelled using the exp-6 intermolecular potential due to Williams[4]. To maintain consistency with the methodology used by Williams to fit the parameters in the above potential, all hydrogen atoms are displaced by 0.1Å along the bond and towards the atom to which they are connected.

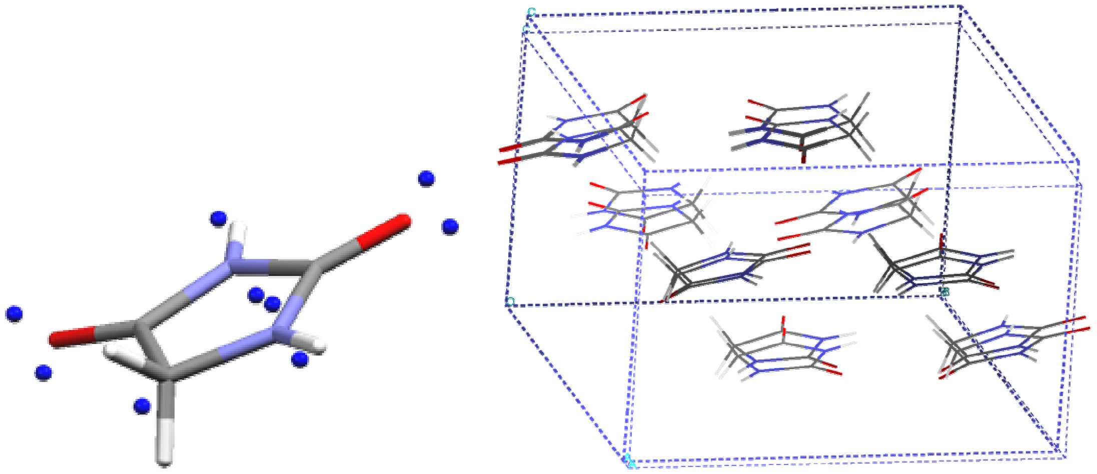
The algorithm has recently been extended to handle flexible molecules with a limited number of torsional degrees of freedom. The molecule is considered as a set of rigid fragments connected via flexible torsion angles. The effects of the latter on the intramolecular energy are described by a multi-dimensional restricted cubic Hermite interpolant constructed on a grid of intramolecular energy values computed using quantum mechanical calculations.

Exactly the same algorithm was applied to all molecules, without the use of any prior information (e.g. relating to similar molecules in the CSD) to influence the search. In all cases, the algorithm identified the experimental crystal structures, ranking them as follows (energy difference above the global minimum shown in brackets): Molecule VIII: 1st (0.00 kJ/mol), IX: 8th (+3.36 kJ/mol); X: >100th (+17.94 kJ/mol); XI: 24th (+0.76 kJ/mol).

Molecule VIII:

A distributed charge model based only on atomic charges was found to exhibit errors of RMS=4.02 kJ/mol and RRMS=6.92%. Introducing 9 satellite charges (2 per oxygen and nitrogen atoms and one for the methyl carbon) reduced these errors to 1.21 kJ/mol and 2.07% respectively.

The global optimisation involved 218985 minimisations with one, and 141564 minimisations with two chemically independent molecules in the asymmetric unit. This identified 143 unique stationary points within 5 kJ/mol from the global minimum. The global minimum was found to be the minimised experimental crystal with an RMS difference for a 12 (16) molecule coordination sphere of 0.392Å (0.475Å), although there was an error of 7.22% in the lattice length l_3. The energy difference from the second most stable crystal was 0.34 kJ/mol.



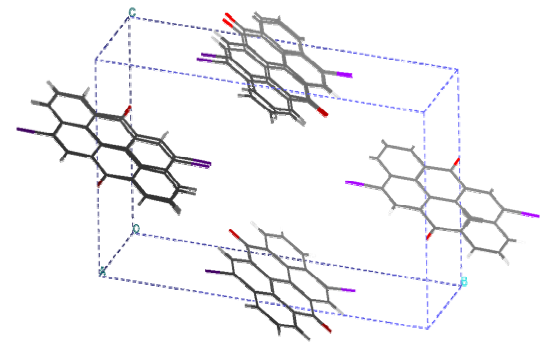
	l_1	l_2	l_3	ω_2
experimental	9.354	12.176	7.229	104.59°
minimised experimental	8.974	12.091	7.751	102.55°
difference	-4.06%	-0.70%	7.22%	-2.04°

Figure 1 Optimal satellite charge model for molecule VIII and overlay of experimental and minimised experimental crystal structures

Molecule IX:

A distributed charge model based only on atomic charges was found to exhibit errors of RMS=7.49 kJ/mol and RRMS=23.85% from an *ab initio* electrostatic field calculated at the SHF/3-21G** level. Introducing 8 satellite charges in the region of the carbonyl and iodine groups reduced this error to 2.59 kJ/mol and 8.25% respectively. A potential dispersion-repulsion potential developed for xenon[5] was used for the interactions of the iodine atoms.

Due the size of the molecule and time limitations, we performed a restricted search with 20701 local minimisations with only one chemically independent molecule in the asymmetric unit. This identified 38 distinct stationary points within 5 kJ/mol from the global minimum. The experimental crystal was identified as the 8th most stable minimum, with an energy of 3.36 kJ/mol above the global minimum. This relatively poor ranking may be attributed to the lack of a sufficiently accurate basis set for the construction of the optimal site charge model and/or the potential model for iodine (the molecular model was sufficiently close to the experimental one, with an RMS error of 0.06Å for the non-hydrogen atoms). However, these inaccuracies did not cause significant deviations during the minimisation of the experimental crystal as can be seen from table 1.



	l_1	l_2	l_3	ω_2
experimental	4.208	20.993	9.286	104.6^o
minimised experimental	4.126	21.044	9.224	102.6^o
difference	-1.95%	0.24%	-0.67%	-2.03^o

Table 1 Overlay and comparison of the lattice constants for the experimental and minimised experimental crystals for molecule IX.

From this evidence, it appears likely that the most stable structures ($C2/c,\ Z'=0.5,\ E=-192.20$ kJ/mol, $P2_1/c,\ Z'=0.5,\ \Delta E=2.12$ kJ/mol and $P2_1/c,\ Z'=1,\ \Delta E=2.87$ kJ/mol) identified by our search method are true minima and not artefacts of the potential model, although more accurate calculations (e.g. anisotropic repulsion for iodine) are needed to verify this.

Molecule X:

This molecule was modelled as a set of 7 rigid fragments connected with the 6 torsion angles $\phi, \omega, \psi, \tau_{1,..,3}$, with all other internal degrees of freedom being fixed at the globally minimal energy conformation values. The six torsion angles were categorised into two independent groups ($\Xi_1 \equiv \{\tau_1, \phi, \omega, \psi\}$ and $\Xi_2 \equiv \{\tau_2, \tau_3\}$), whose contributions to the intramolecular energy were considered to be additive. The contributions of the two groups were approximated by restricted cubic Hermite interpolants computed over grids of, respectively, 30×30×6×6 and 36×36 points. The interpolant matched the intramolecular energy and its gradients with respect to the torsion angles computed at each point on the grid at the SCF/6-31G** level (or lower for points that were sufficiently high in intramolecular energy).

The intermolecular electrostatic interactions were modelled with a single set of atomic charges fitted to a set of 11 conformations spanning the $\phi, \omega, \psi, \tau_{1,..,3}$ torsion angle space that gives rise to energetically reasonable intramolecular energies. This appears to be a good approximation for this particular molecule.

Our search method was successful in identifying the experimental crystal but gave it a very unfavourable rank due to overestimation of the intramolecular energy; for the minimised experimental crystal, this was +18.01 kJ/mol above the gas-phase level. The overestimation was caused by the fact that, in our quantum mechanical calculations of intramolecular energy, all internal degrees of freedom other than the 6 flexible torsion angles were held constant at their gas-phase values. These inaccuracies also gave rise to large errors when the experimental crystal was locally minimised, as shown in the following table:

	Experimental	Minimised	Difference	Difference when accurate intramolecular and electrostatic potentials are used
	Lattice parameters			
l_1	12.569	13.019	3.58 %	-0.86%
l_2	4.853	4.980	2.61%	1.43%
l_3	17.266	17.117	-0.86 %	0.76%
ω_2	99.16^o	96.20^o	-2.96^o	-0.46^o
	Conformational			

φ (C-N-C-C)	140.59^{o}	133.43^{o}	-7.16^{o}	-1.23^{o}
ω (O-C-N-C)	-6.59^{o}	-1.98^{o}	4.61^{o}	1.46^{o}
ψ (H-C-C-O)	-34.00^{o}	-20.44^{o}	13.56^{o}	2.28^{o}
τ_1 (H-C-C-C)	138.26^{o}	131.01^{o}	-7.25^{o}	-1.90^{o}
τ_3 (C-C-N-O)	51.95^{o}	60.62^{o}	8.67^{o}	0.03^{o}
τ_2 (O-N-C-C)	22.91^{o}	29.17^{o}	6.26^{o}	-0.47^{o}

Table 2 Comparison of experimental and minimised experimental crystal lattice and conformation parameters for molecule X

In fact, changes in the torsion angles due to the crystalline environment cause very slight changes in these other degrees of freedom. Albeit unimportant for the geometric aspects of molecular packing, these changes result in a significant decrease in intramolecular energy. To confirm this, we minimised the experimental molecular conformation by fixing the torsion angles $\phi, \omega, \psi, \tau_{1,..,3}$ to their experimental values. The resulting molecular conformation has an RMS discrepancy from the experimental one of 0.031Å for the non-hydrogen atoms. The energy at the optimal point was found to be only +7.28 kJ/mol higher than that of the gas-phase conformation.

As a final check, the minimised experimental crystal was re-computed using atomic charges computed at the experimental molecular conformation (RMS=3.32 kJ/mol and RRMS=4.46%) and the interpolation grid constructed with the rest of the internal degrees of freedom fixed at the *ab initio* optimised conformation with $\phi, \omega, \psi, \tau_{1,..,3}$ at the experimental values. The errors in the lattice energy minimisation of the experimental crystal were greatly reduced, with maximum error of 1.43% in the lattice lengths and 2.28^{o} in the torsion angles (cf. last column of table above and figure below).

The above analysis points towards two main changes in our method. First, the atomic charges need to be expressed as interpolants of the flexible torsion angles[6]. Secondly, the computation of the intramolecular energy at each point in the interpolation grid of the torsion angles needs to take account of the relaxation of the other internal degrees of freedom. However, it does not appear to be necessary to include the latter directly as additional decision variables in the crystal lattice energy minimisation.

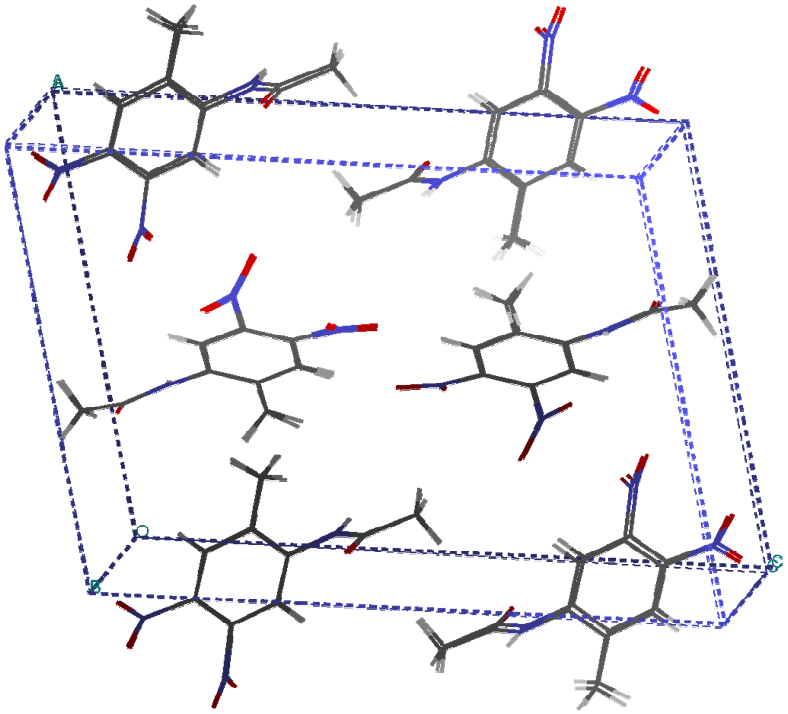


Figure 2 Overlay of experimental and minimised experimental crystal for molecule X for an accurate intramolecular potential

Molecule XI:

A distributed charge model based only on atomic charges was found to exhibit errors of RMS=5.65 kJ/mol and RRMS=22.58%. Introducing satellite charges associated with all carbon and nitrogen atoms reduced these errors to 1.24 kJ/mol and 4.95% respectively.

Our search procedure identified a relatively large number of crystal structures within 1 kJ/mol of the global minimum. The minimised experimental structure was ranked as 24th, at only 0.76 kJ/mol above the global minimum and with an RMS difference for a 16 molecule coordination sphere of 0.257 Å. The *ab initio* molecular structure used in the search was found to be similar to the experimental one, with RMS overlaps for the non-hydrogen atoms of the two molecules in the asymmetric unit of 0.024 Å and 0.015 Å respectively.

An examination of the Hessian of the lattice enthalpy for the minimised experimental point established that this had a small negative eigenvalue (-0.09), which may indicate an unstable structure. Expansion of the minimised structure to P1, perturbation along the direction of the eigenvector corresponding to the negative eigenvalue and subsequent re-minimisation led to a minimum that was only 0.001 kJ/mol more stable. Examination of this point with PLATON revealed that it was a $P\bar{1}$, $Z' = 4$ structure; this was not submitted in the extended list because it was beyond the rules of the blind test. A comparison of the experimental with the two minimised experimental structures is shown in table 3.

Structure	Lattice Energy (kJ/mol)	Volume per molecule (Å^3)	l_1	l_2	l_3	ω_1	ω_2	ω_3
Experimental	-38.472	94.163	9.507	9.122	9.790	90.00^o	117.47^o	90.00^o
Minimised experimental ($P2_1/c$)	-39.284	96.115	9.796	8.954	10.025	90.00^o	119.02^o	90.00^o
Minimised experimental ($P1$)	-39.285	96.121	9.798	8.956	10.030	90.16^o	119.06^o	91.44^o

Table 3 Minimisation of the experimental crystal with and without space group constraints for molecule XI

References:

1. P.G. Karamertzanis and C.C. Pantelides, "*Ab initio* Crystal Structure Prediction – I. Rigid molecules," *J. Comput. Chem.*, **26**, 304-324, 2005.

2. P.G. Karamertzanis and C.C. Pantelides, "Optimal Site Charge Models for Molecular Electrostatic Potentials," *Mol. Sim*., **30**, 413-436, 2004

3. I.M. Sobol', "On the distribution of points in a cube and the approximate evaluation of intergrals," *Comput. Math. Math. Phys*., **7**, 86, 1967

4. D.E. Williams, "Improved intermolecular force field for molecules containing H, C, N and O atoms, with application to nucleoside and peptide crystals," *J. Comp. Chem.*, **22**, 1154, 2001

5. D.E. Williams, "Direct calculation of crystalline thermal expansion and molecular reorientation from nonbonded interatomic potential anharmonicity and thermal amplitudes," *Acta Cryst.* ,**A28**, 84, 1972.

6. P.G. Karamertzanis, "Prediction of Crystal Structure of Molecular Solids," PhD thesis, University of London (2004).

Price (IX), Lewis(VIII), Torrisi (XI), Nowell(X)

Methodology

The main distinguishing feature of our approach was the use of accurate distributed multipole models for the electrostatic contribution to the lattice energy. The rigid molecular model was obtained by ab initio optimisation at the MP2 6-31G** level (SCF 6-311G* for IX). The corresponding charge distribution was subjected to a Distributed Multipole Analysis (Stone and Alderton 1985) to obtain the sets of atomic charges, dipoles, quadrupoles, octupoles and hexadecapoles used in the electrostatic model. All other terms in the potential were assumed to be in the isotropic atom-atom *exp-6* form, using empirically fitted parameters (Cox, Hsu et al. 1981; Williams and Cox 1984; Coombes, Price et al. 1996) with variations for molecules VIII and IX noted below. A set of 1500 to 3000 close packed crystal structures for each rigid conformer was generated using the latest versions of MOLPAK (Holden, Du et al. 1993) kindly supplied by the Maryland group during the course of this blind test. These were used as starting points for minimisations using DMAREL(Willock, Price et al. 1995), using the recent version in which the minimisation is constrained by the initial spacegroup, but the eigen-values of the Hessian are examined to ensure that a true minimum has been found.

The rigid body second-derivative matrix for all the low energy crystal structures found by this approach, (and submitted in the extended lists), was used to estimate their elastic constants (Day, Price et al. 2001) and *k*=0 phonons (Day, Price et al. 2003). These were used to consider whether any crystals were particularly susceptible to shear (i.e. had a very low eigen-value of the shear submatrix) or had low frequency modes that would give distinctive thermal effects. In some cases, the attachment energy model was used to predict the crystal morphologies and the relative volume growth rates of the crystallites from vapour. This evolving methodology of generating low energy crystal structures and comparing their properties has recently been applied to barbituric acid (Lewis, Tocher et al. 2004), chlorothalonil (Tremayne, Grice et al. 2004) and p-dichlorobenzene (Day and Price 2003).

Molecule VIII - Lewis (This extended list starts with the three submissions, then the structures with the optimised conformation and then the other structures with the planar conformation, starting at 28.)

Following our experience in discovering a new polymorph of barbituric acid (Lewis, Tocher et al. 2004), which contained both an envelope and planar conformation of the molecule, we used both the ab initio optimised and a planar transition state conformation of hydantoin in the search procedure. The MP2 estimate is that the planar molecule has ΔE_{intra} ~ 0.3 kJ mol^{-1} and so it is not surprising that the molecule is planar in the experimental structure. Consideration of the crystal structures of alloxan, barbituric acid, cyanuric acid, parabanic acid and urazole (Lewis, Tocher et al. in preparation) led to the use of a reduced C atom repulsion, (A_{CC} = 277180 kJ mol^{-1}). The MOLPAK search used for hydantoin included the latest extension of the range of packing types, including P2/c.

When the abstract was released, it was noted that there were no C2/c structures with a low energy, and we wrongly assumed that the search had failed to find the experimental structure. However, since the majority of the low energy structures had the experimental hydrogen-bonding motif, we were pleased that we had correctly predicted the observed ribbons of dimers. We submitted two predictions from the search using the ab-initio optimised molecule, the global minimum (U_{latt}= -97.8 kJ mol^{-1}), and a low energy structure with a particularly good growth rate and a different hydrogen bonding motif. The third choice was from the search with the planar molecule, and was the optimal structure on properties with the experimental hydrogen-bonding motif. A submission based purely on energy would have been successful. The global minimum energy structure (U_{latt}= -98.2 kJ mol^{-1}) with the planar conformer (PriceLewis.VIII.28) was a $P2_1/n$ structure with Z'=2 (found by symmetry lowering a Pbcn structure) which had a good overlay (Rms_{12}=0.42) with the experimental structure. Similarly the second lowest energy structure in the search with the non-planar molecule (PriceLewis.I.4) was a reasonable prediction (Rms_{12}=0.93), though again not strictly in the correct space group. We therefore conclude that our methodology was successful in predicting the hydrogen-bonding motif and a very close match to the correct co-ordination of the molecule as a minimum in the lattice energy, close to the global minimum, even allowing for the conformational distortion energy.

Molecule XI Torrisi

There was not the time to develop a specific potential for azetidine, which given the weakness of the intermolecular forces and the sensitivity to the hydrogen atom positions and interactions, would have been necessary for a proper prediction study. The MOLPAK/DMAREL search showed a large number of structures between the global minimum at -47.6 kJ mol^{-1} and -44.0 kJ mol^{-1} , all of which showed some type of close N-H…N contacts. There was then a significant energy gap of almost 7 kJ mol^{-1} before a plethora of structures with the N-H groups well separated. From the low energy group, which was submitted as the extended list, the global minimum in the lattice energy was submitted as the first prediction. The second choice was a structure, which was predicted to grow almost twice as fast as the global minimum and the majority of the other structures. The third was chosen on the basis of being the densest, with yet another different N-H…N motif.

When the experimental structure was released, it was obvious that as a Z=8, Z'=2 structure that it could not have been found by our search procedure, and indeed none of the low energy structures even approximated it. A minimisation starting from the experimental crystal structure using the ab initio optimised molecular structure produced a transition state in the $P2_1/c$ spacegroup at -46.85 kJ mol^{-1}, giving a true minimum in P-1 at -46.89kJ mol^{-1} . The reproduction of the crystal structure was poor even using the experimental molecular structure (with the bonds to hydrogen adjusted to standard neutron values). Overall, although the restricted nature of the search guaranteed that we would not find this experimental structure, our intermolecular potentials and the use of a rigid molecule are not particularly well suited for this molecule.

Molecule IX Price

The strong orientation dependence of iodine-iodine van der Waals contacts (Nyburg and Faerman 1985) shows that it is highly probable that the charge distribution around iodine is sufficiently anisotropic that this would need to be represented in the repulsion, as well as in the electrostatic model. Unfortunately, the methods that we have been developing for establishing non-empirical anisotropic atom-atom potentials (Day and Price 2003) could not be applied to iodine because of the number of electrons involved. Empirical testing of the potential also proved problematic, as the most similar molecules with crystal structures in the Cambridge Structural Database, namely, IONTRQ, IOANTQ, IANTRQ01, did not have the hydrogen atoms determined. We experimented with different iodine potentials, including trying to extrapolate the anisotropy from the anisotropy of the van der Waals radii (Nyburg and Faerman 1985) and the anisotropic potential for Cl (Day and Price 2003) and for Br developed for the previous blind test. It was concluded that we did not yet have the resources to develop an iodine intermolecular potential that could be used with any degree of confidence for this molecule. The potential used had the SCF 3-11G* DMA with the standard FIT repulsion dispersion parameters, supplemented by the I..I potential A_{II}=322110.3 kJ mol^{-1}, B_{II}=2.86 $Å^{-1}$, C_{II}=10986.6 kJ mol^{-1} $Å^6$ derived the MMFF force-field (Halgren 1992), with heteroatomic parameters given by the usual combining rules.

The search with this potential yielded a global minimum of -206 $kJmol^{-1}$, which was significantly, more stable than the next distinct structure, at -198 kJ mol^{-1}. These were submitted as predictions, with the third structure being selected from the plethora of structures with energies < -190 kJ mol^{-1} in the extended list, on the basis of properties. The minimum corresponding to the experimental structure was found at -189 kJ mol^{-1}, and so would have been within the top 40 structures in energy. The model potential reasonably reproduced the experimental structure, with the largest cell deviation being 2.2%, though clearly the relative energy to other hypothetical structures was in serious error. This may be mainly due to the lack of repulsion anisotropy, as the exponential nature of exchange-repulsion can lead to very large energy differences for small changes in separation

Molecule X Nowell

The approach used here was to use rigid molecule structures with accurate estimates of ΔE_{intra}, and the corresponding electrostatic model to avoid the errors implicit in assuming that the molecular charge distribution was independent of conformation (Price and Stone 1992). A previous study of aspirin (Ouvrard and Price, 2004), using the low energy ab initio optimised conformations of the molecule, had successfully found the known crystal structure which has a metastable molecular conformer, very close to the minimum in $E_{tot}=U_{latt}+\Delta E_{intra}$. For molecule X, a relaxed SCF scan varying ϕ suggested that there were three conformational energy minima. The MP2 optimised structure ($\phi=-3°$ $\tau_2=40°$, and $\tau_3=41°$) was used in the search, together with two local minima ($\phi=-121°$, $\tau_2=41°$, and $\tau_3=37°$ with $\Delta E_{intra}=4.7$ kJ mol^{-1}, and $\phi=120°$, $\tau_2=43°$and $\tau_3=37°$ with $\Delta E_{intra}=4.1$ kJ mol^{-1}). Following these results, some additional conformations with ϕ constrained to -5°, ±10°and ± 20°and the rest of the molecule optimised were used. We has expected to have further computational resources available to significantly increase the number of rigid conformations that could be tried, but these failed to materialise within the time scale of the blindtest. This was particularly frustrating when the searches with the molecule deviating significantly from planar, ($\phi\geq +10°$) produced low energy crystal structures which clearly showed elongated N-H…O=C hydrogen bonds, suggesting that if only a suitable conformation could be found that would allow the molecules to approach to form such bonds at their optimal length, this would generate a particularly low energy crystal structure.

The release of the experimental structure immediately highlighted the problem of defining which aspects of the molecular structure have to be constrained to evaluate ΔE_{intra} and reproduce the observed crystal structure. (Just taking the experimental structure, even adjusting the bond lengths to hydrogen to neutron values, gives an unphysically large ΔE_{intra} because of differences in bond lengths and angles that are due to a combination of experimental and ab initio error.) If all the torsion angles are constrained at the experimental values, $\Delta E_{intra}=13.7$ kJ mol^{-1}, and a search with this conformation gave a good reproduction of the experimental structure as the global minimum in $U_{latt}=-129$ kJ mol^{-1}. However, if only ϕ, τ_2 and τ_3 were constrained ΔE_{intra} =

5.0 kJ mol^{-1} but the lattice energy was less favourable (U_{latt}=-114 kJ mol^{-1}) and the reproduction of the experimental structure significantly worsened, probably because the methyl group torsion τ_1 plays a role in allowing the hydrogen bond formation. Thus, if we had been able to do a sufficiently fine and exhaustive search varying all six torsion angles, our method for evaluating the energies would have placed the experimental structure within about 5 kJ mol^{-1} of the global minimum in E_{tot}, with most of the uncertainty arising from ΔE_{intra} .

Coombes, D. S., Price, S. L., et al. (1996). *J. Phys. Chem.* 100: 7352-7360.

Cox, S. R., Hsu, L. Y., et al. (1981). *Acta Cryst. A* 37: 293-301.

Day, G. M. and Price, S. L. (2003). *J. Am. Chem. Soc.* 125: 16434-16443.

Day, G. M., Price, S. L., et al. (2001). *Cryst. Growth Des.* 1: 13-26.

Day, G. M., Price, S. L., et al. (2003). *J. Phys. Chem. B* 107: 10919-10933.

Halgren, T. A. (1992). *J. Am. Chem. Soc.* 114: 7827-7843.

Holden, J. R., Du, Z. Y., et al. (1993). *J. Comput. Chem.* 14: 422-437.

Lewis, T. C., Tocher, D. A., et al. (2004). *Cryst. Growth Des.* 4, 979-987.

Lewis, T. C., Tocher, D. A., et al. *Cryst. Growth Des. submitted.*

Nyburg, S. C. and Faerman, C. H. (1985). *Acta Cryst. B* 41: 274-279.

Ouvrard, C. and Price, S. L. (2004) *Cryst. Growth Des. ASAP.*

Price, S. L. and Stone, A. J. (1992). *J. Chem. Soc. Farad. Trans.* 88: 1755-1763.

Stone, A. J. and Alderton, M. (1985). *Mol. Phys.* 56: 1047-1064.

Tremayne, M., Grice, L., et al. (2004). *J. Amer. Chem. Soc.* 126 :7071-7081.

Williams, D. E. and Cox, S. R. (1984). *Acta Cryst. B* 40: 404-417.

Willock, D. J., Price, S. L., et al. (1995). *J. Comput. Chem.* 16: 628-647.

CRYSTALG (Scheraga H.A., Arnautova Y.A)

Molecular geometries and partial charges

Molecular geometries were obtained from DFT calculations (DFT B3LYP/6-311+G) using the GAUSSIAN program [1]. In the case of target IX, the geometry optimization was carried out with iodine atoms replaced by chlorine. Chlorine was changed back to iodine afterwards, and the C-I bond lengths were adjusted to reproduce the average experimental C-I bond length (2.1 Å) from the Cambridge Structural Database (CSD)[2]. The N=O bond lengths in molecule X were adjusted to the average experimental value for aromatic nitrocompounds (1.23 Å) [3]. We also evaluated the relative stabilities of cis and trans conformations of molecule XI. The trans conformation was more stable and it was chosen for further calculations.

The atomic charges were obtained by fitting to the molecular electrostatic potential calculated ab initio (HF 6-31G** for targets VIII, X, XI and HF 6-31G** for C, H, O atoms plus HF 3-21G* for iodine for target IX). The resp method [4a], implemented in the AMBER 6.0 program [4b], was used for fitting charges. The partial atomic charges for compound X and the molecules analogous to it were obtained by fitting to the electrostatic potential of the most stable (ab initio calculated) conformations.

Optimization of lattice energy

Program: CRYSTALG [5]

The CRYSTALG program predicts crystal structures by global optimization of lattice energy without assuming any symmetry information [all searches are carried out in the space group P1 with a different number of molecules in the unit cell (Z)]. Z is a parameter, i.e. several runs for different values of Z are carried out, and the value of Z that leads to the lowest energy per molecule is selected. In the current version of the program, we use a global optimization method, Conformation-Family Monte Carlo (CFMC) [5,6]. The CFMC method can be considered as an extension of the Monte Carlo plus Minimization (MCM) method [7]; however, the CFMC method uses the whole family of conformations for a Monte Carlo step instead of a single conformation.

Both rigid and flexible molecular models can be treated using CRYSTALG [5].

The lattice energy is assumed to be a sum of pairwise interatomic interactions, and includes electrostatic, nonbonded, and torsional terms. The Buckingham 6-exp potentials were used for calculations of targets VIII, IX, X, and XI. Parameters for less common atom types present in molecules IX and X (for example, for iodine, oxygen and nitrogen in nitro group) were obtained using the VMC (Vector Monte Carlo) and the force minimization procedures [8]. The torsional energy was calculated using a third-order Fourier expansion with coefficients obtained by fitting the torsional energy to the difference between ab initio (or experimental) and molecular mechanics (sum of nonbonded and electrostatic) energy barriers and symmetry. The energies of all structures generated during the search were minimized with the Secant Unconstrained Minimization Solver (SUMSL) [9] minimizing procedure.

The global search was carried out for Z = 2, 4, 8 for targets VIII, X, and XI and, because of the large size, for Z = 2, 4 for target IX. In the case of a chiral molecule, two types of runs were carried: one with the L (or R) enantiomer and the other one with the racemic mixture.

Evaluation of force field and the search method

Crystal structure prediction calculations for the target molecules were carried out in two steps.

First, a search of the CSD was carried out in order to find experimental crystal data for molecules analogous to the targets. The molecules were used to evaluate the quality of the potentials and effectiveness of the search method (CFMC). The potential parameters for the atom types present in aliphatic and aromatic hydrocarbons [10], amines and amides [preliminary values which were updated later (11)] were obtained in our group earlier as a part of the work on development of all-atom potentials for amino acids and proteins. Our force field did not contain nonbonded parameters for the oxygen and nitrogen in the nitro group (molecule X) and for oxygen, iodine, and carbonyl carbon in molecule IX. The experimental crystal data for the analogous molecules were also used for deriving (or refining) potential parameters for these atom types. The CSD reference codes and chemical names of the test molecules selected for each target are shown in Tables 1-4.

Second, global energy optimizations were carried out for each target molecule using the potential parameters tested and refined using the analogous molecules.

Molecule VIII

Preliminary calculations carried out for the molecules from Table 1 showed that the nonbonded parameters obtained earlier for amides did not perform very well for the molecules analogous to the first target. That is why we used them only as starting values for the VMC parameter optimization [8]. Final values of the potential parameters are given in Table 5.

To evaluate the accuracy of the optimized parameters, we carried out local and global energy minimizations for the test molecules. Results of these calculations are given in Table 1. The nonbonded parameters, when used for local energy minimization of the experimental structures of PARBAC, BEPNIT, ALOXAN11, and DHURAC10, led to small deviations in the unit cell parameters (less than 2.3 %) and in the positions and orientations of the molecules in the unit cell. The local energy minimization resulted in some changes in the H-bond geometry. Thus, the decrease in H-bond lengths was 0.036(0.047) Å for BEPNIT and 0.013(0.022) Å for DHURAC. In the case of PARBAC, the H-bonds became longer by 0.04 Å. The H-bonds are more linear in the experimental structures than in the minimized ones. The biggest change was observed for PARBAC where the OHN angle decreased from ~170^0 to 162^0. All minimized experimental structures were found by global search as the lowest minima of the refined potentials. The results obtained for the test molecules suggested that our search method and the refined nonbonded parameters could be used for prediction of crystal structure of molecule VIII.

Molecule IX

Since our force field did not contain nonbonded parameters for iodine and the types of oxygen and carbonyl carbon present in this molecule, the force minimization and the VMC procedures [8] were used to obtain them. Three groups of molecules (Table 2) were used for deriving the missing nonbonded parameters. The first group included molecules containing only H, C, and O atoms and was used for deriving parameters for oxygen and carbonyl carbon. The H-, C-, and I-containing molecules constituted the second group used for deriving nonbonded parameters for iodine, whereas, the third group included the molecules containing the same types of atoms as molecule IX. Since the number of crystal structures analogous to molecule IX in the CSD is very limited and the accuracy of the experimental data is low (see R-factors in Table 2), the last group of the test molecules was used mainly for the parameter evaluation.

I…X interactions are highly anisotropic due to the anisotropic electron distribution around the halogen atom causing a decreased repulsive wall in the direction of the extended C-I axis [12,13]. This effect has to be taken into account to describe accurately crystals of iodine-containing molecules. Since molecule IX was the last target that we treated and we were running out of time, we did not want to introduce a different (anisotropic) form of potential for iodine. Instead, the potential parameters for molecule IX were derived in the following way: first, potential parameters for all atom types except aromatic hydrogen and carbon (which were obtained earlier [10] and were fixed) were derived using geometric combination rules; second, the parameters for I…I and I…O interaction were optimized without using combination rules and keeping the parameters for all other types of interactions fixed.

The final values of the nonbonded parameters obtained for the test molecules and used for crystal structure prediction of molecule IX are given in Table 5. In general the best set of parameters, which we managed to obtain, provided good results for the H-, C- and O-containing molecules but did not perform very well for the molecules from the second and third groups. In most cases, the minimized experimental structures were not global minima (column 7, Table 2) and the changes in the unit cell parameters caused by the energy minimization were quite large (columns 3-5, Table 2). The derived potential parameters were used for predicting the crystal structure of molecule IX; however, the results obtained for the analogous molecules suggested that a different form of potential may be necessary for accurate description of iodine.

Molecule X

Molecule X consists of three types of functional groups: benzene ring, amide, and nitro groups. At the beginning of the CSP04 blind test we had a force field including nonbonded parameters for the atom types in the first two groups (Table 7), as well as the corresponding torsional parameters. We selected a set of nitrobenzene molecules (Table 3) to derive the missing parameters. The force minimization and the modified VMC procedure (allowing parameter optimization for flexible molecules) were applied. ACANIL02, ACTOLD and DIXDUJ crystal structures (Table 3) were used for evaluation of the final parameter set. The nonbonded parameters used for molecule X are given in Table 7. Table 8 contains experimental and ab initio values of rotational barriers (column 4) used for deriving torsional parameters for molecule X, as well as the empirical rotational barriers (column 3) obtained using these parameters.

The results of local and global energy minimizations carried out for the test molecules from Table 3 showed that our parameters are quite accurate (small values of Δ, θ, t, and ΔTors) and are able to predict the experimental structure as one of the lowest energy minima.

Molecule XI

In the case of the molecules analogous to molecule XI, the nonbonded parameters derived earlier for amines worked well (Table 4). Average structural deviations Δ's did not exceed 3.3 % (BETLER). The experimental geometry of the H-bonds was well preserved after the local energy minimization, i.e. the bond lengths changed by less than 0.06 Å (except some H-bonds in TIMCHX which became 0.14 Å longer) and the values of the NHN angles decreased by less than 1^0. The minimized experimental structures of BETLER, TIMCHX, and piperidine were found as the lowest energy minima, whereas the experimental structure of ZBCNON was only 0.54 kJ/mol less stable than the corresponding global minimum. Based on these results we decided to use the nonbonded parameters (Table 9) obtained earlier (without refining them) for prediction of molecule XI.

Results of the predictions

Molecule VIII. Since this molecule is chiral (according to the ab initio calculation), two types of global search runs (with one and two enantiomers, respectively) were carried out. The experimental structure was not present on the list of three predictions that we submitted but the structure found as minimum 5 by global search corresponded to the experimental structure minimized with our potentials. The potential reproduced the experimental structure, including the

H-bond geometry (Fig.1a), fairly well (RMSD=1.26 Å); however, it was predicted as 1.79 kJ/mol less stable than the global minimum. The main difference between the experimental structure and the hypothetical structures with lower energies (minima 1-4) is in the geometry of H-bond motif. All four low-energy structures can be characterized by the same H-bond motif (Fig. 1b), which is quite different from the one present in the experimental structure (Fig.1a). The bond lengths were shorter in the hypothetical structures (by ~0.06 Å) and the OHN angles were closer to 180^0 (~162^0 vs. ~156^0 in the experimental structure). The COH angles were ~160^0 which is at variance with the 135^0 angle in the experimental structure and with the expected positions of the oxygen electron pairs.

The results show that the search method worked well; however, the potential used for molecule VIII was not accurate enough and led to the generation of low-energy structures with rather unphysical H-bonds.

Molecule IX. Crystal structure prediction for molecule IX was carried out using our improved potential. Although, the list of the three predictions submitted for this molecule did not contain the experimental structure, the local energy minimization showed that minimum 10 found as a result of the global search corresponds to the experimental structure minimized with our potential. The RMS deviation from the experimental structure is small (0.34 Å), but the relatively high energy of the minimized experimental structure (4.5 kJ/mol above the global minimum) suggests that a better potential (in particular, for iodine) is necessary to predict this target successfully. As in the case of molecule VIII, the search method worked well.

Molecule X. Our CRYSTALG program enables us to conduct a global search for flexible molecules. All six torsional angles of molecule X were allowed to vary during the global search. The experimental structure was not found as result of the search. The local energy minimization with our potential caused large changes in the unit cell parameters of the experimental structure (for example, the c parameter increased by almost 1 Å). At the same time, the molecular conformation stayed close to the experimental one with the biggest change (12°) observed for the orientation of the methyl group. The minimized structure had a much higher lattice energy compared to the lowest minimum found, with the main difference coming from the torsional term, which was much higher in the case of the minimized experimental structure.

In general, the hypothetical structures with low energies were quite different from the experimental one. Most of them contained chains of molecules connected by nitro…amide H-bonds, whereas the molecules in the experimental structure were linked by amide…amide H-bonds.

Since the minimized experimental structure has such a high energy, it is not possible to draw any conclusions regarding the efficiency of the search method. The most probable reason for the unsuccessful prediction for molecule X is an unbalance between the nonbonded, electrostatic and torsional parts of the potential (i.e., the potential is wrong). However, the possibility that the experimental structure is not the thermodynamically most stable form cannot be ruled out completely.

Molecule XI. The calculations carried out for molecule XI yielded many structures within a relatively small energy range. All structures were built of hydrogen-bonded chains of molecules. Our search found many structures with more than one independent molecule in the unit cell; however, none of these hypothetical structures resembled the experimental one. Minimization of the experimental structure led to small changes in the unit cell parameters, but the initial space group symmetry, $P2_1/c$, Z=8, was lost. The space group of the minimized structure was $P2_1$, Z=8. The experimental H-bond motif was preserved, although the N…H distances were more than 0.1 Å longer in the minimized structure.

As in the case of molecule XI, it is difficult to say whether the search method worked well. It is worth noting that a global search for Z=8 without using any space group information is already quite challenging, and the existence of many almost equienergetic minima and relatively high energy of the experimental structure made it even more difficult.

The results obtained for molecule XI may suggest that the potential used for the prediction was not perfect. On the other hand, the temperature of the X-ray data collection was close to the melting temperature for molecule XI; thus, it can be expected that thermal effects are important. Local energy minimization of the experimental structure led to a change in the crystal symmetry from $P2_1/c$ to a lower-symmetry space group with lower symmetry ($P2_1$, Z=8) that may be a manifestation of some disorder in the experimental structure.

References.

GAUSSIAN Gaussian 98, Revision A.7, Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb M.A.; Cheeseman, J.R.; Zakrzewski, V.G.; Montgomery, J.A. Jr.; Stratmann, R.E.; Burant, J.C.; Dapprich, S.; Millam, J.M.; Daniels, A.D.; Kudin, K.N.; Strain, M.C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D.K.; Rabuck, A.D., Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R.L.; Fox, D.J., Keith, T.;. Al-Laham, M.A.; Peng, C.Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. Gaussian, Inc., Pittsburgh PA, 1998.

Allen, F.H.; Kennard, O.; Taylor, R. Acc. Chem. Res. 1983, 16, 146.

Arnautova E.A.; Pivina, T.S.; Gladkikh, O.P; Vilkov, L.V. J. Mol. Struct. 1996, 374, 137.

(a) Bayly, C.I.; Cieplak, P.; Cornell; W.D. and Kollman, P.A. J. Phys. Chem. 1993, 97, 10269.

(b) Case, D.A.; Pearlman, D.A.; Caldwell, J.W.; Cheatham III, T.E.; Ross, W.S.; Simmerling, C.L.; Darden, T.A.; Merz, K.M.; Stanton, R.V.; Cheng, A.L.; Vincent, J.J.; Crowley, M.; Tsui, V.; Radmer, R.J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G.L.; Singh, U.C.; Weiner, P.K. and Kollman, P.A. 1999, AMBER 6, University of California, San Francisco.

Pillardy, J.; Arnautova, Y.A.; Czaplewski, C.; Gibson, K.D.; Scheraga, H.A. Proc. Nat. Acad. Sci. USA. 2001, 98, 12351.

Pillardy, J.; Czaplewski, C.; Wedemeyer, W.J. and Scheraga, H.A. (2000). Helv. Chim. Acta, 83, 2214.

(a) Li, Z.; Scheraga, H.A. Proc. Natl. Acad. Sci. U.S.A. 1987, 84, 6611.

(b) Li, Z.; Scheraga, H.A. J. Mol.Struct. (THEOCHEM) 1988, 179, 333.

Arnautova, Y.A.; Pillardy, J.; Czaplewski, C.; Scheraga, H.A. J. Phys. Chem. B. 2003, 107, 712.

Gay, D.M. ACM Trans. Math. Software 1983, 9, 503.

Arnautova, Y.A.; Jagielska, A.; Pillardy, J.; Scheraga, H.A. J. Phys. Chem. B 2003, 107, 7143.

Jagielska. A., Arnautova, Y.A., Scheraga, H.A. J. Phys. Chem. B. 2004(in press)

Price, S.L.; Stone, A.J.; Lucas, J.; Rowland, R.S.; Thornley, A.E. J. Am. Chem. Soc. 1994, 116, 4910.

Lommerse, J.P.M.; Stone, A.J.; Taylor, R.; Allen, F.H J. Am. Chem. Soc. 1996, 118, 3108.

Table 1. Results of the evaluation of potential parameters for the molecules analogous to compound VIII

CSD ref. code, name	Space group	Δ^a	θ^b	t^c	E^d	ΔE^e
PARBAC Parabanic acid	$P2_1/n$, Z=4	1.7	1.57	0.031	-108.32	0.0
BEPNIT 5,5-Dimethylhydantoin	$P2_12_12_1$, Z=4	2.1	1.95	0.022	-95.35	0.0
ALOXAN11 aloxan	$P4_12_12$, Z=4	2.3	2.18	0.019	-131.13	0.0
DHURAC10 Dihydrouracil	$P2_1/c$, Z=4	2.1	0.57	0.054	-98.62 (-103.81)	0.0

[a] Average deviation in unit cell parameters (in percent) calculated using the following formula:

$$\Delta = 100\% \cdot \sum_{i=1}^{N} \left| \frac{x_{exp}^i - x_{m.exp}^i}{x_{exp}^i} \right| \Big/ N$$

, where x^i are the unit cell parameters; N is number of the unit cell parameters that changed as a result of the energy minimization; [b] Deviation of molecular orientations between the experimental and minimized experimental structures (degrees); [c] Average deviation in positions of molecules in the experimental and the minimized experimental structures (fractional coordinates); [d] Lattice energy (kJ/mol). The experimental sublimation enthalpies (if available) are given in parenthesis; [e] Difference ($E_{m.e.}$-minE) between the energy of the minimized experimental structure and the energy of the lowest minimum found (kJ/mol) by global energy minimization.

Table 2. Results of the evaluation of potential parameters for the molecules analogous to compound IX [a]

CSD ref. code, name	Space group	Δ	θ	t	E	ΔE
H, C, O						
BNZQUI03 p-Benzoquinone	$P2_1/c$, Z=2	2.7	12.1	0.020	-75.3 (-66.8)	0.0
BODCOM Naphthalene-1,4,5,8-tetrone	$P2_1/n$, Z=2	0.9	3.75	0.019	-124.7	0.0
NAPRON Naphthanthrone	$P2_12_12_1$, Z=4	7.6	12.1	0.050	-124.4	-
H, C, I						
ZZZPRO05 1,4-di-iodobenzene	Pbca, Z=4	1.6	2.49	0.009	-53.2 (63.6)	0.0
ZZZPRO07 1,4-di-iodobenzene	Pccn, Z=4	2.4	4.30	0.012	-52.0	1.2
HIBENZ01 hexaiodobenzene	$P2_1/n$, Z=2	3.5	4.19	0.027	-62.9	0.6
TECWUT 9,10-di-iodoanthracene	P-1, Z=2	0.7	3.29	0.029	-90.9	1.1
H, C, O, I						
INAPQU 3-Iodonaphtho-1,4-quinone	$P2_1/c$, Z=4 (R=9%)	2.1	4.29	0.035	-88.0	1.6
IANTRQ 1,5-di-iodoanthraquinone	$P2_1/c$, Z=2 (R=17.5%)	3.5	5.36	0.066	-115.5	4.1
IONTRQ 1-iodo-anthraquinone	$P2_1$, Z=2 (R=13.0%)	2.5	4.93	0.043	-116.3	0.0

[a] For symbols and units, see footnotes to Table 1.

Table 3. Results of the evaluation of potential parameters for the molecules analogous to compound X [a]

CSD ref. code	space group	Δ	θ	t	ΔTors[b]	E	ΔE
NITRBE01 nitrobenzene	$P2_1/c$, Z=4	3.9	6.28	0.018	0.5	-65.7	0.0
1,2-dinitrobenzene [c]	$P2_1/c$, Z=4	1.7	5.71	0.026	3.7	-92.8 (-87.9)	0.2
DNBENZ10 1,3-dinitrobenzene	$Pbn2_1$, Z=4	1.9	4.69	0.019	7.4	-85.98 (-87.0)	0.8
1,4-dinitrobenzene [d]	$P2_1/n$, Z=2	3.6	4.69	0.014	2.0	-91.4 (-92.9)	0.0
BEYDOY 1,2,4-trinitrobenzene	$Pca2_1$, Z=4	2.5	3.19	0.054	4.8	-109.8	2.0
ACANIL02 Acetanilide	Pbca, Z=8	2.8	0.96	0.015	3.2	-83.4 (-80.8)	0.0
ACTOLD p-Acetotoluidine	$P2_1/c$, Z=4	1.1	2.56	0.012	3.9	-84.8	0.0
DIXDUJ o-nitroacetanilide	$P2_1/n$, Z=4	3.4	8.31	0.025	8.4	-89.1	0.0

[a] For symbols and units, see footnotes to Table 1. [b] Average deviation between torsional angles in the experimental and the minimized experimental structures (degrees). [c] Herbstein, F.H.; Kapon, M. Acta Cryst. 1990, B46, 567. [d] Tonogaki, M.; Kawata, T.; Ohba, S. Acta Cryst. 1993, B49, 1031.

Table 4. Results of the evaluation of the potential parameters for the molecules analogous to molecule XI [a]

CSD ref. code, name	space group	Δ	t	θ	E	ΔE
BETLER 2,5-Dimethylpiperazine	$P2_1/c$, Z=2	3.3	0.017	3.20	-101.46	0.0
ZBCNON 3-Azabicyclo[3.2.2]nonane	Aba2, Z=8	2.1	0.022	8.46	-67.53 (57.74)	0.54
TIMCHX cis-(1,2:3,4:5,6)-Tri-imino-cyclohexane	$P2_1/c$, Z=4	2.1	0.012	1.62	-139.24	0.0
piperidine [b]	$P2_1/c$, Z=4	1.2	0.024	2.01	-58.95	0.0

a) For symbols and units, see footnotes to Table 1; b) Parkin, A.; Oswald, I.D.H.; Parsons, S. Acta Cryst. 2004, B60, 219.

Table 5. Nonbonded parameters used for crystal structure prediction of molecule VIII

	A, $kJ\cdot mol^{-1}\cdot Å^{6}$	B, $kJ\cdot mol^{-1}$	C, $Å^{-1}$
aliphatic H	58.12	6559.3	3.81
amide H [a]	67.11	8628.2	4.75
aliphatic C	2722.32	129156.3	3.16
amide C [a]	65.10	87463.6	4.49
amide N [a]	1314.69	290864.1	3.71
amide O [a]	1971.33	62625.7	3.38

[a] Preliminary parameters obtained for amides and refined using the VMC method. The updated values for amides are presented in [13]

Table 6. Nonbonded parameters used for crystal structure prediction of molecule IX

	A, $kJ\cdot mol^{-1}\cdot Å^{6}$	B, $kJ\cdot mol^{-1}$	C, $Å^{-1}$
H…H	115.02	11190.1	3.88
H…C	494.17	44011.5	3.63
H…C’ [a]	620.69	51036.4	3.65
H…O	257.99	39833.4	3.93
H…I	1227.63	121172.4	3.51
C…C	2123.09	173099.6	3.39
C…C’	2666.71	200728.2	3.40
C…O	1108.34	156665.7	3.69
C...I	5274.27	476575.6	3.27
C’...C’	3349.54	232766.8	3.41
C’...O	1392.14	181671.4	3.69
C’...I	6624.74	552642.4	3.28
O...O	578.61	141792.4	3.99
O...I	2335.01	259786.7	3.56
I...I	13102.49	1207502.8	3.14

[a] the C’ atom type was used for carbonyl carbon

Table 7. Nonbonded parameters used for crystal structure prediction of molecule X

	A, kJ·mol^{-1}·Å^{6}	B, kJ·mol^{-1}	C, Å^{-1}
aliphatic H	58.12	6559.3	3.81
aromatic H	174.68	14192.5	3.88
amide H [a]	21.05	8443.3	4.86
aliphatic C	2722.32	129156.3	3.16
aromatic C	1864.59	210708.3	3.39
amide C [a]	0.00	30282.5	3.58
amide N [a]	787.85	354528.3	3.75
nitro N	4075.76	273435.7	3.53
amide O [a]	11053.96	22524.9	2.48
nitro O	497.64	19254.3	3.26

[a] Preliminary values obtained earlier for amides. The updated parameters for amides are presented in [13]

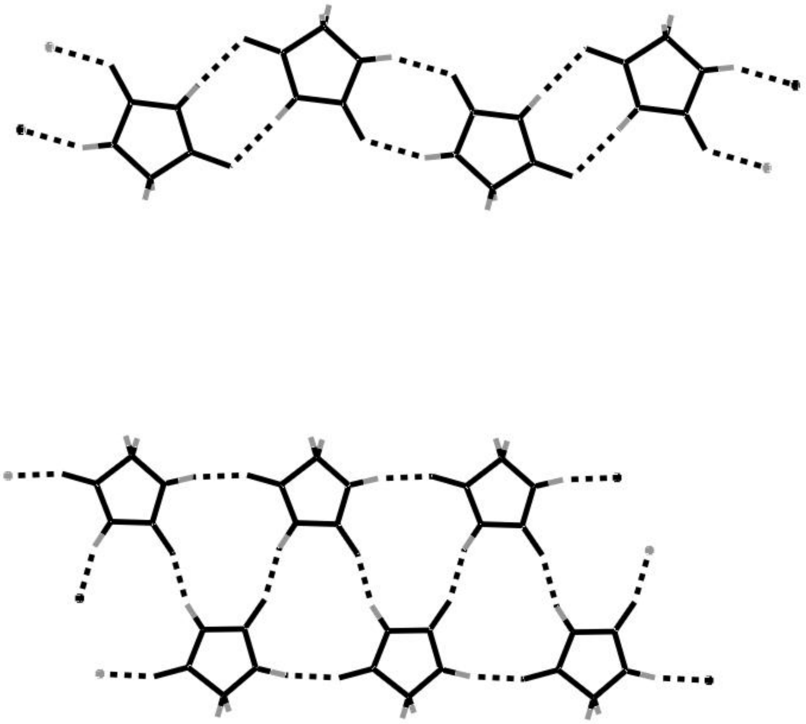
Table 8. Rotational barriers for molecule X

torsional angle		V, kJ/mol	
		calc.	exp.
τ_2, τ_3	exp.	12.3	12.6±6.3; 11.7; 13.8
τ_1	ab initio	0.0	~ 0.0
ϕ	ab initio	17.2	n/a
ω	ab initio	108.2; cis-trans~10.9	94.6-97.1; cis-trans 10.5-12.6
ψ	exp.	0.6	0.6

Table 9. Nonbonded parameters used for crystal structure prediction of molecule XI

	A, kJ·mol^{-1}·Å^{6}	B, kJ·mol^{-1}	C, Å^{-1}
aliphatic H	58.12	6559.3	3.81
amines H [a]	19.16	9825.3	4.75
aliphatic C	2722.32	129156.3	3.16
amide N[a]	21142.17	18558.1	2.09

[a] Preliminary parameters obtained earlier for amines. The updated values are presented in [13]



a)

b)

Figure 1. H-bond motifs predicted for Molecule VIII: a) the experimental structure and minimum 5 found by global search; b) minima 1-4 found by global search.

Martin U. Schmidt (Calculations by CRYSCA)

Institute for Inorganic and Analytical Chemistry, University of Frankfurt, Marie-Curie-Str. 11, D-60439 Frankfurt am Main, Germany, e-mail: m.schmidt@chemie.uni-frankfurt.de

Overview:

Molecule VIII: not treated, because crystal data were partially published.

Molecule IX: not treated, because I had access to the experimental data

Molecule X: - Force field development for the intramolecular torsion
- Lattice energy minimisations
- Experimental structure found, but with very high ranking

Molecule XI: - Lattice energy minimisations
- Experimental structure not found, because no search with Z'=2
- Post-minimisation of the experimental structure: Good energy and ranking

General procedure:

Lattice energy minimisations

All lattice energy minimisations were done with our program CRYSCA (Schmidt, 1995; Schmidt & Englert, 1996; Schmidt, 1999).

The program CRYSCA ("Crystal Structure Calculations") performs global energy optimisations for flexible molecules, starting from random crystal structures. The starting set consists of several thousand up to more than 100 000 crystal structures (for X) with random values for lattice parameters, orientation and position of the molecules as well as for the intramolecular degrees of freedom. The user selects, which intramolecular degrees of freedom (distances, angles or rotations) are to be considered. All starting values are inside sensible ranges. The crystal symmetry is included from the beginning. All crystallographic symmetries can be handled, including scarce space groups, molecules on special positions, supersymmetries, disorders etc.. For the crystal structure predictions, where the space group is not known, all common space groups are tested separately. If the molecule contains intramolecular symmetry, additional runs with molecules on special positions can be carried out. For the blind test, the energy minimisations were carried out in the most popular space groups with one independent molecule per asymmetric unit.

The energy is minimised by a special steepest-descent algorithm. The minima are sorted according to energy, and checked for higher symmetries, meaningful molecular conformations and reliable intermolecular interactions. In several cases, higher symmetries are found.

Generally the packings having the lowest energies are regarded as 'predicted' crystal structures.

Force field and charges

The lattice energy is calculated by CRYSCA by the formula

$$E = \frac{1}{2} \sum_{i} \sum_{j} \left(-A_{ij} r_{ij}^{-6} + B_{ij} e^{-C_{ij} r_{ij}} + \frac{1}{4\pi\varepsilon\varepsilon_0} \frac{q_i q_j}{r_{ij}} \right) + E_{intramol}$$

where A, B, and C are empirical van der Waals parameters (see Table 1), q resembles the atomic charge, and r_{ij} stands for the interatomic distance between the atoms i and j. The dielectric constant ε is set to 1.0.

Table 1. Parameters for the van der Waals potential (Schmidt, 1995; Schmidt & Englert, 1996; Schmidt, 1999)

Atoms	*A* (Å^6·*kJ*/*mol*)	*B* (*kJ*/*mol*)	*C* (Å^{-1})
C···C	2377.0	349908	3.60
C···H	523.0	36677	3.67
H···H	144.2	11104	3.74
N···N	1240.7	201191	3.78
N···C	1483.6	247571	3.73
N···H	407.4	64467	4.00
O···O	1242.6	372203	4.18
O···C	1718.6	360883	3.89
O···H	423.3	64288	3.96
O···N	1241.6	273649	3.98
H[a]···H[a]	0	0	0
H[a]···H	144.2	11104	3.74
H[a]···C	523.0	36677	3.67
H[a]···N	0	0	0
H[a]···O	0	0	0

[a]) hydrogen of OH and NH groups

The charges are calculated for compound X by the EHT charge-iteration procedure using the program ICON (Howell *et al.*, 1977). Normal EHT charges are known to be by far too high. The EHT charge-iteration procedure had been developed in order to get more reliable charges. In this procedure, the diagonal Hamilton matrix elements are given by H_{ii} = -VSIE(q) where VSIE(q) is the valence state ionization energy of orbital i when the atom has total charge q. The off-diagonal elements H_{ij} are calculated in the usual manner using the formula of Wolfsberg and Helmholz (1952). The VSIE(q) functions are assumed to be of the form

$$\mathrm{VSIE}(q) = a\,q^2 + b\,q + c$$

where a, b, and c are parameters which depend on the atoms and on the orbital. We used the a, b, and c values given by McGlynn et al. (1972), which were recommended also for the use in the ICON program (Howell *et al*., 1977).

The resulting charges are more or less comparable to the Gasteiger charges. Test calculations on small organic and organometallic molecules show, that molecular dipole moments are reproduced well in most cases (Schmidt, 1995).

These charges from the charge-iteration EHT procedure have to be scaled by 1.1 to fit with the CRYSCA can der Waals potential (Schmidt, 1995).

In order to get a better description of hydrogen bond interactions for compound X, the charge separation in the NH group was increased manually: The charges on the nitrogen and hydrogen atom were enhanced by -0.2 *e* and +0.2 *e*, resp.

For XI, the charges were calculated by the Gasteiger method. Since the charge on the acid hydrogen seemed to be too low (+0.1222), the charges on the nitrogen and hydrogen atom were enhanced by -0.1 *e* and +0.1 *e*, resp.

Molecule VIII:

Calculations on molecule VIII were not performed, because the (partial) publication of the experimental crystal structure was distributed before I started the calculations.

Molecule IX:

This molecules was not treated because:

(1) Lack of reliable parameters for iodine atoms

(2) I had access to the experimental data: The X-ray structure determination was carried out by E.F. Paulus, my former colleage from Hoechst AG. It was me, who submitted the experimental data to A.L. Speck as a candidate molecule for the blind test.

I regarded this molecule on one hand as very difficult for most groups (because of the iodine atoms), but on the other hand as easy, because it is a rigid, flat molecule having 2/m symmetry; thus surely occuping a crystallographic inversion centre in the solid state.

If I tried to predict the crystal structure, I would probably have restricted the search to *P*-1, *Z*=1, $P2_1/c$, *Z*=2 and *Pbca*, *Z*=4 (which are the most common space groups for molecules with inversion centres), and additionally *Pnma*, *Z*=4 (for the case, that the mirror plane is kept in the crystal structure).

Molecule X:

Molecular geometry

For molecule X extensive CSD searches were carried out in order to get knowledge of the preferred conformations of molecules of this type. Additionally ab initio calculations were carried out on a high level (B3LYP/6-311+G(d,p) // B3LYP/6-31G(d)) by Dr. Thomas Müller (Institute for Inorganic and Analytical Chemistry, University of Frankfurt) using the Gaussian package.

Finally the following geometry was used:

- Bond lenghts and angles were taken from CSD. The geometry was idealised. The C-H bond lengths were set to 1.04 Å, the N-H distance set to 1.01 Å, because these values were used in the fitting of the force field parameters by Williams (1967).

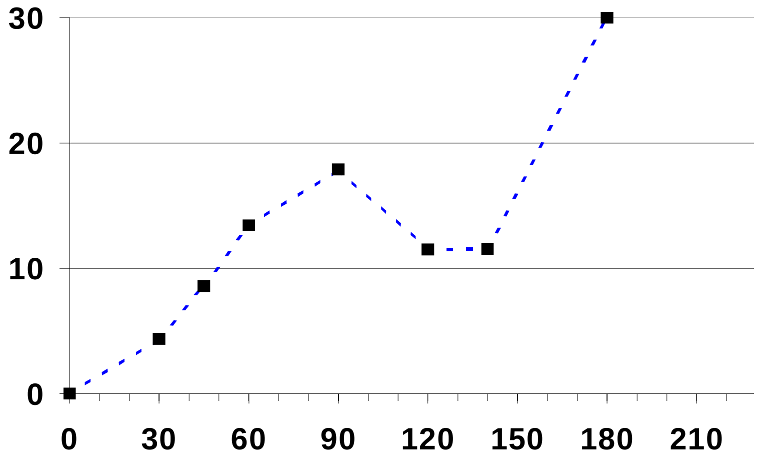
- The torsion angle ψ (H-CH$_2$-C-N) was set to 0. The torsion angle τ_1 (H-CH$_2$-C2-C1) was set to 180°. Both settings were done according to CSD searches and ab initio calculations.

- For the rotation around the peptide bond (ω), the CSD clearly showed, that only a planar *trans* conformation of the CO and the NH group has to be considered, thus ω was set to 180°.

The parameters ψ, τ_1, and ω were kept fixed in the lattice energy minimisations.

Development of an intramolecular potential for the rotation around φ (Phenyl-amin bond)

For the rotation around the phenyl – amin bond (φ) CSD searches lead to the result, that in the case of molecules with an *ortho*-methyl group on the benzene ring the torsion angle φ (CH-C-N-CO) can adopt all values between -144 and +144 degree; values around +/- 180° are not possible because of steric hindrance between the CO group and the methyl group. This result was confirmed by ab the initio calculations with fixed values of φ (the geometries were optimized on B3LYP/6-31G(d) level) followed by energy calculations on B3LYP/6-311+G(d,p) level):



E(*kJ/mol*) vs. torsional angle φ

These data points were fitted by a 6-term cosine series:

$E(\varphi) = 16.602 - 5.778 \cos(\varphi) - 4.902 \cos(2\varphi) - 7.753 \cos(3\varphi) + 4.046 \cos(4\varphi) - 1.316 \cos(5\varphi) - 0.594 \cos(6\varphi)$ [*kJ*/*mol*]

Since several years CRYSCA is able to handle such analytical intramolecular potentials, including analytical calculation of the first derivatives for the use in the optimisation procedure.

Development of an combined intramolecular potential for τ_2 and τ_3 (Nitro groups)

In the case of the two nitro groups, the CSD shows clearly that the torsions described by the torsion angles τ_2 (C3-C4-N-O) and τ_3 (C4-C5-N-O) are not independent: A single nitro group tends to be coplanar to the benzene ring, but two neighbouring nitro groups cannot simultaneously be in-plane. If one nitro group is coplanar to the benzene ring ($\tau = 0°$), the other is almost perpendicular (τ about 80°) and vice versa. A situation with both nitro groups in intermediate conformation (both τ about 30 to 45°) is possible as well. In all conformations the sum of both torsion angles remains approximately constant: $\tau_2 + \tau_3 \approx 75°$.

The ab initio calculations (B3LYP/6-31G(d)//B3LYP/6-311+G(d,p)) show, that for a molecule in the gas phase the torsion angles τ_2 and τ_3 additionally depend on the value of the rotation angle φ of the acetylamino substituent (!):

$\varphi = 0°$: $\tau_2 = 31°$, $\tau_3 = 47°$

$\varphi = 60°$: $\tau_2 = 38°$, $\tau_3 = 37°$

$\varphi = 120°$: $\tau_2 = 41°$, $\tau_3 = 34°$

The rule $\tau_2 + \tau_3 \approx 75°$ holds as well, and a slight preferrence of valules between 30 and 45 ° against 0 / 80° is visible, but the latter condition seems to be weak.

This energy hypersurface was manually fitted using a combined potential for τ_2 and τ_3:

$E(\tau_2, \tau_3) = 0.02 \cdot (\tau_2 + \tau_3 - 75°)^2 + 2 + \cos(4\cdot\tau_2) + \cos(4\cdot\tau_3)$ [*kJ*/*mol*]

This potential was programmed into CRYSCA, including analytical first derivatives.

Additional intramolecular van der Waals or Coulomb interactions were not used, since they would disturb the manually fitted potential.

Lattice energy minimisations

Calculations were performed in $P1$, $P\text{-}1$, $P2_1$, Pc, Cc, $P2_1/c$, $C2/c$, $P2_12_12_1$, $Pna2_1$, and $Pbca$, each with one independent molecule per asymmetric unit.

The minima having lowest energies are:

Table 2.

Rank	SG, Z	E / kJ/mol	a / Å	b / Å	c / Å	β/°	V/Å^3	ϕ / °	τ2 / °	τ3 / °	H bonds
1	$P2_12_12_1$ Z=4	-198.86	6.682	9.642	15.526	90	1000.2	-5.92	20.84	52.05	bifurcated to two NO_2
2	$P2_1/c$ Z=4	-195.87	7.848	12.770	11.415	120.95	981.1	11.79	-92.20	162.80	N-H...O=N
3	$P2_1/c$ Z=4	-195.46	9.599	10.799	9.816	105.92	978.6	-5.06	51.50	24.35	N-H...O=C
4	*Pbca* Z=8	-194.33	9.245	17.270	12.465	90	1990.1	26.73	42.58	24.11	N-H...O=C

All low-energy structures were manually inspected for reliable conformations and intermolecular interactions. Structure 1 seemed somehow suspicious because of the bifurcated hydrogen bond, but the calculated energy was considerably lower than for structures 2 to 4. The structure with energy rank 3 has a very long hydrogen bond (H...O = 2.42 Å, angle (N-H...O) = 156°), and was regarded as unreliable. The structures with energy ranks 1, 2, and 4 were submitted as “predictions”.

Comparison with the experimental structure

The experimental structure was found in the list of predicted structures, but with a very high ranking and a large energy difference (E = -171.13, ΔE = 27.74 kJ/mol) to the global minimum:

Table 3.

	SG, Z	a / Å	b / Å	c / Å	β/°	V/Å^3	ϕ / °	τ2 / °	τ3 / °	N...O in H br.
Predicted	$P2_1/c$ Z=4	13.016	4.847	19.634	125.60	1007.0	-52.19	14.20	60.60	2.832
Transformed to $P2_1$/n	$P2_1/n$ Z=4	13.016	4.847	16.042	95.67	1007.0	-52.19	14.20	60.60	2.832
Exp.	$P2_1/n$ Z=4	12.569	4.853	17.266	99.16	1039.8	-41.44	25.84	51.95	2.864

The geometric accuracy of the calculated structure is fair; the geometry of the hydrogen bond is excellent. But the energy difference is too high. The reason remains unclear. (The intramolecular energy ($E(\varphi) + E(\tau_2, \tau_3)$) is much smaller than 27.7 kJ/mol; the charges are not so different from charges used in other force fields; and the van der Waals force field of

CRYSCA gives generally good results). A problem might be the parametrisation of the intermolecular interactions of the nitro group.

Table 4. Atomic coordinates and charges of the predicted packing (high ranking), which turned out to correspond to the experimental structure of molecule X.

Coordinates are given for the $P2_1/n$ setting (a = 13.016 Å, b=4.847 Å, c=16.042 Å, β = 95.67°)

Atom	x	y	z	charge
C1	0.01009	-0.17983	0.27323	0.11632
C2	0.04924	-0.31179	0.20624	0.06570
C3	0.00447	-0.26670	0.12508	0.06247
C4	-0.07943	-0.08960	0.11092	0.14367
C5	-0.11855	0.04254	0.17793	0.14625
C6	-0.07381	-0.00270	0.25906	0.06936
H3	0.03374	-0.36555	0.07495	0.01511
H6	-0.10311	0.09601	0.30917	0.02030
C7	0.13990	-0.50354	0.22071	0.02731
H7a	0.12060	-0.67774	0.23756	0.00737
H7b	0.20387	-0.38352	0.27099	0.00737
H7c	0.15802	-0.58210	0.16365	0.00274
N1	0.05630	-0.22741	0.35563	-0.21264
C8	0.09052	-0.03793	0.41342	0.17087
O8	0.08564	0.21142	0.40123	-0.35582
H1	0.06463	-0.42691	0.37321	0.27724
C9	0.13624	-0.15375	0.49593	0.04741
H9a	0.22377	-0.06398	0.52671	0.01491
H9b	0.08325	-0.10922	0.52429	0.01491
H9c	0.13330	-0.36807	0.49377	0.01124
N4	-0.12411	-0.04859	0.02409	0.19828
O41	-0.09999	-0.20896	-0.02919	-0.27115
O42	-0.18334	0.14393	0.00917	-0.27115
N5	-0.20697	0.23170	0.16616	0.20737
O51	-0.28869	0.14236	0.13331	-0.25771
O52	-0.19472	0.46961	0.18975	-0.25771

Molecule XI:

Molecular geometry

CSD searches gave no definite results for the molecular geometry of an unsubstituted azetidine molecule. Many structures in the database have bulky substituents, which can lead to distortions of the 4-membered ring and do lead to an increase of bond lengths. Also the position of the hydrogen attached to the nitrogen atom was not clear (axial or equatorial). Hence we calculated the molecular geometry by ab initio methods on 6-31++G(d,p) level. The NH group turned out to prefer the equatorial orientation. Finally, the C-H and N-H bond lengths set to 1.04 Å, and 1.01 Å, resp.. The geometry was kept fixed in the lattice energy minimisations.

Lattice energy minimisations

Calculations were performed in *P*1, *P*-1, *P*2_1, *P*c, *Cc*, *P*2_1/*c* , *C*2/*c*, *P*$2_12_12_1$, *Pna*2_1, and *Pbca*, each with one independent molecule per asymmetric unit.

The man power did not allow for performing calculations with Z’ > 1, although this is possible with CRYSCA.

The resulting list of possible crystal structures contains many minima within a small energy range (10 minima within 1 kJ/mol above the global minimum, 14 minima within 2 kJ/mol).

All low-energy structures were manually inspected for reliable conformations and intermolecular interactions. As expected, they all have hydrogen bonds of the type N-H...N.

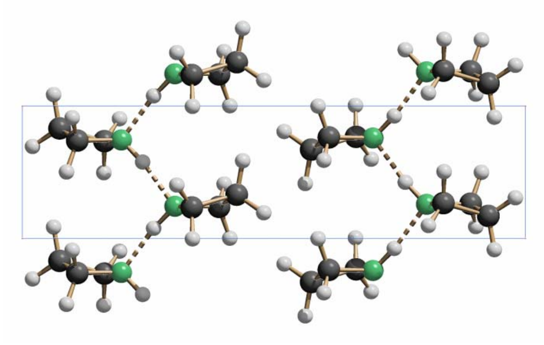
The minima with lowest energies are:

Table 5.

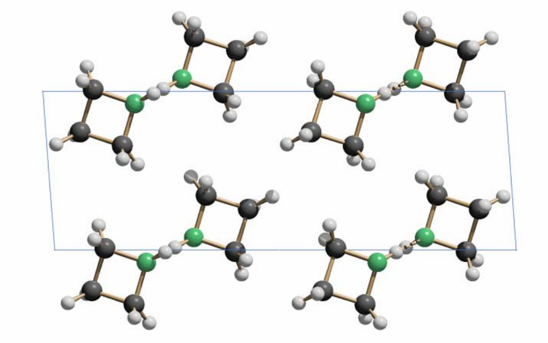
Rank	E/kJ/mol	SG	Z	*a* / Å	*b* / Å	*c* / Å	β / °	V/Å^3	d(H...N)/Å
1	-41.88	*P* 2_1/*c*	4	5.424	4.148	15.770	94.24	88.5	1.77
2	-41.85	*P* 2_1/*c*	4	8.460	4.111	9.964	90.52	86.6	1.95
3	-41.77$^{(a)}$	*P* 2_1	2	5.416	4.141	7.925	93.88	88.7	1.77
Experimental:		*P* 2_1/*c*	8	9.507	9.122	9.790	117.47	94.2	(2.27,2.30)$^{(b)}$
Experimental structure optimized by CRYSCA using our molecular geometry and FF:									
~15	-39.87	*P* 2_1/*c*	8	9.821	8.633	9.966	120.99	90.6	1.92,1.99

(a) Comment: Polar structure. The Coulomb energy was calculated by direct summation (box of 11 · 201 · 11 unit cells). The total Coulomb energy depends on the crystal shape (probably needles) and on surface charges.
(b) N...H distances not comparible because of shorter N-H bond distance.

The structures with energy ranks 1 to 3 were submitted as guesses for the crystal structure prediction test. In structure 1 the molecules are arranged in zigzag chains along the 2_1 axes:



Minimum 1, view direction [100]



Minimum 1, view direction [010]

The packings of structures 2 and 3 are similar to the packing of structure 1.

Comparison with the experimental structure

The experimental structure could was not found, because we made no search with Z'=2.

A post-minimisation of the experimental structure, using the same geometry and force field as in the blind prediction step lead to a packing with an energy only 2.01 kJ/mol above the global minimum, and a good geometric accuracy. This packing would have had a ranking of about 15.

Hence, CRYSCA would have been able to predict this structure as a possible crystal structure, if the search had been performed with Z'=2.

Table 6. Post-minimization of the experimental structure, using the same molecular geometry and force field as for the blind prediction.

Atom	x	y	z	charge
Molecule 1:				
N1	0.34283	0.18157	0.61828	-0.4165
H2	0.43520	0.12787	0.70862	0.2222
C3	0.17758	0.13958	0.57606	-0.0037
H4	0.16478	0.02185	0.60656	0.0428
H5	0.12383	0.21957	0.61854	0.0428
C6	0.30771	0.15146	0.45557	-0.0037
H7	0.34399	0.03823	0.43767	0.0428
H8	0.35022	0.23999	0.40982	0.0428
C9	0.12660	0.15514	0.39957	-0.0281
H10	0.05457	0.06288	0.32273	0.0293
H11	0.07370	0.26638	0.35166	0.0293
Molecule 2:				
N31	0.36673	0.51733	0.63755	-0.4165

H32	0.36026	0.40195	0.64541	0.2222
C33	0.25052	0.61826	0.65114	-0.0037
H34	0.13703	0.56334	0.61429	0.0428
H35	0.29635	0.67058	0.76480	0.0428
C36	0.32622	0.58786	0.48414	-0.0037
H37	0.24137	0.52262	0.38155	0.0428
H38	0.42781	0.61726	0.47495	0.0428
C39	0.24413	0.72384	0.52011	-0.0281
H40	0.12774	0.75773	0.42667	0.0293
H41	0.32068	0.82355	0.56800	0.0293

References:

Schmidt, M.U. (1995). *Kristallstrukturberechnungen metallorganischer Molekülverbindungen*, Verlag Shaker, Aachen, Germany, ISBN 3-8265-0588-3.

Schmidt, M.U. (1999). *Crystal Engineering: From Molecules and Crystals to Materials* (Ed: D. Braga, F. Grepioni, A. G. Orpen), Kluwer Academic Publishers, Dordrecht, pp. 331-348.

Schmidt, M.U. and Englert, U. (1996). *J. Chem. Soc. Dalton Trans.* 1996, 2077-82.

Williams (1967). *J. Chem. Phys.* 47, 4680-4.

Wolfsberg, M., Helmholz, L. (1952). *J. Chem. Phys.* 20, 837-43.

Howell, J., Rossi, A., Wallace, D., Haraki, K. & Hoffmann, R. (1977). Program ICON (FORTICON Version 8, QCPE program no. 517), Quantum Chemical Programs Exchange 11, 344.

B. Schweizer

Methodology used for prediction mol VIII, X and XI. Molecular dimensions were taken from results of an ab initio calculation (Gaussian98, 6-31*). Atomic point charges were used to fit the electrostatic potential. The ZipPromet module of OPiX (A. Gavezzotti, OpiX, A computer program package for the calculation of intermolecular interactions and crystal energies, University of Milano, 2004) was used to search structures in P-1 , P21 P21/c, C2/c, P212121 and Pbca all with Z'=1. The 2000 to 5000 starting structures in each spacegroup were minimized with Minop for 100 cycles removing duplicates and those 30 kJ higher in energy as the minimum (60 kJ for first round) after every round. The refinement was terminated close to the energy minimum when no significant change in energy was observed.

For molecule X a CSD search showed 3 structures with $\phi = 0$ of 5 hits, so this conformation was selected for the Gaussian98 geometry optimization. The torsion angles in mol X: $\phi= 0$, $\omega= 180$, $\psi= 180$, $\tau1 = 180$, $\tau2 = 30$, $\tau3=47$.

P. Verwer

Crystal structure predictions for molecules VIII, IX, and XI, using the implementation of the Polymorph Predictor in Materials Modeling (Accelrys).

Software/hardware

In the previous blind tests, CSP1999 and CSP2001, the Polymorph Predictor module in MSI's Cerius^2 software was used. For CSP2004 it was decided to use the successor to Cerius^2, called Materials Modeling, release 3.0.1 (January 2003) [1], which should provide (at least) equivalent functionality, while running on PC-hardware. The calculations were carried out on the 40 node, 80 CPU linux cluster of the Institute of Pharmaceutical Innovation (IPI) at Bradford University, UK, in collaboration with F.J.J. Leusen (IPI) and C. Liang (Accelrys).

Simulations

The method used for crystal structure predictions in Materials Modeling is essentially identical to the approach that was used in Cerius^2, and which has been described in detail elsewhere [2]. Structures are generated in a Monte Carlo simulation, for a given space group and number of molecules in the asymmetric unit. Ranking of the generated structures is based on the Molecular Mechanics energies of the structures after optimization of the cell parameters and atomic coordinates (i.e. allowing for flexible molecules).

Because none of the molecules in this test is chiral, their crystal structures can in principle be in any space group. For practical reasons, only the 12 most commonly occurring space groups were searched, i.e. $P2_1/c$, P-1, $P2_12_12_1$, $P2_1$, C2, Pbca, Pnma, $Pna2_1$, Pbcn, P1, $Pca2_1$, and Cc. Although runs with $Z' \neq 1$ were planned, technical difficulties did not allow us enough time to carry out these runs; for the same reason, no predictions for molecule X were done. Thus, only runs for VIII, IX, and XI were done, in 12 space groups, with $Z'=1$. Runs were repeated until no new structures were generated that would rank in the top 10 according to their packing energy. Finally, the symmetry of the structures having the lowest energy was checked: in many cases, the intra-molecular symmetry present, or symmetry caused by the positioning of the molecules leads to a space group of higher symmetry. For example, structures for molecule VIII predicted in $P2_1$ can in fact have $Cmc2_1$ symmetry, and $Pna2_1$ can sometimes be reduced to Pnma.

In the previous blind tests, the Dreiding 2.21 force field [3] has been used. Roughly speaking, this force field ranked the correct structure in the top ten in about half the cases. Although this has certainly not been a disappointing result, for the current blind test it was decided to use the COMPASS force field version 99.01 [4] whenver possible, as it is specifically directed at solid state systems. COMPASS comes with its own atomic charges. In cases where COMPASS did not have parameters for all atom types present in the molecule, Dreiding 2.21 was used, and atomic charges were obtained by fitting to the electrostatic potential as obtained from a quantum mechanical calculation. In the molecular mechanics optimizations, electrostatic and van der Waals interactions were calculated using Ewald summation.

Compound VIII

Structures were predicted in the 12 most common space groups, with Z'=1. Because the COMPASS force field lacks parameters for some atom types, the Dreiding 2.21 force field was used, in combination with point charges fit to the electrostatic potential using the ChelpG method [5], based on a HF 6-31+G* geometry optimization carried in the Gaussian94 program [6]. Two runs were carried out, the last one producing no more low-energy structures.

Compound IX

Structures were predicted in the 12 most common space groups, with Z'=1. Because the COMPASS force field lacks parameters for Iodine, the Dreiding 2.21 force field was used, in combination with charges fit to the electrostatic potential using the ChelpG method, based on a HF 3-21G geometry optimization carried in the Gaussian94 program.

In initial runs rather short C…I were observed. The bonded distance for aromatic C to I was therefore increased from 0.205 to 0.210 nm, and the van der Waals radius of I was increased from 0.415 to 0.500 nm. Two prediction runs were carried out in each space group.

Compound X

In view of time constraints it was decided that no meaningful predictions could be carried out for this rather flexible molecule.

Compound XI

Structures were predicted in the 12 most common space groups, with Z'=1, using the COMPASS force field. Three prediction runs were carried out for $P2_1/c$, Cc, and Pbcn, two runs for the 9 other space groups.

References:

[1] Materials Studio Modeling, Accelrys Inc., 9785 Scranton Road, San Diego, CA.

[2] Leusen, F. J. J., Wilke, S., Verwer, P., & Engel, G. E. (1999). *Implications of Molecular and Materials Structure for New Technologies*, NATO Science Series E, Vol 360, edited by J. A. K. Howard, F. H. Allen & G. P. Shields, pp 303-314. Dordrecht: Kluwer Academic Publishers.

[3] Mayo, S. L. Olafson, B. D. & Goddard III, W. A. (1990). *J. Phys. Chem.* **94**, 8897-8909.

[4] Sun, H; Rigby, D., Spectrochimica Acta (A) , 1997, 53, 1301.

[5] Breneman, C. M., & Wiberg, K. B. (1990). *J. Comput. Chem.* **11**, 361-373.

[6] Frisch, M. J., Trucks, G. W., Schlegel, H. B., Gill, P. M. W., Johnson, B. G., Robb, M. A., Cheeseman, J. R., Keith, T., Petersson, G. A., Montgomery, J. A. Raghavachari, K., Al-Laham, M. A., Zakrzewski, V. G., Ortiz, J. V., Foresman, J. B., Cioslowski, J., Stefanov, B. B., Nanayakkara, A., Challacombe, M., Peng, C. Y., Ayala, P. Y., Chen, W., Wong, M. W., Andres, J. L., Replogle, E. S. Gomperts, R., Martin, R. L., Fox, D. J., Binkley, J. S., Defrees, D. J., Baker, J., Stewart, J. P., Head-Gordon, M., Gonzalez, C., & Pople, J. A. (1995). *Gaussian94*, Revision D.4.Gaussian Inc., Pittsburg PA, USA.