research papers

Blind crystal structure prediction of a novel second polymorph of 1-hydroxy-7-azabenzotriazole

Harriott Nowell,[a,b]
Christopher S. Frampton,[c] **Julie Waite**[c] **and Sarah L. Price**[a]*

[a]Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, England, [b]Diamond Light Source Ltd, Diamond House, Chilton, Didcot, Oxfordshire OX11 0DE, England, and [c]Pharmorphix Ltd, 250 Cambridge Science Park, Milton Road, Cambridge CB4 0WE, England

Correspondence e-mail: s.l.price@ucl.ac.uk

The commercially available peptide coupling reagent 1-hydroxy-7-azabenzotriazole has been shown to crystallize in two polymorphic forms. The two polymorphs differ in their hydrogen-bonding motif, with form I having an $R_2^2(10)$ dimer motif and form II having a $C(5)$ chain motif. The previously unreported form II was used as an informal blind test of computational crystal structure prediction for flexible molecules. The crystal structure of form II has been successfully predicted blind from lattice-energy minimization calculations following a series of searches using a large number of rigid conformers. The structure for form II was the third lowest in energy with form I found as the global minimum, with the energy calculated as the sum of the *ab initio* intramolecular energy penalty for conformational distortion and the intermolecular lattice energy which is calculated from a distributed multipole representation of the charge density. The predicted structure was sufficiently close to the experimental structure that it could be used as a starting model for crystal structure refinement. A subsequent limited polymorph screen failed to yield a third polymorphic form, but demonstrated that alcohol solvents are implicated in the formation of the form I dimer structure.

1. Introduction

Crystal structure prediction (CSP) is of great importance for industries involved in the manufacture of small organic molecular materials in the solid state (Bernstein, 2002; Price, 2004). The development of techniques for CSP of flexible molecules is necessary as industrially relevant molecules usually have some degree of flexibility. A flexible molecule can distort away from the optimal (gas phase) conformation if the resulting intermolecular interactions in the solid-state crystal structure give rise to a sufficiently low lattice energy to compensate for the intramolecular energy penalty associated with the suboptimal conformer. Flexible-molecule CSP is particularly challenging as it is necessary to evaluate the intramolecular energy penalty in addition to the usual challenge of finding and calculating the global minimum in lattice energy, which remains difficult (Gavezzotti, 2002). The low success rate for the flexible molecules included in the 'blind tests' of CSP (Lommerse *et al.*, 2000; Motherwell *et al.*, 2002; Day *et al.*, 2005) organized by the Cambridge Crystallographic Data Centre (CCDC) is, therefore, perhaps unsurprising.

Flexible-molecule CSP can be approached using empirical atom–atom force fields to simultaneously model the intra- and intermolecular interactions (Verwer & Leusen, 1998), but this can sometimes lead to non-physical molecular distortion (Brodersen *et al.*, 2003). The huge computer resources required to implement recently proposed methods, in which

high-quality *ab initio* force fields are used to refine both intra- and intermolecular interactions during optimization (van Eijck *et al.*, 2001), mean that this is only really practicable for a limited range of low-energy structures of small molecules in the foreseeable future.

The approach taken here is to perform a large number of searches, each with a different high-quality *ab initio* rigid conformer, and is similar to the approach used recently in another 'blind' prediction of the flexible molecule piracetam (Nowell & Price, 2005). Approximately 750 high-density structures are lattice-energy minimized for each conformer and structures from all searches are collated and ranked in terms of the sum of their intra- and intermolecular energies,

$$E_{\text{tot}} = \Delta E_{\text{intra}} + U_{\text{latt}}. \tag{1}$$

The most stable structures are generally those in which the conformation allows the formation of efficient hydrogen bonding in a close-packed lattice.

The novel crystal structure, form II, of 1-hydroxy-7-azabenzotriazole (HOAt), kept unpublished until now, was used as a suitable test case for a blind test of flexible-molecule CSP. The experimental team at Pharmorphix simply indicated that they had a new polymorph of HOAt and the computational team at UCL conducted the search for its structure knowing only the atomic connectivity. The structure, which is reported herein, was predicted as the third-lowest-energy crystal structure.

HOAt, Fig. 1, is of limited flexibility, with one variable torsion angle, C1—N1—O1—H1, denoted θ. Form I crystallizes in space group *C*2/*c* with $R_2^2(10)$ dimers forming antiparallel chains of molecules, while form II crystallizes in the space group *P*2$_1$/*c* with parallel molecules linked by *C*(5) hydrogen-bond chains. The room-temperature structure of form I has been reported previously (Hoffmann *et al.*, 1999); the structures of form I and form II reported herein are both low-temperature structures (123 K). The unit cells were also measured at room temperature to check for any phase changes that might have occurred on cooling.

1.1. Method

1.1.1. Experimental methods.
Colourless prisms of form I were obtained by re-crystallization of HOAt from a 1:1 aceto-nitrile:ethanol solution at 278 K. Colourless needles of form II
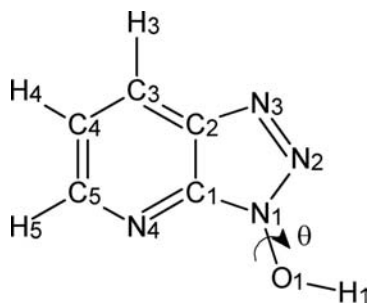
Figure 1
1-Hydroxy-7-azabenzotriazole (HOAt) with the atomic-numbering scheme.

were grown from a 2:1 acetonitrile:ethanol solution at 278 K. Crystal data for both forms were measured at low temperature (123 K) on a Bruker SMART 1K CCD diffractometer equipped with an Oxford Cryosystems Cryostream cooler with graphite-monochromated Mo *K*α radiation (λ = 0.71073 Å). The structures were solved and refined with *SHELXTL* (Bruker-AXS, 2000) with all H atoms freely refined. The crystal data for both crystal forms are summarized in Table 1.[1]

After the calculations had been analysed, a polymorph screen was carried out, starting with the commercially available (form II) material (Acros). Maturation crystallizations (three heat/cool cycles of 4 h at 278 K and 4 h at 323 K, over a 24 h period) were carried out from 22 different solvents using 25 mg ± 2 mg of HOAt and 1 ml of the selected solvents which included alcohols, ethers, ketones, hydrocarbons, dichloromethane, DMF and water. After maturation the solids were filtered and analysed by polarized light microscopy and X-ray powder diffraction. The filtrates were allowed to evaporate and any solids remaining were again analysed by polarized light microscopy and X-ray powder diffraction.

1.1.2. Computational methods.
The known form I structure was used to test the model and determine the quality of crystal structure reproduction that could be expected from the searches. (Corresponding calculations were performed on form II after the completion of the informal blind test.) The crystal structure was lattice-energy minimized using the experimental molecular structure (expminexp calculation) with H-atom positions corrected to values expected from neutron diffraction (Allen *et al.*, 1987). Subsequently, the molecular structure was *ab initio* optimized, to approximate the gas-phase structure, and used in a lattice-energy minimization of form I to understand the effect of conformational changes on the crystal structure (exptminopt calculation). This was repeated with θ constrained to the value observed in form I (expmincon calculation). The MP2 level of theory and a 6-31G(d,p) basis set within *GAUSSIAN* (Frisch *et al.*, 2004) were used to calculate the gas-phase conformers and perform a potential-energy surface (PES) scan of θ to survey the gas-phase intramolecular energy dependence. A 360° PES scan was performed using a series of partial optimizations with θ constrained and the rest of the molecule optimized, using a step size of 5°. This scan procedure confirmed that the ring remained sufficiently planar such that the same results could be expected for 180–360°, as found for 180–0°.

A search was carried out for each of a series of partially optimized conformers covering the θ range 0–180°, in 10° steps (a total of 19 searches). The associated distributed multipole analysis (DMA; Stone & Alderton, 1985) was obtained for each conformer using the *GDMA* algorithm (Stone, 1999). A *MOLPAK* search (Holden *et al.*, 1993) was performed for each rigid conformer producing densely packed crystal structures in eight common space groups (15 *MOLPAK* packing types); *P*$\bar{1}$, *P*2$_1$/*c*, *P*2$_1$2$_1$2$_1$, *Pbcn*, *Pbca*, *Cc*, *C*2 and

Table 1
Experimental data.

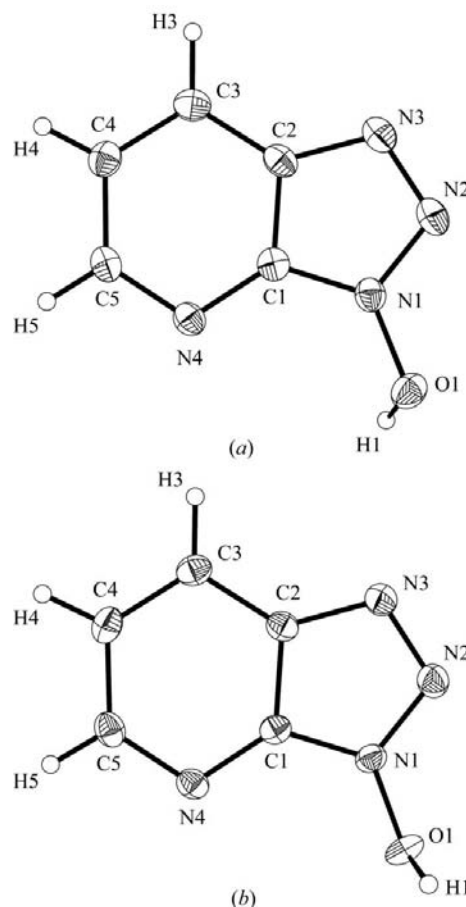|	Form I	Form II
Crystal data		
Chemical formula	$C_5H_4N_4O$	$C_5H_4N_4O$
M_r	136.12	136.12
Cell setting, space group	Monoclinic, $C2/c$	Monoclinic, $P2_1/c$
Temperature (K)	123 (1)	123 (1)
a, b, c (Å)	22.0427 (6), 3.9311 (1), 14.2340 (4)	3.7487 (1), 20.8881 (5), 7.0982 (2)
β (°)	114.8850 (10)	96.2570 (10)
V (Å^3)	1118.89 (5)	552.50 (3)
Z	8	4
D_x (Mg m^{-3})	1.616	1.636
Radiation type	Mo $K\alpha$	Mo $K\alpha$
No. of reflections for cell parameters	3604	4010
θ range (°)	3.0–29.1	3.0–29.1
μ (mm^{-1})	0.12	0.12
Crystal form, colour	Prism, colourless	Needle, colourless
Crystal size (mm)	$0.35 \times 0.30 \times 0.25$	$0.35 \times 0.10 \times 0.05$
Data collection		
Diffractometer	SMART 1 K CCD area detector	Bruker 1 K CCD area detector
Data collection method	Narrow-frame ω scans	Narrow-frame ω scans
Absorption correction	None	None
No. of measured, independent and observed reflections	4460, 1140, 1038	4697, 1133, 1049
Criterion for observed reflections	$I > 2\sigma(I)$	$I > 2\sigma(I)$
R_{int}	0.017	0.017
θ_{max} (°)	26.4	26.4
Range of h, k, l	$-26 \Rightarrow h \Rightarrow 26$	$-4 \Rightarrow h \Rightarrow 4$
	$-4 \Rightarrow k \Rightarrow 4$	$-26 \Rightarrow k \Rightarrow 25$
	$-17 \Rightarrow l \Rightarrow 17$	$-8 \Rightarrow l \Rightarrow 8$
Refinement		
Refinement on	F^2	F^2
$R[F^2 > 2\sigma(F^2)]$, $wR(F^2)$, S	0.030, 0.090, 1.00	0.033, 0.091, 1.00
No. of reflections	1140	1133
No. of parameters	108	107
H-atom treatment	Mixture of independent and constrained refinement	Mixture of independent and constrained refinement
Weighting scheme	$w = 1/[\sigma^2(F_o^2) + (0.0628P)^2 + 0.6435P]$, where $P = (F_o^2 + 2F_c^2)/3$	$w = 1/[\sigma^2(F_o^2) + (0.0595P)^2 + 0.2P]$, where $P = (F_o^2 + 2F_c^2)/3$
$(\Delta/\sigma)_{max}$	<0.0001	<0.0001
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.28, −0.16	0.26, −0.22
Extinction method	SHELXTL	None
Extinction coefficient	0.0104 (16)	

Computer programs used: *SMART* (Bruker AXS, 1998*a*), *SHELXTL* (Bruker AXS, 2000), *SADABS* (Bruker AXS, 1998*b*).

$C2/c$. The densest 50 structures from each packing type were lattice-energy minimized using *DMAREL* (Willock *et al.*, 1995) and the DMA description of the conformer. The electrostatic contribution to U_{latt} included all terms in the atom–atom multipole series up to R^{-5}, with charge–charge, charge–dipole and dipole–dipole terms calculated by Ewald summation. The remaining terms were calculated by direct summation up to a molecule–molecule separation of 15 Å. The

empirical repulsion–dispersion potential given in (2) was used to model the non-electrostatic contribution, where atom i in molecule 1 is of type ι and atom k in molecule 2 is of type κ. This model intermolecular potential has been widely used in crystal-structure prediction studies, and its suitability for HOAt was verified by testing its ability to reproduce form I by lattice-energy minimization. Parameters for C, N, O, H($-$C) (Williams & Cox, 1984; Cox *et al.*, 1981) and the polar hydrogen, H($-$O) (Coombes *et al.*, 1996), were taken from empirically derived potentials,

$$U = \sum_{\iota \in 1, \kappa \in 2}(A_{\iota\iota}A_{\kappa\kappa})^{1/2}\exp[-(B_{\iota\iota} + B_{\kappa\kappa})R_{ik}/2] - (C_{\iota\iota}C_{\kappa\kappa})^{1/2}/R_{ik}^6.$$

(2)

All *DMAREL* lattice-energy minimizations calculate the second-derivative properties, so that any structures that had not reached a true minimum, usually because they were revealed to be transition states, could be discarded. Duplicate structures apparent by consideration of U_{latt} and reduced cell parameters calculated using *PLATON* (Spek, 2002) from each search were removed. The use of rigid conformers meant that minima from different searches could represent the same crystal structure, because if the intramolecular energy could have been optimized the two crystal structures would have

Figure 2
Displacement ellipsoid diagrams for (*a*) form I and (*b*) form II of HOAt.

Table 2
Selected hydrogen-bond information for form I and form II of HOAt.

$D-$H$\cdots A$	$d(D-$H) (Å)	$d($H$\cdots A)$ (Å)	$d(D\cdots A)$ (Å)	$\angle(D-$H$-A)$ (°)
Form I				
O1$-$H1$\cdots$N4^{i}	0.962 (18)	1.655 (18)	2.6167 (12)	178.4 (16)
C3$-$H3$\cdots$N3^{ii}	0.984 (15)	2.626 (15)	3.5685 (16)	160.3 (11)
C4$-$H4$\cdots$N2^{iii}	0.954 (15)	2.702 (14)	3.3092 (15)	122.1 (10)
C5$-$H5$\cdots$N2^{iii}	0.938 (14)	2.737 (14)	3.3033 (15)	119.7 (10)
Form II				
O1$-$H1$\cdots$N4^{iv}	0.96 (2)	1.71 (2)	2.6564 (13)	170.7 (19)
C3$-$H3$\cdots$N3^{v}	0.981 (16)	2.503 (16)	3.4582 (16)	164.6 (12)
C4$-$H4$\cdots$N2^{vi}	0.955 (16)	2.670 (16)	3.3715 (16)	130.7 (12)
C5$-$H5$\cdots$N2^{vii}	0.930 (17)	2.712 (16)	3.2572 (16)	118.3 (12)

Symmetry codes: (i) $-x + 1, y, -z + \frac{3}{2}$; (ii) $-x + \frac{1}{2}, -y - \frac{3}{2}, -z + 1$; (iii) $x, -y - 1, z + \frac{1}{2}$; (iv) $x, -y + \frac{3}{2}, z - \frac{1}{2}$; (v) $-x + 2, -y + 2, -z$; (vi) $x + 1, y, z + 1$; (vii) $x, y, z + 1$.

become the same. It was therefore necessary to classify all structures into structure types, where structures belonging to a particular structure type have similar packing motifs and hydrogen bonds despite different molecular conformations. This classification was performed using *Mercury* (Bruno *et al.*, 2002; Macrae *et al.*, 2006) for visual comparison and *COMPACK* (Chisholm & Motherwell, 2005) for automated pairwise comparison of molecular coordination spheres.

2. Results

2.1. Experimental X-ray diffraction results

2.1.1. Form I. The molecular structure of form I is shown in Fig. 2(*a*). The overall conformation of the molecule is planar and a least-squares fit to the nine atoms of the ring system (C1 to C5, N1 to N4) yields an r.m.s. deviation of 0.0113 Å, with the maximum deviation being -0.017 (1) Å for atom C4. Atom O1 lies -0.1924 (13) Å below this calculated plane and the hydroxyl H1 atom lies 0.616 (18) Å above the plane. The flexible torsion angle C1$-$N1$-$O1$-$H1 (θ) was determined to be 72 (1)°. A view of the crystal packing down the *b* axis of the cell is shown in Fig. 3(*a*) and selected hydrogen-bond information is given in Table 2. The primary hydrogen-bond interaction is the formation of an $R_2^2(10)$ dimer from two identical hydroxyl O$-$H to pyridyl$-$N interactions (O1$-$H1$\cdots$N4). The hydrogen-bonding network is augmented by three weaker C$-$H$\cdots$N interactions. The first interaction is the formation of an $R_2^2(8)$ dimer originating from a pair of identical aromatic C$-$H to triazole$-$N interactions (C3$-$H3$\cdots$N3). The remaining two interactions allow for the formation of a weak chain of molecules parallel to the *c* axis of the unit cell, originating from aromatic C$-$H to triazole$-$N interactions (C4$-$H4$\cdots$N2 and C5$-$H5$\cdots$N2). Overall these interactions have the effect of producing antiparallel chains of molecules along the *c* direction of the unit cell which are cross-linked by hydrogen bonds to form a two-dimensional network in the *ac* plane.

2.1.2. Form II. The molecular structure of form II is shown in Fig. 2(*b*). The overall conformation of the molecule is similar to that of form I in that it is planar, and a least-squares fit to the nine atoms of the ring system (C1 to C5, N1 to N4) yields an r.m.s. deviation of 0.0058 Å, with the maximum deviation being -0.010 (1) Å for the N1 atom. The O1 atom lies 0.0893 (13) Å above this calculated plane and the hydroxyl H1 atom lies -0.807 (22) Å below the plane. The flexible torsion angle C1$-$N1$-$O1$-$H1 (θ) was determined to be -97(1)°. A view of the crystal packing down the *a* axis of the cell is shown in Fig. 3(*b*) and selected hydrogen-bond information is given in Table 2(*b*). The primary hydrogen-bond interaction is the formation of a $C(5)$ chain of molecules parallel to the *c* axis of the unit cell from the hydroxyl O$-$H to the pyridyl N (O1$-$H1$\cdots$N4). The hydrogen-bond network is again supported by weaker C$-$H$\cdots$N interactions, which are
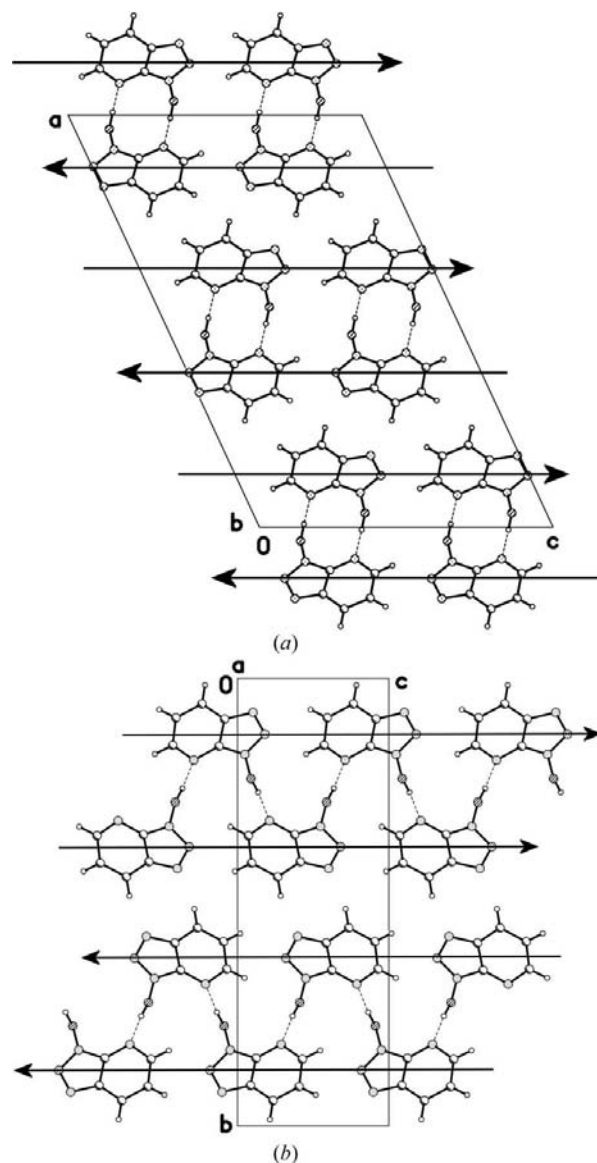
Figure 3
(*a*) Form I and (*b*) form II HOAt crystal structures. Note the antiparallel and parallel chains in the form I and II structures, respectively. OH$\cdots$N hydrogen bonds are shown as dotted lines.
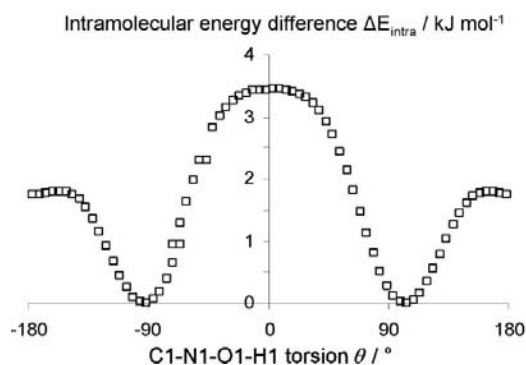
Table 3
Calculated structures corresponding to the two observed forms of HOAt.

The unit-cell parameters for the calculated structures given in italics are percentage differences from the observed values.

	θ (°)	a (Å)	b (Å)	c (Å)	β (°)	U_{latt} (kJ mol^{-1})	Volume per molecule (Å^3)
Form I							
Observed	71.3	22.0493 (11)	3.9336 (18)	14.2346 (14)	114.893 (5)		139.99
expminexp	71.3	−2.99	4.82	1.38	−1.18	−115.5	145.86
expminopt	97.8	−4.22	10.39	1.00	0.48	−109.2	148.83
expmincon	71.3	−6.05	8.93	0.42	−2.53	−112.5	147.08
Lowest energy match in search	80	8.87	−7.94	−0.08	6.08	−112.2	147.73
Closest r.m.s. match in search	50	−5.53	4.76	0.95	−4.36	−111.0	145.01
Form II							
Observed	97.8	3.750 (3)	20.858 (2)	7.0965 (17)	96.37 (3)°		137.91
expminexp	97.8	5.24	−0.69	−0.12	−1.11	−112.3	144.24
expminopt	97.8	6.38	−0.72	−0.63	−1.72	−110.5	145.14
expmincon	97.8	6.37	−0.73	−0.61	−1.71	−110.6	145.13
Lowest energy match in search	110	3.20	−0.76	3.90	3.70	−111.0	145.45
Closest r.m.s. match in search	100	5.65	−0.75	0.25	−0.77	−110.7	145.17

very similar to those found in form I. The first is the formation of the weak $R_2^2(8)$ dimer originating from a pair of identical aromatic C—H to triazole–N interactions (C3—H3···N3). The second and third, although similar in nature to the aromatic C—H to triazole–N interactions, C4—H4···N2 and C5—H5···N2, found in form I, differ in the sense of their directionality. These interactions link the molecules into stacks along the *a* direction of the unit cell, giving rise to the short ring–ring contacts C1—C3 (3.377 Å) and C2—C3 (3.391 Å) between the rings in the stack. Overall this packing motif differs from that of form I in that the chains formed along the *c* direction of the unit cell and linked by O—H···N bonds are parallel rather than antiparallel and the hydrogen-bonding network is three-dimensional rather than two-dimensional. It is interesting that all of the C—H···N interactions in form II are slightly shorter than those found in form I, giving the impression of a slightly stronger lattice. This is also reflected in the calculated densities of the two materials, with form II having the slightly higher density of 1.636 g cm^{-3} relative to that of 1.616 g cm^{-3} calculated for form I.

2.2. Computational results

2.2.1. The isolated molecule potential energy surface of θ.
The conformational potential energy surface, shown in Fig. 4, is symmetrical about $\theta = 0°$ because of the planarity of the ring and so only the range $0 < \theta < 180°$ was considered. The least favourable gas-phase conformation ($\Delta E_{intra} \simeq 3.5$ kJ mol^{-1}) has $\theta = 0°$, with the O—H bond eclipsing the electron density in the C1—N1 bond. From the peaks in the PES at $\theta = \pm 180°$, an eclipse of the O—H bond and the N1—N2 bond is less unfavourable ($\Delta E_{intra} \simeq 1.7$ kJ mol^{-1}), implying that the proton in relatively close proximity to N4 and the bulk of the molecule is more destabilizing. The *ab initio* global-minimum conformer has $\theta = -97.8°$ with the O—H bond almost normal to the ring plane, but slightly displaced away from C1. An overlay of the gas-phase global-minimum conformer with the two observed conformers is shown in Fig. 5. Because the entire conformational potential energy surface is within 3.5 kJ mol^{-1} of the global minimum, it was considered inappropriate to rule out any region of θ as energetically unfeasible, in that the intramolecular energy cost of distorting from the gas-phase minimum is small enough that the formation of sufficiently stabilizing hydrogen bonds could produce a stable packing

Figure 4
Conformational energy scan of θ for an isolated HOAt molecule.
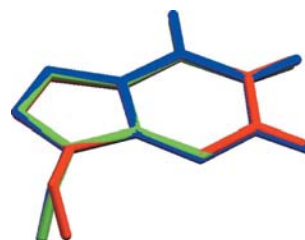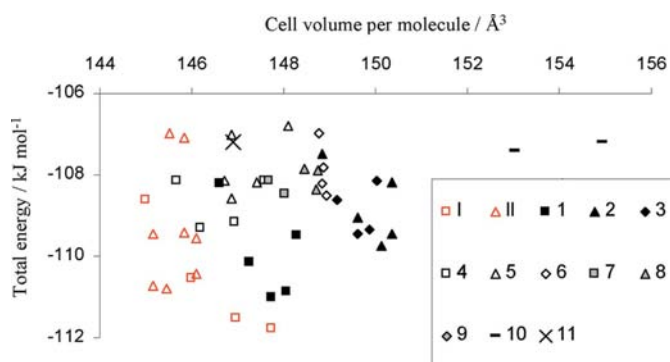
Figure 5
An overlay of the gas-phase optimized conformer (blue) with the observed form I (red) and II (green) conformers.

arrangement. Consequently, searches were carried out using conformers representing the entire range of values of θ.

2.2.2. Sensitivity of lattice-energy minima to molecular conformation. Both observed crystal structures of HOAt gave lattice-energy minima using the experimental molecular structure (expminexp calculations, see Table 3) which were in good agreement with the observed structure, indicating the model potential and modelling method to be adequate. The crystal structures were subsequently lattice-energy minimized using the gas-phase optimized conformer (the expminopt calculations), in which $\theta = -97.8°$, and also the partially optimized conformer, in which θ was constrained to the experimental values (the expmincon calculations; $\theta = -71.3$ in form I and $-97.8°$ in II). Note that the conformer in form II closely resembles the gas-phase optimized conformer with equal θ and an r.m.s. deviation of 0.016 Å for non-H atoms (*cf.* 0.073 Å for form I), see Fig. 5, and so the expmincon and expminopt calculations for form II result in essentially the same crystal structure. The form II conformer also resembles the optimized conformer in that O1 is almost coplanar with the azabenzotriazole ring, while in the experimental form I conformer O1 deviates from the plane of the ring. The lattice-energy minimizations indicate that form I is more stable than form II by no more than a few kJ mol^{-1} at 0 K.

The discrepancies between lattice-energy minimized structures and observed structures are greater when using the *ab initio* optimized and constrained conformers rather than experimental conformers, but follow the same trends. Calculated form I structures have underestimated and overestimated a and b cell lengths, respectively. This results in an underestimation of the distance between neighbouring dimers in the a direction and an overestimation of the distance between stacked dimers in the b direction. The lengths of $R_2^2(8)$ and $C(5)$ hydrogen bonds in forms I and II, respectively, are overestimated, resulting in the density of all calculated structures being underestimated, with the volume per molecule typically 7 Å^3 per molecule greater than in the observed structures. The very close agreement between the gas-phase optimized conformer and the form II conformer (Fig. 5) leads

to good agreement between the lattice-energy minimized form II structures and observed form II; the biggest difference is that the a cell length is overestimated by about 0.2 Å in the calculated structures, resulting in a larger distance between stacked hydrogen-bonded chains.

Despite the agreement between optimized and experimental form II conformers, the expminopt and expmincon minimizations for both forms resulted in crystal structures that were markedly worse than the expminexp calculations, see Table 3, suggesting considerable sensitivity of the crystal structure and lattice energy to the precise molecular conformation. The form I reproductions using optimized conformers give structures of similar quality (both considerably worse than the expminexp structures), despite more than a 20° difference in θ between the constrained optimized conformer ($\theta = 71.3°$) and the freely optimized conformer ($\theta = 97.8°$), suggesting that small deviations from the exact experimental conformer in the position of the O and ring atoms (Fig. 5) have an effect comparable with larger deviations in the hydroxyl proton positions on the crystal structure.

2.2.3. Search results. The plot in Fig. 6 shows that there is a range of thermodynamically feasible polymorphs and the lowest-energy structure from each structure type is given in Table 4. (These structures in *SHELX* .res format are available from the authors.) Note that the same structure types are found for a range of conformers; typically a difference of up to 40° in θ can lead to effectively equivalent crystal structures that would be expected to optimize to the same structure if it were possible to optimize the conformation and the packing simultaneously.

The most stable structure type from the search corresponds to observed form I; form II was found in the search to be 1.0 kJ mol^{-1} less stable. A unit-cell overlay of the predicted form II crystal structure and the experimental structure is shown in Fig. 7; the r.m.s. deviation for a 15-molecule coordination sphere of the predicted and observed structures is 0.24 Å for non-H atoms. The quality of the prediction is further evidenced by the isotropic refinement of the predicted structure coordinates against single-crystal diffraction data, which converged in just six cycles to the experimental crystal structure. There is one hypothetical structure type (1) found in the search with E_{tot} in between the two observed forms, and
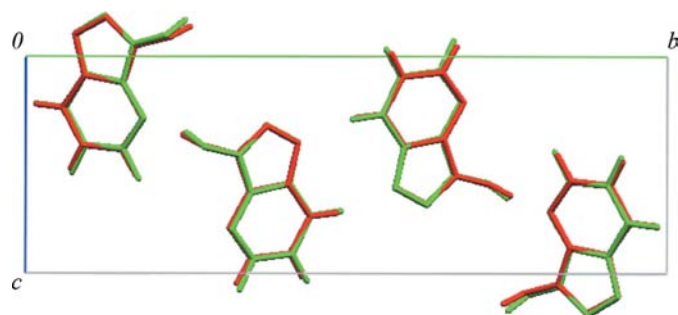

Figure 6
Crystal energy landscape for HOAt. Each lattice-energy-optimized crystal structure is denoted by a symbol which represents the structure type (Table 4, penultimate column), with the different minima within a structure type having different values of θ. Red symbols and black symbols correspond to the observed and hypothetical crystal structures, respectively.

Figure 7
Unit-cell overlay of form II (red) with the predicted crystal structure (green).

Table 4
Low-energy crystal structures of HOAt from the computational search.

ID	θ (°)	Space group	Cell volume (Å^3)	Energy (kJ mol^{-1})			Acceptor and graph-set motif involving O1—H1	Acceptor and graph-set motif involving C3—H3	Structure type†	Chains‡
				ΔE_{intra}	U_{latt}	E_{tot}				
de41	80	$C2/c$	147.73	0.45	−112.23	−111.78	N4 $R_2^2(10)$	N3 $R_2^2(8)$	Form I	a
ai9	100	$P2_1/c$	147.75	0.01	−111.03	−111.03	N4 $C(5)$	N3 $R_2^2(8)$	1	a
fa40	110	$P2_1/c$	145.45	0.20	−111.02	−110.81	N4 $C(5)$	N3 $R_2^2(8)$	form II	p
fc48	100	$P2_1/c$	150.12	0.01	−109.74	−109.73	N4 $C(5)$	N3 $R_2^2(8)$	2	a
az48	100	$P2_12_12_1$	149.61	0.01	−109.47	−109.46	N4 $C(5)$	N3 $C(4)$	3	a
aq38	110	$P2_12_12_1$	146.2	0.20	−109.50	−109.30	N4 $C(5)$		4	a
am6	90	$P2_1/c$	146.88	0.09	−108.66	−108.57	N4 $R_2^2(10)$§		5	a
da28	70	Cc	148.92	1.04	−109.55	−108.52	N3 $C(5)$	N4 $C(5)$	6	p
dd25	70	$C2/c$	148.03	1.04	−109.53	−108.49	N4 $R_2^2(10)$		7	a
ak35	100	$P2_1/c$	148.71	0.01	−108.36	−108.35	N4 $C(5)$		8	a
dc21	70	$C2/c$	154.08	1.04	−109.25	−108.22	N4 $R_2^2(10)$	N3 $C(4)$	9	a
dc43	60	$C2/c$	153.02	1.72	−109.15	−107.43	N4 $R_2^2(10)$	N3 $R_2^2(8)$	10	a
am46	110	$P2_1/c$	146.89	0.20	−107.42	−107.22	N3 $C(5)$		11	a

† Only the lowest-energy structure of each type is given. Fig. 6 shows other structures which have similar packing motifs and hydrogen bonds despite having different conformations (θ) classified by structure type. ‡ a and p denote antiparallel and parallel chains, respectively. § The shortest hydrogen bond is O1—H1$\cdots$N4 $R_2^2(10)$, but H1 is oriented between N4 and N3 making a $C(5)$ motif possible.

ten other hypothetical structure types (2–11) less stable than the two observed forms and within 5 kJ mol^{-1} of the E_{tot} global minimum.

Hydrogen-bond interactions in the calculated structures are detailed in Table 4; in all cases the major hydrogen-bond motif (Etter, 1990; Etter *et al.*, 1990) involves O1 as the donor, and either $R_2^2(10)$ dimers with N4 as the acceptor or $C(5)$ chains with either N3 or N4 as the acceptor. The molecules in the $R_2^2(10)$ dimers are related by inversion resulting in antiparallel chains of molecules, as in form I, see Fig. 3(*a*). All but two of the $C(5)$ structures exhibit antiparallel chains; the two exceptions, form II and structure type 6, have parallel chains with N4 and N3 as the acceptors, respectively. Many calculated structures display additional weaker hydrogen-bond interactions, often in the form of C3—H3$\cdots$N3 $R_2^2(8)$ dimers; interactions involving C3—H3 are listed in Table 4. Herringbone packing is exhibited to some degree in all low-energy calculated structures.

3. Discussion

The successful prediction of form II as the third most stable structure and form I as the global minimum indicates that the use of a large number of searches, each with a different rigid conformer and associated intramolecular energy, is a realistic approach to CSP for molecules with some degree of flexibility. The adequacy of the step size used here is demonstrated by the duplication of effectively equivalent structures arising from different conformers. Energetic sensitivity to exact molecular geometry is apparent, however, in the sensitivity of the lattice-energy minima to conformation (Table 3), and it would therefore be desirable to implement intramolecular relaxation within the lowest-energy crystal structures and increase the confidence in ranking the predicted structures. The conformational potential energy surface of HOAt

in the gas phase is relatively flat and so the entire 180° range was considered. It is likely that consideration of a smaller range, and therefore fewer searches, would be necessary for molecules with higher energy barriers between conformations. The availability of grid computer facilities (Butchart *et al.*, 2003) with the parallelization of *MOLPAK* and *DMAREL* jobs over a large number of computer nodes means that searches finish within a few hours and so the consideration of a large number of conformers is feasible. The benefit of infrastructure such as this for this type of crystal-structure prediction study is clear.

The quality of the predicted structure is highlighted by the ease with which its refinement against form II single-crystal data rapidly converged to match the observed structure. The use of predicted crystal structures to aid solution from powder diffraction data has been found useful (Tremayne *et al.*, 2004; Schmidt *et al.*, 2005). Certainly the simulated powder pattern for the predicted form II structure bears more resemblance to the simulated powder pattern of the experimental form II structure than to all other structures included in the low-energy list of suggested crystal structures. However, in this case the discrepancy between the simulated powder diffraction patterns for predicted and observed form II, resulting from differences in the lattice parameters (>3% shifts in unit-cell lengths *a* and *c*), suggests that the refinement of the predicted structure from X-ray powder diffraction data would be more difficult.

The possible existence of additional polymorphs of HOAt cannot be ruled out on thermodynamic grounds because of the variety of low-energy structures in Fig. 6 and Table 4. Structure 6 is one of the hypothetical structures that stands out as an intuitively reasonable crystal structure which is sufficiently different from the known polymorphs that there would be a significant kinetic barrier to transformation; it has zigzag O1—H4$\cdots$N3 $C(5)$ chains that have the possibility of further stabilization by weaker C3—H3$\cdots$N4 $C(5)$ interactions, the

zigzag chains stack giving rise to a feasible herringbone pattern. However, the solvent polymorph screen found no evidence of any further polymorphs and no solvates, although transient solvates could have led to the amorphous materials obtained from *n*-butanol and tertiary butyl methyl ether. The analysed solids were predominantly form II with the exception of those crystallized from alcohols (other than methanol), which resulted in form I. The filtrate residues gave varying mixtures of form I and form II. Further polymorphic forms cannot be completely excluded given that this polymorph screen did not include other methods of crystallization such as rapid cooling, high-pressure methodologies or the presence of different template molecules.

The crystal structures of HOAt forms I and II are inherently very different and solid-state transformation between the two forms is therefore considered unlikely. The change from parallel chains of molecules linked by $C(5)$ hydrogen bonds in form II to the antiparallel chains linked by dimer interactions in form I would cause huge disruption to the lattice and involve a considerable energy barrier. No evidence for a transformation was seen on cooling the crystals on the diffractometer, nor in differential scanning calorimetry, prior to the explosive decomposition of the samples. This chemical degradation also prevented the determination of the relative stability of the two forms and whether they are monotropically or enantiotropically related. Since the majority of maturation experiments led to form II, it is probably more stable at room temperature. The lattice-energy calculations indicate that form I is probably more stable at 0 K. However, all the available evidence points to very small differences in the thermodynamic stability of the two forms.

Thus, kinetic factors must play a major role in determining whether form I or form II is formed in a solvent crystallization, and why more of the other computed low-energy crystal structures are found. It seems likely that there is dimer association or emerging chain formation during the early stages of nucleation that provide templates for the growth of I or II, respectively. It may well be that the growth unit for form I is the $R_2^2(10)$ dimer. The observation that the filtrate residues were mixtures of forms I and II suggests that both the dimer and the chain fragments are present in solution for nucleation. Since form II was used as a starting material for the polymorph screen, the re-crystallization solvent would need to break up the chain structure in order for dimers to be formed. It is notable that larger alcohol molecules appear to encourage crystallization of form I; perhaps they are most effective at disrupting the OH$\cdots$N hydrogen-bond chain, facilitating dimer formation and thus providing the opportunity for form I to appear. The fact that the harvested solids from non-alcohohic solutions were predominantly form II is consistent with the fact that the form II chain fragment still existed in solution to act as a template for the immediate growth of form II. It is conceivable that further polymorphs of HOAt might be found if an appropriate template were present, but the combination of the computational and limited screening results make it very unlikely that any further polymorphs would be significantly more thermodynamically stable than the known forms.

4. Conclusions

The previously unreported conformational polymorph of form II 1-hydroxy-7-azabenzotriazole has been quantitatively predicted computationally from the molecular diagram. The combination of experimental results and computational prediction provides for a better understanding of the polymorphism of 1-hydroxy-7-azabenzotriazole. There are a number of possible hydrogen-bond motifs that pack with competitive energies for modest conformational changes in an approximate 40° range in the hydroxyl torsional angle. It seems likely that the initial association of the molecules within the solvent plays a major role in determining the polymorphic outcome of crystallization experiments.

References

Allen, F. H., Kennard, O., Watson, D. G., Brammer, L., Orpen, A. G. & Taylor, R. (1987). *J. Chem. Soc. Perkins Trans. 2*, pp. S1–S19.

Bernstein, J. (2002). *Polymorphism in Molecular Crystals*. Oxford: Clarendon Press.

Brodersen, S., Wilke, S., Leusen, F. J. J. & Engel, G. (2003). *Phys. Chem. Chem. Phys.* **5**, 4923–4931.

Bruker AXS (1998*a*). *SMART*, V5.054. Bruker AXS, Madison, Wisconsin, USA.

Bruker AXS (1998*b*). *SADABS*, V2.05. Bruker AXS, Madison, Wisconsin, USA.

Bruker-AXS (2000). *SHELXTL*, V6.14. Bruker AXS, Madison, Wisconsin, USA.

Bruno, I. J., Cole, J. C., Edgington, P. R., Kessler, M., Macrae, C. F., McCabe, P., Pearson, J. & Taylor, R. (2002). *Acta Cryst.* B**58**, 389–397.

Butchart, B., Chapman, C. & Emmerich, W. (2003). *Proceedings of the UK e-Science All Hands Meeting*, edited by S. J. Cox. UK: EPSRC.

Chisholm, J. A. & Motherwell, S. (2005). *J. Appl. Cryst.* **38**, 228–231.

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

Cox, S. R., Hsu, L. Y. & Williams, D. E. (1981). *Acta Cryst.* A**37**, 293–301.

Day, G. M. *et al.* (2005). *Acta Cryst.* B**61**, 511–527.

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

Etter, M. C. (1990). *Acc. Chem. Res.* **23**, 120–126.

Etter, M. C., MacDonald, J. C. & Bernstein, J. (1990). *Acta Cryst.* B**46**, 256–262.

Frisch, M. J. *et al.* (2004). *GAUSSIAN03*, Revision C.02. Gaussian Inc., Wallingford, Connecticut, USA.

Gavezzotti, A. (2002). *CrystEngComm.* **4**, 343–347.

Hoffmann, F., Kolbe, A. & Griehl, C. (1999). *J. Mol. Struct.* **476**, 289–294.

Holden, J. R., Du, Z. Y. & Ammon, H. L. (1993). *J. Comput. Chem.* **14**, 422–437.

Lommerse, J. P. M., Motherwell, W. D. S., Ammon, H. L., Dunitz, J. D., Gavezzotti, A., Hofmann, W. M., Leusen, F. J. J., Mooij, W. T. M.,

Price, S. L., Schweizer, B., Schmidt, M. U., van Eijck, B. P., Verwer, P. & Williams, D. E. (2000). *Acta Cryst.* B**56**, 697–714.

Macrae, C. F., Edgington, P. R., McCabe, P., Pidcock, E., Shields, G. P., Taylor, R., Towler, M. & van de Streek, J. (2006). *J. Appl. Cryst.* **39**, 453–457.

Motherwell, W. D. S. *et al.* (2002). *Acta Cryst.* B**58**, 647–661.

Nowell, H. & Price, S. L. (2005). *Acta Cryst.* B**61**, 558–568.

Price, S. L. (2004). *Adv. Drug Deliv. Rev.* **56**, 301–319.

Schmidt, M., Ermrich, M. & Dinnebier, R. E. (2005). *Acta Cryst.* B**61**, 37–45.

Spek, A. L. (2002). *PLATON*. University of Utrecht, The Netherlands.

Stone, A. J. (1999). *GDMA*, 1.0 ed. University of Cambridge, UK.

Stone, A. J. & Alderton, M. (1985). *Mol. Phys.* **56**, 1047–1064.

Tremayne, M., Grice, L., Pyatt, J. C., Seaton, C. C., Kariuki, B. M., Tsui, H. H. Y., Price, S. L. & Cherryman, J. C. (2004). *J. Am. Chem. Soc.* **126**, 7071–7081.

Verwer, P. & Leusen, F. J. J. (1998). *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz & D. B. Boyd, pp. 327–365. New York: Wiley-VCH.

Williams, D. E. & Cox, S. R. (1984). *Acta Cryst.* B**40**, 404–417.

Willock, D. J., Price, S. L., Leslie, M. & Catlow, C. R. A. (1995). *J. Comput. Chem.* **16**, 628–647.