Acta Crystallographica Section B
Structural Science
ISSN 0108-7681

Structure prediction as a tool for solution of the crystal structures of metallo-organic complexes using powder X-ray diffraction data

Andrew D. Bond* and William Jones

University of Cambridge, Department of Chemistry, Lensfield Road, Cambridge CB2 1EW, England

Correspondence e-mail: adb29@cam.ac.uk

Received 21 September 2001
Accepted 15 November 2001

A simulated-annealing direct-space approach has been applied to predict the crystal structures of a series of metallo-organic complexes containing Zn, Cu and Ni. The prediction methodology generates a set of energetically reasonable crystal structures among which the actual structure is present, but it is not always possible to specify unambiguously the known crystal structure solely on the basis of energy. In each case, however, the ambiguity may be resolved by recourse to laboratory powder X-ray diffraction (PXRD) data. In this manner, structure prediction is shown to be a powerful tool for structure solution using PXRD data, with the additional advantage that indexing of the PXRD profile is not required at the outset.

1. Introduction

Recent years have heralded an explosion in the study of the molecular solid state, and the field of crystal engineering in particular has assumed major prominence (Desiraju, 1989; Jones, 1997). This explosion is undoubtedly related to major advances in techniques for crystal structure determination and, in particular, developments in the technique of choice, single-crystal X-ray diffraction. The emergence of CCD area detectors coupled with huge advances in computing power have served to make single-crystal X-ray diffraction a rapid and generally straightforward analytical technique. In addition, low-temperature data collection has become standard, improving the precision of general structure determinations and facilitating the study of crystals that may be unstable (or indeed liquid) at room temperature (see for example, Davies & Bond, 2001). Such are the improvements in instrumentation that the limiting factor is now commonly the production of suitable single crystals of a material – a chemical problem rather than a technical one. In cases where very small, but still single, crystals are the best that can be obtained, exploitation of sources of additional brightness (*i.e.* synchrotron sources) can facilitate structure solution (Clegg, 2000). The fact remains, however, that, in many instances, systems of chemical interest do not readily form single crystals suitable for X-ray diffraction even when analysed with a synchrotron source. In these cases, diffraction data collected from crystalline powders are the primary source of information from which crystal structures may be deduced, and solution of the crystal structures of molecular materials from powder X-ray diffraction (PXRD) data is therefore an area of considerable current interest (for a recent review, see Harris *et al.*, 2001).

The inherent compression of the three-dimensional distribution of diffracted intensities generated from a single crystal into one dimension in the PXRD profile leads to a consider-

able loss of information as a result of peak overlap. This information loss can be particularly severe for PXRD data collected using laboratory instruments, and an inability to extract a sufficient number of unambiguously assigned peak intensities generally prohibits application of *ab initio* reciprocal-space methods for structure solution (*i.e.* Patterson and direct-phasing methods). Alternative direct-space methods have emerged in recent years in which structure models are postulated in real space and optimized against the whole PXRD profile (Harris *et al.*, 1994; Tremayne *et al.*, 1997). Structure solution then becomes a matter of global optimization of a multi-dimensional hypersurface comprising the numerous variables that describe the crystal structure against a function describing the overall fit to the PXRD profile. This function can utilize either a fit to extracted intensities [commonly derived from Pawley or Le Bail algorithms (Pawley, 1981; Le Bail *et al.*, 1988)] or a fit to the whole profile (usually quantified by R_{wp}) (Harris *et al.*, 1994). Several approaches have been applied to this optimization problem, including grid-search procedures (Hammond *et al.*, 1997), simulated-annealing (David *et al.*, 1998) and genetic algorithm protocols (Shankland *et al.*, 1997; Harris *et al.*, 1998), with significant successes for organic molecules of considerable complexity. Extension of these methods to metallo-organic complexes has also been considered, and a recent report of the crystal structure of an Fe(III) porphyrin solved by simulated-annealing direct-space methods from synchrotron PXRD data is notable (Pagola *et al.*, 2000).

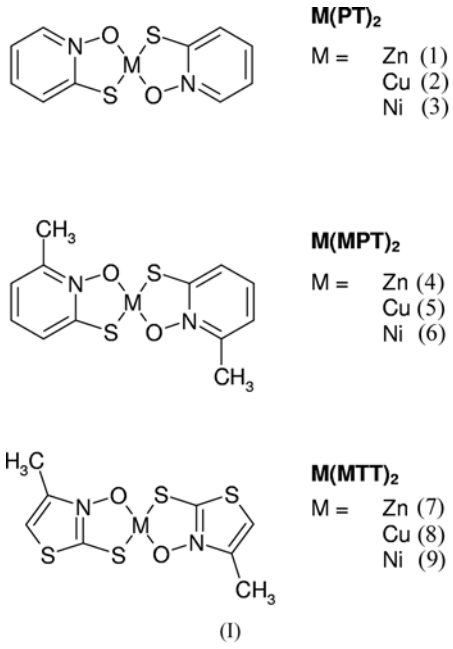
Despite these considerable successes, however, direct-space methods cannot necessarily be considered a panacea for crystal structure solution from PXRD data since one major assumption is implicit in their use: the crystal system and unit-cell parameters must be deduced at the outset by indexing the PXRD profile. Failure to index successfully the PXRD profile renders direct-space methods impracticable. Indexing is not always straightforward, particularly for PXRD data collected from laboratory instruments, and many structure solution attempts do not progress beyond this preliminary stage.

A concurrent avenue of research, and arguably the ultimate goal in molecular solid-state chemistry, is *ab initio* prediction of crystal structure from molecular structure. A universal structure prediction procedure would essentially supersede all problems of structure solution and facilitate the prediction of the bulk properties of molecular materials prior to their synthesis. A recent evaluation of the state of this art, however, suggests that prediction methodology is somewhat less advanced than that for structure solution (Lommerse *et al.*, 2000). Arguably the most successful current approach may be compared with the direct-space methods employed for structure solution from PXRD data: trial structures are postulated in real space and global optimization is performed against an empirical function that describes the crystal's energy (Karfunkel & Gdanitz, 1992). A considerable advantage of structure prediction techniques compared with those for structure solution is that in the case of prediction it is not necessary to specify *a priori* the unit-cell parameters. As a result, global optimization is considerably more complex, but genuine *ab initio* structure prediction (specifying only the molecular connectivity at the outset) may be facilitated in favourable cases where the molecular geometry is relatively rigid and known in advance with some degree of confidence (see, for example, Karfunkel & Gdanitz, 1992; Schmidt & Englert, 1996; Beyer *et al.*, 2001). It is common, however, for numerous structures to be predicted within a small range of calculated energies such that, in general, absolute specification of the crystal structure is not possible, *i.e.* the energy function does not contain sufficient information to specify the crystal structure unambiguously (Gavezzotti & Filippini, 1996). This factor, together with an inability to account for kinetic aspects of crystallization, means that current approaches for structure prediction may only be considered to provide sets of energetically reasonable crystal structures among which the actual structure is likely to be present.

The similarity between direct-space methodologies for structure solution and structure prediction has prompted suggestions that the PXRD profile and the energy function may be combined to yield a hybrid function against which to optimize. The premise is straightforward: if neither the PXRD profile nor the energy function contain sufficient information to specify the structure, combination of the two sources of information may facilitate structure solution. In general, the combined optimization of two fitness functions weakens the minima that belong to only one of the two functions and strengthens minima belonging to both functions. For crystal structure determination, therefore, successful optimization of a combined PXRD profile/energy function has a considerably greater chance of locating the correct crystal structure. The process is one of structure *solution* rather than *prediction* since recourse must be made to experimental PXRD data. Combined direct-space methods of this type have been proposed recently by Putz *et al.* (1999) and by Lanning *et al.* (2000). Despite the unquestionable elegance of these unified approaches, the fundamental difficulty remains: it is necessary to specify at the outset the crystal system and unit-cell parameters, *i.e.* it is assumed that the PXRD profile may be indexed successfully. It is clear that, in cases where this is not possible, a structure prediction approach is most desirable.

In this work, we describe an approach that involves application of a structure prediction technique followed, in a subsequent step, by consideration of PXRD data in order to resolve any ambiguity arising from the prediction of several structures representing local minima of the energy function. In this manner, structure prediction becomes a tool for structure solution using PXRD data, which is of particular use where indexing is problematic. The approach has been applied to determine the crystal structures of a series of divalent complexes of the cyclic thiohydroxamic acids, pyrithione (PT), methylpyrithione (MPT) and methlythiazolethione (MTT) with zinc, copper and nickel [see (I)]. The methodology is applicable to general metallo-organic complexes incorporating relatively rigid bidentate ligands, and, indeed, all molecular materials for which the molecular conformation in the crystal can be specified with some degree of confidence at the outset.

2. Crystal structures of (1)–(9)



(I)

The crystal structures of the PT complexes (1), (2) and (3) have been reported previously (Barnett *et al.*, 1977; Chen *et al.*, 1991; Bond *et al.*, 2001). (1), bis[1-hydroxypyridine-2(1*H*)-thionato]zinc(II), adopts a dimeric structure in the solid state in which zinc adopts a coordination geometry part way between trigonal–bipyramidal and square–pyramidal (Fig. 1*a*). Both (2) and (3), bis[1-hydroxypyridine-2(1*H*)-thionato]-copper(II) and bis[1-hydroxypyridine-2(1*H*)-thionato]nickel(II), respectively, adopt square–planar geometries, with a *trans* and *cis* arrangement of the PT ligands, respectively. For (2) and (3), molecules are linked by an $R_2^2(8)$ C—H···O hydrogen-bond motif, forming extended chains in the case of the *trans* molecule (2) and discrete dimers for the *cis* molecule (3) (Figs. 1*b* and 1*c*). In the case of (2), the molecules stack along a short axis (∼4 Å) with additional close intermolecular Cu···S contacts such that the structure may be considered to contain stacks of edge-sharing Jahn–Teller-distorted CuO_2S_4 octahedra (Fig. 1*d*). The crystal structures of the MPT complexes (4), (5) and (6) have been reported previously by West *et al.* (1998). (4), bis[1-hydroxy-6-methylpyridine-2(1*H*)-thionato]zinc(II), adopts a monomeric structure with Zn in a tetrahedral coordination geometry (Fig. 2*a*); (4) is prevented from dimerization in a manner similar to (1) by the steric influence of the methyl substituent. (5) and (6), bis[1-hydroxy-6-methylpyridine-2(1*H*)-thionato]copper(II) and bis[1-hydroxy-6-methylpyridine-2(1*H*)-thionato]nickel(II), respectively, are isostructural and contain *cis* square–planar molecules linked into dimers by intermolecular *M*···O contacts (Fig. 2*b*). We have reported the crystal structures of the MTT complexes previously (Bond & Jones, 2001). In the solid state, (7) and (8) adopt tetrahedral and *trans* square–planar molecular geometries, respectively. (7), bis[3-hydroxy-4-methyl-

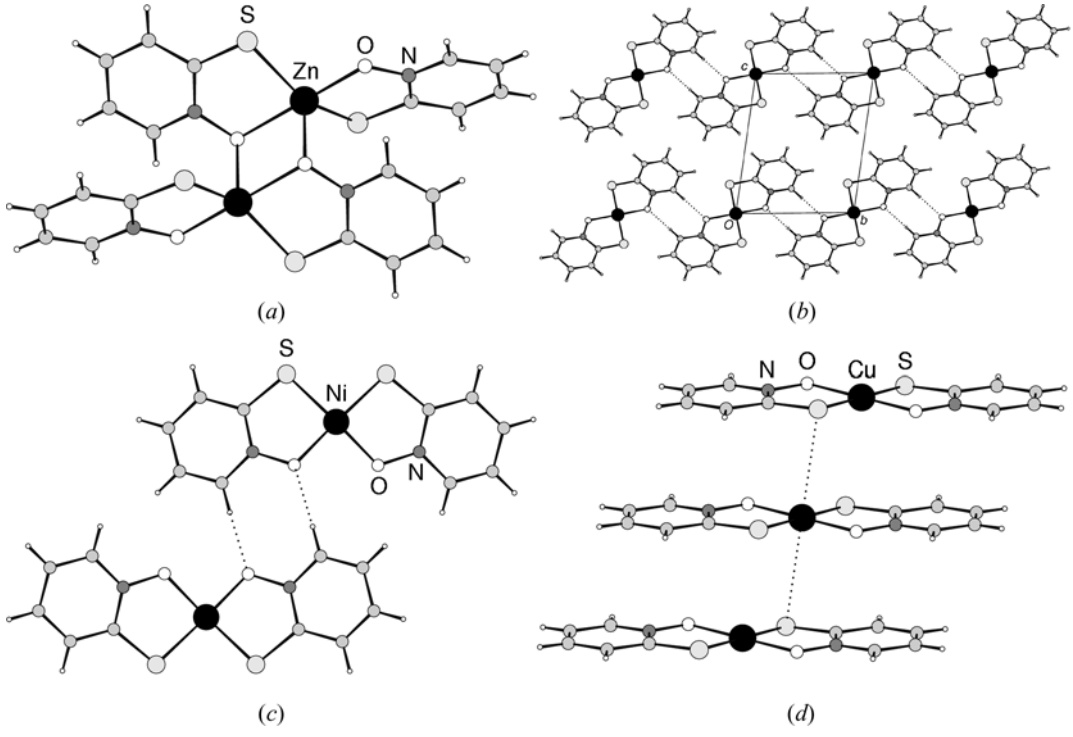
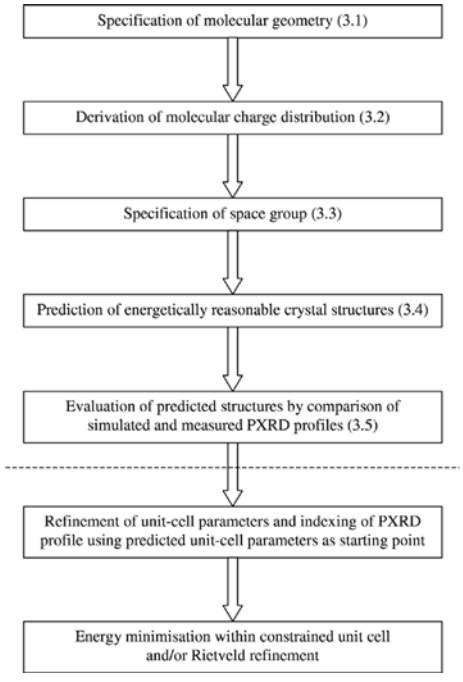


Figure 1
(*a*) Dimer unit in the crystal structure of (1). (*b*) Projection onto (100) of the crystal structure of (2) showing extended chains linked by $R_2^2(8)$ motifs comprising C—H···O interactions. (*c*) Dimers in the crystal structure of (3) linked by $R_2^2(8)$ motifs comprising C—H···O interactions. (*d*) Stacks containing short intermolecular Cu···S contacts in the crystal structure of (2).

1,3-thiazole-2(3*H*)-thionato]zinc(II), is isostructural with the analogous MPT complex (4) (Fig. 3*a*). In (8), bis[3-hydroxy-4-methyl-1,3-thiazole-2(3*H*)-thionato]copper(II), stacks of molecules exist similar to those observed in (2) but, in this case, additional intermolecular Cu···O contacts give rise to stacks of edge-sharing Jahn–Teller-distorted CuO_4S_2 octahedra (Figs. 3*b* and 3*c*). (9), bis[3-hydroxy-4-methyl-1,3-thiazole-2(3*H*-thionato]nickel(II), forms a trimeric structure in which Ni atoms adopt both six- and five-coordination (Fig. 3*d*).

3. Methodology

A schematic description of the methodology employed is given below.



(The dotted line represents the transitionfrom structure solution to structure refinement.) Structure predictions were performed using MSI's (now Accelrys) *Polymorph Predictor* (Leusen, 1996; Leusen *et al.*, 1999), based on the methodology of Karfunkel & Gdanitz (1992); this package has been shown to be one of the most successful currently available (Lommerse *et al.*, 2000). Prediction of the crystal structures of organometallic molecules using an empirical energy function has been shown previously to be feasible by Schmidt & Englert (1996). In Schmidt's methodology, steepest-descent minimization is employed for a number of randomly generated crystal structures. This corresponds to picking random positions on the energy hypersurface and optimizing to reach the nearest local minimum. Location of the global minimum relies, therefore, on generation of a random starting model sufficiently close to the true structure that the minimization starts within the valley that will lead to the lowest point on the hypersurface. The simulated-annealing (SA) optimization protocol employed here is somewhat more powerful in that it facilitates 'uphill' movement as well as simple descent, such that the system is able to escape from local minima in the search for the global minimum (Kirkpatrick *et al.*, 1983). Extensive description of the methodology employed in *Polymorph Predictor* has been published previously (Karfunkel & Gdanitz, 1992), but specific procedures employed for application to the metallo-organic complexes of PT, MPT and MTT are outlined below.

3.1. Specification of the molecular geometry

For the purposes of validating the methodology, predictions were performed for complexes whose crystal structures were known previously. The molecular geometries were taken, therefore, directly from the known crystal structures and treated as rigid bodies in the prediction process. It is common practice to separate the molecular conformation problem from the crystal packing problem in this manner; specification of the molecular unit as a rigid body removes all degrees of freedom associated with the intramolecular geometry, reducing considerably the complexity of the global optimization. For an '*ab initio*' prediction, where only the molecular connectivity is known at the outset, it is necessary to derive the molecular conformation using molecular mechanics or, more commonly, semi-empirical molecular-orbital methods (see, for example, Aakeröy *et al.*, 1998; Beyer *et al.*, 2001). It has also been noted, however, that molecular conformations in crystals may differ from those calculated in the gas phase owing to the influence of intermolecular interactions in the solid state (Starbuck *et al.*, 1999). It is generally advisable, therefore, to consider other comparable molecules in the Cambridge Structural Database (CSD) for derivation of the molecular

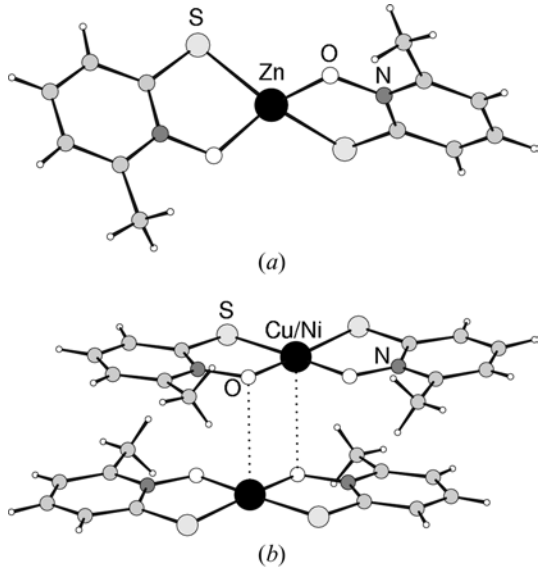


Figure 2
(*a*) Tetrahedral molecular unit in (4). (*b*) Molecules of (5) linked into dimers by *M*···O interactions. The crystal structure of (6) is comparable.

geometry. For the complexes presented here, the molecular units employed in the structure predictions are in most cases of a 'standard' MX_2 type (*i.e.* tetrahedral or *cis*/*trans* square–planar), where *X* represents a relatively rigid bidentate ligand. The methodology is, therefore, readily applicable for general MX_2 complexes. For the 'non-standard' cases (1) and (9), prediction of the molecular structures would be very much reliant upon observation of similar structures in the CSD.

3.2. Choice of force field and description of the molecular charge distribution

For any prediction methodology seeking general applicability, it is desirable to employ a generic force field – transferability *is* an essential feature. The Dreiding generic force field has been shown to be adequate for general organic molecules and was employed here with default parameters (Mayo *et al.*, 1990). It has been suggested that atomic charges derived from semi-empirical molecular-orbital calculations are suitable in conjunction with this force field (Karfunkel & Gdanitz, 1992), and this has also been our experience with treating organic molecules (Bond *et al.*, 2000). Atomic charges were derived therefore from the electrostatic potential using the *MOPAC* package (Stewart, 1990), employing the AM1 model (Dewar *et al.*, 1985). It should be noted that neither the Dreiding force field nor the AM1 model incorporate default parameters for 3*d* elements other than zinc (Stewart & Rzepa, 1999). Since it was not our intention to undertake extensive force-field parameterization, the complexes of nickel and copper were treated simply by describing the metal atom as zinc. This significant approximation is justified subsequently by the resulting success of the structure predictions and may be rationalized by considering that the role of the metal atom in determining the crystal structures is negligible. The Ewald summation technique was employed to accelerate convergence of the electrostatic terms.

3.3. Specification of the space group

In principle, specification of the space group is not a limiting feature of the prediction methodology since it is possible to search in all 230 groups. Such a search would, of course, be computationally extravagant and, in practice, well known space-group statistics limit the search initially to several of the most probable space groups (Padmaja, *et al.*, 1990; Baur & Kassner, 1992). It is also possible in theory to perform

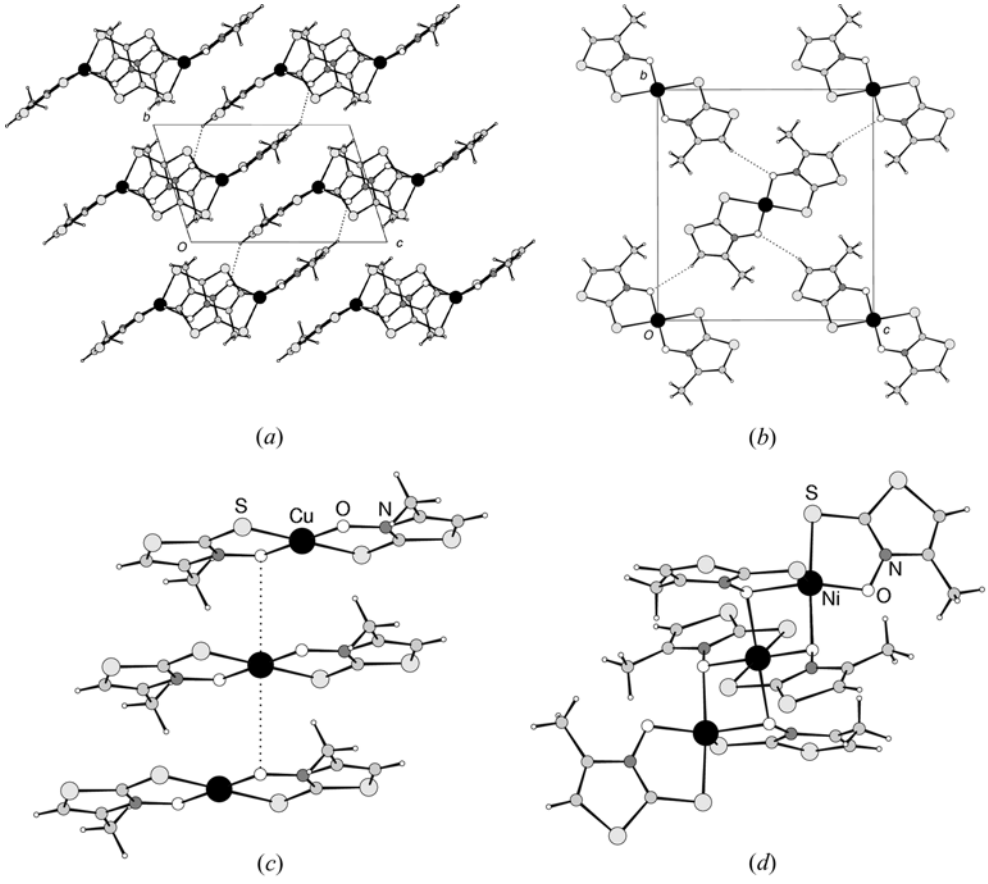


Figure 3
(*a*) Projection onto (100) of the crystal structure of (7) with C—H···O interactions indicated by dotted lines. The crystal structure of (4) is comparable. (*b*) Projection onto (100) of the crystal structure of (8) with C—H···O interactions indicated by dotted lines. (*c*) Intermolecular Cu···O contacts in the crystal structure of (8). (*d*) Trimer unit in the crystal structure of (9).

Table 1
Details of energy minimization for the observed single-crystal structures of (1)–(9).

(i) Observed model, no minimization. (ii) Minimized model, unit-cell parameters relaxed within the constraints of the crystal system. Energies are quoted as lattice binding energies (*i.e.* exothermic quantities).

				a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)	E (kJ mol^{-1})
(1)	$Zn(PT)_2$	$P2_1/c$	(i)	8.355	10.148	13.673	90	96.82	90	−108.0
			(ii)	8.310	10.224	14.024	90	96.81	90	−111.4
(2)	$Cu(PT)_2$	$P\bar{1}$	(i)	4.055	7.572	9.063	81.10	86.19	85.20	−125.0
			(ii)	4.143	7.714	9.144	85.77	83.90	90.45	−134.4
(3)	$Ni(PT)_2$	$P2_1/n$	(i)	8.769	13.627	9.857	90	97.76	90	−130.9
			(ii)	8.778	13.747	9.915	90	96.80	90	−133.3
(4)	$Zn(MPT)_2$	$P\bar{1}$	(i)	7.974	7.921	11.779	104.50	93.45	104.50	−136.8
			(ii)	7.963	7.925	11.743	103.80	92.64	104.62	−138.3
(5)	$Cu(MPT)_2$	$Pbca$	(i)	15.663	13.914	12.984	90	90	90	−134.2
			(ii)	15.331	14.414	12.938	90	90	90	−136.8
(6)	$Ni(MPT)_2$	$Pbca$	(i)	15.602	13.606	13.348	90	90	90	−137.7
			(ii)	15.398	13.656	13.350	90	90	90	−140.5
(7)	$Zn(MTT)_2$	$P\bar{1}$	(i)	7.622	7.695	11.779	105.15	96.21	106.59	−126.6
			(ii)	7.739	7.736	12.005	106.26	93.39	104.55	−140.6
(8)	$Cu(MTT)_2$	$P2_1/n$	(i)	4.086	12.757	11.988	90	97.01	90	−161.1
			(ii)	4.147	13.107	11.837	90	95.95	90	−164.3
(9)	$Ni(MTT)_2$	$P\bar{1}$	(i)	10.067	10.167	11.184	92.71	114.00	112.92	−85.2
			(ii)	10.140	10.084	11.208	93.44	113.09	111.65	−89.3

predictions in $P1$ with numerous symmetry-independent molecules; any higher space-group symmetry should be reproduced in the resulting model. This approach is generally impractical, however, since it increases considerably the number of independent variables such that global optimization rapidly becomes intractable. For the complexes treated here, space groups were known initially from single-crystal studies and predictions were performed only in these known space groups. It may be noted that each complex adopts one of the top seven statistically observed space groups. In the cases where the molecular unit is centrosymmetric, searches were also performed in the corresponding subgroup with the centre of symmetry removed (*i.e.* $P1$ for $P\bar{1}$, $P2_1$ for $P2_1/c$), but using the *same* molecular unit. The higher space symmetry should then be reproduced in the predicted structures.

3.4. Structure prediction

Predictions were performed using the default SA parameters of *Polymorph Predictor*. In each case, the predicted structures from the initial SA run were subjected to cluster analysis (in which structures deemed to be identical are grouped together and only one of the set, that with the lowest energy, is retained), followed by local minimization. The cluster analysis was repeated for the minimized set and the resulting structures constituted the final set of predicted structures.

3.5. Evaluation of the predicted structures

In general, a large number of structures (in some cases >100) may be predicted within a relatively small range of calculated energies. To facilitate identification of the correct crystal structure, a PXRD profile was simulated from each predicted model, and this was compared with the PXRD profile simulated from the appropriate single-crystal structure. It is sufficient simply to perform this comparison 'by eye', since the PXRD profiles generated from the predicted models generally differ significantly, and visual evaluation of the best fit is straightforward. For unknown structures, of course, the comparison must be performed against PXRD data measured for the material of interest. When assessing the goodness of the fit, it is essential to appreciate that small differences in unit-cell parameters mean that peak positions do not line up exactly; these differences may be attributed to minor inadequacies in the force field coupled with the extreme complexity of the hypersurface. In a subsequent step, the parameters and peak indexing deduced from the predicted structure may be used as a starting point for unit-cell refinement against measured PXRD data. Energy minimization of the predicted structure within a unit cell constrained to these refined dimensions will produce a structure with a considerably better overall fit to the PXRD profile. Ultimately, the structure may be subjected to full Rietveld refinement (Rietveld, 1967, 1969). This final step provides confirmation that the minimum in the energy function is also a minimum in the fit to the PXRD profile.

4. Results

4.1. Validation of the force field and description of the molecular charge distribution

The validity of the force field and the description of the molecular charge distribution may be tested by steepest-descent minimization of the observed crystal structures, confirming that these represent (at least local) minima of the energy function. Results for the minimization of (1)–(9) are listed in Table 1. In general, the lattice parameters of the minimized models are in good agreement with those of the observed structures, indicating that the force field provides a satisfactory empirical description of the intermolecular forces. It is noteworthy that the worst agreement is observed for

Table 2
Comparison of the single-crystal structures and predicted structures for (1)–(9) (overlays of the structures are available as supplementary material).

(i) Single-crystal structure. (ii) Predicted model.

				a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)	N†	Energy range (kJ mol^{-1}) Min	Max	Energy of correct predicted model (kJ mol^{-1})	Energy rank of correct model‡
(1)	$Zn(PT)_2$	(i)	$P2_1/c$	8.355 (1)	10.148 (1)	13.673 (1)	90	96.82 (1)	90					
		(ii)	$P2_1$	8.351	10.369	14.032	90	98.48	90	144	−110.0	−63.4	−110.0	1
(2)	$Cu(PT)_2$	(i)	$P\bar{1}$	4.055 (1)	7.572 (1)	9.063 (1)	81.10 (1)	86.19 (1)	85.20 (1)					
		(ii)	$P1$	4.111	7.759	9.214	80.95	81.91	84.40	100	−131.4	−127.8	−131.4	1
(3)	$Ni(PT)_2$	(i)	$P2_1/n$	8.769 (4)	13.627 (4)	9.857 (3)	90	97.76 (3)	90					
		(ii)	$P2_1/n$	9.052	13.284	10.013	90	99.02	90	194	−122.5	−108.6	−117.9	9
(4)	$Zn(MPT)_2$	(i)	$P\bar{1}$	7.974 (1)	7.921 (1)	11.779 (1)	104.50 (1)	93.45 (1)	104.50 (1)					
		(ii)	$P\bar{1}$	7.963	7.925	11.743	103.80	92.62	104.58	83	−138.3	−77.0	−138.3	1
(5)	$Cu(MPT)_2$	(i)	$Pbca$	15.663 (6)	13.914 (5)	12.984 (5)	90	90	90					
		(ii)	$Pbca$	15.196	14.642	12.915	90	90	90	62	−130.3	−83.5	−130.3	1
(6)	$Ni(MPT)_2$§	(i)	$Pbca$	15.602 (8)	13.606 (4)	13.348 (8)	90	90	90					
		(ii)	$Pbca$	15.196	14.642	12.915	90	90	90	62	−130.3	−83.5	−130.3	1
(7)	$Zn(MTT)_2$	(i)	$P\bar{1}$	7.622 (1)	7.695 (1)	11.779 (1)	105.15 (1)	96.21 (1)	106.59 (1)					
		(ii)	$P\bar{1}$	7.703	7.973	12.079	108.36	93.91	106.35	53	−132.0	−80.8	−127.0	5
(8)	$Cu(MTT)_2$	(i)	$P2_1/n$	4.086 (1)	12.757 (1)	11.898 (1)	90	97.01 (1)	90					
		(ii)	$P2_1$	4.154	12.981	11.867	90	96.04	90	47	−151.0	−109.1	−151.0	1
(9)	$Ni(MTT)_2$	(i)	$P\bar{1}$	10.067 (5)	10.167 (8)	11.184 (4)	92.71 (6)	114.00 (4)	112.92 (5)					
		(ii)	$P1$	10.052	10.299	11.100	93.62	112.65	112.50	14	−92.0	−57.8	−92.0	1

† N denotes the number of structures in the final clustered set; these have the energy range indicated in the subsequent column. ‡ The lowest-energy structure has rank 1, the next lowest has rank 2 *etc.* § Since the metal atom in each case is reassigned as Zn and (5) is isostructural with (6), prediction of the structure of (6) is equivalent to that of (5).

$Cu(MPT)_2$, (5), where intermolecular Cu···O contacts exist in the crystal structure. These interactions lie roughly parallel to the (100) plane, and the change in the b parameter on minimization is particularly large (0.5 Å). This suggests that the intermolecular Cu···O interactions are not modelled particularly well. Better agreement is observed on minimization of the analogous $Ni(MPT)_2$ structure, (6), suggesting that the corresponding Ni···O interactions are modelled somewhat more satisfactorily.[1] This level of agreement between the observed and minimized structure of (5) may be contrasted with the particularly good fit for $Ni(MTT)_2$, (9), for example, where the Ni atoms are essentially encased by organic ligands. These observations provide a quantitative illustration of the intuitive notion that the model provides the best fit where the metal atoms have little or no involvement in intermolecular interactions.

4.2. Results of the structure predictions

A summary of the results from the structure predictions is given in Table 2. In each case, the correct crystal structure was present among the predicted models.[2] For all except (3) and (7), the correct structure was predicted as the global minimum, *i.e.* the energy function alone facilitates determination of the correct crystal structure. For the centrosymmetric molecular units in (1), (8) and (9), the correct crystal structure was predicted in the subgroup of the actual space group with the centre of symmetry removed. In the predicted models, the molecular centre coincides with the crystallographic centre of symmetry in the higher space group such that the correct space-group symmetry is reproduced. For (3) and (7), the correct crystal structure is not predicted to be the global minimum in the energy function, *i.e.* the energy function alone does not contain sufficient information to deduce the crystal structure correctly. In these cases, however, simulation of a PXRD profile for each predicted model facilitates rapid identification of the correct structure using the visual comparison. The PXRD profiles of the ten lowest-energy structures predicted for (3) and (7) are shown in Figs. 4 and 5, respectively, and the correct crystal structures may be identified as the ninth lowest-energy model for (3) and the fifth lowest-energy model for (7).[3] PXRD profiles simulated from the known structure and these predicted models together with overlays of the predicted and observed structures are shown in Figs. 6 and 7. In each case, the unit-cell parameters of the predicted models are sufficiently close to the correct parameters to facilitate subsequent refinement against the PXRD

[1] The fact that the isostructural (5) and (6) (when both described as zinc complexes) do not converge to exactly the same structure on minimization (Table 1) illustrates the extreme complexity of the energy hypersurface. This result does not present a problem for the structure prediction procedure, however, since the two structures would be grouped together and deemed identical during cluster analysis.

[2] Overlays of the observed and predicted structures for (1)–(9) and their simulated PXRD profiles are included as supplementary material, available from the IUCr electronic archives (Reference: AN0589). Services for accessing these data are described at the back of the journal.

[3] For (7), the PXRD profiles simulated from the fifth and eighth lowest-energy predicted structures are extremely similar. Examination of the models reveals that they are essentially identical: for the eighth lowest-energy model, a = 7.747, b = 7.833, c = 12.145 Å, α = 108.33, β = 95.46, γ = 103.83°. The small differences in the unit-cell parameters of the two models, however, render them different within the tolerances employed for cluster analysis; wider tolerances would lead to the two models being deemed identical. Re-clustering is not necessary here since either model would be suitable for subsequent Rietveld analysis. The fifth lowest-energy model was selected as the correct one on the basis of its lower energy.

data, and the overlays of the predicted and observed crystal structures show them to be effectively identical. Thus, combination of the information present in the energy function and the PXRD profile facilitates absolute determination of the crystal structure in each case. Recourse to the PXRD profile to distinguish between the predicted structures means that the

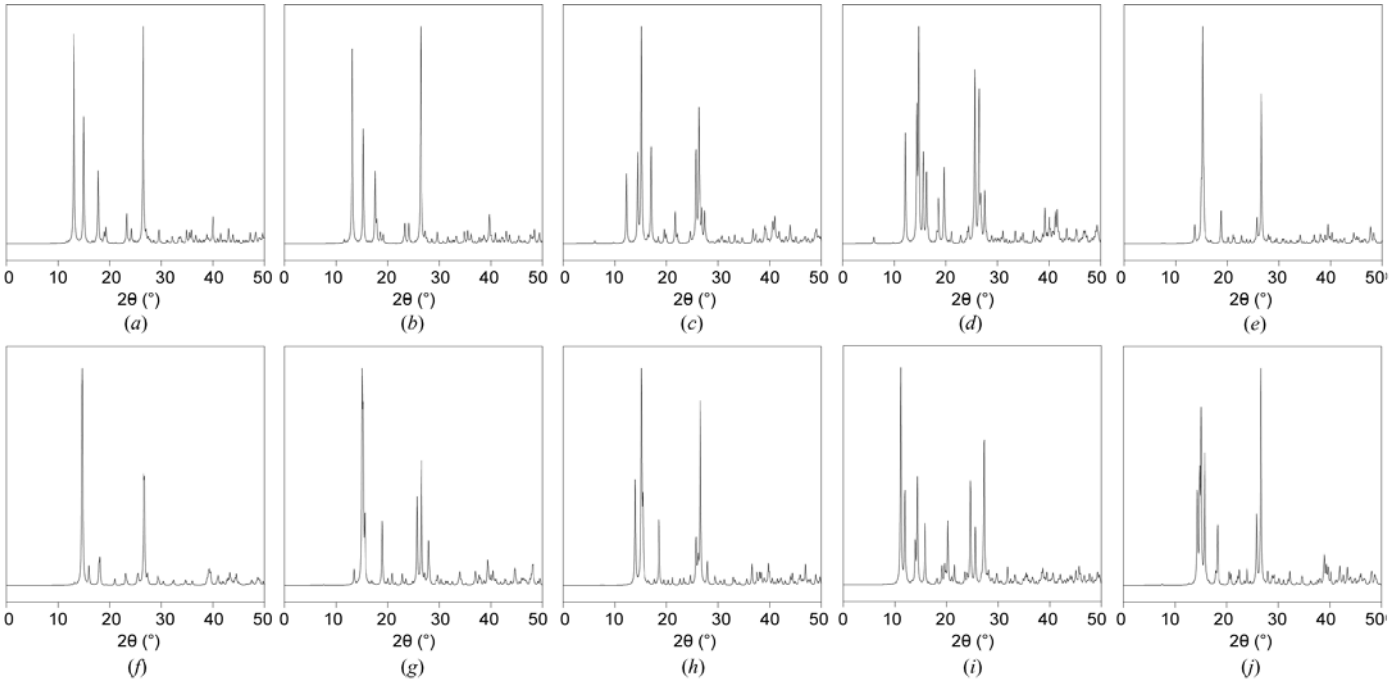


Figure 4
PXRD profiles simulated from the ten lowest-energy structures predicted for (3). (*a*) Model 1; Lattice binding energy $E = -122.5$ kJ mol^{-1}. (*b*) Model 2; $E = -122.2$ kJ mol^{-1}. (*c*) Model 3; $E = -121.5$ kJ mol^{-1}. (*d*) Model 4; $E = -120.7$ kJ mol^{-1}. (*e*) Model 5; $E = -119.8$ kJ mol^{-1}. (*f*) Model 6; $E = -119.5$ kJ mol^{-1}. (*g*) Model 7; $E = -119.5$ kJ mol^{-1}. (*h*) Model 8; $E = -118.7$ kJ mol^{-1}. (*i*) Model 9; $E = -117.9$ kJ mol^{-1}. (*j*) Model 10; $E = -116.9$ kJ mol^{-1}.

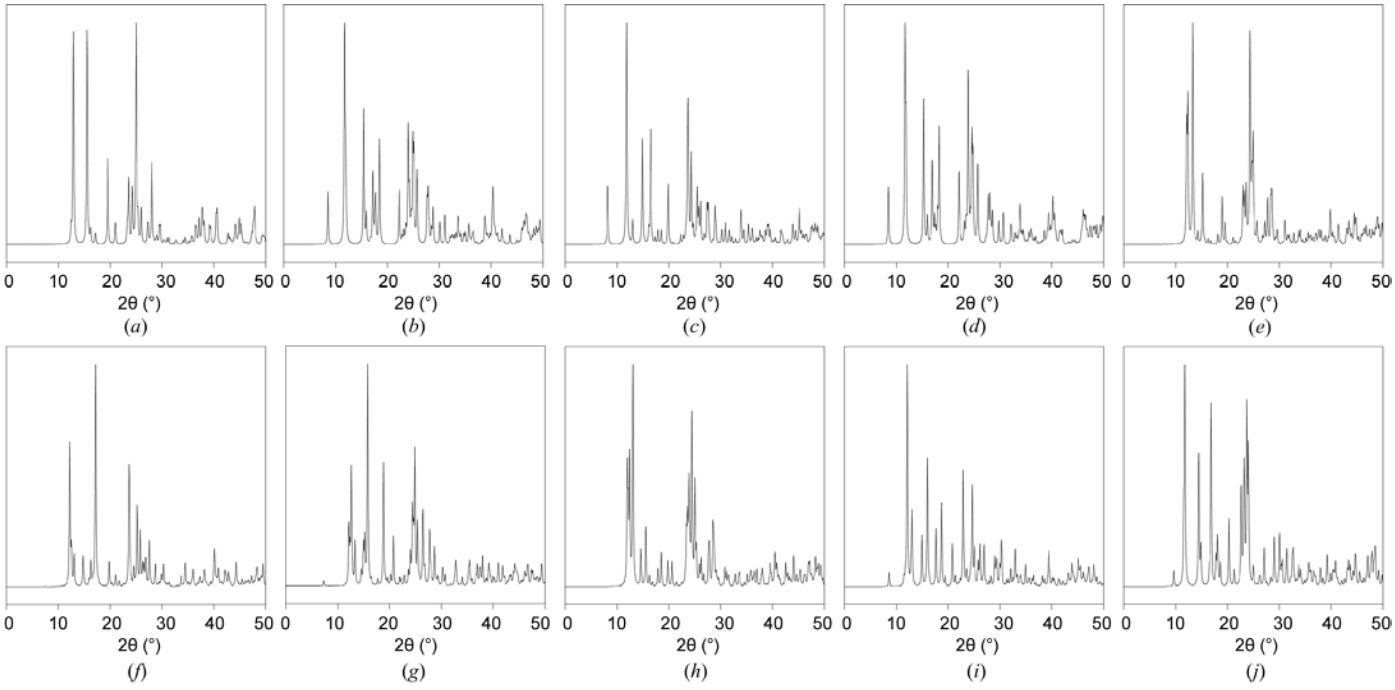


Figure 5
PXRD profiles simulated from the ten lowest-energy structures predicted for (7). (*a*) Model 1; Lattice binding energy $E = -132.0$ kJ mol^{-1}. (*b*) Model 2; $E = -130.0$ kJ mol^{-1}. (*c*) Model 3; $E = -127.8$ kJ mol^{-1}. (*d*) Model 4; $E = -127.4$ kJ mol^{-1}. (*e*) Model 5; $E = -127.0$ kJ mol^{-1}. (*f*) Model 6; $E = -126.3$ kJ mol^{-1}. (*g*) Model 7; $E = -125.9$ kJ mol^{-1}. (*h*) Model 8; $E = -124.5$ kJ mol^{-1}. (*i*) Model 9; $E = -124.2$ kJ mol^{-1}. (*j*) Model 10; $E = -123.9$ kJ mol^{-1}.

procedure may no longer be described as *prediction*, but rather structure *solution*, with the additional advantage that *a priori* knowledge of the unit-cell parameters is not required.

5. Discussion

It is evident that the direct-space approach for crystal structure solution from PXRD data and the structure prediction methodology applied here are entirely analogous, differing only in the nature of the function against which optimization is performed. For both approaches, the key criterion for success is the balance between the amount of reliable information that may be specified at the outset and the information content of the relevant function. Two approaches therefore present themselves to maximize the chances of success: either the extent of the information specified at the outset is increased (by specifying the correct molecular conformation, for example), or the available information against which to optimize may be increased (by combination of the energy function and the PXRD profile). These two cases are analogous, respectively, to the introduction of constraints and restraints, familiar in least-squares refinement of single-crystal data sets where the data-to-parameter ratio is low.

The initial manifestation of this information balance is the requirement to specify the molecular connectivity and (to a greater or lesser extent) the molecular conformation for both structure solution from PXRD data and for structure prediction – global optimization of the positions of each individual atom in a molecular crystal is far too complex to be viable. The next most important information that can be specified at the outset is the space-group and unit-cell parameters. Specification of the symmetry operators and the size and shape of the unit cell reduces the number of independent variables that must be optimized, reducing considerably the dimensionality of the hypersurface against which optimization is performed. For structure solution, the unit-cell parameters are derived at the outset by indexing the PXRD profile and are subsequently constrained during the global optimization procedure. Global optimization for direct-space structure solution from PXRD data is, therefore, considerably less complex than that for structure prediction where unit-cell parameters must be optimized simultaneously. It is for this reason that direct-space structure solution is always preferable to prediction when the PXRD profile can be indexed unambiguously. The inherent simplification of the global optimization means that it becomes viable to treat simultaneously a considerable degree of conformational flexibility, and simultaneous optimization of

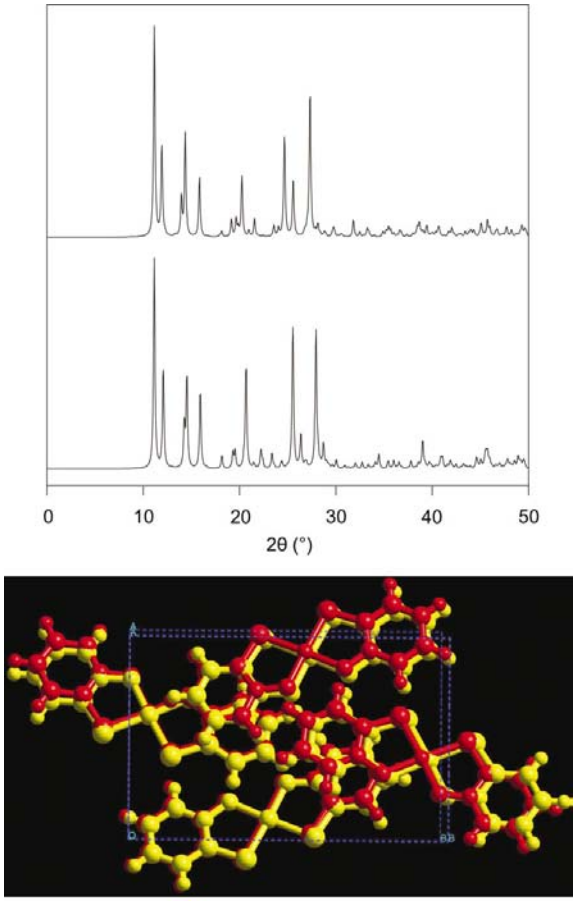


Figure 6
PXRD profiles simulated from the known structure (bottom) and ninth lowest-energy predicted model (top) for (3) and overlay of predicted (red) and observed (yellow) structures.

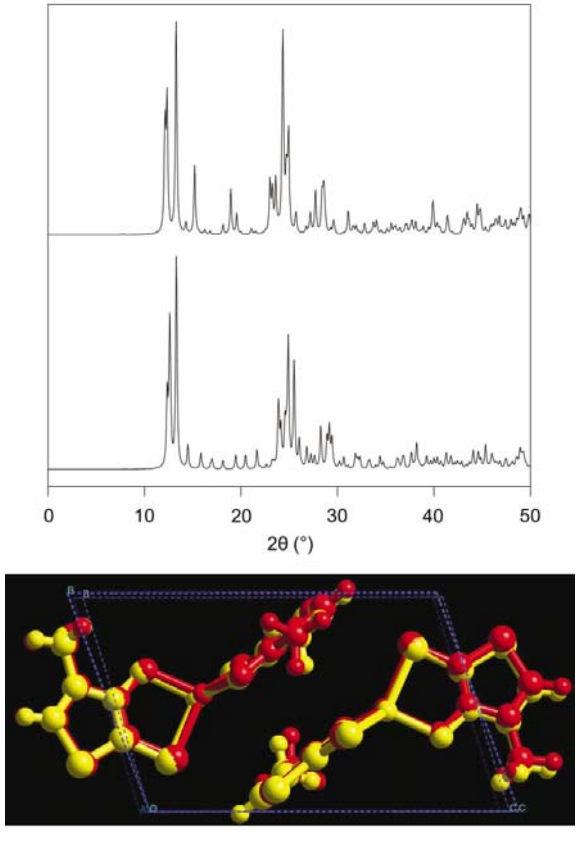


Figure 7
PXRD profiles simulated from the known structure (bottom) and fifth lowest-energy predicted model (top) for (7) and overlay of predicted (red) and observed (yellow) structures.

intramolecular and intermolecular geometry is commonplace in solution procedures. Such an approach is generally less viable for prediction, although examples have been reported in which a limited number of intramolecular degrees of freedom have been optimized simultaneously during the prediction procedure.[4] In many instances, however, the molecular conformation problem is separated from the crystal packing problem by specification of the molecule as a rigid body. With this approach, it may be possible to allow the molecular conformation to relax also in a final minimization step, but this cannot be used to overcome gross errors in the initial molecular conformation since location of the correct structure on minimization requires that the initial structure be located within the basin of attraction of the global minimum, *i.e.* the structure obtained from the SA run must be sufficiently close to the global minimum for minimization to locate it. This is unlikely to be the case where the molecular conformation used in the SA run is grossly in error. As a consequence, the prediction approach is more reliant on reliable specification of the molecular conformation at the outset.

While it is generally accepted that genuine *ab initio* structure prediction is a distant goal, the results presented here suggest that current methodology is somewhat more effective than may be perceived when the limitations concerning the assessment of the predicted structures solely on the basis of energy are understood. The energy function implemented is an empirical approximation to the true energy of the crystal, parameterized against structures which themselves are subject to experimental uncertainties. As a result, the predicted structures will always be subject to a degree of uncertainty, *i.e.* the energy function simply does not contain sufficient information to specify the correct crystal structure unambiguously (Gavezzotti & Filippini, 1996). The relatively simple prediction approach employed here (*i.e.* the generic nature of the force field, the simple approach to the treatment of the molecular charge distribution, and the significant approximation involved in treating the metal atom) is sufficient to provide the correct structure among the low-energy structures, and utilization of additional information in the form of the (indexed or unindexed) PXRD profile facilitates discrimination between them. It is noted that the molecules treated here lack conformational flexibility and that their molecular conformations are therefore readily predictable; the technique is undoubtedly limited for molecules where extensive conformational flexibility is possible. Nonetheless, the degree of success is encouraging. In seven of the nine structures considered, the correct crystal structure is predicted as the global minimum in the energy function. For the two molecules where this is not the case, the correct structure is one of the top ten lowest-energy structures. While it is fair to say, therefore, that existing methodology may not be adequate for reliable crystal structure prediction, it does in fact provide a powerful tool for structure solution which is especially useful in cases where unambiguous indexing of the PXRD profile is not possible.

[4] In the 1999 CCDC blind test (Lommerse *et al.*, 2000), several methods were employed for compound III (with two intramolecular rotational degrees of freedom) in which the molecular conformation was flexible during the prediction procedure. One of these, *UPACK* (van Eijck & Kroon, 1999), provided a successful structure prediction solely on the basis of energy, and the other four methods provided the correct structure within their set of lowest-energy structures (energy ranks between 1 and 32; Schmidt, 2001).

We thank the EPSRC and Avecia Ltd for funding *via* a CASE studentship to ADB, and MSI Inc. (now Accelrys) for provision of a demonstration version of *Polymorph Predictor*.

References

Aakeröy, C. B., Nieuwenhuyzen, M. & Price, S. L. (1998). *J. Am. Chem. Soc.* **120**, 8986–8993.

Barnett, B. L., Kretschmar, H. C. & Hartman, F. A. (1977). *Inorg. Chem.* **16**, 1834–1838.

Baur, W. H. & Kassner, D. (1992). *Acta Cryst.* B**48**, 356–369.

Beyer, T., Day, G. M. & Price, S. L. (2001). *J. Am. Chem. Soc.* **123**, 5086–5094.

Bond, A. D., Feeder, N., Teat, S. J. & Jones, W. (2000). *Tetrahedron*, **56**, 6617–6624.

Bond, A. D., Feeder, N., Teat, S. J. & Jones, W. (2001). *Acta Cryst.* C**57**, 1157–1158.

Bond, A. D. & Jones, W. (2001). *J. Chem. Soc. Dalton Trans.* pp. 3045–3051.

Chen, X.-T., Hu, Y.-H., Wu, D.-X., Weng, L.-H. & Kang, B.-S. (1991). *Polyhedron*, **10**, 1651–2657.

Clegg, W. (2000). *J. Chem. Soc. Dalton Trans.* pp. 3223–3232.

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

Davies, J. E. & Bond, A. D. (2001). *Acta Cryst.* E**57**, o947–o949.

Desiraju, G. R. (1989). *Crystal Engineering: The Design of Organic Solids.* Amsterdam: Elsevier.

Dewar, M. J. S., Zoebisch, E. G., Healy, E. F. & Stewart, J. J. P. (1985). *J. Am. Chem. Soc.* **107**, 3902–3909.

Eijck, B. P. van & Kroon, J. (1999). *J. Comput. Chem.* **20**, 799–812.

Gavezzotti, A. & Filippini, G. (1996). *J. Am. Chem. Soc.* **118**, 7153–7157.

Hammond, R. B., Roberts, K. J., Docherty, R. & Edmondson, M. (1997). *J. Phys. Chem. B*, **101**, 6532–6536.

Harris, K. D. M., Johnston, R. L. & Kariuki, B. M. (1998). *Acta Cryst.* A**54**, 632–645.

Harris, K. D. M., Tremayne, M. & Kariuki, B. M. (2001). *Angew. Chem. Int. Ed. Engl.* **40**, 1626–1651.

Harris, K. D. M., Tremayne, M., Lightfoot, P. & Bruce, P. G. (1994). *J. Am. Chem. Soc.* **116**, 3543–3547.

Jones, W. (1997). *Organic Molecular Solids: Properties and Applications*, edited by W. Jones, pp. 149–199. New York: CRC Press.

Karfunkel, H. R. & Gdanitz, R. J. (1992). *J. Comput. Chem.* **13**, 1171–1183.

Kirkpatrick, S., Gelatt, C. D. J. & Vecchi, M. P. (1983). *Science*, **220**, 671–680.

Lanning, O. J., Habershon, S., Harris, K. D. M., Johnston, R. L., Kariuki, B. M., Tedesco, E. & Turner, G. W. (2000). *Chem. Phys. Lett.* **317**, 296–303.

Le Bail, A., Duray, H. & Fourquet, J. L. (1988). *Mater. Res. Bull.* **23**, 447–452.

Leusen, F. J. J. (1996). *J. Cryst. Growth*, **166**, 900–903.

Leusen, F. J. J., Wilke, S., Verwer, P. & Engel, G. E. (1999). *Nato Science Series E*, Vol. 360, edited by J. A. K. Howard, F. H. Allen & G. P. Shields, pp. 303–314. Dordrecht: Kluwer Academic Publishers.

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

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

Padmaja, N., Ramakumar, S. & Viswamitra, M. A. (1990). *Acta Cryst.* A**46**, 725–730.
Pagola, S., Stephens, P. W., Bohle, D. S., Kosar, A. D. & Madsen, S. K. (2000). *Nature (London)*, **404**, 307–310.
Pawley, G. S. (1981). *J. Appl. Cryst.* **14**, 357–361.
Putz, H., Schön, J. C. & Jansen, M. (1999). *J. Appl. Cryst.* **32**, 864–870.
Rietveld, H. M. (1967). *Acta Cryst.* **22**, 151–152.
Rietveld, H. M. (1969). *J. Appl. Cryst.* **2**, 65–71.
Schmidt, M. U. (2001). Personal communication.
Schmidt, M. U. & Englert, U. (1996). *J. Chem. Soc. Dalton Trans.* pp. 2077–2082.
Shankland, K., David, W. I. F. & Csoka, T. (1997). *Z. Kristallogr.* **212**, 550–552.
Starbuck, J., Docherty, R., Charlton, M. H. & Buttar, D. (1999). *J. Chem. Soc. Perkin Trans. 2*, pp. 677–691.
Stewart, J. J. P. (1990). *J. Comp. Aid. Mol. Design.* **4**, 1–45.
Stewart, J. J. P. & Rzepa, H. S. (1999). http://home.att.net/~mopacmanual/node347.html#6151.
Tremayne, M., Kariuki, B. M. & Harris, K. D. M. (1997). *Angew. Chem. Int. Ed. Engl.* **36**, 770–772.
West, D. X., Brown, C. A., Jasinski, J. P., Heathwaite, R. M., Fortier, D. G., Staples, R. J. & Butcher, R. J. (1998). *J. Chem. Cryst.* **28**, 853–860.