COMMUNICATION

ChemComm
www.rsc.org/chemcomm

Differential evolution: crystal structure determination of a triclinic polymorph of adipamide from powder diffraction data

Colin C. Seaton and Maryjane Tremayne*
School of Chemical Sciences, University of Birmingham, Edgbaston, Birmingham, UK B15 2TT. E-mail: m.tremayne@bham.ac.uk

Received (in Cambridge, UK) 15th January 2002, Accepted 7th March 2002
First published as an Advance Article on the web 22nd March 2002

The crystal structure of a previously unknown triclinic polymorph of adipamide has been solved from laboratory X-ray powder diffraction data using a new direct space global optimisation method based on differential evolution.

The study of crystalline polymorphism in organic materials continues to attract considerable academic and industrial attention, although true understanding of the aspects controlling this phenomenon still requires full structural characterisation in each case. However, the conditions used to prepare many polymorphs, in particular metastable forms, often yield materials that occur only as polycrystalline powders and hence structural details must be obtained from powder diffraction studies.[1,2]

The structure solution of organic materials from powder diffraction data is a rapidly expanding field, driven recently by the development of direct space methods of structure solution.[3] These approach structure solution by generation of trial crystal structures based on the known chemical connectivity of the material, and assessment of the fitness of each structure by comparison with the experimental data. Global optimisation methods such as Monte Carlo,[2,4,5] simulated annealing[6,7] or genetic algorithms[8,9] are then used to locate the global minimum corresponding to the structure solution. In this paper we report the first application of a new global optimisation technique based on differential evolution (DE), to structure solution from powder diffraction data. DE is an evolutionary algorithm that is both relatively simple and easy to implement, and offers robust searching of minima.[10,11]

Like genetic algorithms, DE maintains a population of trial structures that are recombined and mutated together over a number of generations until convergence upon the global minimum is achieved. However, the processes used to achieve this are markedly different. In genetic algorithms, a series of recombination and mutation steps are performed on randomly selected members of the population and from this collection the new population is probabilistically selected. In a DE population, each child is created from randomly selected members of the population by the summation of their differences weighted according to the amount of recombination and mutation required. Hence the recombination and mutation are performed in a single step, generating the new population in a deterministic manner by comparison of the child with its parent, where the superior of the two is added to the new population. Only four parameters are required to control the DE calculation; the population size N_p, the total number of generations G_{max}, and K and F, used to dictate the level of recombination and mutation respectively. K and F can take any value between 0 and 1, with $K = 0$ corresponding to mutation only, and small F resulting in little disruption to the population.

For structure solution, each member of the population is a trial structure described by a list of elements: the position (x,y,z) and the orientation (θ, ϕ, γ) of the molecular model in the unit cell, and the conformation of the molecule defined by variable torsion angles $(\tau_1 \cdots \tau_n)$. Each of these elements has an associated upper and lower bound, which is checked by the DE algorithm when trial structures are generated. If the value of any of these elements exceeds the corresponding bounds, it is reset to a median value between the parent and the boundary. This procedure allows the incorporation of geometrical limits *i.e.* prior knowledge of areas of molecular conformation, while enhancing the efficiency of the search rather than disrupting the natural optimisation pathways.

The powder diffraction pattern was indexed giving a triclinic unit cell, consistent with the presence of a single molecule in the asymmetric unit. The space group was initially assumed to be $P1$ and the whole molecule (excluding amino Hs) used as the model for structure solution (Scheme 1). The molecule was constructed using standard bond lengths and angles and described by eight elements: three angles used to define the overall orientation of the molecule, and five flexible torsion angles to define the conformation (with an upper and lower bound of 360° and 0° in each case).

Scheme 1 Structural model of adipamide used for DE structure solution. Variable torsion angles are indicated by arrows.

The DE structure solution calculation was run several times with parameters $K = 1$, $F = 0.3$, $N_p = 220$, $G_{max} = 200$. In each case, the same structure solution was located with $R_{wp} \approx 11.4\%$ ($R_{wp} \approx 40\%$ for randomly generated structures) (Fig. 1). This structure was used as the starting point for Rietveld refinement (final agreement factors $R_{wp} = 9.77\%$, $\chi^2 = 7.81$). A comparison between the position of the adipamide molecule found in the DE solution and the refined structure (Fig. 2) shows how effectively the DE method locates the solution corresponding to a global minimum in R_{wp}. The minimum, maximum and

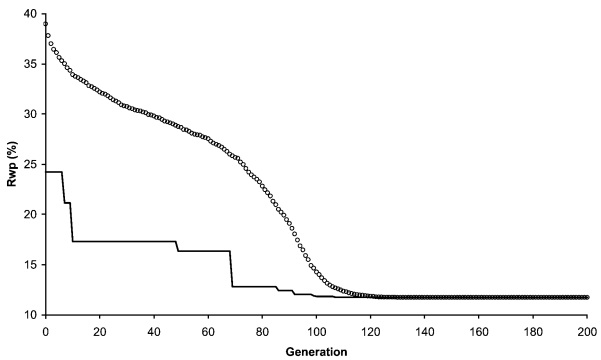

Fig. 1 DE progress plot showing the best R_{wp} (line) and mean R_{wp} (open circles) for each generation.

Fig. 2 Comparison between the position of the adipamide molecule found in the DE solution (black) and the final refined solution (grey).

DOI: 10.1039/b200436d

mean distances between pairs of corresponding atoms are 0.09, 0.36 and 0.20 Å for non-H atoms, and 0.32, 0.79 and 0.48 Å for methylene Hs. Analysis of the conformation of the refined structure suggested that the molecule may lie on an inversion centre, and that the space group is $P\bar{1}$. ^{13}C MASNMR confirmed the presence of only three crystallographically distinct C environments, and hence the structure was modified and successfully re-refined in the correct space group† (Fig. 3).

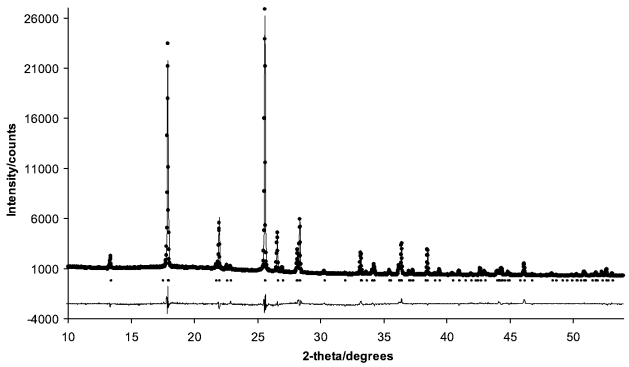

Fig. 3 Final observed (circles), calculated (solid line) and difference (below) X-ray powder diffraction profile for the final Rietveld refinement of adipamide form 2. Reflection positions are also marked.

The crystal structures of both adipamide polymorphs contain molecules that are effectively planar (except methylene Hs) and form continuous sheets generated by N–H···O hydrogen bonds. In the case of the triclinic form, the amino N acts as a double hydrogen bond donor; *via* one amino H to carbonyl O producing centrosymmetric amide dimers with the distinctive $R^2_2(8)$ motif, and *via* the second amino H to another carbonyl O generating a C(4) chain running in the [100] direction. Combination of these two motifs generates a secondary network of $R^2_4(8)$ rings forming a ladder of alternating ring motifs.[12] These ladders are linked through the molecules themselves to form continuous hydrogen bonded sheets stacked parallel to the (011) plane, revealing an intermolecular network typical of primary diamides (Fig. 4). Similar C(4) chains are observed in the monoclinic form,[13] although propagation about the 2_1 axis generates a ladder of $R^2_3(8)$ rings (rather than dimers), forming sheets parallel to the (101) plane. Although the presence of both polymorphs was originally reported *via* the identification of distinct crystal morphologies,[13] this structural study should enable an investigation of the exact nature of the relationship between these polymorphs, *i.e.* solvent dependency and thermodynamic stability.

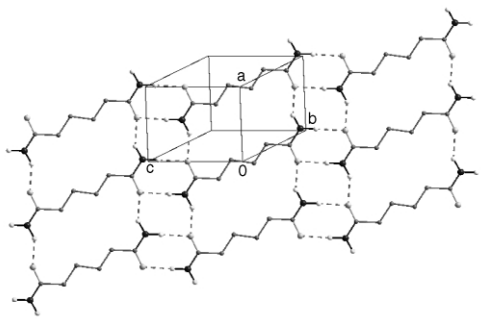

Fig. 4 View of the crystal structure of adipamide form 2 showing a hydrogen-bonded sheet parallel to the (011) plane. N–H···O hydrogen bonds are shown as dashed lines (other H atoms are omitted for clarity).

We have demonstrated the application of a powerful new global optimisation technique to direct space structure solution from powder diffraction data through the elucidation of the second polymorph of adipamide. In our tests, the differential evolution algorithm has proved robust, reliable and easy to use when applied to both rigid and flexible systems from conventional laboratory powder data, in this case affected by preferred orientation. Further optimisation of our control parameters has the potential to eliminate the need for multiple runs, resulting in a highly efficient evolutionary structure solution process.

We thank the Royal Society (URF to MT) and GlaxoSmithKline for their support.

Notes and references

† *Crystal data for adipamide* $C_6H_{12}N_2O_2$: M_r = 144.16, triclinic, a = 5.1097(2), b = 5.5722(2), c = 7.0472(3) Å, α = 69.577(1), β = 87.120(3), γ = 75.465(3)°, V = 181.87(2) Å^3, space group $P\bar{1}$ (no. 2), Z = 1, D_c = 1.243(1) g cm^{-3}, T = 273 K. CCDC 177914. See http://www.rsc.org/suppdata/cc/b2/b200436d/ for crystallographic files in .cif or other electronic format.

Data collection and Rietveld refinement: the sample was purchased from Aldrich and used as supplied. The powder diffraction data ($4 \leqslant 2\theta \leqslant 54°$ in 0.019° steps over 1 h) were collected on a Bruker-AXS D5000 using Ge monochromated Cu-Kα_1 radiation and a linear PSD covering 8° in 2θ. All atom positions (except the amide H atoms in calculated positions) were refined using geometric restraints and isotropic displacement parameters (refined for non-H only) constrained by atom type. A preferred orientation parameter was also refined in the [100] direction: final ratio = 1.102 (confirmed by comparison of data collected in disc and capillary geometries). Final refinement gave R_{wp} = 8.23%, R_p = 5.91%, χ^2 = 5.45 for 73 reflections and 43 parameters.

1 G. A. Stephenson, *J. Pharm. Sci*., 2000, **89**, 958.
2 E. J. Maclean, M. Tremayne, B. M. Kariuki, K. D. M. Harris, A. F. M. Iqbal and Z. Hao, *J. Chem. Soc., Perkin Trans. 2*, 2000, 1513.
3 K. D. M. Harris, M. Tremayne and B. M. Kariuki, *Angew. Chem., Int. Ed*., 2001, **40**, 1626.
4 K. D. M. Harris, M. Tremayne, P. Lightfoot and P. G. Bruce, *J. Am. Chem. Soc*., 1994, **116**, 3543.
5 M. Tremayne and C. Glidewell, *Chem. Commun*., 2000, 2425.
6 W. I. F. David, K. Shankland and N. Shankland, *Chem. Commun*., 1998, 931.
7 Y. G. Andreev, P. Lightfoot and P. G. Bruce, *J. Appl. Crystallogr*., 1997, **30**, 294.
8 K. D. M. Harris, R. L. Johnston and B. M. Kariuki, *Acta Crystallogr., Sect. A*, 1998, **54**, 632.
9 E. Tedesco, G. W. Turner, K. D. M. Harris, R. L. Johnston and B. M. Kariuki, *Angew. Chem., Int. Ed*., 2000, **39**, 4488.
10 R. Storn and K. V. Price, *J. Global Optimization*, 1997, **11**, 341.
11 K. V. Price, *New Ideas in Optimisation*, D. Corne, M. Dorigo and F. Glover, McGraw-Hill, London, UK, 1999, p. 77.
12 J. Bernstein, R. E. Davis, L. Shimoni and N.-L. Chang, *Angew. Chem., Int. Ed. Engl*., 1995, **34**, 1555.
13 M. Hospital and J. Housty, *Acta Crystallogr*., 1966, **20**, 626.