Computationally Assisted Structure Determination for Molecular Materials from X-ray Powder Diffraction Data

Robert B. Hammond and Kevin J. Roberts*,†

Centre for Molecular and Interface Engineering, Department of Mechanical and Chemical Engineering, Heriot Watt University, Riccarton, Edinburgh EH14 4AS, U.K.

Robert Docherty and Michael Edmondson

Zeneca Specialties Research Centre, Hexagon House, Blackley, Manchester M9 8ZS, U.K.

Received: May 7, 1997[⊗]

For organic molecular materials, the key step in solving crystal structures from high-resolution powder diffraction data is the generation of reliable trial structures for final refinement. In this paper we demonstrate an efficient new methodology for generating trial structures that predicts the correct molecular packing in a crystal lattice given the unit cell dimensions and space group. The method uses a systematic search of intermolecular, atom−atom interactions to examine and rank all possible packing arrangements. The approach is demonstrated for a number of systems with increasing degrees of search complexity including indigo, 6,13-dichlorotriphendioxazine, phenanthrene, paracetamol, and benzophenone. The final example is of particular importance as it illustrates the first application of this type of methodology to a conformationally flexible system.

1. Introduction

The link between the molecular structure and the solid state packing of organic materials is the foundation upon which solid state chemists build structure activity relationships. The favored method of structure elucidation is single-crystal X-ray diffraction, which has recently, for most small and medium size molecules, become an almost routine analytical technique. Evidence the Cambridge Crystallographic Database,[1] where the number of entries has risen to 160 000, the quality of structures reported has improved, and the complexity of systems investigated has increased. However, it is not always possible to produce single crystals of sufficient size and quality, in the desired polymorphic form, for conventional X-ray analysis. This is of particular relevance in the area of "effect" chemical species such as pharmaceuticals, agrochemicals, pigments, and organic photoconductors. In the absence of crystallographic information, our understanding of how the physical, chemical, and biochemical properties of these materials are related to the solid state structure remains, at best, tentative.

When single-crystal X-ray diffraction studies cannot be applied, structural information has to be generated by other techniques such as powder diffraction studies[2,3] or theoretical approaches.[4−6] Structure determination from powder diffraction data for molecular materials is hindered by the difficulties in generating trial structures for final Rietveld refinement[7] (*i.e. where to start?*). Ab initio prediction of crystal structures remains an admirable long term goal, and much progress has been made in recent years.[4−6] However, to date such methods are still plagued by difficulties including the location of global energy minima, force field accuracy, and the description of the electrostatic interactions. Given a novel material, these problems combine to make it difficult to know when the correct packing arrangement has been obtained (*i.e. when to finish!*). A further problem that has hindered ab initio crystal structure prediction is the use of "in vacuo" optimized molecular structures as the input probe for all the search methods. In certain approaches[5] the molecular structure is allowed to relax as part of the final lattice energy minimization procedures, but the possible changes in conformation are not assessed during the search procedure. In reality, the most common approach for conformationally flexible molecules is to run separate searches for each conformation.

Overall what is required is a pragmatic combination of both theoretical and experimental approaches[8,9] that exploits the complementary aspects of the two methods to overcome their inherent individual limitations. Theoretical methods provide a route to generate the good trial structures necessary for Rietveld fitting methods[10] and molecular constraints for the refinement. The use of the diffraction trace as an experimental observable reduces, in principle, some of the difficulties that have plagued the routine application of ab initio prediction methods.

Structure determination using powder X-ray diffraction for molecular materials can be considered as a three-stage process:

(1) indexing the diffraction pattern, which provides information on the unit cell and space group symmetry;

(2) trial structure generation, postulating the best potential packing arrangements;

(3) structural refinement using the Rietveld method. It is important to remember that the Rietveld method is a refinement process and that successful implementation relies on good starting structures.[13] It is accepted that in the case of molecular materials the generation of sufficiently good starting stuctural models is the key to determination.

In previous work, diffraction pattern decomposition into "single-crystal-like" integrated intensities and subsequent application of Patterson or direct methods have been used to generate approximate trial structures, but this approach has limited applicability for low∼symmetry organics as the overlap between adjacent reflections makes assignment of individual intensities problematic.[10] Recently the elegant application of Metropolis Monte Carlo methods to generate trial structures has been demonstrated.[2] In that approach random movements of

† Also affiliated with CCLRC Daresbury Laboratory, Warrington, WA4 4AD.

⊗ Abstract published in *Advance ACS Abstracts,* July 1, 1997.

S1089-5647(97)01535-6 CCC: $14.00

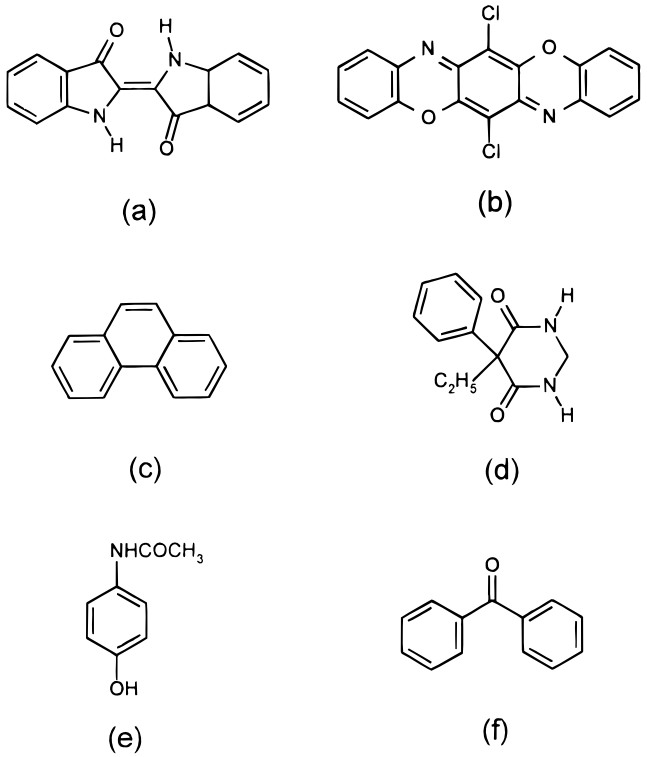

Figure 1. Molecular structure of the materials selected for study: (a) indigo; (b) 6,13-dichlorotriphendioxazine; (c) phenanthrene; (d) primidone; (e) paracetamol; (f) benzophenone.

molecules or molecular fragments within a fixed cell are carried out according to a Metropolis Monte Carlo algorithm. A move is accepted or rejected on the basis of the fit of a simulated diffraction pattern, for a postulated structure, with the observed X-ray diffraction data.

In the approach described here we have utilized calculated molecular structures and lattice energies to rank systematically all potential packing arrangements. This has the advantage that it gives a truly *independent* measure of the quality of a postulated structure. The benefits of this approach to structure determination from powders have recently been demonstrated in the case of the photosensitive X-form of metal-free phthalocyanine. The initial structure postulated purely from the powder diffraction data using a Monte Carlo type approach[11] had an R_{wp}-factor around 20%, some improbable bond lengths, and a physically unrealistic lattice energy. Application of our techniques[12] produced a structure with a much better fit to the observed pattern, bond lengths within acceptable limits,[14] and a lattice energy consistent with the other known polymorphs.

In this paper we demonstrate the general applicability of the systematic search procedure for the efficient generation of reliable trial structures not only for rigid systems but also for materials where conformational flexibility is important. It is not the purpose of this paper to describe individual structure determinations in detail but rather to illustrate the speed and reliability of our trial structure generation method for a number of test cases.

2. Computational Methodology

The systematic search approach uses a grid-based search of rotations and translations to assess *all* the possible packing arrangements. In the past, systematic searching methods were often regarded as too time-consuming for practical application in this area. In this work, however, the use of in-built chemical sense and calculated lattice energies to decide upon the

TABLE 1: Summary Data for the Performance of the Systematic Search Approach on Six Trial Materials That Exhibit Varying Degrees of Molecular Complexity[a]

material	crystal structure data	reference crystal structure	distance lower limit [Å]	E_{latt}(exp) [kcal/mol]	E_{latt}(calc) − published crystal structure [kcal/mol]	E_{latt}(calc) − systematic search structure [kcal/mol]	number of configurations	reference force field	CPU time
indigo	$P2_1/c$, $a = 9.24$ Å, $b = 5.77$ Å, $c = 12.22$ Å, $\beta = 117.0°$, $z = 2$	20	1.9	−31.8	−32.5	−31.2	46K	18	18
6,13-dichloro-triphendioxazine	$P2_1/c$, $a = 8.72$ Å, $b = 4.89$ Å, $c = 17.15$ Å, $\beta = 97.87°$, $z = 2$	8	2.1	n.a.[b]	−47.2	−40.7	46K	19	8
phenanthrene	$P2_1$, $a = 8.47$ Å, $b = 6.17$ Å, $c = 9.47$ Å, $\beta = 98.01°$, $z = 2$	21	2.1	−22.9	−19.8	−17.6	4.67M	18	53
primidone	$P2_1/c$, $a = 12.25$ Å, $b = 7.09$ Å, $c = 14.81$ Å, $\beta = 117.8°$, $z = 4$	22	1.7	n.a.	−25.6	−22.4	5.83M	18	65
paracetamol	$Pcab$, $a = 11.81$ Å, $b = 17.16$ Å, $c = 7.39$ Å, $z = 8$	23	1.8	n.a.	−19.3	−17.4	5.83M	18	264
benzophenone	$P2_12_12_1$, $a = 10.3$ Å, $b = 12.15$ Å, $c = 8.00$ Å, $z = 4$	24	2.1	−23.9	−24.5	−19.9	64.9M	19	313

[a] The lower lattice energy cutoff used was −10 kcal/mol. CPU times are in minutes. Experimental lattice energies after ref 25. [b] n.a. = not available.

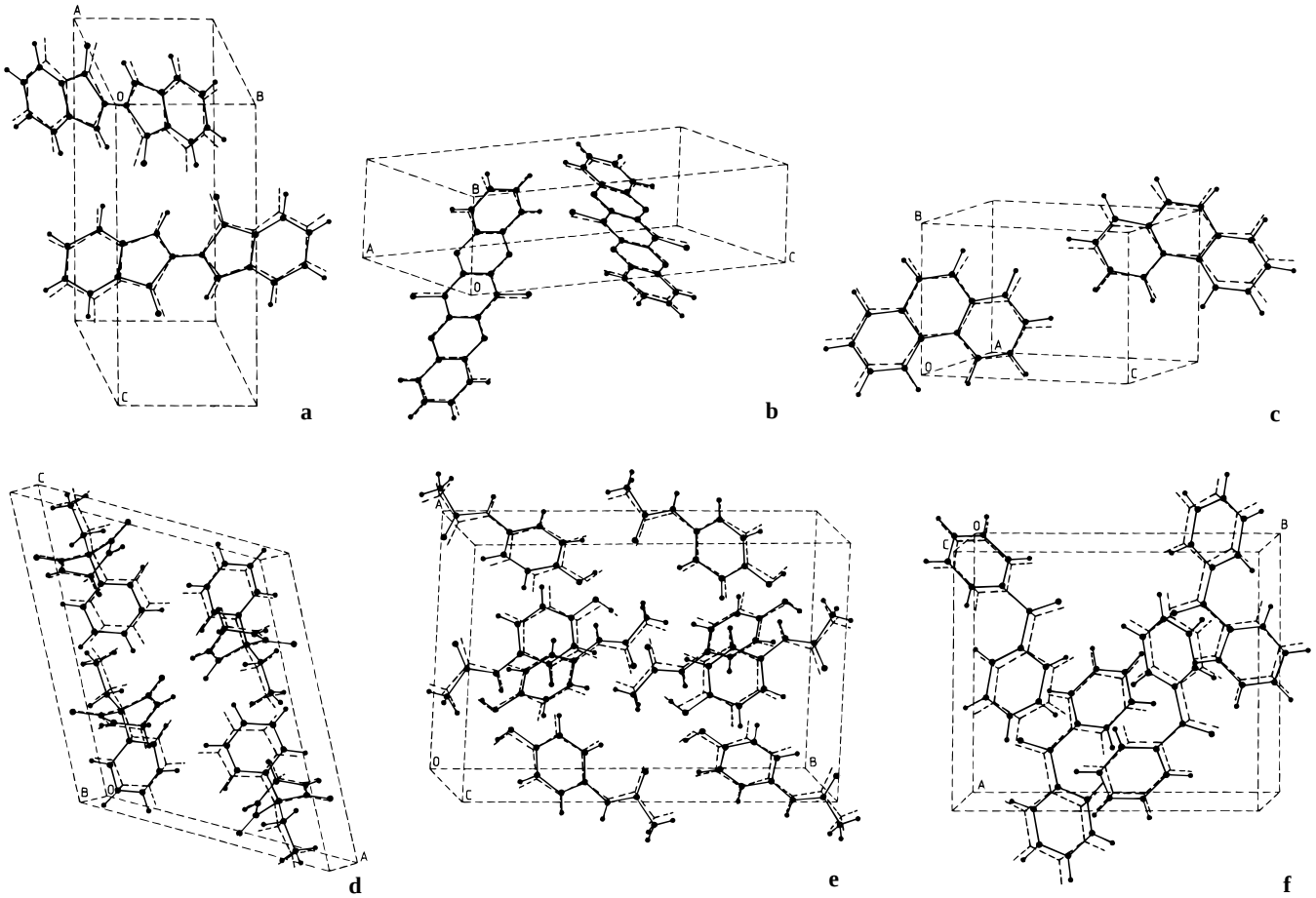

Figure 2. Results of the systematic search (full lines, crystallographic structure; broken lines, computed structure): (a) indigo; (b) 6,13-dichlorotriphendioxazine; (c) phenanthrene; (d) primidone; (e) paracetamol; (f) benzophenone.

suitability and ranking of potential arrangements is used to provide a fast, efficient, and reliable filter for the selection of trial structures. The philosophy is to use a twin track approach, ranking trial structures both on lattice energy and goodness of fit to the observed X-ray diffraction data.

The lattice energy E_{latt} (often referred to as the crystal binding or cohesive energy) can be calculated for molecular materials by summing all the interactions between a central molecule and all the surrounding molecules. Each intermolecular interaction can be considered to consist of the sum of the constituent atom−atom interactions. If there are n atoms in the central molecules and n' atoms in each of the N surrounding molecules, then a lattice energy can be calculated by the eq 1 shown below. In most cases n and n' will be equal, but in the case of molecular complexes they may differ. V_{kij} is the interaction potential between atom i in the central molecule and atom j in the kth surrounding molecule.

$$E = \frac{1}{2}\sum_{k=1}^{N}\sum_{i=1}^{n}\sum_{j=1}^{n'} V_{kij} \qquad (1)$$

In the implementation of the method described here, a truncated lattice is constructed during the search based upon reduced radial cutoff distances. Previous calculations have shown that, for electronically neutral molecular materials, the first coordination sphere contributes the majority of the lattice energy; e.g. in urea the first coordination sphere of six hydrogen-bonded molecules accounts for 82% of the lattice energy.[15]

The fit of the simulated and observed powder diffraction patterns is assessed using eq 2. Initially a scale factor, λ, is refined through minimizing the weighted, w_i least-squares difference between observed, y_i, and calculated, y_{ci}, diffraction intensities. Using this scale factor, the residual of the fit between the calculated and observed intensities S_y is evaluated.

$$S_y = \sum_i w_i(y_i - \lambda y_{ci})^2 \qquad (2)$$

Given the large number of possible configurations that have to be checked for a system with three translational and three rotational degrees of freedom (never mind any conformational flexibility!), a truncated lattice energy calculation, or diffraction fit determination, will still be so time-consuming as to make the calculations impracticable. An initial screen has to be used to cut down the number of possibilities that are to be assessed on lattice energy and diffraction pattern fit. In our implementation a distance-based cut-off criteria is used as the preliminary screen of possible structures. The use of distance-based search algorithms is well-known in the generation 3D molecular structures from 2D databases.[16] We shall demonstrate how effective these simple concepts can also be in a trial structure generation procedure. By considering the approach distances of nonbonded atoms, it is easy to reject improbable arrangements during cell construction. Only "sensible" structures will pass through this screen. This reduces considerably the number of structural configurations that have to be checked by lattice energy calculations and/or fit to the observed X-ray diffraction pattern. It is this screen that enables systematic searches on quite fine grids to be undertaken in reasonable time scales.

In this paper we illustrate a limited examination of conformational flexibility, a capability integrated within the search procedure. While this is a limited example with only two

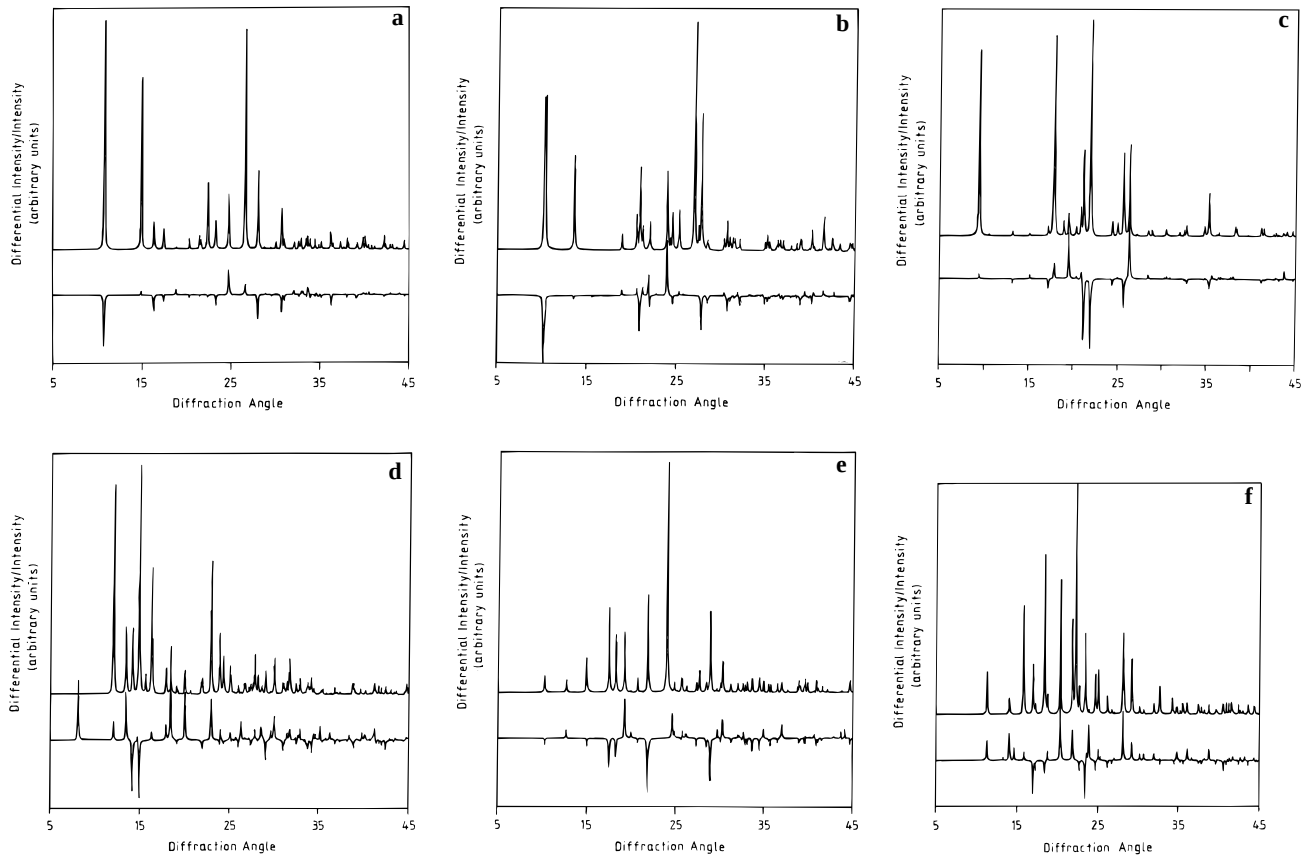

Figure 3. Plots of simulated X-ray powder diffraction data based on the best trial structures (upper) and difference plots generated by subtraction of simulated diffraction data for the crystallographic structure from the systematic search structure (bottom): (a) indigo; (b) 6,13-dichlorotriphendioxazine; (c) phenanthrene; (d) primidone; (e) paracetamol; (f) benzophenone.

conformationally flexible bonds, it does demonstrate that the approach is viable and offers considerable promise. The concept of a fast prescreening procedure to significantly reduce the total number of potential trial configurations is, once again, employed. Taking an isolated molecule, all possible conformations are ranked in terms of an increasing intramolecular potential energy difference compared with the most stable conformation. Van der Waals and electrostatic terms together with a molecular mechanics type, torsional, potential energy term are used in the energy calculations. The calculated difference can be regarded as an energy penalty that must be offset by a corresponding improvement in the lattice energy. Accordingly, as has been recently demonstrated,[17] it is possible to set quite a tight rejection criteria for the selection of conformations based upon this energy difference. Only those conformations that pass this energy screen are subsequently considered in the systematic search.

3. Results

To demonstrate the application of this methodology, we describe six different systems; the molecular structures are drawn in Figure 1. The first two, indigo and 6,13-dichlorotriphendioazine, represent systems where only rotational space is explored. The other examples that include translational search space are phenanthrene, paracetamol, and primidone (in the case of phenanthrene, space group symmetry results in only two degrees of freedom). These encompass a range of different intermolecular interaction types and exhibit varied crystal chemistry. In all cases the experimental unit cell parameters were taken, and the position of the asymmetric unit was randomized in the cell. The systematic search procedure was then implemented, and the results are described. In all but one case the molecular structure was held fixed at the experimental conformation. In the final example, benzophenone, the two torsions describing the phenyl rings were initialized such that the molecule was planar, a deliberately unfavorable arrangement.

In routine applications of the search method, both diffraction data and lattice energy are used to rank trial structures. For the present purpose of validating the approach, however, it was only necessary to compare the top trial structure, ranked on lattice energy, directly with the experimental crystal structure. In one case, 6,13-dichlorotriphendioxazine, experimental diffraction data were available allowing a comparison of the separate ranking criteria.

All calculation were carried out using a Silicon Graphics Indigo^2 workstation with a R4400 processor.

(a) Indigo is a well-known vat dye used in the coloration of textiles, especially blue jeans. The intermolecular packing for indigo consists of $\pi-\pi$ stacks that are held together by intermolecular hydrogen bonds. These are somewhat longer than ideal because of the competing intramolecular hydrogen bond. The crystal structure was taken, and, with the center of symmetry retained, the orientation of the molecule was completely randomized. The systematic search routine was applied in steps of 5° in all three degrees of rotational freedom.

(b) 6,13-Dichlorotriphendioxazine is the basic chromophore unit of a number of commercially important dyestuffs. The main features of the crystal packing are $\pi-\pi$ stacks along the *b*-axis and nonbonded Cl--H contacts at 2.8 Å forming dimers in the *ac*-plane. The systematic search routine was applied in steps of 5° in all three degrees of rotational freedom. In this example the best trial configurations selected after ranking, separately, on lattice energy and fit to the experimental diffraction data were coincident.

(c) Phenanthrene was selected as an example of a typical planar aromatic hydrocarbon. Here the crystal packing shows a classical herringbone arrangement with the molecular planes forming an angle of 58°. Given the space group, it was only necessary to explore translation space along the crystallographic *a*- and *c*-axes. The systematic search routine was applied in steps of 10° in all three degrees of rotational freedom and translational steps of 0.05 in fractional coordinates.

(d) Primidone is a pharmaceutical material. The main feature of the crystal packing is chains of molecules joined by hydrogen bonds alternating between dimers, with a C═O···NH distance of 2.04 Å, and a single hydrogen bond each to two further molecules, with a C═O···NH distance of 1.94 Å. The systematic search routine was applied in steps of 10° in all three degrees of rotational freedom and translational steps of 0.05 in fractional coordinates.

(e) Paracetamol is a well-known pharmaceutical material. The crystal packing features hydrogen-bonded chains parallel to the *b*-axis, with a C═O···OH distance of 1.91 Å, and NH···OH interactions with a nonbonded distance of 2.21 Å. The systematic search routine was applied in steps of 10° in all three degrees of rotational freedom and translational steps of 0.05 in fractional coordinates.

(f) Benzophenone was selected as a test example for a systematic search to include two torsional degrees of freedom. To undertake this search it was necessary to employ a force field parametrized for treating both intramolecular and intermolecular interactions, and the generic Dreiding force field[20] was selected. The torsion angles defining the orientations of the phenyl moieties were treated in 5° steps in the range 0−90°, a total of 361 conformations. Conformations with a calculated energy difference of less than 2.5 kcal/mol, with respect to the most stable conformation, were selected for inclusion in the systematic search. A total of 89 conformations were, therefore, selected. The molecular orientation was iterated in 10° rotational steps, and molecular position of 0.1 fractional coordinate steps.

Structural details of all these trial cases together with summary results are given in Table 1. It should be noted that neither the experimental crystal structures nor the best trial structures found in the searches were subjected to energy minimization prior to comparison. In Figure 2 the best trial packing arrangement, based on energy ranking, is overlaid against the experimental crystal structure for all the molecules examined. A comparison of the simulated diffraction pattern for the best trial structure and a plot of the calculated difference between diffraction traces for the trial and experimental structures are given in Figure 3. In all cases the fit between the molecular packing predicted on the basis of the systematic search with that based on the published crystal structure is excellent.

4. Discussion and Conclusions

The results from the systematic search algorithm demonstrate the efficiency, effectiveness, and validity of this approach for generating trial structures to facilitate structure solution of organic molecular materials from powder diffraction data. The advantages of the method are:

(i) that all local energy minima, within the fixed unit cell, are elucidated;

(ii) that the trial structures with the lowest lattice energies can be cross referenced with the available experimental diffraction data;

(iii) that the lattice energy calculation provides an independent evaluation of the quality of the postulated structure. These features are combined here in an automated and quantitative manner, to establish the best starting configuration for structure solution.

The results for benzophenone, having two torsional degrees of freedom in the search space, are particularly significant. Without reference to the torsion angles from the known crystal structure, the systematic search algorithm correctly identified the packing configuration and molecular conformation in the solid state. This methodology, therefore, offers the possibility of treating conformationally flexible molecules on a reasonable time scale, without the need for a series of examinations, in separate searches, of many low-energy, gas phase conformers. A very much wider range of materials become amenable to treatment by our algorithm given the possibility of conformational searching. In future publications we will report the application of the methodology described here in further solutions, from powder diffraction data, of previously unknown crystal structures.

Acknowledgment. We gratefully acknowledge Zeneca Specialities for the financial support of this work as well as the EPSRC for provision of computational facilities (Grants GR/H/40891, GR/J/40891, GR/J/44711, GR/K/63160) and for the support of a Senior Fellowship to one of us (K.J.R.).

References and Notes

(1) Cambridge Structural Database, Cambridge Crystallographic Data Centre, University Chemical Laboratory, Lensfield Road, Cambridge CB2 2EW.

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

(3) Fitch, A. N.; Jobic, H. *J. Chem. Soc., Chem. Commun.* **1993**, 1516.

(4) a) Gavezzotti, A. *Acc. Chem. Res.* **1994**, *27*, 309. b) Gavezzotti, A.; Filippini, G. *J. Am. Chem. Soc.* **1996**, *118*, 7153.

(5) Karfunkel, R.; Gdanitz, R. J. *J. Comput. Chem.* **1992**, *13*, 1771.

(6) Chaka, A. M.; Zaniewski, R.; Youngs, W.; Tessier, C.; Klopman, G. *Acta Crystallogr.* **1996**, *B52*, 165.

(7) Rietveld, H. M. *J. Appl. Crystallogr.* **1969**, *2*, 65.

(8) Fagan, P. G.; Hammond, R. B.; Roberts, K. J.; Docherty, R.; Chorlton, A. P.; Jones, W.; Potts, G. D. *Chem. Mater.* **1995**, *7*, 2322.

(9) Fagan, P. G.; Roberts, K. J.; Docherty, R.; Chorlton, A. P.; Jones, W.; Potts, G. D. *Mol. Cryst. Liq. Cryst.* **1994**, *248*, 277.

(10) Cheetham, A. K. *The Rietveld Method*; Young, R. A., Ed.; Oxford University Press: Oxford 1993; Chapter 15.

(11) Oka, K.; Okada, O. *J. Imaging Sci. Technol.* **1993**, *37*, 13.

(12) Hammond, R. B.; Roberts, K. J.; Docherty, R.; Edmondson, M.; Gairns, R. *J. Chem. Soc., Perkin Trans. 2* **1996**, 1527.

(13) Newsam, J. M.; Deem, M. W.; Freeman, C. M. *Accuracy In Powder Diffraction II*; NIST: Gaithersburg, MD, May 1992.

(14) Allen, F. H.; Kennard, O.; Watson, D. G.; Brammer, L.; Orpen, G. A.; Taylor, R. *J. Chem. Soc., Perkin Trans. 2* **1987**, *S2*.

(15) Docherty, R.; Roberts, K. J.; Saunders, V.; Black, S. N.; Davey, R. J. *Faraday Discuss.* **1993**, *95*, 11.

(16) Blaney, J. M.; Crippen, G. M.; Dearing, A.; Dixon, J. S. DGEOM: QCPE program no 590; Quantum Chemistry Program Exchange, Department Of Chemistry, Indiana University, Bloomington, IN, 47405.

(17) Allen, F. H.; Harris, S. E.; Taylor, R. *J. Comput. Aided Mol. Design* **1996**, *10*, 247.

(18) Momany, F. A.; Carruthers, L. M.; McGuire, R. F.; Scherega, H. A. *J. Phys. Chem.* **1974**, *78*, 1595.

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

(20) von Eller, H. *Bull. Soc. Chim. Fr.* **1955**, 1433.

(21) Kay, M. I.; Okaya, Y.; Cox, D. E. *Acta. Crystallogr.* **1971**, *B27*, 26.

(22) Yeates, D. G. R.; Palmer, R. A. *Acta Crystallogr.* **1975**, *B31*, 1077.

(23) Haisa, H.; Kashino, S.; Maeda, H. *Acta Crystallogr.* **1974**, *B30*, 2590.

(24) Fleicsher, E. B.; Sung, N.; Hawkinson, S. *J. Phys. Chem.* **1968**, *72*, 4311.

(25) Cox, J. D.; Pilcher, G. *Thermochemistry of Organic and Organometallic Compounds*; Academic Press: New York, 1970.