

Predicting stoichiometry and structure of solvates†

Aurora J. Cruz-Cabeza,^a Shyam Karki,^b László Fábián,^a Tomislav Friščić,^b Graeme M. Day^b and William Jones^{*b}

Received (in Cambridge, UK) 2nd November 2009, Accepted 5th February 2010

First published as an Advance Article on the web 26th February 2010

DOI: 10.1039/b922955h

We demonstrate that crystal structure prediction calculations can be used to predict both the stoichiometry and structure of multicomponent molecular crystals. The methods are used here to determine the structure of a recently discovered acetic acid solvate of theobromine.

Methodologies for the computational prediction of crystal structures have developed considerably in the last decade or so.¹ Crystal Structure Prediction (CSP) calculations, based on global minimisation of the lattice energy, have been successfully applied to many small rigid molecules,² some challenging flexible molecules^{3,4} and a small number of salts,^{5,6} cocrystals¹ and solvates,^{7,8} including hydrates.^{9,10} These calculations allow new perspectives into the understanding of multicomponent crystals,¹¹ their formation,^{7,12–14} their structures and stoichiometries.⁸

The use of neat and liquid-assisted grinding (LAG) methods¹⁵ for cocrystal and solvate synthesis has many advantages over traditional solution crystallisation methods: they are faster, cleaner and more successful at finding new crystal forms. The structures of the resulting crystalline powders, however, can be difficult to determine if subsequent seeded crystallisations do not afford single crystals. In cases where structure solution from powder diffraction data is not straightforward, the synergy between LAG and CSP calculations can provide a solution. However, when the stoichiometry of the resulting multicomponent crystal is unknown, this adds an additional challenge for the computational methods: what ratio of components will be preferred?

In a previous study, we rationalised the stoichiometry observed for an acetic acid (AcOH) solvate of urea with the aid of computational methods.⁸ The structures predicted by CSP methods provided the basis for an energetic analysis of the possible predicted systems, and their hydrogen bond motifs, to point to the observed stoichiometry as the most likely. Following our previous lines of experimental research with the xanthines (caffeine, theophylline and theobromine),^{16–18} here we report the synthesis of an AcOH solvate of theobromine (tb) by LAG, the prediction of its stoichiometry and the solution of its structure with the help of computational methods.

The grinding of theobromine with AcOH† yielded a PXRD pattern which did not correspond to any of the previously known solid forms. The obtained PXRD patterns suggested that the crystalline solids obtained by grinding 1 : 1 and 1 : 2 equivalent amounts of tb and AcOH were identical. However, we found difficulties in experimentally determining the stoichiometric ratio of the resulting material due to its poor stability.§ After failing to grow a single crystal of the material or to obtain powder diffraction data of high enough quality for structure solution, we decided to approach the problem from a fully computational perspective. The challenge of this system lies not only in the prediction of its structure but also its stoichiometry. Stoichiometric preferences are difficult to predict and can change within a family of solvates (Fig. 1). For example, both 1 : 1 and 1 : 2 stoichiometries are observed for the known caffeine : AcOH structures¹⁹ whereas only a 1 : 2 structure is known for urea : AcOH.

We proceeded to run CSP calculations for tb : AcOH in both 1 : 1 and 1 : 2 stoichiometries. The molecules were treated as rigid throughout the calculations and we used the methodologies described in the ESI† and in our previous work.⁸ Since the computational demand of CSP for multicomponent systems is very high, we only generated structures in the two most relevant space groups for AcOH solvates ($P2_1/c$ and $P\bar{1}$).²⁰ Around 40 and 90 crystal structures were generated within 10 kJ mol⁻¹ of the global minimum for the 1 : 1 and 1 : 2 stoichiometries, respectively.

We took the most stable crystal structure from our predicted 1 : 1 and 1 : 2 tb : AcOH systems and proceeded to compare the relative stability of the stoichiometric possibilities.

Lattice energies are calculated using periodic boundary conditions and are normally expressed as kJ mol⁻¹ of independent molecules in the asymmetric unit (Z'') of the crystal structure under consideration. When dealing with crystals with the same components but different stoichiometries, lattice energies cannot be compared directly as Z'' , and therefore the total composition of the system, varies. For this reason we compared the relative energies of the different crystalline products

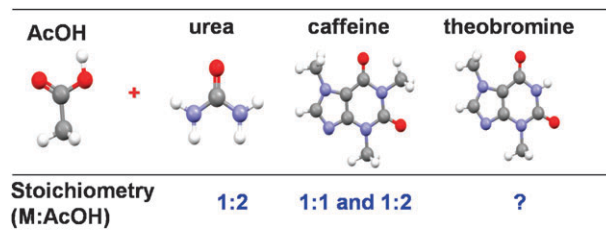


Fig. 1 Acetic acid systems compared in this study.

^a The Pfizer Institute for Pharmaceutical Materials Science, The Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge, CB2 1EZ, UK. E-mail: cruz@ccdc.cam.ac.uk

^b Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK. E-mail: wj10@cam.ac.uk; Tel: +44 (0)1223 336468

† Electronic supplementary information (ESI) available: Details of the Rietveld refinement procedure. CCDC 753170 and 753171. For ESI and crystallographic data in CIF or other electronic format see DOI: 10.1039/b922955h

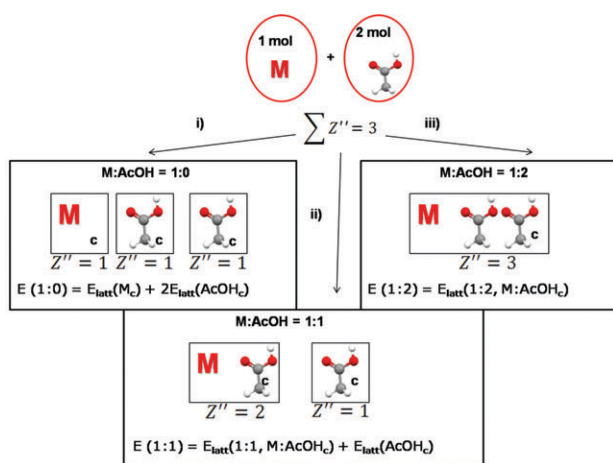


Fig. 2 Crystal systems and energy expressions compared for a constant stoichiometry of 1 mol of main component (M) and 2 mols of AcOH ($\sum Z'' = 3$).

that can result from a fixed starting reactant stoichiometry (or fixed number of total molecules ($\sum Z''$)). The energy expressions and the three different crystal systems that can be compared for a constant stoichiometry of 1 mol of main component (M) and 2 mols of AcOH are depicted in Fig. 2. For the first possibility (i) no solvate forms, whereas for (ii) and (iii) there is formation of the 1 : 1 and the 1 : 2 M : AcOH solvates respectively. Our energy model is only based on the lattice energies of the crystalline materials (Fig. 2) and the calculations do not include any temperature. Furthermore, we only consider the bulk energies of the different phases, and assume that interfacial interactions between any phases that are in contact (such as crystal–crystal and crystal–solvent interactions) will not significantly affect the relative ordering of the various possibilities.

We illustrate the results in plots of total energy *versus* solvate stoichiometry for the three molecules in Fig. 1: urea (Fig. 3a), caffeine (Fig. 3b) and tb (Fig. 3c). The lattice energies of the experimental structures (crystalline tb, caffeine,²¹ AcOH and caffeine : AcOH structures¹⁹) were calculated by geometry optimising the known structures with the same energy model used in the CSP calculations,[¶] whereas the lattice energies of the hypothetical 1 : 1 and 1 : 2 solvates were taken from the CSP calculation results (using the global lattice

energy minima of the 1 : 1 and 1 : 2 predictions). The urea plot was taken from our previous work (Fig. 3a).⁸ The plots in Fig. 3 illustrate the energetics of these three M : AcOH systems when no solvate forms (1 : 0), when the 1 : 1 solvate forms (1 : 1) or when the 1 : 2 solvate forms (1 : 2). For example, in the urea : AcOH system (Fig. 3a), formation of a 1 : 2 solvate is energetically favoured over formation of a 1 : 1 solvate (by ~ 4 kJ mol⁻¹) or no formation of solvate at all (by almost 10 kJ mol⁻¹), in good agreement with the experimental observations.⁸ Interestingly for the caffeine system (Fig. 3b), the formation of both the 1 : 1 solvate and the 1 : 2 solvate is isoenergetic and, therefore, it is perhaps not surprising that both solvates are observed.¹⁹ Finally for our true prediction (Fig. 3c), these calculations suggest that formation of the 1 : 1 tb : AcOH solvate is favoured over the formation of the 1 : 2 solvate (by ~ 4.4 kJ mol⁻¹) or no formation of solvate at all (by ~ 23 kJ mol⁻¹). Furthermore, we note that the energy associated with solvate formation in the tb system (23 kJ mol⁻¹) is comparable to that of caffeine (25 kJ mol⁻¹). Knowing, therefore, that the 1 : 1 stoichiometry is favoured for our target molecule (tb), we further analysed our CSP results.

The lattice energies of the computationally generated 1 : 1 tb : AcOH crystal structures are plotted against their packing coefficients in Fig. 4. Each data point in the plot corresponds to a distinct crystal structure with different properties, including density and lattice energy. The most common hydrogen bond motif found in the low energy predicted structures consists of a doubly hydrogen bonded tb dimer to which two AcOH molecules are hydrogen bonded, one at each side (Fig. 4). Structures containing this hydrogen bonding are represented as green circles in Fig. 4; the five lowest energy structures all contain this tetramer, with the global minimum energy prediction (Fig. 4 and 5) over 2 kJ mol⁻¹ more stable than any other structure.

Knowing the stoichiometry and crystal structure of the most likely solvate, we simulated its PXRD pattern and compared it to the one obtained experimentally (Fig. 6). The good agreement between the predicted and experimental PXRD patterns allowed Rietveld refinement of the predicted structure against the experimental diffraction pattern.^{||} The successful completion of the refinement confirmed the correctness of the predicted structure.^{**}

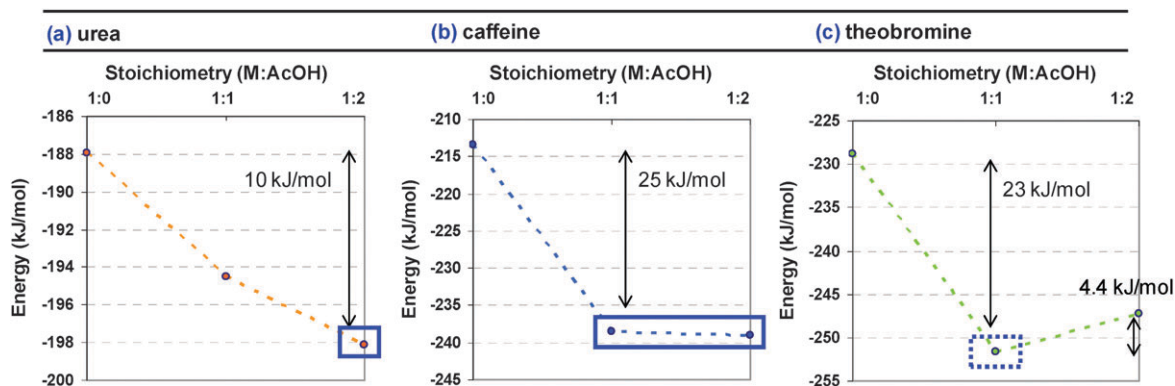


Fig. 3 Energy *versus* solvate stoichiometry plots for the systems (a) urea : AcOH, (b) caffeine : AcOH and (c) tb : AcOH with total 1 : 2 constant stoichiometries.

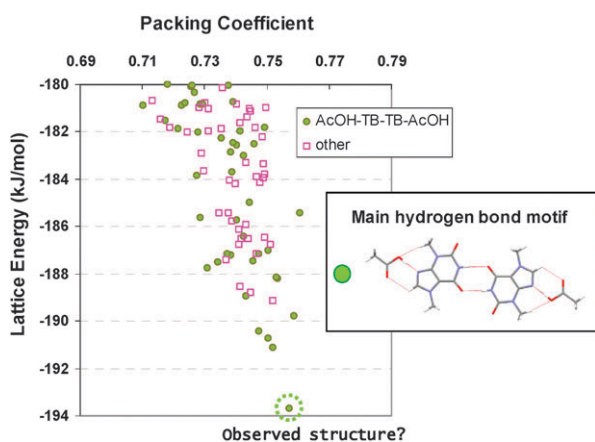


Fig. 4 Summary of the 1 : 1 tb : AcOH CSP results.

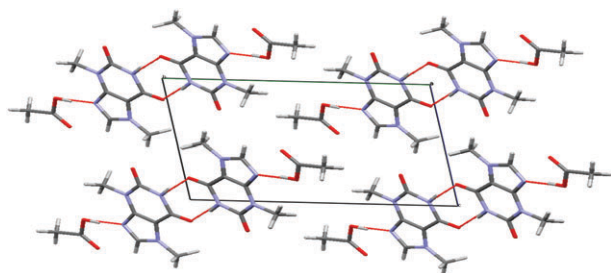


Fig. 5 Unit cell representation of the 1 : 1 tb : AcOH solvate.

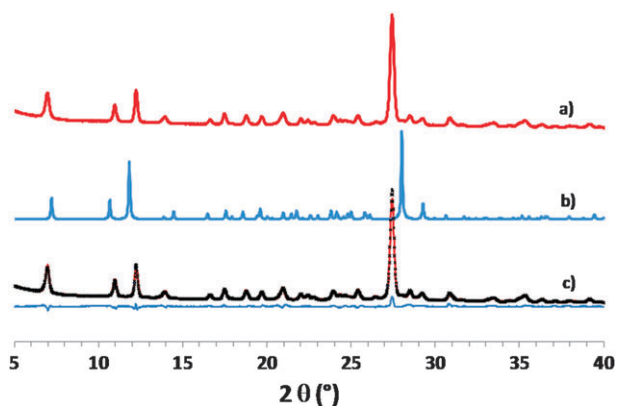


Fig. 6 Experimental PXRD pattern obtained at room temperature by grinding tb and AcOH (a), simulated PXRD pattern ($\lambda = 1.5418 \text{ \AA}$) for the predicted structure of 1 : 1 tb : AcOH at the global minimum (b) and Rietveld refinement (c).

Global lattice energy searches have been used to assess the possible crystal structures of theobromine with acetic acid in different stoichiometric ratios. The total energy of each possible stoichiometry, relative to the pure components, is shown to correctly predict the stoichiometry of the material that is produced in a solid state grinding experiment. The computer-generated structures are used, in conjunction with powder X-ray diffraction data, to determine the crystal structure of the solvate. The example presented here demonstrates the potential role of computational methods in predicting and characterising multicomponent molecular materials.

Notes and references

‡ All the grinding experiments were performed in stainless steel jars being shaken with two stainless steel balls (7 mm diameter), at a rate of 30 Hz for 30 minutes using a Retsch MM200 Mixer Mill. The same crystalline solid was obtained by grinding theobromine (150 mg) with 1 or 2 molar equivalents of AcOH.

§ The crystalline material obtained by grinding readily decomposes leaving crystalline theobromine behind.

¶ A classical forcefield calculation was used consisting of the W99 parameters for the evaluation of the intermolecular van der Waals interactions and an atomic multipole model for the evaluation of the electrostatics. For AcOH, both polymorphic structures were minimised but only the lattice energy of the most stable form was used.

|| PXRD data were collected at room temperature on a Philips X'Pert Pro diffractometer using Ni-filtered Cu-K α radiation ($\lambda = 1.5418 \text{ \AA}$, $2\theta = 5.016^\circ$ – 65.436° , step size = 0.01675°). Crystal data: $\text{C}_9\text{H}_{12}\text{N}_4\text{O}_4$, $M = 240.22$, $a = 5.2185(4) \text{ \AA}$, $b = 8.2901(8) \text{ \AA}$, $c = 12.9358(11) \text{ \AA}$, $\alpha = 95.827(6)^\circ$, $\beta = 96.854(6)^\circ$, $\gamma = 100.333(8)^\circ$, $P1$, $Z = 2$, 393 reflections, $R_{\text{wp}} = 0.0458$, $R_{\text{exp}} = 0.0503$, $R_1 = 0.0734$, $\chi^2 = 0.898$.

** A Compack comparison²² between the predicted and refined structures gives an RMS deviation in atomic positions of 0.307 \AA for 20 molecules.

- G. M. Day, T. G. Cooper, A. J. Cruz-Cabeza, K. E. Hejczyk, H. L. Ammon, S. X. M. Boerrigter, J. S. Tan, R. G. Della Valle, E. Venuti, J. Jose, S. R. Gadre, G. R. Desiraju, T. S. Thakur, B. P. van Eijck, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, M. A. Neumann, F. J. J. Leusen, J. Kendrick, S. L. Price, A. J. Misquitta, P. G. Karamertzanis, G. W. A. Welch, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, J. van de Streek, A. K. Wolf and B. Schweizer, *Acta Crystallogr., Sect. B: Struct. Sci.*, 2009, **65**, 107.
- G. M. Day, W. D. S. Motherwell and W. Jones, *Cryst. Growth Des.*, 2005, **5**, 1023.
- P. G. Karamertzanis and C. C. Pantelides, *Mol. Phys.*, 2007, **105**, 273.
- G. M. Day, W. D. S. Motherwell and W. Jones, *Phys. Chem. Chem. Phys.*, 2007, **9**, 1693.
- P. G. Karamertzanis and S. L. Price, *J. Phys. Chem. B*, 2005, **109**, 17134.
- F. J. J. Leusen, *Cryst. Growth Des.*, 2003, **3**, 189.
- A. J. Cruz Cabeza, G. M. Day, W. D. S. Motherwell and W. Jones, *J. Am. Chem. Soc.*, 2006, **128**, 14466.
- A. J. Cruz-Cabeza, G. M. Day and W. Jones, *Chem.–Eur. J.*, 2008, **14**, 8830.
- B. P. van Eijck and J. Kroon, *Acta Crystallogr., Sect. B: Struct. Sci.*, 2000, **56**, 535.
- A. T. Hulme and S. L. Price, *J. Chem. Theor. Comput.*, 2007, **3**, 1597.
- M. Polito, E. D'Oria, L. Maini, P. G. Karamertzanis, F. Grepioni, D. Braga and S. L. Price, *CrystEngComm*, 2008, **10**, 1848.
- S. Mohamed, D. A. Tocher, M. Vickers, P. G. Karamertzanis and S. L. Price, *Cryst. Growth Des.*, 2009, **9**, 2881.
- P. G. Karamertzanis, A. V. Kazantsev, N. Issa, G. W. A. Welch, C. S. Adjiman, C. C. Pantelides and S. L. Price, *J. Chem. Theor. Comput.*, 2009, **5**, 1432.
- N. Issa, P. G. Karamertzanis, G. W. A. Welch and S. L. Price, *Cryst. Growth Des.*, 2009, **9**, 442.
- T. Friscic, A. V. Trask, W. Jones and W. D. S. Motherwell, *Angew. Chem., Int. Ed.*, 2006, **45**, 7546.
- A. V. Trask, W. D. S. Motherwell and W. Jones, *Cryst. Growth Des.*, 2005, **5**, 1013.
- A. V. Trask, W. D. S. Motherwell and W. Jones, *Int. J. Pharm.*, 2006, **320**, 114.
- S. Karki, L. Fabian, T. Friscic and W. Jones, *Org. Lett.*, 2007, **9**, 3133.
- A. V. Trask, J. van de Streek, W. D. S. Motherwell and W. Jones, *Cryst. Growth Des.*, 2005, **5**, 2233.
- A. J. Cruz Cabeza, E. Pidcock, G. M. Day, W. D. S. Motherwell and W. Jones, *CrystEngComm*, 2007, **9**, 556.
- C. W. Lehmann and F. Stowasser, *Chem.–Eur. J.*, 2007, **13**, 2908.
- J. A. Chisholm and S. Motherwell, *J. Appl. Crystallogr.*, 2005, **38**, 228.