Solving molecular crystal structures from laboratory X-ray powder diffraction data with *DASH*: the state of the art and challenges

Alastair J. Florence,[a]* Norman Shankland,[a,b] Kenneth Shankland,[c] William I. F. David,[c] Elna Pidcock,[d] Xuelian Xu,[a] Andrea Johnston,[a] Alan R. Kennedy,[e] Philip J. Cox,[f] John S. O. Evans,[g] Gerald Steele,[h] Stephen D. Cosgrove[h] and Christopher S. Frampton[i]

[a]Department of Pharmaceutical Sciences, University of Strathclyde, 27 Taylor Street, Glasgow G4 0NR, UK, [b]CrystallografX Ltd, 38 Queen Street, Glasgow G1 3DX, UK, [c]ISIS Facility, Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, UK, [d]Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK, [e]Department of Pure and Applied Chemistry, University of Strathclyde, Glasgow G1 1XL, UK, [f]School of Pharmacy, The Robert Gordon University, Schoolhill, Aberdeen AB10 1FR, UK, [g]University of Durham, Department of Chemistry, University Science laboratories, South Road, Durham DH1 3LE, UK, [h]AstraZeneca R&D Charnwood, Loughborough, Leicestershire LE11 5RH, UK, and [i]Pharmorphix Ltd, 250 Cambridge Science Park, Milton Road, Cambridge CB4 0WE, UK. Correspondence e-mail: alastair.florence@strath.ac.uk

The crystal structures of 35 molecular compounds have been redetermined from laboratory monochromatic capillary transmission X-ray powder diffraction data using the simulated-annealing approach embodied within the *DASH* structure solution package. The compounds represent industrially relevant areas (pharmaceuticals; metal coordination compounds; nonlinear optical materials; dyes) in which the research groups in this multi-centre study are active. The molecules were specifically selected to form a series within which the degree of structural complexity (*i.e.* degrees of freedom in the global optimization) increased systematically, the degrees of freedom increasing with increasing number of optimizable torsion angles in the structural model and with the inclusion of positional disorder or multiple fragments (counterions; crystallization solvent; $Z' > 1$). At the lower end of the complexity scale, the structure was solved with excellent reproducibility and high accuracy. At the opposite end of the scale, the more complex search space offered a significant challenge to the global optimization procedure and it was demonstrated that the inclusion of modal torsional constraints, derived from the Cambridge Structural Database, offered significant benefits in terms of increasing the frequency of successful structure solution by restricting the magnitude of the search space in the global optimization.

1. Introduction

Global optimization methods for crystal structure determination from powder diffraction data (SDPD) have become widely available in recent years and have successfully been applied to solve the structures of organic (Harris & Cheung, 2003; Johnston *et al.*, 2004; Rukiah *et al.*, 2004; Zaske *et al.*, 2004), inorganic (Deem & Newsam, 1989; Edgar *et al.*, 2002; Reinaudi *et al.*, 2000) and organometallic (Ivashevskaja *et al.*, 2002; Dinnebier *et al.*, 2000) materials, to cite but a few examples. The basis of global optimization strategies has been fully described elsewhere (Shankland & David, 2002) and software implementing global optimization methods is now widely available [*e.g. DASH* (David *et al.*, 2001), *ESPOIR* (Le Bail, 2001), *FOX* (Favre-Nicolin & Cerny, 2002), *PowderSolve* (Engel *et al.*, 1999), *TOPAS* (Coelho, 2003)].

It is the application of global optimization methods to structure determination from data collected on standard, widely available, laboratory diffractometers that concerns us here. Specifically, the aim is to quantify the accuracy of a series of crystal structures solved from laboratory X-ray powder diffraction (XRPD) data using the simulated-annealing (SA) approach (David *et al.*, 1998) implemented in the *DASH* structure solution package and to investigate factors influencing the chances of successful structure solution.

research papers

1.1. Data quality

To maximize the chances of successfully and accurately solving crystal structures from laboratory XRPD data, the following data requirements should be addressed: accurate measurement of reflection positions and intensities; high angular resolution (*i.e.* small FWHM) and spatial resolution (*ca* 2 Å or better); good signal-to-background ratios across the full pattern and minimal preferred orientation (PO) effects. These requirements are best achieved in the laboratory with the sample mounted in a rotating capillary and the data collected in transmission geometry using monochromatic Cu $K\alpha_1$ radiation. Linear one-dimensional position-sensitive detectors (PSDs) combine excellent angular resolution with favourable count rates; recent developments in solid-state PSDs offer the prospect of even greater improvements in performance with respect to background, sensitivity and data acquisition rates.
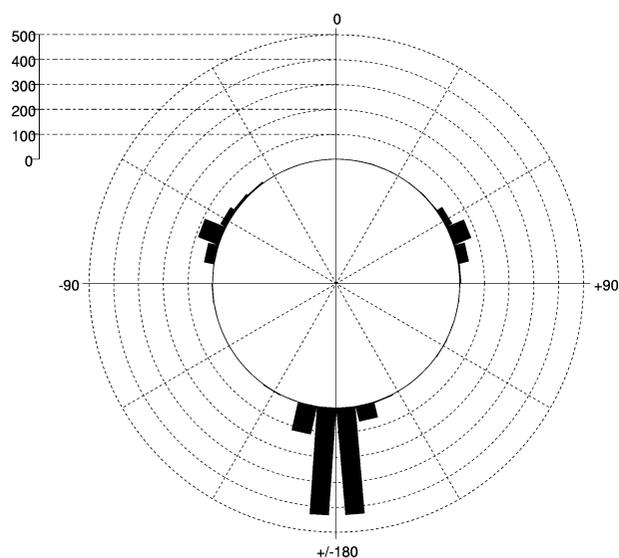
Figure 1
A polar plot showing the torsion-angle values obtained from a search of the CSD for the fragment CH_2–CH_2–CH_2–$C(=O)$. All C atoms were defined as acyclic. The distribution shows three clear modes (*i.e.* a trimodal distribution) centred on *ca* 60, 180, −60°, with the largest number of structures within the distribution adopting a *trans* conformation at this torsion angle, *i.e.* in the mode centred around 180°.

1.2. Maximizing the chances of success

There are a range of strategies which can be generally applied to maximize the chances of successfully solving a crystal structure from laboratory XRPD data, including those summarized in Table 1. Of particular interest in the study of complex structures is the incorporation of prior chemical information in the form of torsion-angle constraints. These constraints do not reduce the number of degrees of freedom (DOF) to be optimized during the search, but do reduce the extent of the search space explored during the SA process, a strategy that has been shown to be highly effective in SDPD (Middleton *et al.*, 2002; torsion-angle constraints derived by solid-state NMR conformational analysis). For the constraints approach to become amenable to routine application, the derivation of the constraints for any given problem has to be as straightforward as possible. Fortunately, this is readily tractable with the Cambridge Structural Database (CSD; Allen, 2002) and single-range torsion-angle constraints (*e.g.* 40 to 160°) derived from the CSD have previously been used to increase the frequency of success in global optimization structure solution (Shankland *et al.*, 1998).

Although it has always been possible to input single-range constraints into the *DASH* program, modal torsion constraints, whereby multiple ranges are defined for each individual torsion angle, offer a more selective means of constraining complex optimization problems. It has been found that, in general, the values of specific torsion angles within crystal structures in the CSD will form distributions that can be classified as uni-, bi- or trimodal. For example, the $C-C-C-C$ torsion angle defined by $CH_2-CH_2-CH_2-C(=O)$ adopts values that fall into three ranges; 40 to 80°, 160 to −160° and −80 to −40° (Fig. 1; see also §3). Accordingly, this torsion angle is classified as 'trimodal', with the middle of each discrete cluster separated by 120°. Thus, a lower and upper bound for a single mode of the torsion angle may be input *via* the *DASH* interface (*e.g.* 40 to 80°) along with the modal type (*i.e.* trimodal), whereupon the program automatically generates the bounds for the other two modes. The SA algorithm then samples torsion-angle values from these three ranges during the search. The ability to sample only relevant regions of torsion-angle space is potentially advantageous in solving crystal structures with a large number of

Table 2
Compound name and molecular formula with the reference code used throughout the text.

Code	Name	Molecular formula
1	Hydrochlorothiazide	$C_7H_8ClN_3O_4S_2$
2	2-Mercaptobenzoic acid	$C_7H_6O_2S$
3	N,N'-Bis[1-pyridin-4-yl-meth-(E)-ylidene]hydrazine	$C_{12}H_{10}N_4$
4	Carbamazepine (β polymorph)	$C_{15}H_{12}N_2O$
5	Dapsone	$C_{12}H_{12}N_2O_2S$
6	Hydroflumethiazide	$C_8H_8F_3N_3O_4S_2$
7	Paracetamol (form I polymorph)	$C_8H_9NO_2$
8	Paracetamol (form II polymorph)	$C_8H_9NO_2$
9	Phenylacetic acid	$C_8H_8O_2$
10	2-(Phenylsulfonyl)acetamide	$C_8H_9NO_3S$
11	Captopril	$C_9H_{15}NO_3S$
12	Methyl 4-[(4-aminophenyl)ethynyl]-benzoate	$C_{16}H_{13}NO_2$
13	$trans$-Dichlorobis(triphenylphos-phine)nickel(II)	$C_{36}H_{30}Cl_2NiP_2$
14	2-(4-Hydroxy-2-oxo-2,3-dihydro-1,3-benzothiazol-7-yl)ethylammonium chloride	$C_9H_{11}N_2O_2S\cdot Cl$
15	Salbutamol	$C_{13}H_{21}NO_3$
16	$trans$-Diisothiocyanatobis(triphenyl-phosphine)nickel(II)	$C_{38}H_{30}N_2NiP_2S_2$
17	Dopamine hydrobromide	$C_8H_{12}NO_2\cdot Br$
18	Methyl 4-{[4-(dimethylamino)phenyl]-ethynyl}benzoate	$C_{18}H_{17}NO_2$
19	cis-Thiothixene	$C_{23}H_{29}N_3O_2S_2$
20	Chlorpropamide	$C_{10}H_{13}ClN_2O_3S$
21	Creatine monohydrate	$C_4H_9N_3O_2\cdot H_2O$
22	1,4-Bis(2-phenethyloxyethanesulfo-nyl)piperazine	$C_{24}H_{34}N_2O_6S_2$
23	Clomipramine hydrochloride	$C_{19}H_{24}ClN_2\cdot Cl$
24	α-Lactose monohydrate	$C_{12}H_{22}O_{11}\cdot H_2O$
25	Promazine hydrochloride	$C_{17}H_{21}N_2S\cdot Cl$
26	Tolbutamide	$C_{12}H_{18}N_2O_3S$
27	Carbamazepine dihydrate	$C_{15}H_{12}N_2O\cdot 2H_2O$
28	Famotidine	$C_8H_{15}N_7O_2S_3$
29	Diltiazem hydrochloride	$C_{22}H_{27}N_2O_4S\cdot Cl$
30	Zopiclone dihydrate	$C_{17}H_{17}ClN_6O_3\cdot 2H_2O$
31	Capsaicin	$C_{18}H_{27}NO_3$
32	Sodium 4-[(E)-(4-hydroxyphenyl)dia-zenyl]benzene sulfonate dihydrate	$C_{12}H_9N_2O_4S\cdot Na\cdot 2H_2O$
33	2-{[3-(2-Phenylethoxy)propyl]sul-fonyl}ethyl benzoate	$C_{20}H_{24}O_5S$
34	S-Ibuprofen	$C_{13}H_{18}O_2$
35	Verapamil hydrochloride	$C_{27}H_{39}N_2O_4\cdot Cl$

Table 3
Instrumental and data collection parameters.

Typical instrument settings	
System	D8 Advance $\theta/2\theta$
Generator	50 kV, 40 mA
Measuring diameter (mm)	435
Radiation (Å)	Cu $K\alpha_1$, $\lambda = 1.54056$ Å
Monochromator	Primary, focusing curved Ge 111
Geometry	Transmission capillary configuration
Sample holder	0.7 mm borosilicate glass capillary
Detector	PSD system MBraun OED-50M
Typical measuring conditions	
Range (° 2θ)	5–65
Step size (° 2θ)	0.0145
Step time (s)	10.0
Total data collection time (h)	ca 10

were lightly ground in an agate mortar and pestle and filled into 0.7 mm borosilicate glass capillaries prior to being mounted and aligned on a Bruker-AXS D8 Advance powder diffractometer (Table 3). Compound **8**, the orthorhombic form of paracetamol (form II), was prepared *in situ* by cooling a molten sample of paracetamol to room temperature inside a 0.7 mm borosilicate glass capillary.

All data were collected at room temperature and can be accessed at http://www.powderdata.info.

3. Data analysis and simulated annealing

Diffraction patterns were indexed using *DICVOL91* (Boultif & Louër, 1991) to obtain lattice parameters that were subsequently refined (Table 4) along with background, zero point, peak shape parameters and reflection intensities in a Pawley fit (Pawley, 1981) using *DASH*. All samples gave sharp diffraction, with good to moderate angular resolution and a mean FWHM = 0.099 ± 0.015 Å (Table 5). Data were truncated as necessary to allow up to 350 reflections to be extracted from each pattern, with the spatial resolution across all of the data sets ranging from 1.44 to 2.18 Å (Table 5).

Z matrices describing the molecular topology of the fragments in each compound were generated automatically from the appropriate reference crystal structure[1] using *DASH*, and all optimizable torsion angles were automatically assigned to vary in the range −180 to 180°. A single O atom was used to approximate each water molecule of crystallization in hydrates **21**, **24**, **27**, **30** and **32** and the *Z* matrices of **27** and **33** were manually altered to accommodate positional disorder.

Global optimization of all external (rotational and translational) and internal (torsion angles) DOF (Table 5) against the extracted intensities was carried out with all *DASH* SA control parameters set to default values. 20 runs with 1×10^7

internal DOF. One important caveat to the general applicability of this approach is the finite possibility that the conformation adopted at a specific torsion angle within the molecule of interest may lie outside the ranges measured from known structures within the CSD.

2. Data collection

The 35 compounds used in the study were selected to cover a wide range of structural complexity, including significant conformational (torsional) flexibility, salts, solvates, positional disorder (**27** and **33**) and $Z' > 1$ (**34**) (Table 2, Fig. 2). A prerequisite for inclusion in the study was the availability of reference crystal structures for the purpose of evaluating the accuracy of the structures solved using the SDPD approach (see §3). All polycrystalline samples (except compound **8**)

---

[1] The recommended, and indeed simplest, way to construct an accurate *Z* matrix is from a related crystal structure, such as a polymorphic, salt or solvated form of the molecule of interest. In the absence of such a structure, mean values for bond lengths, covalent bond angles and non-optimizable torsion angles are preferably extracted from the CSD (Allen, 2002) and input into the *Z* matrix. This requirement to input the chemical formula and connectivity of fragments is, in fact, one of the drawbacks of global optimization methods for solving unknown crystal structures.
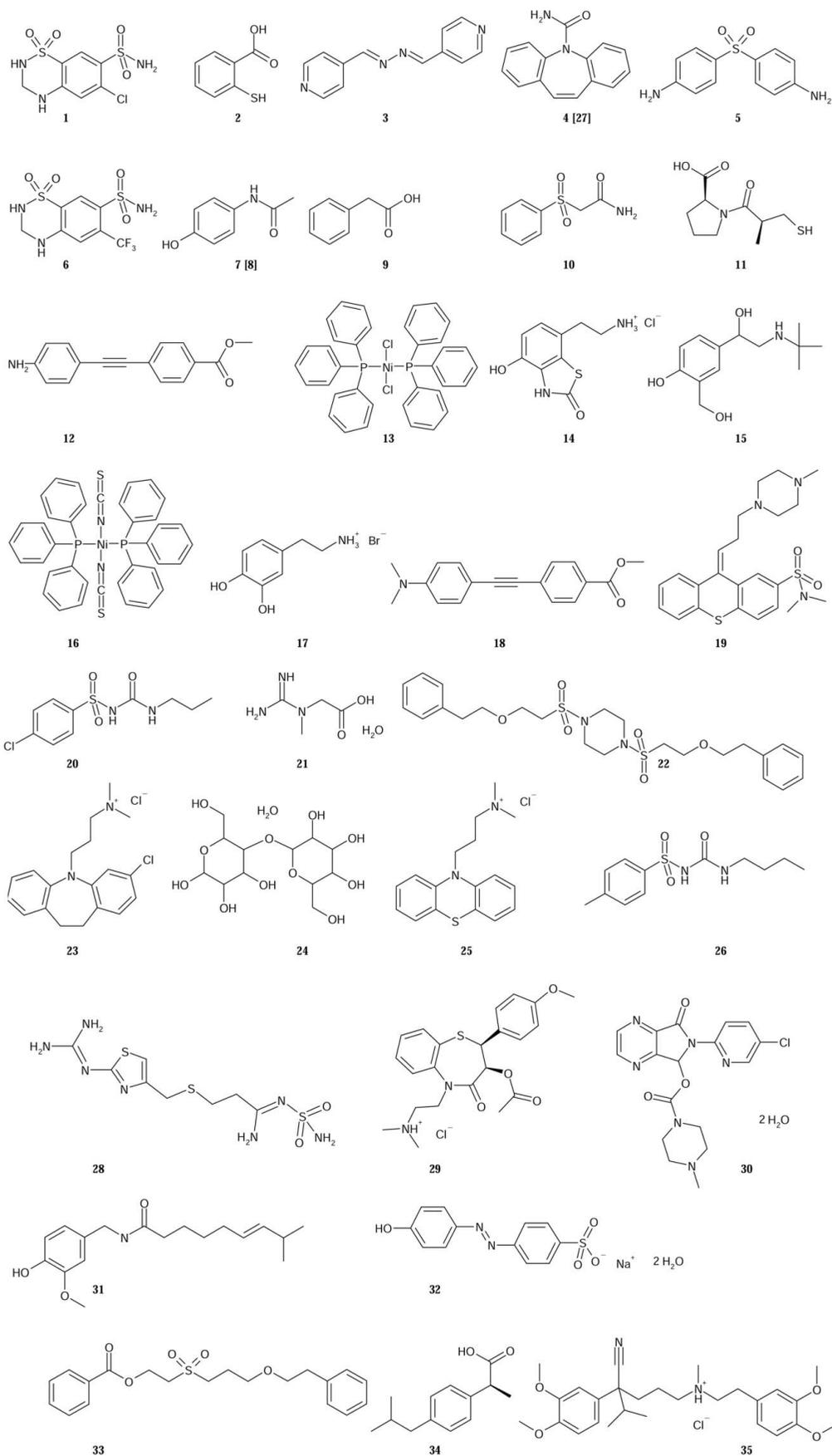
Figure 2
Molecular structures of compounds **1–35**.

Table 4
Space group and refined unit-cell parameters (this work) for compounds **1–35**.

The last column identifies the reference crystal structures (typically CSD refcode/CCDC deposition number) used to calculate the RMSD values in Table 5. The structures of **6**, **14**, **22**, **27**, **28** and **33**, in CIF format, can be obtained free of charge *via* http://www.ccdc.cam.ac.uk/conts/retrieving.html (or from the CCDC, 12 Union Road, Cambridge CB2 1EZ, UK; fax: +44 1223 336033; e-mail: deposit@ccdc.cam.ac.uk). For the other structures, the individual citations corresponding to each refcode are given in Table 8.
Pawley χ^2 is the profile χ^2 for the Pawley fit as described in the *DASH* manual: $\chi^2 = \left\{ \sum_i^N w_i[y_i(\mathrm{obs}) - y_i(\mathrm{calc})]^2 \right\} / (N - P + C)$, where $y_i(\mathrm{obs})$ is the observed intensity at the ith step in the powder diffraction pattern, $y_i(\mathrm{calc})$ is the associated calculated intensity; $w_i = 1/\sigma_i^2$, where σ_i is the standard deviation of the observed intensity at that point. The summation is performed over all N data points; $(N - P + C) = $ (number of data points) − (number of parameters) + (number of parameter constraints).

Code	Space group	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)	Pawley χ^2	Reference
1	$P2_1$	7.400	8.506	10.006	90	111.72	90	2.18	HCSBTZ
2	$P2_1/c$	7.885	5.976	14.949	90	100.48	90	2.84	ZZZLWW01
3	$P2_1/c$	3.848	11.005	12.727	90	92.38	90	4.25	LIZCUS
4	$P2_1/n$	7.537	11.157	13.918	90	92.87	90	2.98	CBMZPN10
5	$P2_12_12_1$	25.538	8.061	5.762	90	90	90	3.10	DAPSUO10
6	$P2_1$	7.636	8.662	9.743	90	110.30	90	3.92	CCDC 198487
7	$P2_1/n$	7.100	9.380	11.708	90	97.42	90	5.42	HXACAN07
8	$Pbca$	17.142	11.822	7.404	90	90	90	2.82	HXACAN08
9	$P2_1/a$	10.226	4.967	14.467	90	99.25	90	7.72	ZZZMLY01
10	$P2_1/c$	8.884	5.408	19.469	90	101.66	90	4.77	Frampton (2004)
11	$P2_12_12_1$	8.810	17.948	6.834	90	90	90	2.65	MCPRPL
12	$P2_1$	7.572	5.908	14.141	90	95.34	90	2.80	Marder (2004)
13	$P2/c$	11.638	8.197	17.388	90	107.03	90	3.93	CLTPNI03
14	$P2_1/a$	7.555	14.640	10.246	90	109.30	90	1.86	CCDC 247129
15	$Pbca$	21.657	8.783	14.555	90	90	90	2.94	BHHPHE
16	$P\bar{1}$	7.958	10.488	11.500	111.10	74.56	92.29	4.36	GEBZUI
17	$Pbc2_1$	10.671	11.459	7.950	90	90	90	7.61	QQQAEJ01
18	$Pna2_1$	6.121	7.472	33.002	90	90	90	2.97	Marder (2004)
19	$P2_1$	10.141	8.695	13.683	90	110.60	90	3.13	THTHXN01
20	$P2_12_12_1$	9.078	5.220	26.658	90	90	90	6.87	BEDMIG
21	$P2_1/c$	12.506	5.046	12.169	90	108.88	90	3.06	CREATH03
22	$P2_1/a$	13.442	5.182	19.796	90	108.74	90	3.10	CCDC 247131
23	$P2_1/c$	15.514	8.610	14.035	90	96.93	90	3.81	CIMPRA
24	$P2_1$	7.937	21.573	4.814	90	109.75	90	2.59	LACTOS10
25	$P2_1/c$	11.806	11.497	13.429	90	111.71	90	2.41	PROMZC01
26	$Pna2_1$	20.218	7.820	9.072	90	90	90	7.67	ZZZPUS02
27	$Cmca$	19.775	4.937	28.719	90	90	90	7.87	CCDC 247132
28	$P2_1/c$	17.767	5.534	18.311	90	123.64	90	2.70	CCDC 198488
29	$P2_12_12_1$	42.190	9.075	6.037	90	90	90	8.29	CEYHUJ01
30	$P2_1/c$	16.479	7.145	17.398	90	108.80	90	3.72	UCUVET
31	$P2_1/c$	12.672	14.980	9.426	90	93.69	90	6.90	Frampton (2004)
32	$Pbcn$	14.591	5.831	32.952	90	90	90	4.39	YAYWUQ
33	$P2_1/n$	5.137	37.934	9.844	90	98.50	90	4.22	CCDC 247130
34	$P2_1$	12.463	8.029	13.538	90	112.93	90	2.09	JEKNOC10
35	$P\bar{1}$	7.089	10.593	19.207	100.11	93.75	101.56	4.70	CURHOM

that the appropriate torsion values from hits were recorded. Hits of the search were viewed in *Vista* (CCDC software) and the torsion-angle ranges to be used in *DASH* were chosen by inspection (Table 6). A modal torsion-angle range was not specified if there were less than 30 observations or if there was no clear distribution.

4. Results and discussion

The crystal structures of all compounds were solved successfully[2] and the correct solution obtained with excellent reproducibility in the majority of cases (Table 5). For compounds **1–27**, with DOF < 15, correct solutions were generated in ~100% of SA runs, with a relatively narrow spread in the $\chi^2_{\mathrm{profile}}$ range observed for any one compound.

For the more complex structures **28–35**, with DOF ≥ 15, the adverse effect of local minima in the agreement hypersurface is reflected in the reduced frequency of success and the accompanying increased spread in $\chi^2_{\mathrm{profile}}$ for a particular compound. The exceptions to this are **30** (DOF = 16; N_{sol} = 17) and **32** (DOF = 18; N_{sol} = 19), which have in common a high degree of planar aromatic structure and a small number of internal DOF (4 and 3 DOF, respectively). This combination of molecular features clearly favoured success in reaching the global minimum in the SA runs.

For the level of success achieved with **28–34**, a batch size of 20 SA runs proved sufficient to solve the structure reproducibly and, therefore, convincingly. The same does not hold true for the most complex example, **35**, which returned only one solution in 20 runs; this aspect of verapamil hydrochloride is discussed further in §4.2.3.

SA moves per run were implemented for each structure determination, with a simplex refinement being executed upon completion. The structure of the best solution (*i.e.* that with the lowest profile χ^2) was overlaid upon the corresponding reference crystal structure and the root mean square displacement (RMSD, Å) calculated for all non-H atoms (Table 5). The majority of the reference data comprised single-crystal structures retrieved from the CSD (Table 4). In instances where the data collection temperatures for the XRPD and reference structures are not matched, the magnitude of the RMSD value necessarily contains a contribution which is attributable to this temperature difference.

SA runs for **28**, **29**, **31**, **34** and **35** were repeated using modal torsional constraints (§1.2). The CSD (November 2002, v5.24) was searched for fragments of molecules corresponding to the torsion angle of interest using *Conquest* (Bruno *et al.*, 2002). The torsion angle was specified as a geometric parameter so

---

[2] A reliable indicator of how close the structure is to the global minimum is obtained by taking the ratio $\chi^2_{\mathrm{profile}} / \chi^2_{\mathrm{Pawley}}$; the smaller the ratio, the more likely it is that the correct solution has been obtained. Favourable values for this ratio typically range from 2 to 10, the higher ratios often indicating that additional details (such as PO or positional disorder) need to be factored into the model. In determining N_{sol} in Table 5, all structures with $\chi^2_{\mathrm{profile}} / \chi^2_{\mathrm{Pawley}} < 10$ were considered to be solved and the solutions were confirmed by subsequent comparison with the known crystal structure.

Table 5
Summary of results of the SA runs for compounds **1**–**35**.

FWHM = average full width at half-maximum of 8 reflections measured in the XRPD data sets out to *ca* 30° 2θ; DOF (ext) = number of optimized external degrees of freedom; DOF (int) = number of optimized internal degrees of freedom; Data range = data range used in the Pawley fit; N_{refs} = number of reflections in the fitted data range; Res. = spatial resolution of data used in the SA runs; Profile χ^2 = range of profile χ^2 values observed at the end of 20 SA runs (χ^2 calculated as per footnote to Table 4); N_{sol} = number of correct structure solutions obtained from a batch of 20 SA runs; χ^2 ratio = $\chi^2_{\text{profile}}/\chi^2_{\text{Pawley}}$ for the best solution, *i.e.* that with the lowest profile χ^2; RMSD = root mean square displacement (as defined in §3) for the best solution.

Code	FWHM (° 2θ)	DOF (ext)	DOF (int)	Data range (° 2θ)	N_{refs}	Res. (Å)	Profile χ^2	N_{sol}	χ^2 ratio	RMSD (Å)
1	0.097	6	1	8.0–55.8	157	1.65	6.75–18.97	20	3.1	0.102
2	0.095	6	1	8.5–64.6	244	1.44	5.41–5.42	20	1.9	0.026
3	0.102	6	1	5.0–54.7	121	1.68	9.69–9.88	20	2.3	0.026
4	0.093	6	1	7.5–56.4	284	1.63	9.04–12.28	20	3.0	0.017
5	0.108	6	2	5.0–56.3	196	1.63	11.45–13.77	20	3.7	0.027
6	0.117	6	2	9.0–45.7	94	1.99	13.62–13.98	20	3.5	0.131
7	0.080	6	2	10.0–64.7	268	1.44	13.56–20.05	20	2.5	0.092
8	0.138	6	2	9.0–64.9	265	1.44	19.26–23.87 †	20	6.8	0.140
9	0.112	6	2	5.0–56.9	171	1.63	22.92–24.26	20	3.0	0.077
10	0.085	6	3	8.0–49.9	153	1.83	9.96–10.24	20	2.1	0.079
11	0.101	6	4	9.0–55.2	169	1.66	12.63–12.73	20	4.8	0.077
12	0.138	6	4	4.0–44.6	97	2.03	4.16–4.82	20	1.5	0.109
13	0.084	6	4	5.0–52.2	288	1.75	8.04–19.45	20	2.0	0.070
14	0.092	6 + 3 ‡	2	5.0–52.1	315	1.75	4.63–7.07	20	2.5	0.058
15	0.095	6	5	6.0–51.8	263	1.76	27.51–29.05	20	9.4	0.088
16	0.108	6	5	6.5–47.9	264	1.90	15.63–18.15	20	3.6	0.058
17	0.104	6 + 3	2	7.0–49.6	101	1.84	34.25–43.23	20	4.5	0.075
18	0.125	6	5	4.0–44.3	98	2.04	5.14–6.05	20	1.7	0.066
19	0.097	6	5	5.0–48.9	214	1.88	22.95–55.94 †	19	7.3	0.129
20	0.097	6	6	5.0–49.2	149	1.85	23.43–24.95	20	3.4	0.070
21	0.088	6 + 3	3	5.0–62.4	232	1.49	15.56–161.00 †§	19	5.1	0.059
22	0.115	6	7	4.0–51.3	242	1.77	7.02–9.50	20	2.3	0.138
23	0.084	6 + 3	4	10.0–49.0	306	1.86	15.47–184.54 §	19	4.1	0.109
24	0.085	6 + 3	4	5.0–49.9	145	1.82	20.21–63.52	19	7.8	0.073
25	0.079	6 + 3	4	6.7–52.0	330	1.76	10.33–12.16	20	4.3	0.107
26	0.111	6	7	7.5–59.9	228	1.54	17.39–72.45	20	2.3	0.127
27	0.092	6 + 3 + 3	2	5.0–60.4	218	1.53	38.05–39.27	20	4.8	0.116
28	0.105	6	9	5.0–48.9	229	1.86	10.00–204.18	8	3.7	0.063
29	0.098	6 + 3	7	4.0–41.5	164	2.18	23.97–229.63	8	2.9	0.118
30	0.086	6 + 3 + 3	4	4.7–50.5	337	1.80	13.07–223.82	17	3.5	0.096
31	0.092	6	11	5.0–51.8	338	1.76	29.13–192.38	5	4.2	0.168
32	0.093	6 + 3 + 3 + 3	3	4.5–56.4	341	1.63	21.90–122.20 §	19	5.0	0.136
33	0.086	6	13	4.0–46.1	302	1.96	17.24–221.32 ¶	3	4.1	0.165
34	0.090	6 + 6	4 + 4	6.0–54.7	321	1.68	8.87–96.95	7	4.2	0.073
35	0.082	6 + 3	13	4.4–42.6	308	2.12	45.02–232.50	1	9.6	0.204

† PO correction included in SA runs for **8** ($r = 1.50$, [001]), **19** ($r = 0.88$, [010]), **21** ($r = 1.06$, [100]).   ‡ Six DOF for the cation and three for the anion in **14**. In other structures comprising >1 fragment, the number of DOF for each fragment that was optimized is identified in the table.   § For **21**, **23** and **32**, the largest profile χ^2 among the 19 correct solutions in each case was 22.79, 27.17 and 36.27, respectively.   ¶ Disordered (half-occupancy) phenyl ring model.

The excellent accuracy of the best SA solution, across the full range of 35 structures, is reflected in each case by the favourable $\chi^2_{\text{profile}}/\chi^2_{\text{Pawley}}$ ratio and the small RMSD. The latter ranged in value from 0.017 Å (**4**) to 0.204 Å (**35**), with a mean across all 35 structures equal to 0.093 ± 0.043 Å. This close agreement arises from the ability of the SA algorithm to 'fine tune' both the internal and external DOF. In each case, the atomic displacements of the best SA solution are within the radius of convergence of a typical Rietveld refinement (see §4.2.3 for the example of the restrained Rietveld refinement of the SA solution for structure **35**).

Critical examination of the fit to the diffraction data returned by the SA process is a key step in the structure determination process. The observation of a high $\chi^2_{\text{profile}}/\chi^2_{\text{Pawley}}$ ratio or significant misfit in any region of the diffraction pattern, or unfavourably short atom–atom contacts, is diagnostic of problems that are best addressed at the structure solution stage, prior to refinement. In practice, this means checking for the possibility of PO and consulting other available experimental data for any evidence of disorder in the structure.

Three representative examples from each of the two populations identified above (**1**–**27** and **28**–**35**) are now considered in §§4.1 and 4.2.

4.1. Compounds 1–27 with <15 DOF

4.1.1. Paracetamol form II. SDPD is a powerful means of solving the structures of metastable phases crystallized *in situ* in a glass capillary (Shankland *et al.*, 2001). In the case of metastable orthorhombic form II paracetamol (**8**), a polycrystalline sample was readily obtained in a capillary by cooling molten paracetamol to room temperature. Unsurprisingly, oriented growth of crystallites within the capillary necessitated a significant March–Dollase correction of intensities for PO in the data ($r = 1.50$, [001]; r determined as an optimizable parameter in the SA runs) (Dollase, 1986). With this PO correction, an accurate structure solution was obtained, with $\chi^2_{\text{profile}}/\chi^2_{\text{Pawley}} = 6.8$ and RMSD = 0.140 Å (*cf.* $\chi^2_{\text{profile}}/\chi^2_{\text{Pawley}} = 22.2$ and RMSD = 0.428 Å for the best SA solution with no PO correction included; Fig. 3).

4.1.2. Promazine hydrochloride. Promazine hydrochloride (**25**) falls in the mid-range of structural complexity represented by compounds **1**–**35**. The diffraction data are particularly high quality (Fig. 4) and yielded the lowest FWHM value in Table 4 (0.079°; *cf.* 0.059° for the 100 reflection of LaB_6 collected under the same conditions). The structure was solved with 100% success, the best SA solution returning $\chi^2_{\text{profile}}/\chi^2_{\text{Pawley}} = 4.3$ and RMSD = 0.107 Å.

4.1.3. cis-Thiothixene. The initial runs on *cis*-Thiothixene (**19**) returned a narrow profile χ^2 range (29.25–29.75; *cf.* $\chi^2_{\text{Pawley}} = 3.13$), but the high $\chi^2_{\text{profile}}/\chi^2_{\text{Pawley}}$ ratio raised the strong suspicion that the global minimum had not been reached. Inclusion of a PO correction in the SA re-runs ($r = 0.88$, [010]) confirmed this to be the case, with the $\chi^2_{\text{profile}}/\chi^2_{\text{Pawley}}$ ratio improving significantly to 7.3 (RMSD = 0.129 Å). A subsequent comparison of the initial structure with the global minimum structure showed that the orientation of the $-N(CH_3)_2$ group (*i.e.* eclipsing $-SO_2$) in the former was incorrect.

4.2. Compounds 28–35 with ≥15 DOF

4.2.1. 2-{[3-(2-Phenylethoxy)propyl]sulfonyl}ethyl benzoate. The single-crystal structure of compound **33** has a rota-

Table 6
CSD searches and modal ranges utilized in the modal constraints within *DASH*.

CSD search. Atom numbers correspond with the schemes used by Golič *et al.* (1989) (**28**); Kojic Prodic *et al.* (1984) (**29**); CCDC 171602 (David *et al.*, 1998) (**31**); Freer *et al.* (1993) (**34**) and Carpy *et al.* (1985) (**35**). The four atoms of the torsion angle were specified as 'cyclic' or 'acyclic' (subscripts c and a, respectively) and the appropriate bond types between the atoms of the torsion angle were also defined in the search ('$\sim$' is an unspecified bond type). In addition, where appropriate, the environment of the torsion-angle atoms was specified to narrow the *Conquest* search to include only very closely related fragments, *e.g.* number of H atoms attached or total number of coordinated atoms (T2 = only two connections to atom allowed; T3 = only three connections to atom allowed; T4 = only four connections to atom allowed).

Mode B = bimodal. Planar torsion-angle ranges (centred around 0° and 180°) are searched by inputting into *DASH* the bounds for a single mode of the torsion angle (*e.g.* −160° to 160°), along with the modal type (bimodal). The program automatically generates the complementary bounds for the other mode (in this case, −20° to 20°). For non-planar torsion angles, inputting, say, 30° to 50°, and specifying 'bimodal' will automatically generate −30° to −50° for the bounds of the other mode.

Mode T = trimodal. The modal type and the bounds for a single mode of a torsion angle (*e.g.* −160° to 160°) are input into *DASH* and the program automatically generates the bounds for the other two modes.

Code	Torsion angle	CSD search	N_{obs}	Range (°)	Mode
28	N1:C1:N3:C2	N_aH_2—C_a=N_a—C_c	51	−160 to 160	B
	S3:N5:C8:C7	S_aO_2—N_a~C_a—C_aH_2	124	60 to 180	B
	C8:C7:C6:S2	C_a—C_aH_2—C_aH_2—S_a	613	−160 to 160	T
	C3:C5:S2:C6	C_c—C_aH_2—S_a—C_aH_2	173	−150 to 150	T
	C7:C6:S2:C5	C_aH_2—C_aH_2—S_a—C_aH_2	233	−150 to 150	T
	N4:C3:C5:S2	N_c—C_c—C_aH_2—S_aT2	49	−150 to 150	T
	O1:S3:N5:C8	O_a=S_a—N_a~C_a	750	−150 to 150	T
	C1:N3:C2:N4	C_a=N_a—C_c=N_c	38	20 to −20	T
	N5:C8:C7:C6	N_a=C_a—C_aH_2—C_aH_2	49	No clear distribution	
29	C17:O2:C2:C1	C_a—O_a—C_aH—C_aH	1566	60 to 180	B
	C12:C10:C1:C2	C_cH—C_c—C_aH—C_aH	986	30 to 150	B
	O3:C17:O2:C2	O=C_a—O_a—C_aH	1958	−20 to 0	B
	C20:C19:N1:C3	C_aH_2—C_aH_2—N_aT3—C_a=O	892	60 to 180	B
	C16:O1:C13:C11	C_aH_3—O_a—C_c—C_cH	4126	−160 to 160	B
	N2:C20:C19:N1	N_aH^+—C_aH_2—C_aH_2—N_aT3	63	−150 to 150	T
	C21:N2:C20:C19	C_a—N_aH^+—C_aH_2—C_aH_2	1504	−150 to 150	T
31	C8:N1:C9:C10	C_aH_2—N_aH—C_a(=O)—C_aH_2	326	−160 to −180	B
	C1:C8:N1:C9	C_c—C_aH_2—N_aH—C_aH_2	82	60 to 125	B
	C15:C14:C13:C12	C_a—C_aH=C_aH—C_aH_2	159	90 to 160	B
	H20:C15:C14:C13	H—C_a=C_aH—C_aH_2	3877	−160 to 160	B
	C7:O1:C3:C2	C_aH_3—O_a—C_c—C_cH	4126	−160 to 160	B
	N1:C9:C10:C11	C—N_aH—C_a(=O)—C_aH_2—C_aH_2	280	60 to 180	B
	C13:C12:C11:C10	C_aH_2—C_aH_2—C_aH_2—C_aH_2	>2000	−150 to 150	T
	C14:C13:C12:C11	C_aH_2—C_aH_2—C_aH_2—C_aH_2	>2000	−150 to 150	T
	C12:C11:C10:C9	C_aH_2—C_aH_2—C_aH_2—C_a(=O) †	1075	−160 to 160	T
	C17:C16:C15:C14	C_a—C_aT4—C_aH=C_a	2501	−30 to 30	T
	C2:C1:C8:N1	C_c—C_c—C_aH_2—N_aHT3	516	No clear distribution	
34	C1:C2:C4:C5	C_c—C_c—C_aHT4—C_aT4	>10000	20 to 120	B
	C11:C10:C7:C6	C_c—C_c—C_aH_2—C_aT4	4087	50 to 130	B
	C12:C11:C10:C7	C_c—C_aH_2—C_aT4—C_aT4	1255	−150 to 150	T
	O1:C1:C2:C4	OH—C(=O)—C_aHT4—C_c	190	−150 to 150	T
35	C14:C13:C12:C11	C_c—C_c—C_aH_2—C_aT4	4087	50 to 130	B
	C2:C1:C7:C8	C_c—C_c—C_aH_2—C_aT4	4087	50 to 130	B
	C27:O4:C17:C18	C_aH_3—O_a—C_c—C_cH	4126	−160 to 160	B
	C26:O3:C16:C15	C_aH_3—O_a—C_c—C_cH	4126	−160 to 160	B
	C20:O2:C5:C6	C_aH_3—O_a—C_c—C_cH	4126	−160 to 160	B
	C19:O1:C4:C3	C_aH_3—O_a—C_c—C_cH	4126	−160 to 160	B
	C12:C11:C10:C9	C_aH_2—C_aH_2—C_aH_2—C_aH_2	>2000	−150 to 150	T
	C23:C22:C12:C11	C_aH_2—C_aH_2—C_aH_2—C_aH_2	>2000	−150 to 150	T
	C8:N1:C9:C10	C_aH_2—N_a[C(H)]—C_aH_2—C_aH_2	>1000	−150 to 150	T
	C9:N1:C8:C7	C_aH_2—N_a[C(H)]—C_aH_2—C_aH_2	>1000	−150 to 150	T
	C13:C12:C11:C10	C_c—C_a(C_2)—C_aH_2—C_aH_2	204	−150 to 150	T
	N1:C8:C7:C1	NHT4—C_aH_2—C_aH_2—C	2090	−150 to 150	T
	N1:C9:C10:C11	NHT4—C_aH_2—C_aH_2—C	2090	−150 to 150	T

† See Fig. 1.

tionally disordered phenyl ring, resulting in four of the C atoms being disordered over two sites. Accordingly, a *Z* matrix allowing for phenyl rotational disorder (19 DOF) was constructed by incorporating two independent half-occupancy phenyl rings.

The structure was solved successfully, the best SA solution yielding $\chi^2_{profile}/\chi^2_{Pawley}$ = 4.1, RMSD = 0.165 Å and accurate orientations for the half-occupancy phenyl rings (Fig. 5). As expected, a simplified model (18 DOF) with a fully ordered phenyl ring gave a solution that was largely correct, except for a compromise in the ring orientation (Fig. 5). Thus, as has been found elsewhere with disordered fragments (Graham *et al.*, 2004; Johnston *et al.*, 2004), a significant improvement in the fit to the diffraction data can be derived by including fractional-occupancy atoms in the global optimization, albeit at the cost of reducing the frequency of success (Table 5; eliminating one DOF in the fully ordered model increased the number of SA solutions with $\chi^2_{profile}/\chi^2_{Pawley}$ < 10 from three to nine).

There is a good correspondence among the top six SA solutions ($\chi^2_{profile}/\chi^2_{Pawley}$ = 4.1–14.2) with regard to the positions of the phenyl ring centroids and the S atoms (Fig. 6) and Fig. 7 confirms that the higher $\chi^2_{profile}/\chi^2_{Pawley}$ ratios reflect a decrease in the accuracy with which the positions of the acyclic backbone atoms (other than the S atom) are located.

In summary, in those instances where an SA search fails to reach an acceptably low $\chi^2_{profile}/\chi^2_{Pawley}$ ratio for any particular compound, it is probable that the regions of relatively high X-ray scattering power, at least, are located reliably, giving a partial structure around which further constrained SA runs may be instigated.

4.2.2. Capsaicin. Capsaicin (**31**; N_{sol} = 5) has two fewer internal DOF than disordered **33** (N_{sol} = 3). As was observed with **33**, the aromatic rings are located with good reproducibility and the higher $\chi^2_{profile}/\chi^2_{Pawley}$ ratios reflect a lower accuracy in locating the acyclic chain atoms (Fig. 8).

4.2.3. Verapamil hydrochloride. The best solution for verapamil hydrochloride (**35**, 22 DOF) yielded $\chi^2_{profile}/\chi^2_{Pawley}$ = 9.6 and RMSD = 0.204 Å. A restrained Rietveld refinement (*TOPAS*; Coelho, 2003) of the SA structure (R_{wp} = 6.3) against the raw data re-fitted to 65° 2θ (Pawley R_{wp} = 2.2) resulted in a significant improvement in accuracy, with a final R_{wp} of 3.1 and a reduced RMSD value of 0.134 Å, only slightly greater than the mean RMSD across all structures. A PO correction in the direction [100] was included in the refinement, the magnitude of
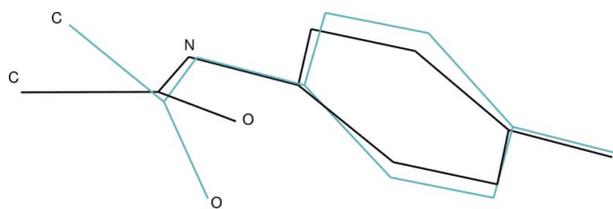
Figure 3
Overlay of the best SA solutions for **8**, with (black) and without (grey) a March–Dollase PO correction of intensities included in the SA searches. Without the PO correction, the aromatic ring is tilted *ca* 10° out of its correct position and the C–O–N plane suffers a rotation of some 37° (H atoms in this and subsequent figures have been omitted for clarity).
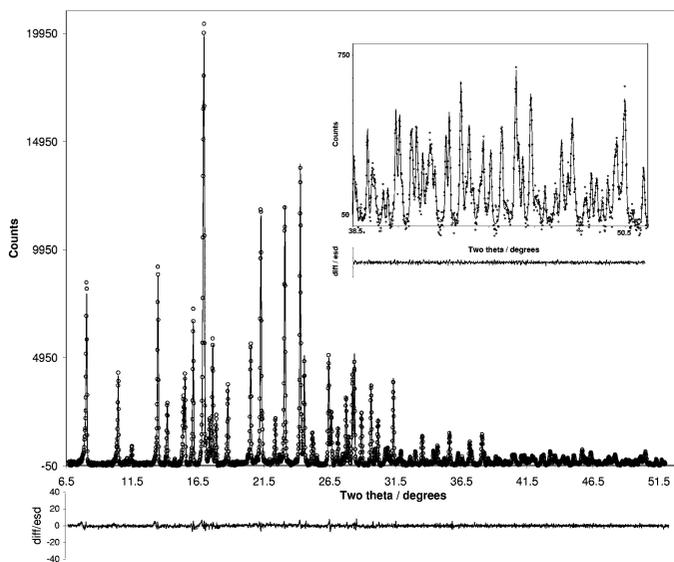
Figure 4
Observed profile (circles), calculated profile (line) and difference plot $[(y_{obs} - y_{calc})/\sigma(y_{obs})]$ of the Pawley fit for **25** (Pawley χ^2 = 2.41) in the range 6.7–51.5° 2θ. Inset: high-angle data in the range 31.5–51.5° 2θ showing the excellent fit to the data out to 1.76 Å resolution.

the correction (r = 0.91) being consistent with mild PO in the sample.

Whilst the single answer required to 'solve the structure' was obtained, the frequency with which low-lying areas of χ^2 space were visited was low (N_{sol} = 1). In such instances, a better sampling of low-lying space can be achieved by increasing the number of SA runs, increasing the number of moves per run and decreasing the cooling rate (Table 1). It is also reasonable to expect that the inclusion of modal torsional constraints will also increase the number of runs successfully locating low-lying regions of the search space (§1.2) and this aspect is reported in §4.2.4.

4.2.4. Results of constrained SA structure determinations. The application of modal torsional constraints increased the frequency with which the structures of **28**, **29**, **31** and **34** were solved (Table 7), yielding solutions with comparable RMSD values to those reported in Table 5. In the case of verapamil hydrochloride (**35**), the first attempt to apply the modal constraints failed to yield any correct structures (lowest profile χ^2 among 20 constrained SA runs equalled 110.51). Repeat

Table 7
Summary of results for SA runs constrained using the modal torsion constraints.

Column headings are as defined in Table 5.

Code	Profile χ^2	N_{sol}	χ^2 ratio	RMSD (Å)
28	9.97–216.23	13	3.7	0.084
29	24.03–194.05	12	2.9	0.093
31	36.83–195.19	8	5.3	0.226
34	11.26–70.22	9	5.4	0.056
35 †	23.20–248.91	2 ‡	4.9	0.092

† Batch size = 50 runs for compound **35** (20 runs for the others). The corresponding batch of 50 unconstrained SA runs for **35** (§4.2.4; cooling rate = 0.01, moves per run = 2 × 10^7) returned one correct solution in a profile χ^2 range of 28.08–230.10. ‡ Profile χ^2 = 23.20 and 28.56.
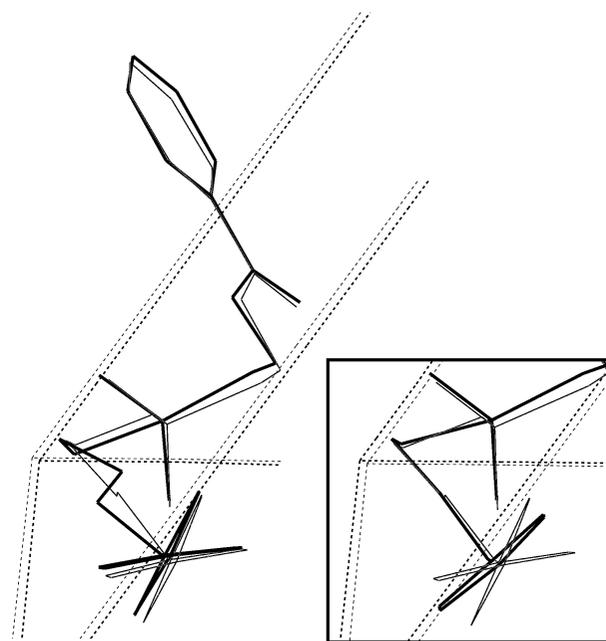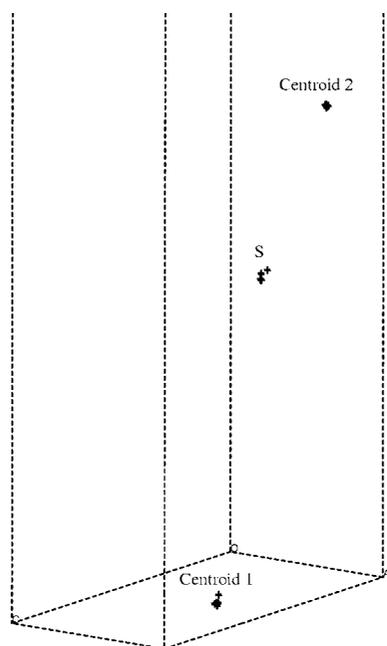

Figure 5
The best SA solution for **33** (black; disordered model) overlaid on the single-crystal structure (grey), showing the excellent agreement between corresponding half-occupancy phenyl ring positions. Inset: with the phenyl ring in the SA model set to full occupancy, the best solution (black) returned $\chi^2_{profile}/\chi^2_{Pawley}$ = 4.4, with the ring positioned approximately midway between the disordered phenyl positions in the single-crystal structure (grey).

batches of unconstrained runs indicated that the frequency of success for **35** was somewhat less than the 5% observed for the initial batch of 20 runs reported in Table 5. Following this realization, the structure was solved reproducibly (in replicate batches) by combining a bigger batch size (minimum of 50 runs) and an increased number of SA moves per run (2 × 10^7) with a reduced SA cooling rate (0.01). Thereafter, the application of modal torsion constraints doubled the frequency with which the structure of **35** was solved, from N_{sol} = 1 to 2.
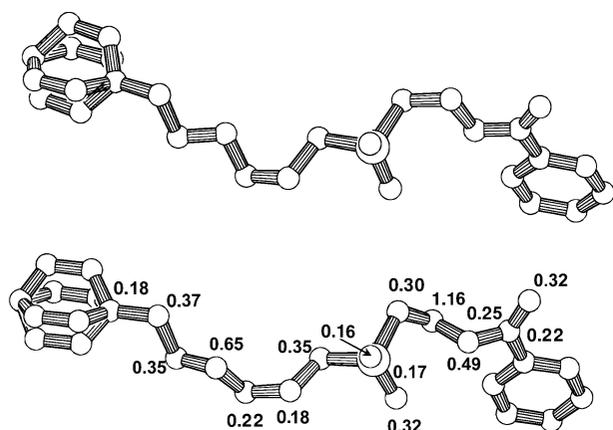
5. Conclusions

At a time when the development of experimental methods for increasing the efficiency and throughput of drug development

Figure 6
Comparison of the positions of the phenyl ring centroids and S atoms in the top six SA solutions for **33** and in the equivalent single-crystal structure (centroid 1 corresponds to the disordered phenyl ring).

Table 8
CSD refcodes and references.

CSD refcode	Reference
HCSBTZ	Dupont & Dideberg (1972).
ZZZLWW01	Steiner (2000)
LIZCUS	Raj *et al.* (2000)
CBMZPN10	Himes *et al.* (1981)
DAPSUO10	Alleaume (1967)
HXACAN07	Nichols & Frampton (1998)
HXACAN08	Nichols & Frampton (1998)
ZZZMLY01	Hodgson & Asplund (1991)
MCPRPL	Fujinaga & James (1980)
CLTPNI03	Brammer & Stevens (1989)
BHHPHE	Beale & Grainger (1972)
GEBZUI	Bamgboye & Sowerby (1986)
QQQAEJ01	Shankland *et al.* (1996)
THTHXN01	David *et al.* (1998)
BEDMIG	Koo *et al.* (1980)
CREATH03	Kato *et al.* (1979)
CIMPRA	Post & Horn (1977)
LACTOS10	Fries *et al.* (1971)
PROMZC01	David *et al.* (1998)
ZZZPUS02	Donaldson *et al.* (1981)
CEYHUJ01	Kojic-Prodic *et al.* (1984)
UCUVET	Shankland *et al.* (2001)
YAYWUQ	Kennedy *et al.* (2001)
JEKNOC10	Freer *et al.* (1993)
CURHOM	Carpy *et al.* (1985)

is a priority within the pharmaceutical industry, valuable savings in time, materials and analysis can be achieved by wider reliance on high-quality laboratory XRPD data and SDPD.

Given XRPD data collected to 2 Å resolution or better, the results of this investigation substantiate the following conclusions: (*a*) structures with <15 DOF present little challenge to the SA process, reproducibly yielding accurate structure solutions; (*b*) for structures with greater complexity

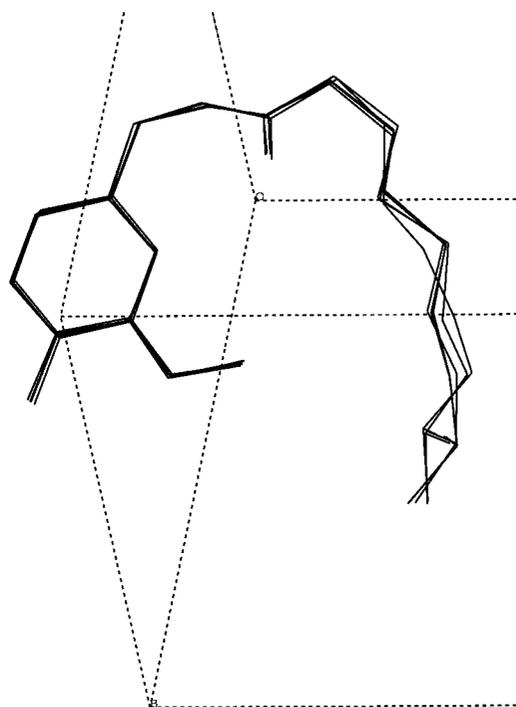(DOF = 15–20), where the preponderance of local minima in the agreement hypersurface reduces the frequency of success, the SA algorithm is still able to locate the global minimum with a reasonable frequency.

Figure 7
Comparison of the best SA solution for **33** (top; $\chi^2_{profile}/\chi^2_{Pawley} = 4.1$) with the fourth-ranked structure in the batch of 20 SA runs (bottom: $\chi^2_{profile}/\chi^2_{Pawley} = 10.9$). The corresponding phenyl ring centroids in each structure are separated by 0.10 Å (disordered ring) and 0.20 Å, with the S atoms 0.17 Å apart. The numerals in the bottom structure indicate the separation (Å) from the corresponding atom in the top structure.

Figure 8
Overlay of the top five SA solutions for **31**, spanning the range $\chi^2_{profile}/\chi^2_{Pawley} = 4.2$–6.3.

It is at higher levels of complexity (DOF > 20) where the greatest challenges remain, with structures such as **35** (DOF = 22) and AR-C69457CC (DOF = 26; Johnston *et al.*, 2004) representing the current state of the art of SDPD from laboratory XRPD data. Modal torsion-angle constraints can significantly increase the frequency of success and offer a convenient means by which to introduce prior chemical knowledge to reduce the size of the search space. Automation of the process of determining the constraints from a knowledge base of molecular geometry (*MOGUL*; Bruno *et al.*, 2004) in the latest version of *DASH* (v3.0) should enable this knowledge to be used more routinely.

One might also consider the application of modified or alternative search algorithms such as parallel tempering (Hansmann, 1997) or hybrid Monte Carlo (Johnston *et al.*, 2002), and/or different evaluation functions (*e.g.* maximum likelihood; Markvardsen *et al.*, 2002). Not all of these approaches have been implemented beyond the proof-of-concept stage in SDPD, and the challenge therefore remains to implement such approaches for routine structure solution.

Powder diffractometers at synchrotron sources continue to offer considerable advantages over laboratory-based instrumentation (*e.g.* increased incident flux and higher instrumental resolution) and these advantages often translate to an increased diffraction information content that allows more complex structures to be determined and refined.

Rietveld refinement of a solved structure can always be recommended, with the caveat that an improved fit to the diffraction data should not be pursued at the expense of chemical sense. This typically means the careful application of restraints, or perhaps rigid-body refinements. The latter option is similar to the simplex refinement at the end of a *DASH* run and thus it is not uncommon to find little improvement on refining an SA solution where the structure concerned is either relatively simple or substantially rigid.

References

Alleaume, M. (1967). Thesis, Bordeaux, France.

Allen, F. H. (2002). *Acta Cryst.* B**58**, 380–388.

Bamgboye, T. T. & Sowerby, D. B. (1986). *Polyhedron*, **5**, 1487–1488.

Beale, J. P. & Grainger, C. T. (1972). *Cryst. Struct. Commun.* **1**, 71–74.

Boultif, A. & Louër, D. (1991). *J. Appl. Cryst.* **24**, 987–993.

Brammer, L. & Stevens, E. D. (1989). *Acta Cryst.* C**45**, 400–403.

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.

Bruno, I. J., Cole, J. C., Kessler, M., Luo, J., Motherwell, W. D. S., Purkis, L. H., Smith, B. R., Taylor, R., Cooper, R. I., Harris, S. E. & Orpen, A. G. (2004). *J. Chem. Inf. Comput. Sci.* In preparation.

Carpy, A., Leger, J.-M. & Melchiorre, C. (1985). *Acta Cryst.* C**41**, 624–627.

Coelho, A. A. (2003). *J. Appl. Cryst.* **36**, 86–95.

David, W. I. F., Shankland, K., Cole, J., Maginn, S., Motherwell, W. D. S. & Taylor, R. (2001). *DASH* user manual. Cambridge Crystallographic Data Centre, Cambridge, UK.

David, W. I. F., Shankland, K. & Shankland, N. (1998). *Chem. Commun.* pp. 931–932.

Deem, M. W. & Newsam, J. M. (1989). *Nature (London)*, **342**, 260–262.

Dinnebier, R. E., Wagner, M., Peters, F., Shankland, K. & David, W. I. F. (2000). *Z. Anorg. Allg. Chem.* **626**, 1400–1405.

Dollase W. A. (1986). *J. Appl. Cryst.* **19**, 267–272.

Donaldson, J. D., Leary, J. R., Ross, S. D., Thomas, M. J. K. & Smith, C. H. (1981). *Acta Cryst.* B**37**, 2245–2248.

Dupont, L. & Dideberg, O. (1972). *Acta Cryst.* B**28**, 2340–2347.

Edgar, M., Carter, V. J., Tunstall, D. P., Grewal, P., Favre-Nicolin, V., Cox, P. A., Lightfoot, P. & Wright, P. A. (2002). *Chem. Commun.* **8**, 808–809.

Engel, G. E., Wilke, S., König, O., Harris, K. D. M. & Leusen, F. J. J. (1999). *J. Appl. Cryst.* **32**, 1169–1179.

Favre-Nicolin, V. & Cerny, R. (2002). *J. Appl. Cryst.* **35**, 734–743.

Frampton, C. S. (2004). Unpublished single-crystal data.

Freer, A. A., Bunyan, J. M., Shankland, N. & Sheen, D. B. (1993). *Acta Cryst.* C**49**, 1378–1380.

Fries, D. C., Rao, S. T. & Sundaralingam, M. (1971). *Acta Cryst.* B**27**, 994–1005.

Fujinaga, M. & James, M. N. G. (1980). *Acta Cryst.* B**36**, 3196–3199.

Golič, L., Djinović, K. & Florjanič, M. (1989). *Acta Cryst.* C**45**, 1381–1384.

Graham, D., Kennedy, A. R., McHugh, C. J., Smith, W. E., David, W. I. F., Shankland, K. & Shankland, N. (2004). *New J. Chem.* **28**, 161–165.

Hansmann, U. H. E. (1997). *Chem. Phys. Lett.* **281**, 140–150.

Harris, K. D. M. & Cheung, E. Y. (2003). *Org. Process Res. Dev.* **7**, 970–976.

Himes, V. L., Mighell, A. D. & DeCamp, W. H. (1981). *Acta Cryst.* B**37**, 2242–2245.

Hodgson, D. J. & Asplund, R. O. (1991). *Acta Cryst.* C**47**, 1986–1987.

Ivashevskaja, S. N., Aleshina, L. A., Andreev, V. P., Nizhnik, Y. P., Chernyshev V. V. & Schenk, H. (2002). *Acta Cryst.* C**58**, m300–m301.

Johnston, J. C., Markvardsen, A. J., David, W. I. F. & Shankland, K. (2002). *Acta Cryst.* A**58**, 441–447.

Johnston, A., Florence, A. J., Shankland, K., Markvardsen, A., Shankland, N., Steele, G. & Cosgrove, S. D. (2004). *Acta Cryst.* E**60**, o1751–o1753.

Kato, Y., Haimoto, Y. & Sakurai, K. (1979). *Bull. Chem. Soc. Jpn*, **52**, 233–234.

Kennedy, A. R., Hughes, M. P., Monaghan, M. L., Staunton, E., Teat, S. J. & Smith, W. E. (2001). *J. Chem. Soc. Dalton Trans.* pp. 2199–2205.

Kojic Prodic, B., Ruzic Toros, Z., Sunjic, V., Decorte, E. & Moimas, F. (1984). *Helv. Chim. Acta*, **67**, 916–925.

Koo, C. H., Cho, S. I. & Yeon, Y. H. (1980). *Arch. Pharm. Res.* **3**, 37–49.

Le Bail, A. (2001). *Mater. Sci. Forum*, **378–381**, 65–70.

Madsen, I. C. & Hill, R. J. (1994). *J. Appl. Cryst.* **27**, 385–392.

Marder, T. B. (2004). Unpublished single-crystal data.

Markvardsen, A. J., David, W. I. F. & Shankland, K. (2002). *Acta Cryst.* A**58**, 316–326.

Middleton, D. A., Peng, X., Saunders, D., Shankland, K., David, W. I. F. & Markvardsen, A. J. (2002). *Chem. Commun.* pp. 1976–1977.

Nichols, C. & Frampton, C. S. (1998). *J. Pharm. Sci.* **87**, 684–693.

Pawley, G. S. (1981). *J. Appl. Cryst.* **14**, 357–361.

Post, M. L. & Horn, A. S. (1977). *Acta Cryst.* B**33**, 2590–2595.

Raj, S. S. S., Fun, H. -K., Zhang, J., Xiong, R. -G. & You, X.-Z. (2000). *Acta Cryst.* C**56**, e274–e275.

Reinaudi, L., Leiva, E. P. M. & Carbonio, R. E. (2000). *J. Chem. Soc. Dalton Trans.* **23**, 4258–4262.

Rukiah, M., Lefebvre, J., Descamps, M., Hemon, S. & Dzyabchenko, A. (2004). *J. Appl. Cryst.* **37**, 464–471.

Shankland, K. & David, W. I. F. (2002). *Structure Determination from Powder Diffraction Data*, edited by W. I. F. David, K. Shankland, L. B. McCusker & Ch. Baerlocher, pp. 252–285. IUCr/Oxford University Press.

Shankland, K., David, W. I. F., Csoka, T. & McBride, L. (1998). *Int. J. Pharm.* **165**, 117–126.

Shankland, N., David, W. I. F., Shankland, K., Kennedy, A. R., Frampton, C. S. & Florence, A. J. (2001). *Chem. Commun.* pp. 2204–2205.

Shankland, K., David, W. I. F. & Sivia, D. (1997). *J. Mater. Chem.* **7**, 569–572.

Shankland, N., Love, S. W., Watson, D. G., Knight, K. S., Shankland, K. & David, W. I. F. (1996). *J. Chem. Soc. Faraday Trans.* **92**, 4555–4559.

Shankland, K., McBride, L., David, W. I. F., Shankland, N. & Steele, G. (2002). *J. Appl. Cryst.* **35**, 443–454.

Steiner, T. (2000). *Acta Cryst.* C**56**, 876–877.

Zachariasen, W. H. & Ellinger, F. H. (1963). *Acta Cryst.* **16**, 369–375.

Zaske, L., Perrin, M.-A., Daiguebonne, C. & Guillou, O. (2004). *Mater. Sci. Forum*, **443–444**, 411–421.