Crystal Structure Solution from Powder X-ray Diffraction Data: The Development of Monte Carlo Methods To Solve the Crystal Structure of the γ-Phase of 3-Chloro-*trans*-cinnamic Acid[†]

Benson M. Kariuki,[‡] Dimple M. S. Zin,[§] Maryjane Tremayne,[§] and Kenneth D. M. Harris*,[‡]

School of Chemistry, University of Birmingham, Edgbaston, Birmingham B15 2TT, U.K., and Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, U.K.

Many important crystalline materials do not form single crystals of sufficient size and/or quality for single-crystal diffraction studies, and in such cases it is essential that the crystal structure can be solved from powder diffraction data; however, there are many difficulties associated with solving crystal structures, ab initio, from powder diffraction data. In this paper, we report the successful application of a Monte Carlo technique to solve the (previously unknown) crystal structure of the γ-phase of 3-chloro-*trans*-cinnamic acid from powder X-ray diffraction data. The "structural fragment" used in the Monte Carlo calculation comprised a rigid *trans*-cinnamic acid molecule (with the chlorine and hydrogen atoms omitted) with its oxygen atoms at a fixed distance from the crystallographic center of symmetry, and with the center of symmetry lying in the molecular plane. The structural fragment was rotated by a random angular displacement around a random axis constrained to pass through the center of symmetry. The "correct" position of this structural fragment was discriminated readily (on the basis of the agreement between experimental and calculated powder X-ray diffractograms) from "wrong" positions sampled during the Monte Carlo calculation, and the "correct" position was then used as the initial structural model in Rietveld refinement and difference Fourier calculations. The paper concludes with a discussion of general considerations relating to the application of the Monte Carlo method for crystal structure solution from powder diffraction data.

1. Introduction

The determination of crystal structures from single-crystal X-ray diffraction data can now, in general, be carried out routinely and straightforwardly. However, a wide range of solids are microcrystalline and are therefore not amenable to investigation using conventional single-crystal X-ray diffraction methods. The availability of approaches for crystal structure determination from powder diffraction data is clearly essential for the structural characterization of such materials.

A major problem limiting the success of conventional approaches for structure solution from powder diffraction data is the requirement to extract the intensities of individual diffraction maxima from the experimental powder diffractogram. In the powder diffractogram, three-dimensional diffraction data are essentially "compressed" into one dimension; as a consequence, peaks in the powder diffractogram generally overlap appreciably, leading to considerable ambiguities in extracting the intensities of individual diffraction maxima.

It is relevant to note that the problems encountered in structure solution from powder X-ray diffraction are particularly severe for "equal atom" organic compounds (i.e., those containing no atom heavier than oxygen), as there is generally little diffraction data at high diffraction angle for such materials. For these "equal atom" structures, our experience has shown that it is necessary to determine a substantial fraction (at least 50%) of the structure in the structure solution calculation, in order for subsequent structure refinement to be successful. In addition, electron density maps produced for these materials (e.g., by the direct methods approach) lack prominent peaks, rendering the identification of atomic positions difficult and unreliable. For these reasons, it is not surprising that only a few successful structure determinations of previously-unknown structures of this type from powder X-ray diffraction data have been reported.

A method based upon a Monte Carlo approach has been developed recently[1,2] for crystal structure solution from powder diffraction data; importantly, this method avoids the need to extract the intensities of individual diffraction maxima directly from the experimental powder diffractogram. In this method, trial structural models are generated *independently* of the experimental

[‡] School of Chemistry, University of Birmingham, Edgbaston, Birmingham B15 2TT, U.K.

[§] Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, U.K.

* Author to whom correspondence should be addressed.

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

(2) Newsam, J. M.; Deem, M. W.; Freeman, C. M. *Accuracy in Powder Diffraction II,* NIST Special Publication, 1992; p 846.
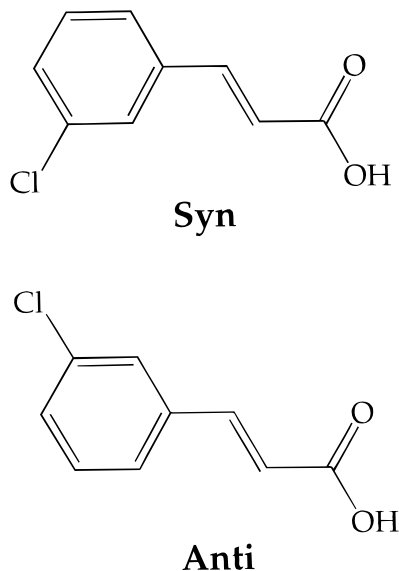
Figure 1. Structural formulas for 3-chloro-*trans*-cinnamic acid in the syn and anti conformations.

powder diffraction data, and the validity of each trial structure is assessed on the basis of the agreement (as defined by the R_{wp} factor (eq 1)) between the experimental diffractogram and the diffractogram calculated for the trial structure:

$$R_{wp} = 100\sqrt{\frac{\sum w_i(y_i - y_{ci})^2}{\sum w_i y_i^2}} \quad (1)$$

where y_i is the intensity of ith observed point in the experimental powder diffractogram, y_{ci} is the intensity of the corresponding point in the calculated powder diffractogram, and w_i is a weighting factor for the ith point.

The trial structural models are generated by random movement of a collection of atoms [denoted the "structural fragment" and generally comprising heavy atoms (dominant X-ray scatterers) and/or well-defined rigid bodies (e.g., aromatic rings)] within the unit cell, and each trial structure is accepted or rejected, via the Metropolis importance sampling algorithm, on the basis of R_{wp}. This approach clearly avoids the problems (encountered in the conventional approaches for structure solution) associated with extracting intensity information directly from the experimental powder diffractogram. Full details of the theory underlying this Monte Carlo approach are given in ref 1.

The best structure solution obtained during the Monte Carlo calculation is then considered further in Rietveld refinement and (if necessary) difference Fourier calculations. The success of this Monte Carlo technique for crystal structure solution from powder diffraction data has been demonstrated[1] by its application to determine the previously known crystal structure of p-CH$_3$C$_6$H$_4$-SO$_2$NHNH$_2$ and the previously unknown crystal structure of p-BrC$_6$H$_4$CH$_2$CO$_2$H.

In this paper, we report the procedure, invoking the Monte Carlo methodology discussed above, that we have developed to solve the crystal structure of the γ phase of 3-chloro-*trans*-cinnamic acid (ClC$_6$H$_4$CHCHCO$_2$H; abbreviated 3-ClCA; Figure 1).

Crystals of *trans*-cinnamic acid and its derivatives can be classified as α-, β-, or γ-types according to their behavior upon exposure to UV radiation.[3] UV irradiation of α-type crystals produces a centrosymmetric (α-truxillic acid) dimer and UV irradiation of β-type crystals produces a mirror-symmetric (β-truxinic acid) dimer; in contrast, no reaction occurs when γ-type crystals are exposed to UV radiation. The single-crystal X-ray diffraction studies of Schmidt[4] demonstrated well-defined correlations between crystal structure and photoreactivity in these materials, with the α-, β-, and γ-type crystals each having a characteristic mode of molecular packing. Furthermore, it has been noted that, in all photoreactive (α- and β-type) crystals, the distance between the C=C bonds of potentially reactive monomer molecules is less than ca. 4.2 Å, whereas the corresponding distance in all photostable (γ-type) crystals is greater than ca. 4.7 Å.

Two crystalline phases (β- and γ-types) of 3-ClCA are known:[4] the β phase undergoes a [2+2] photodimerization reaction upon UV irradiation, whereas the γ phase is photostable under UV irradiation. The crystal structure of the β phase of 3-ClCA is known.[5] Although an early paper[4] reported lattice parameters and a space group assignment for the γ-phase, the actual crystal structure was not determined previously.

2. Experimental Section

The γ-phase of 3-ClCA was prepared by crystallisation from ethanol, with the solution cooled from 65 to 15 °C over a period of 8 days and left at 15 °C for 5 days before collecting the crystals. The sample remained unchanged (as assessed by powder X-ray diffraction) upon UV irradiation, consistent with the assignment of the sample as a photostable γ-phase. Note that it was originally reported[6] that the β-phase of 3-ClCA is obtained by crystallization from ethanol and the γ-phase of 3-ClCA is obtained by crystallization from acetic acid. However, our work suggests that the β-phase is obtained by crystallization from acetic acid and the γ-phase is obtained by crystallization from ethanol. This is in agreement with other work[5] which reported preparation of the β phase by crystallization from acetic acid.

The powder X-ray diffractogram of the γ-phase of 3-ClCA (for a ground sample in a capillary tube) was recorded at ambient temperature on a Siemens D5000 diffractometer (operating in transmission mode using Ge-monochromatized Cu Kα_1 radiation). The data were recorded for 2θ in the range 4−70° in steps $\Delta(2\theta) = 0.02°$, and the total data collection time was 22 h.

3. Structure Solution and Refinement

The program ITO[7] was used to index the first 20 lines in the powder X-ray diffractogram, giving a monoclinic unit cell [lattice parameters (refined values from the Rietveld refinement calculation): $a = 12.400(1)$ Å, $b = 4.9560(4)$ Å, $c = 13.943(1)$ Å, $\beta = 94.265(3)°$]. Systematic absences were compatible with space group $P2_1/a$, requiring one 3-ClCA molecule in the asymmetric unit. The lattice parameters and space group assignment are in agreement with those reported previously.[4]

In the Monte Carlo structure solution calculation, it is desirable to limit the number of structural variables (degrees of freedom) considered, as this reduces consid-

(3) Schmidt, G. M. J. *Pure Appl. Chem.* **1971**, *27*, 647.
(4) Schmidt, G. M. J. *J. Chem. Soc.* **1964**, 2014.
(5) Kanao, S.; Kashino, S.; Haisa, M. *Acta Crystallogr., Sect. C* **1990**, *46*, 2436.
(6) Cohen, M. D.; Schmidt, G. M. J.; Sonntag, F. I. *J. Chem. Soc.* **1964**, 2000.
(7) Visser, J. W. *J. Appl. Cryst.* **1969**, *2*, 89.

erably the amount of computer time required (see also the discussion in section 4). For this reason, the following structurally plausible constraints were imposed in our structure solution calculations on 3-ClCA.

(i) The 3-ClCA molecule was assumed to be planar, as observed in the crystal structures of other derivatives of *trans*-cinnamic acid, including the β phase of 3-ClCA.[5]

(ii) The structure was assumed to be composed of centrosymmetric pairs of 3-ClCA molecules, hydrogen bonded to each other through the carboxylic acid groups [$R_2^2(8)$ hydrogen bond motifs[8–10]], as commonly observed in the crystal structures of carboxylic acids. The lengths of the two C−O bonds within the carboxylic acid group were taken to be equal (1.26 Å) in the structure solution calculation.

(iii) The chlorine atom was not included in the structure solution calculation because it is not known a priori whether it is in the meta position that is syn or anti (see Figure 1) with respect to the trans olefinic group and the carboxylic acid group. This represents an interesting case of crystal structure solution in which the heaviest atom in the structure (i.e., the strongest X-ray scatterer) has actually been omitted from the calculation [it should be noted that the structural fragment used in the Monte Carlo calculation does, however, represent 74% of the total scattering power of the molecule]. The fact (vide infra) that the crystal structure has been solved successfully, despite omitting the strongest X-ray scatterer from the calculation, suggests that, *when necessary* (see section 4), this approach may be used more widely in applying the Monte Carlo method. Note that, alternatively, the chlorine atom could have been included in the structural fragment, but this would have required rotation about the C(olefin)−C(phenyl) bond to be considered in the Monte Carlo calculation to allow both syn and anti conformations to be sampled. However, as this introduces an additional degree of freedom into the Monte Carlo calculation, it would require substantially more computer time than the approach (omitting the chlorine atom from the structural fragment) reported here.

Thus, the structural fragment used in the Monte Carlo calculation comprised a rigid *trans*-cinnamic acid molecule (with no hydrogen atoms) with its oxygen atoms at a fixed distance from the crystallographic center of symmetry, and with the center of symmetry lying in the molecular plane. The structural fragment was rotated by a random angular displacement (maximum allowed angular displacement = ±9°) around a random axis constrained to pass through the center of symmetry. The value of the parameter S (defined in ref 1) was fixed at 3 throughout the Monte Carlo calculation. The calculation was performed for 1000 Monte Carlo moves, and the best structure solution corresponded to $R_{wp} = 47.8\%$. Note that the structures with the next three lowest values of R_{wp} (all below 49%) all represented essentially the same position of the structural fragment. R_{wp} was around 53% for the majority of the structures (corresponding to "wrong" positions of the structural fragment) sampled in the Monte Carlo calculation.
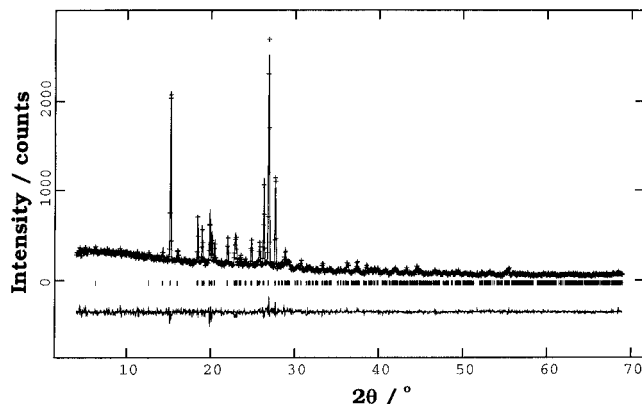

Figure 2. Experimental (+), calculated (solid line), and difference (bottom) powder diffraction profiles for the Rietveld refinement of the γ-phase of 3-ClCA. Reflection positions are marked. The calculated powder diffraction profile is for the final refined crystal structure, details of which are given in Table 1.
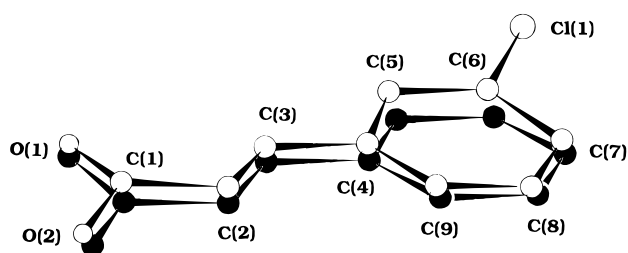

Figure 3. Comparison of the position of the structural fragment obtained in the best structure solution from the Monte Carlo calculation, overlayed on the position of the 3-ClCA molecule in the final refined crystal structure.

Rietveld refinement calculations were then carried out using the GSAS program package,[11] taking the structure solution corresponding to $R_{wp} = 47.8\%$ as the starting structural model. Difference Fourier synthesis yielded an unambiguous position for the chlorine atom in the meta position that is anti with respect to the trans olefinic group and the carboxylic acid group. Indeed, it is clear from visual examination of the structure that, on steric considerations, the chlorine atom could not be accommodated in the meta position that is syn with respect to the trans olefinic group and the carboxylic acid group. The chlorine atom was introduced in this position for the subsequent Rietveld refinement calculations. Soft constraints were applied to the geometry of the molecule, and hydrogen atoms were added to the structure in positions consistent with standard geometry (although the positions of the hydrogen atoms were not refined). For the non-hydrogen atoms, isotropic atomic displacement parameters were fixed at physically-plausible values. For the final refined structure, $R_{wp} = 8.4\%$ and $R_p = 7.2\%$. The experimental diffractogram, the diffractogram calculated for the fully refined structure, and the difference between these diffractograms are shown in Figure 2. Figure 3 shows the position of the molecule in the final refined structure overlaid on the position of the structural fragment determined from the Monte Carlo structure solution calculation. The largest shift in the position of any atom between the structure solution and the final refined structure is 0.37 Å for C(6).

(8) Etter, M. C. *Acc. Chem. Res.* **1990**, *23*, 120.
(9) Etter, M. C.; MacDonald, J. C.; Bernstein, J. *Acta Crystallogr., Sect. B* **1990**, *46*, 256.
(10) Bernstein, J.; Davis, R. E.; Shimoni, L.; Chang, N. L. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 1555.

(11) Larson, A. C.; Von Dreele, R. B. Los Alamos Laboratory Report No. LA-UR-86-748, 1987.

Table 1. Final Refined Fractional Coordinates and Isotropic Atomic Displacement Parameters from the Rietveld Refinement Calculations for the γ-Phase of 3-ClCA

atom	x/a	y/b	z/c	$U_{iso}/Å^2$
Cl	0.3866(6)	−0.188(2)	−0.4606(7)	0.090
O(1)	0.507(1)	−1.263(5)	−0.089(2)	0.050
O(2)	0.385(2)	−1.348(4)	0.005(1)	0.050
C(1)	0.424(3)	−1.199(7)	−0.060(3)	0.025
C(2)	0.353(3)	−0.984(5)	−0.106(3)	0.025
C(3)	0.386(2)	−0.862(5)	−0.182(2)	0.025
C(4)	0.329(4)	−0.648(6)	−0.240(3)	0.025
C(5)	0.383(3)	−0.541(8)	−0.314(3)	0.025
C(6)	0.325(4)	−0.333(8)	−0.361(2)	0.025
C(7)	0.219(4)	−0.246(5)	−0.348(3)	0.025
C(8)	0.175(2)	−0.364(8)	−0.271(3)	0.025
C(9)	0.232(4)	−0.556(7)	−0.219(2)	0.025

Table 2. Bond Lengths and Bond Angles in the Crystal Structure of the γ-Phase of 3-ClCA Refined from Powder Diffraction Data
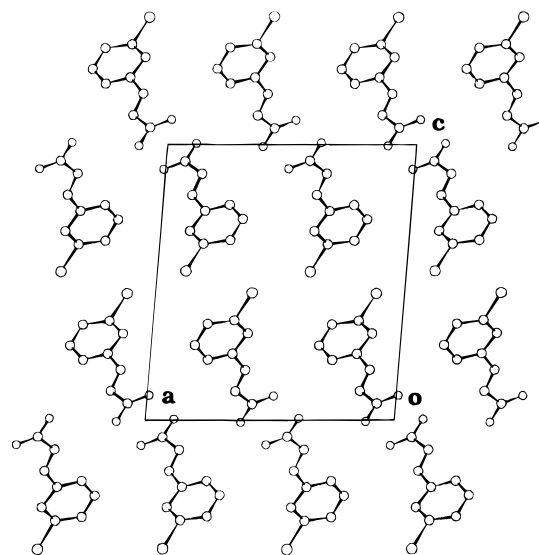
bond lengths (Å)		bond angles (deg)	
Cl−C(6)	1.79(3)	O(1)−C(1)−O(2)	118(3)
O(1)−C(1)	1.18(4)	O(1)−C(1)−C(2)	124(4)
O(2)−C(1)	1.30(4)	O(2)−C(1)−C(2)	118(3)
C(1)−C(2)	1.50(3)	C(1)−C(2)−C(3)	118(4)
C(2)−C(3)	1.32(4)	C(2)−C(3)−C(4)	128(4)
C(3)−C(4)	1.48(3)	C(3)−C(4)−C(5)	117(5)
C(4)−C(5)	1.37(4)	C(3)−C(4)−C(9)	121(5)
C(4)−C(9)	1.34(4)	C(5)−C(4)−C(9)	122(3)
C(5)−C(6)	1.39(3)	C(4)−C(5)−C(6)	112(4)
C(6)−C(7)	1.41(4)	Cl−C(6)−C(5)	115(4)
C(7)−C(8)	1.36(4)	Cl−C(6)−C(7)	116(4)
C(8)−C(9)	1.37(3)	C(5)−C(6)−C(7)	129(3)
		C(6)−C(7)−C(8)	114(2)
		C(7)−C(8)−C(9)	120(3)
		C(4)−C(9)−C(8)	124(4)

In the crystal structure (Table 1 and 2; Figure 4), molecules of 3-ClCA form columns parallel to the *b* axis. The columns are arranged such that the hydrogen-bonded carboxylic acid dimers are sandwiched between layers of chlorine atoms. Within each column, the molecular planes are parallel, with the C=C bonds of adjacent molecules separated by 4.96 Å. This distance is too large for topochemical [2+2] dimerization to occur,[3,4] consistent with the classification of this material as a photostable γ-phase.

4. Concluding Remarks

The crystal structure solution of the γ-phase of 3-ClCA reported in this paper represents a significant success in the development of the Monte Carlo method for crystal structure solution from powder diffraction data, and the knowledge gained from this study represents an important advance toward our ultimate goal of developing optimum and efficient strategies for the application of the Monte Carlo technique. Progress in achieving this goal will emanate, to a large extent, through experience gained in attempting to apply the technique to a wide variety of different types of structural problem.

It is important to note that the optimum strategy for applying the Monte Carlo method in crystal structure solution from powder diffraction data will depend on the specific structural problem under investigation. The efficiency with which the correct structure is obtained depends critically on the optimum definition of the structural fragment, and the optimum choice of how to handle the structural fragment during the calculation;

Figure 4. Final refined crystal structure of the γ-phase of 3-ClCA (hydrogen atoms not shown) viewed along the *b* axis.

in both aspects, the optimum approach may vary considerably from one type of structural problem to another. In defining the optimum choice of structural fragment, there are two major considerations: (a) the structural fragment should represent a sufficiently large fraction of the scattering matter in the unit cell, such that the correct position of the structural fragment gives rise to a significantly better agreement between the calculated and experimental powder diffractograms than wrong positions of the structural fragment; (b) the number of degrees of freedom in the structural fragment should be as low as possible in order to ensure rapid and efficient propagation of the Monte Carlo procedure.

As an illustration of the compromise that must be reached between (a) and (b), the number of degrees of freedom required in the Monte Carlo structure solution calculation reported here was reduced by omitting the strongest X-ray scatterer (the chlorine atom) from the structural fragment (see section 3). As a consequence, the structure solution calculation involved the location of a partial structural model containing *only* carbon and oxygen atoms, despite the fact that the diffraction data contain a significant contribution (ca. 18%) from scattering by the chlorine atom. Thus, the Monte Carlo approach can be successful even when strong X-ray scatterers present in the structure are omitted from the structural fragment used in the calculation, provided the structural fragment used constitutes a sufficiently high proportion of the total scattering matter in the asymmetric unit.

It is also important to recognize that, *when justified*, structurally plausible constraints should be imposed on the structural fragment during the Monte Carlo structure solution calculation. As an example, the work reported here made use of the knowledge that pairs of carboxylic acid molecules tend to be related across crystallographic centers of symmetry. As a consequence of imposing this constraint in the Monte Carlo structure solution calculation for 3-ClCA, it was necessary to consider only rotation of the structural fragent, rather than both translation and rotation; the number of degrees of freedom defining the system would have been greater by a factor of 2 if both translation and rotation had been considered. Reducing the number of degrees

of freedom in this way substantially decreases the computer time required to explore, under the Monte Carlo algorithm, the parameter space available to the structural fragment.

In addition to highlighting specific progress in the development and application of the Monte Carlo tech- nique, the work reported here also conveys, in more general terms, some of the current scope and possibili- ties for enabling molecular crystal structures to be solved from powder diffraction data.