Z. Kristallogr. Suppl. **26** (2007) 45-51

Aspects of validation in the structure determination of organic materials from powder X-ray diffraction data

Kenneth D.M. Harris*

School of Chemistry, Cardiff University, Park Place, Cardiff CF10 3AT, Wales
* E-mail: HarrisKDM@cardiff.ac.uk

Keywords: powder X-ray diffraction, validation, molecular solids, solid state NMR spectroscopy, neutron diffraction

Abstract. Techniques for structure determination of organic materials from powder X-ray diffraction data are now applied widely in many different fields of scientific research. Nevertheless, structure determination from powder diffraction data is far from an automatic, "black-box" process, and it is essential that the results obtained from such analysis are subjected to adequate scrutiny before they are assigned to be valid and correct. Here we consider two aspects of validation relevant to this field: (i) validation of the structural model used in direct-space structure solution calculations, and (ii) validation of the final structure obtained from Rietveld refinement. The advantages of exploiting information obtained from other experimental and computational techniques are emphasized.

Introduction

Structure determination of organic molecular materials from powder X-ray diffraction data is nowadays carried out extensively by researchers in both academic and industrial sectors [1-7]. For structure determination of organic materials, traditional structure solution strategies have been used successfully in many cases, although most recent work in this field has employed direct-space techniques for structure solution. Indeed, the recent upsurge of activity in this field has coincided with the development of the direct-space strategy for structure solution, for which a range of different computational implementations are now available. Although such techniques are now readily accessible and fairly straightforward to use, it is essential that the structural results are subjected to adequate scrutiny before they can be assigned as incontrovertibly correct. In this regard, there are two aspects of validation to consider: (i) validating the use of an appropriate structural model in the direct-space structure solution calculation, and (ii) validating the final structure obtained from the Rietveld refinement calculation. Here we consider both of these aspects of validation, focusing specifically on structure determination of organic molecular materials and taking illustrative examples from our own research in this field. Particular emphasis is given to the use of information obtained from other experimental and computational techniques as an important component of the validation process.

In the direct-space strategy [8], trial structures are generated independently of the experimental powder X-ray diffraction pattern, and the quality of each trial structure is assessed by direct comparison between the powder diffraction pattern calculated for the trial structure and the experimental powder diffraction pattern. This comparison is quantified using an appropriate figure-of-merit (in our work, the weighted powder profile R-factor R_{wp} is used). The direct-space strategy handles structure solution as a global optimization problem, with the aim of finding the trial structure that corresponds to optimal agreement (lowest R_{wp}) between calculated and experimental powder diffraction patterns, and is equivalent to exploring a hypersurface $R_{wp}(\mathbf{\Gamma})$ to find the global minimum, where $\mathbf{\Gamma}$ represents the set of variables that define the structure. In principle, any technique for global optimization may be employed, and most work in this field has used either Monte Carlo/Simulated Annealing or Genetic Algorithm techniques. Our current work in this field is focused on the development and implementation of the Genetic Algorithm technique [9,10]. Conventionally, the structural variables in the set $\mathbf{\Gamma}$ comprise, for each molecule in the asymmetric unit, the position $\{x, y, z\}$ and orientation $\{\theta, \varphi, \psi\}$ of the whole molecule, and a set of n variable torsion angles $\{\tau_1, \tau_2, ..., \tau_n\}$ to define the molecular conformation. An important feature of the direct-space strategy is that it incorporates reliable prior knowledge of molecular geometry (i.e. bond lengths, bond angles, geometries of aromatic rings, etc) directly within the structure solution calculation.

Aspects of validation prior to direct-space structure solution are focused on: (i) establishing the correct representation of molecular geometry to be used in the direct-space structure solution calculation, and (ii) establishing independent evidence for the correct number of molecules in the asymmetric unit. Aspects of validation after Rietveld refinement are focused on: (i) establishing whether the quality of agreement between experimental and calculated powder X-ray diffraction patterns is sufficiently good to give confidence that the refined structure is correct, (ii) assessing whether the refined structure is chemically and structurally sensible, (iii) assessing whether there is evidence for disorder in the structure, and (iv) assessing whether the use of powder X-ray diffraction data alone provides an adequate description of the structure or whether complementary techniques are required to resolve specific questions or ambiguities.

As highlighted below, other experimental and computational techniques may be exploited to provide validation on specific structural aspects, including solid state NMR spectroscopy, energy calculations (either on individual molecules or crystal structures), vibrational spectroscopies, and techniques of thermal analysis (e.g. DSC and TGA).

Validation before direct-space structure solution

First we consider the assignment of the structural model prior to direct-space structure solution. In general, the identity of the molecule(s) present in the crystalline phase and the stoichiometry (e.g. in the case of a solvate or co-crystal phase) may be established readily from the application of standard analytical techniques for determining the chemical composition of materials, such as solution state NMR (upon dissolution of the crystalline material), solid state NMR and vibrational spectroscopies, etc. Another issue concerns details of the molecular geometry in the material of interest, recognizing that some molecules may be able to adopt different tautomeric forms. The structure determination of red fluorescein [11] from powder X-ray diffraction data represented an early example of the use of solid state NMR spectroscopy in conjunction with direct-space structure solution, based on the fact that high-

resolution solid state ^{13}C NMR readily distinguishes the different tautomeric forms of this molecule. Clearly the use of a reliable representation of molecular geometry in direct-space structure solution calculations may be crucial for obtaining a successful structure solution.
Second, we consider assignment of the number of molecules in the asymmetric unit. Following unit cell determination, the number of molecules in the unit cell can generally be deduced straightforwardly from density considerations, but, in general, such information does not lead to a unique assignment of the number of molecules in the asymmetric unit and/or the space group. In such cases, solid state NMR spectroscopy can often provide valuable independent insights regarding the number of molecules in the asymmetric unit, based on the fact that the high-resolution solid state ^{13}C NMR spectrum of an organic material should contain one peak for each crystallographically distinguishable carbon atom in the structure (although we note that, in practice, the actual number of *observed* peaks may be less than this number due to accidental peak overlap). Thus, following the assignment of each peak in the ^{13}C NMR spectrum to a specific carbon environment in the molecule, it is generally straightforward to establish whether there are one, two, or more molecules in the asymmetric unit, or only a fraction of the molecule (indicating that the molecule resides on a special position).

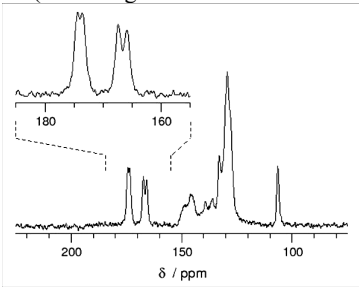

Figure 1. High-resolution solid state ^{13}C NMR spectrum of the 1:1 co-crystal containing benzoic acid and pentafluorobenzoic acid. The two peaks at ca. 174 ppm represent the carboxylic acid group of PFBA; the two peaks at ca. 167 ppm represent the carboxylic acid group of BA.

As an example [12], the high-resolution solid state ^{13}C NMR spectrum (Figure 1) of the 1:1 co-crystal formed between benzoic acid (BA) and pentafluorobenzoic acid (PFBA) is found to contain two peaks for the carbon atom of the carboxylic acid group of BA and two peaks for the carbon atom of the carboxylic acid group of PFBA, leading to the conclusion that there are two molecules of BA and two molecules of PFBA in the asymmetric unit. Systematic absences in the powder X-ray diffraction pattern indicate that the structure is C-centred and has a c-glide plane, and density considerations suggest that there are eight molecules of each type in the unit cell. The solid state NMR data therefore point towards Cc as the correct space group (with two molecules of BA and two molecules of PFBA in the asymmetric unit), rather than C2/c (with one molecule of BA and one molecule of PFBA in the asymmetric unit). An example in which high-resolution solid state ^{13}C NMR provides support for a molecule residing on a special position (in this case a 2-fold axis, with the asymmetric unit comprising half the molecule) has been encountered in the structure determination of an early-generation dendrimeric material [13].

Aspects of validation after structure refinement

To assess the quality of the final structure obtained in the structure determination process, it is important to give careful consideration to the difference profile obtained in the Rietveld refinement calculation (which represents the difference between the experimental and calculated powder X-ray diffraction patterns). Clearly, the difference profile should not contain any significant discrepancies, and any minor discrepancies that do exist (i.e. discrepancies that are higher than the noise level in the experimental data) must be properly understood before the Rietveld fit can be regarded as acceptable. An important issue in this regard is to compare the difference profile obtained in the Rietveld refinement with the difference profile obtained for the same experimental powder X-ray diffraction pattern in the pattern-decomposition/profile-fitting stage of the structure determination process (the two most commonly used techniques for which are those developed by Pawley [14] and Le Bail [15]). Importantly, the pattern-decomposition/profile-fitting procedure establishes an upper limit to the quality of fit that could be obtained in a Rietveld refinement calculation for the same experimental powder X-ray diffraction pattern (and for the same 2θ range). Thus, the aim of the Rietveld refinement should be to strive to obtain a quality of fit (assessed from the difference profile) that is as close as possible to that obtained in the pattern-decomposition/profile-fitting calculation. If the fit obtained in the Rietveld refinement is significantly worse, it is probably an indication that the refined structure is incorrect, or at least that some aspect of the true structure is not adequately described in the structural model.

An example of the importance of applying rigorous scrutiny before accepting the results from a Rietveld refinement calculation concerns the structure determination of 3,5-bis(3,4,5-trimethoxybenzyloxy)benzyl alcohol (BTBA) from powder X-ray diffraction data [16]. Following structure solution by the direct-space Genetic Algorithm technique, Rietveld refinement gave the fit shown in Figure 2a. For this Rietveld refinement (R_{wp} = 8.25 %), the difference profile might at first glance appear to be relatively flat, and could therefore be quite readily misinterpreted as representing a correct, fully refined crystal structure. However, the discrepancies between experimental and calculated powder X-ray diffraction patterns in the difference plot are greater than those in the Le Bail fit (Figure 2b; R_{wp} = 2.99 %), suggesting that the refined structure, while probably substantially correct, is still not acceptable. Furthermore, it was found that the structure did not contain any O–H···O hydrogen bonding, which was somewhat surprising (although not impossible) for a molecule that has a single hydrogen bond donor (OH group) and several potential hydrogen bond acceptors (oxygen atoms of the $COCH_3$, $COCH_2$ and OH groups). Difference Fourier analysis suggested some "missing" electron density in the structure, and following further analysis (primarily by liquid state ^{1}H NMR and TGA) it was discovered that the material was actually a monohydrate of BTBA. Further Rietveld refinement was then carried out following addition of the water molecule to the structural model, leading to a significantly improved quality of fit (Figure 2c; R_{wp} = 4.33 %) which was considered to be acceptably close to the quality of fit obtained in the Le Bail fitting procedure (Figure 2b). Furthermore, the position of the water molecule in the crystal structure (Figure 2d) was found to correspond to a structurally sensible hydrogen bonded array (Figure 2e) involving the water molecules and the OH groups of the BTBA molecules.

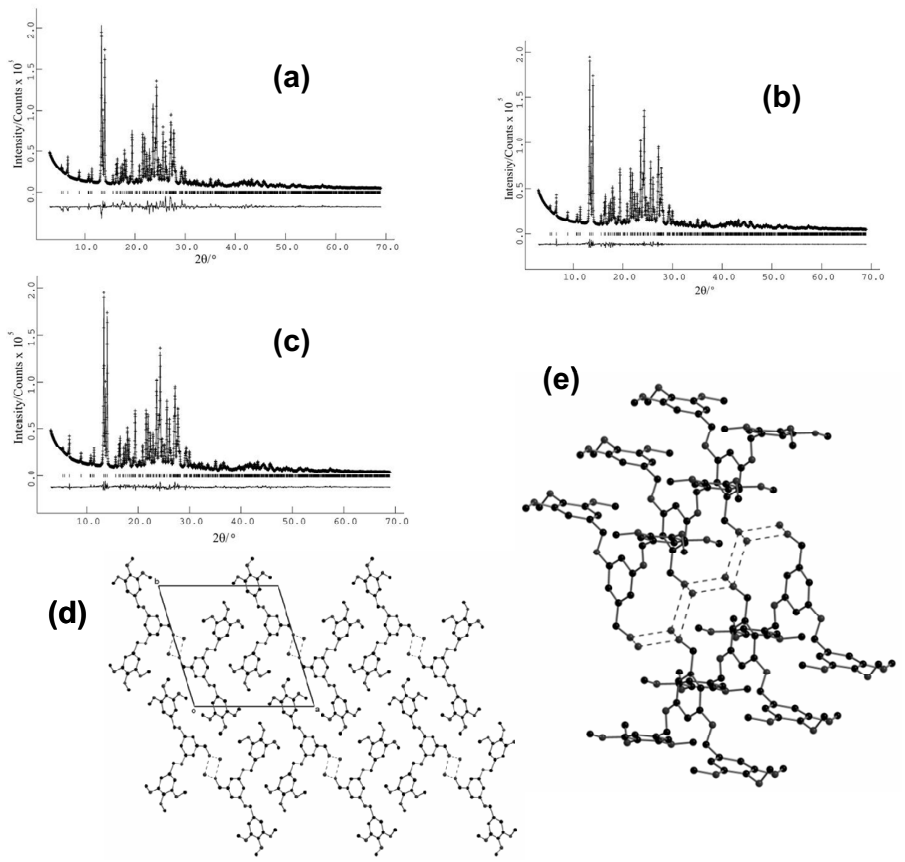

Figure 2. (a) Rietveld fit for the structural model comprising BTBA only. (b) Fit obtained in the Le Bail fitting procedure. (c) Rietveld fit for the structural model comprising BTBA and water. (d) Final refined crystal structure of BTBA monohydrate (dotted lines indicate O–H···O interactions in the extended hydrogen bonded array, which runs into the page). (e) Crystal structure of BTBA monohydrate viewed in a direction that shows the extended hydrogen bonded array (dotted lines indicate O–H···O interactions). Hydrogen atoms are omitted for clarity.

Next, we consider the use of other sources of information to provide additional support for specific structural features. As an example [17], in the structure determination of 1-formyl-2,4,6-trimethoxybenzene and 1-acetyl-2,4,6-trimethoxybenzene from powder X-ray diffraction data, DFT calculations were used to provide independent evidence to confirm that the structures determined correspond to energetically accessible conformations of the molecules. Another example [17] addressed the question of whether an OH group is involved in O–H···O hydrogen bonding by employing solid state ^{2}H NMR spectroscopy (on the material in which the OH group is deuterated – i.e. OD), based on the fact that the ^{2}H NMR quadru-

pole coupling constant (measured directly from the ^{2}H NMR spectrum) is sensitive to details of hydrogen bonding geometry.

Another aspect of validation concerns the detection of disorder in the crystal structure. In cases of severe disorder, it would generally be necessary to incorporate an appropriate description of the disorder in both the structure solution and refinement stages, and thus a good fit in the Rietveld refinement would be obtained only for a suitably accurate description of the disorder. However, in cases for which the disorder concerns only a small feature of the structure, it may be possible to obtain a Rietveld fit that is nearly acceptable using an ordered model. This situation was encountered in the structure determination of the β polymorph of *p*-formyl-*trans*-cinnamic acid from powder X-ray diffraction data [18]. The structure was solved and refined as an ordered structure, leading to a good quality of fit in the Rietveld refinement (R_{wp} = 3.27 %). However, the high-resolution solid state ^{13}C NMR spectrum for this material showed evidence for disorder of the formyl group (with the remainder of the structure ordered), and inferred that there are two components with relative populations of *ca*. 69 % and 31 %. In subsequent Rietveld refinement, the disorder was modelled in terms of two orientations of the formyl group (the two components differ by 180° rotation about the C–C bond that links the formyl group and the aromatic ring). Rietveld refinement of the disordered model gave rise to an improved fit (R_{wp} = 2.87 %), with refined occupancies for the two orientations of the formyl group of 59 % and 39 %, in relatively close agreement with the relative populations established from the solid state ^{13}C NMR data.

Finally, we consider a case in which structure determination from powder X-ray diffraction data alone could not elucidate all aspects of the crystal structure. In spite of the historical importance of ammonium cyanate $NH_4^+OCN^-$ (first studied by Friedrich Wöhler over 170 years ago), structure determination by single-crystal X-ray diffraction has never been reported, as this material can only be prepared as a microcrystalline powder. Structure determination from powder X-ray diffraction data established the positions of the non-hydrogen atoms [19], but could not reliably distinguish the correct orientation of the ammonium cation. In the structure, the nitrogen atom of the ammonium cation resides at the centre of a nearly "cubic" arrangement of O and N atoms (from cyanate anions), which occupy alternate corners of the "cube". Two plausible orientations of the ammonium cation may be proposed, in one case forming four N–H···O hydrogen bonds and in the other case forming four N–H···N hydrogen bonds. Early computational studies (periodic Hartree-Fock calculations) favoured the structure with N–H···O hydrogen bonding [19], although the results were later found to be very basis-set dependent, and subsequent computational studies based on periodic DFT calculations favoured the structure with N–H···N hydrogen bonding. To resolve this issue, neutron powder diffraction studies were carried out [20] on the deuterated material $ND_4^+OCN^-$ (actually *ca*. 81 % D, 19 % H). The neutron diffraction results definitively support the structure with N–D···N hydrogen bonding, with no detectable extent of disorder between the N–D···O and N–D···N hydrogen bonding arrangements. Results from solid state ^{15}N NMR studies [20] are consistent with this assignment of the hydrogen bonding arrangement; in particular, the temperature dependence of the ^{15}N NMR chemical shift of the cyanate anion is consistent with the changes that would be expected for the case in which the cyanate nitrogen atom is involved in N–H···N hydrogen bonding.

References

1. Harris, K.D.M., Tremayne, M., Kariuki, B.M., 2001, *Angew. Chemie Int. Ed.*, **113**, 1674.
2. Chernyshev, V.V., 2001, *Russian Chem. Bull.*, **50**, 2273.
3. David, W.I.F., Shankland, K., McCusker, L.B., Baerlocher, C. (Editors), 2002, *Structure Determination from Powder Diffraction Data* (Oxford University Press/International Union of Crystallography).
4. Favre-Nicolin, V., Černý, R., 2004, *Z. Kristallogr.*, **216**, 847.
5. Tremayne, M., 2004, *Phil. Trans. Roy. Soc.*, **362**, 2691.
6. Altomare, A., Caliandro, R., Camalli, M., Cuocci, C., Giacovazzo, C., Moliterni, A.G.G., Rizzi, R., Spagna, R., Gonzalez-Platas, J., 2004, *Z. Kristallogr.*, **216**, 833.
7. Černý, R., 2006, *Croat. Chem. Acta*, **79**, 319.
8. Harris, K.D.M., Tremayne, M., Lightfoot, P., Bruce, P.G., 1994, *J. Am. Chem. Soc.*, **116**, 3543.
9. Kariuki, B.M., Serrano-González, H., Johnston, R.L., Harris, K.D.M., 1997, *Chem. Phys. Lett.*, **280,** 189.
10. Harris, K.D.M., Johnston, R.L., Kariuki, B.M., 1998, *Acta Crystallogr.*, **A54**, 632.
11. Tremayne, M., Kariuki, B.M., Harris, K.D.M., 1997, *Angew. Chemie Int. Ed. Engl.*, **36**, 770.
12. Albesa-Jové, D., Kariuki, B.M., Kitchin, S.J., Grice, L., Cheung, E.Y., Harris, K.D.M., 2004, *ChemPhysChem*, **5**, 414.
13. Pan, Z., Xu, M., Cheung, E.Y., Harris, K.D.M., Constable, E.C., Housecroft, C.E., 2006, *J. Phys. Chem. B*, **110**, 11620.
14. Pawley, G.S., 1981, *J. Appl. Crystallogr.*, **14**, 357.
15. Le Bail, A., Duroy, H., Fourquet, J.L., 1988, *Mater. Res. Bull.*, **23**, 452.
16. Pan, Z., Cheung, E.Y., Harris, K.D.M., Constable, E.C., Housecroft, C.E., 2005, *Cryst. Growth Des.*, **5**, 2084.
17. Pan, Z., Xu, M., Cheung, E.Y., Platts, J.A., Harris, K.D.M., Constable, E.C., Housecroft, C.E., 2006, *J. Solid State Chem.*, **179**, 3214.
18. Meejoo, S., Kariuki, B.M., Kitchin, S.J., Cheung, E.Y., Albesa-Jové, D., Harris, K.D.M., 2003, *Helv. Chim. Acta*, **86**, 1467.
19. Dunitz, J.D., Harris, K.D.M., Johnston, R.L., Kariuki, B.M., MacLean, E.J., Psallidas, K., Schweizer, W.B., Tykwinski, R.R., 1998, *J. Am. Chem. Soc.*, **120**, 13274.
20. MacLean, E.J., Harris, K.D.M., Kariuki, B.M., Kitchin, S.J., Tykwinski, R.R., Swainson, I.P., Dunitz, J.D., 2003, *J. Am. Chem. Soc.*, **125**, 14449.

Acknowledgements. I am grateful to EPSRC for supporting our research in the field covered by this article. The contributions of other research group members and collaborators mentioned in the references are also gratefully acknowledged.