ELSEVIER

International Journal of Pharmaceutics 139 (1996) 9–14

international journal of pharmaceutics

Unit cell determination of sotalol hydrochloride by powder X-ray diffraction

K. Shankland*, D.S. Sivia

ISIS Facility, Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, UK

Received 6 February 1996; accepted 17 April 1996

Abstract

The unit cell dimensions of sotalol hydrochloride have been determined by powder diffraction using X-ray data collected on a laboratory source. A Bayesian method for estimating the number of reflections contributing to any given peak has been used to provide the best estimates of the reflection positions for input into a powder indexing program. The method should help extend the range of applicability of laboratory powder X-ray diffraction to problems involving large, low symmetry unit cells and is equally applicable to synchrotron X-ray and neutron diffraction data. Sotalol hydrochloride has a monoclinic unit cell of volume 3165 $Å^3$ and is thus a valid representative of a large subclass of organic compounds of pharmaceutical interest.

Keywords: Powder X-ray diffraction; Sotalol hydrochloride; Indexing; Bayesian probability theory

1. Introduction

Powder diffraction, both of X-rays and of neutrons, has been used for many years to provide 'fingerprints' which can be used to identify compounds or particular polymorphic forms, and to determine the sizes and shapes of unit cells and their associated space groups (Klug and Alexander, 1973). With the advent of the Rietveld method (Rietveld, 1969) came a realisation that the information content of a powder pattern could be routinely used to allow the refinement of structural parameters such as atomic positions and site occupancies. More recently, methods have been developed which allow the extraction of structure factors from powder diffraction data, and structures can now be solved ab initio from powder data in a manner analogous to that of solving a structure from a single crystal X-ray data set (Cheetham, 1995). However, very few molecular organic crystals, which tend to crystallise in low symmetry space groups, have had their structures solved by this route. Fig. 1, which

* Corresponding author.

PII S0378-5173(96)04563-2

shows a typical powder X-ray diffraction pattern for a molecular organic crystal, illustrates why this is the case. The pattern is complex, with many overlapping peaks. The drop in scattered intensity with increasing angle, attributable to the Lorentz-polarisation factor, Debye–Waller factor and form factor fall off (Cullity, 1978), is exacerbated by a lack of strong X-ray scatterers in the compound. Retrieval of information beyond about 40° 2θ (equivalent to a minimum interplanar spacing of 2.25 Å) is problematic, making ab initio structure solution very difficult. Nevertheless, a few notable successes have been reported with both synchrotron and laboratory X-ray sources (Cernik et al., 1991; Tremayne et al., 1992).

Information about the size and the shape of the unit cell is obtained by indexing the powder diffraction pattern. The methodology for indexing has been known for many years (Klug and Alexander, 1973) but the advent of powerful computers has allowed the implementation of numerically intensive techniques which make the task routine in many cases. However, the success of the indexing process always depends upon having accurate estimates of reflection positions. For a well calibrated powder diffractometer, this is not usually a problem. Difficulties do arise though, when the measured diffraction peaks are broad, closely spaced, possess an unusual shape or suffer from poor counting statistics. Some or all of these phenomena may be present in powder diffraction data collected from organic compounds, and an awareness of the problems means that the sample can be prepared and data collected in such a manner as to minimise any adverse effect on the indexing process. However, the spacing of the reflections, which is a function of the size and shape of the unit cell, is fixed and one is often left asking the question 'how many reflections lie under this measured peak?'. A quantitative method which answers this question and returns the best estimates of the reflection positions is a powerful tool for indexing problems. We show here the application of such a method to laboratory X-ray diffraction data collected from a powder sample of the known structure sotalol hydrochloride.

2. Materials and methods

A thin layer of sotalol hydrochloride powder was lightly dusted through a 300-μm sieve onto a zero-background silicon plate. The plate was secured in a sample holder and a diffraction pattern measured using a Siemens D5000 X-ray diffractometer arranged in Bragg–Brentano geometry. The generator was set to 40 kV, 30 mA using a copper target and a monochromator to give CuKα_1 radiation of 1.5406 Å. Constant slit sizes were used with a 0.1-mm slit on the detector assembly. The sample was scanned in steps of 0.02° 2θ over a range of 9.82–22.46° 2θ, with a collection time of 15.0 s at each point. The diffraction data collected are shown in Fig. 2.

The powder pattern was first analysed using a Bayesian model selection procedure (Sivia et al., 1993). This entails a generalised form of model fitting that is appropriate when the number of parameters to be optimised is not known. In particular, it allows for both a quantitative assessment of the number of Bragg peaks N that are present in the data and an estimate of their positions and amplitudes. With suitable simplifying approximations, the analysis reduces to conventional least-squares when the value of N is known. When it is not, probability theory shows that the best estimate of N is given by a natural balance between the requirement to account for the data adequately and an intrinsic preference for simpler models; it captures the essence of Ockham's razor

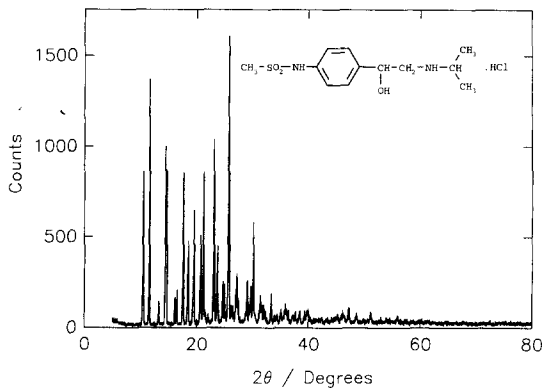

Fig. 1. Powder diffraction data for sotalol hydrochloride at λ = 1.5406 Å.

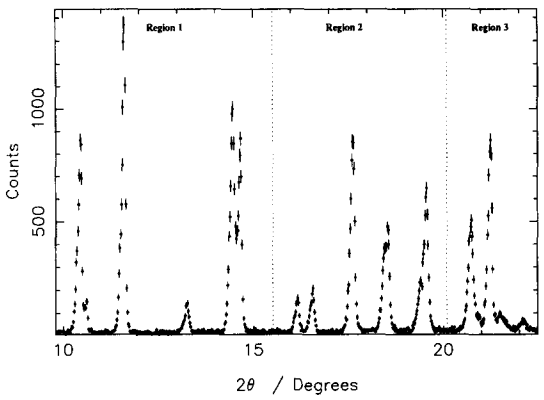

Fig. 2. Powder diffraction data used to index the unit cell for sotalol hydrochloride. The regions used in the peak fitting process are delineated by the dotted lines.

that the least number of reflections possible should be assumed. The model selection analysis is formally summarised by the conditional probability distribution function prob(N | Data, I), where '|' means 'given' and I represents all the relevant background information, such as the simplifying approximations and a knowledge of the peak shape corresponding to a single Bragg reflection. The latter was obtained from a maximum entropy fit of a smooth nonparametric line to the relatively sharp peak at $\sim 11.6°$ 2θ (with the assumption that this signal is due to a single reflection) giving a resolution function which was used to analyse subsequent areas of the pattern. Data were analysed in three adjacent regions with N being increased until a maximum in the value of prob(N | data, I) was passed.

3. Results

Fig. 3 shows the logarithm of prob(N | Data, I) for the three data regions; a characteristic 'Ockham hill' can be seen in each case. That is to say, there is (a) a rapid rise in $\log_{10}$[prob(N | Data, I)] with increasing N, as reflections required to fit the strongest features of the data are added; (b) a gently rising or almost flat region where weaker features in the data are being fitted; (c) a fall in $\log_{10}$[prob(N | Data, I)] with increasing N, as the fit to the data continues to improve (i.e. χ^2 is falling) but the model has become unnecessarily complicated and is being penalised in probabilistic terms for this additional complexity.

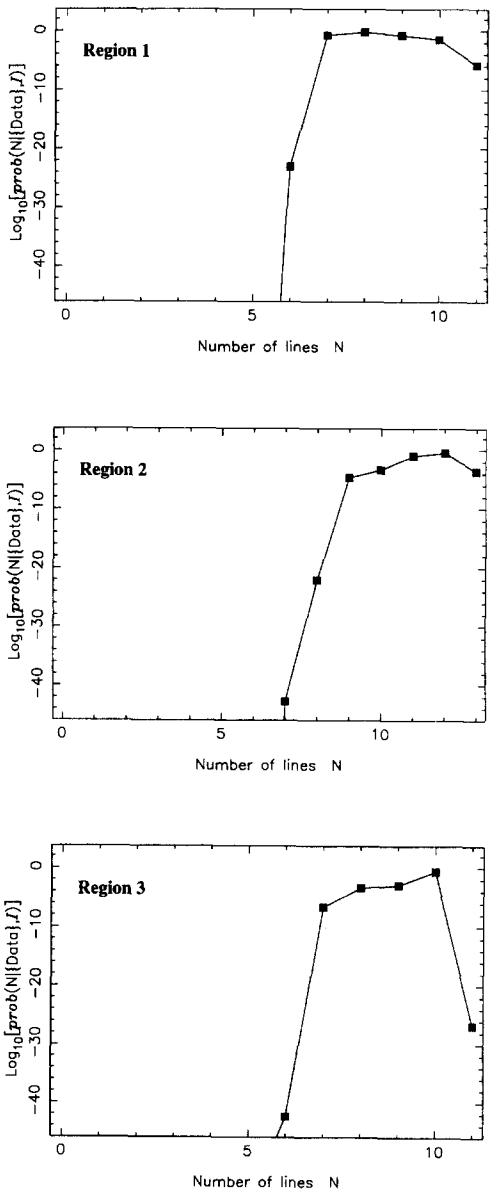

Fig. 3. Plots of the posterior probability, $\log_{10}$[prob(N | Data, I)] versus the number of reflections fitted to the diffraction data for each of the regions indicated in Fig. 2. Maxima are present at $N = 8$, $N = 12$ and $N = 10$, respectively.

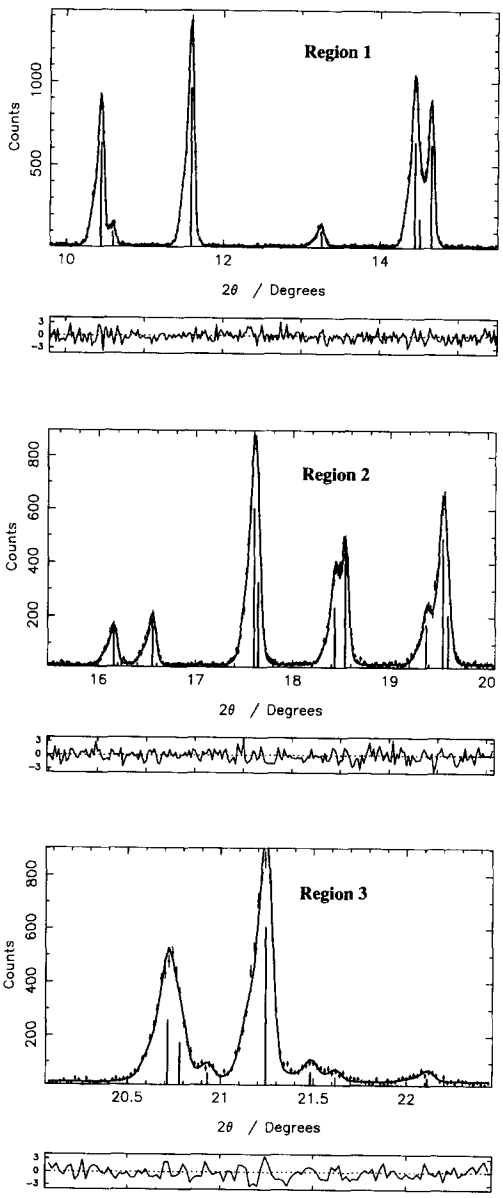

Fig. 4. Fits to the diffraction data regions indicated in Fig. 2, using the resolution function derived from the peak at $\sim 11.6°$ 2θ and reflection positions corresponding to values of $N = 7$, $N = 9$ and $N = 7$, respectively. Vertical lines indicate the reflection positions and the normalised residuals are plotted below each fit.

Conservative estimates of the 'correct' number of reflections in regions 1, 2 and 3 were taken as 7, 9 and 7, respectively. Fig. 4 shows the least-squares fits corresponding to these values of N, and Table 1 lists the reflection positions obtained from the fit. The first 20 reflection positions were used as input to the DICVOL91 indexing program (Boultif and Louër, 1991). In searching for cells of volume up to 4000 Å^3 in cubic, tetragonal, hexagonal, orthorhombic and monoclinic crystal systems, DICVOL91 reported only one solution which predicted the input reflection positions adequately, with figures of merit $M(20) = 7.6$ and $F(20) = 24.1$. This solution corresponds to a monoclinic system with $a = 15.366$ Å, $b = 13.343$ Å, $c = 15.275$ Å, $\beta = 91.37°$ and $V = 3131$ Å^3. This is close to the unit cell of the C2/c structure of sotalol hydrochloride reported from a single crystal experiment (Gadret et al., 1976), which has dimensions $a = 15.350$ Å, $b = 13.480$ Å, $c = 15.300$ Å, $\beta = 91.45°$ and $V = 3165$ Å^3.

Table 2 gives a comparison of the input reflection positions in regions 1 and 2 with the calculated positions based on the unit cell given by DICVOL91. Sixteen of the 19 reflections present in the region 10–20° 2θ were recovered from the data.

4. Discussion

Obtaining a unit cell for sotalol hydrochloride using powder X-ray diffraction on a laboratory source is not straightforward for several reasons. The 152 possible reflections in the first 30° 2θ for

Table 1
Reflection positions for sotalol hydrochloride, obtained from least-squares fits to regions 1, 2 and 3 of the diffraction data using $N = 7$, $N = 9$ and $N = 7$, respectively

Region	Reflection positions (° 2θ)		
1	10.444 ± 0.001	10.598 ± 0.002	11.593 ± 0.001
	13.255 ± 0.003	14.445 ± 0.002	14.509 ± 0.007
	14.654 ± 0.001		
2	16.156 ± 0.002	16.556 ± 0.002	17.600 ± 0.001
	17.645 ± 0.002	18.432 ± 0.003	18.539 ± 0.001
	19.372 ± 0.003	19.544 ± 0.003	19.599 ± 0.006
3	20.714 ± 0.002	20.779 ± 0.003	20.928 ± 0.009
	21.241 ± 0.001	21.483 ± 0.010	21.617 ± 0.017
	22.112 ± 0.009		

Table 2
A comparison of the input reflection positions from regions 1 and 2 of the diffraction data with the true positions calculated from the cell suggested by DICVOL91

Measured lines	Calculated reflection positions	
Position (° 2θ)	Reflection	Position (° 2θ)
10.444	−1 1 1	10.435
10.598	1 1 1	10.588
Not found	2 0 0	11.511
11.593	0 0 2	11.581
13.255	0 2 0	13.260
14.445	−1 1 2	14.434
14.509	0 2 1	14.476
14.654	1 1 2	14.655
16.156	−2 0 2	16.158
16.556	2 0 2	16.552
17.600	−2 2 0	17.594
17.645	0 2 2	17.639
18.432	−2 2 1	18.446
18.539	−3 1 0	18.545
Not found	2 2 1	18.621
Not found	−3 1 1	19.315
19.372	−1 1 3	19.398
19.544	3 1 1	19.565
19.599	1 1 3	19.648

a cell of this size in space group P2/m are reduced to only 64 by the systematic absences of C2/c, but useful reflections for the purposes of indexing are lost in the process, particularly the (1 0 0), (0 1 0) and (0 0 1) reflections. In addition, the similarity of the a and c axes and the proximity of β to 90° means that pairs of reflections such as the (2 0 0) and (0 0 2) appear closely spaced in the diffraction pattern. It is also apparent from Fig. 4 that the peak shape is somewhat asymmetric, possibly due to slight precession in the sample. Nevertheless, reflection positions were recovered with sufficient accuracy by the Bayesian model fitting procedure for the unit cell to be successfully indexed.

Knowing the full structure from the single crystal study, one can assess how well the analysis performed. Only three reflections up to 20° 2θ were not located. Of these, the (2 2 1) is very weak, the weak (−3 1 1) is part of a quartet of overlapping reflections around 19.5° 2θ, and the weak (2 0 0) lies close to the strong (0 0 2). It is now apparent that the resolution function used to fit the diffraction data was based on the (2 0 0)/(0 0 2) doublet at ~11.6° 2θ. Whilst this means that it is not a correct description of a single reflection as we had hoped, the overall analysis still succeeds because the weak (2 0 0) makes almost no difference to the shape of the very strong (0 0 2). It is important to remember that it is not possible to know a priori which peak, if any, represents an isolated single reflection. Only an assessment of the peak width relative to other peaks in the pattern, and the increased likelihood of finding an isolated reflection at low values of 2θ, can be used as a guide in selecting a strong peak with which to generate the resolution function. Viewed in this context, the peak at ~11.6° 2θ is the only suitable choice in this particular diffraction data set.

In estimating the number of reflections present in each region of the data, we did not choose the most probable values of N. It is clear from Fig. 3 that each 'probability hill' has a point beyond which the relative gains in $\log_{10}[\text{prob}(N \mid \text{Data}, I)]$ are much less for each increment of N. Up to these points at $N = 7$, $N = 9$ and $N = 7$, respectively, it is the major features of the diffraction profile that are being fitted, and so it is intuitively sensible to start the indexing process using positions derived from these fits. Beyond these points, it is the weaker features in which we have less confidence that are being fitted. Had it not been possible to find a unit cell using the first set of positions chosen, positions derived from more probable values of N could then have been used.

There is no doubt that the higher instrumental resolution and intensity available at a good synchrotron X-ray source would have made indexing the powder diffraction pattern of sotalol hydrochloride even easier, via better counting statistics, splitting of overlapping peaks and detection of weak diffraction features. These in turn would have allowed us to define a more appropriate resolution function and hopefully obtain sharper maxima in the $\log_{10}[\text{prob}(N \mid \text{Data}, I)]$ versus N plots. Nevertheless, we have shown here that a laboratory diffractometer can be used to index a complex diffraction pattern derived from a large volume unit cell in a low symmetry space group.

The problem of handling overlapping reflections in the diffraction data is eased by the use of the Bayesian model selection procedure, but the general rule of 'better data, better results' still holds. The diffraction data presented here were collected in a short time with a sample taken straight 'from the bottle', in order to assess the performance of the model selection procedure with the sort of data routinely collected for simple polymorph identification purposes. Experience has shown us that whether the diffraction data is collected at a laboratory X-ray, synchrotron X-ray or neutron source, time spent on (a) ensuring that the measured diffraction peaks are as narrow as possible, by recrystallisation of the sample if necessary and (b) collecting the data long enough to ensure good counting statistics is usually repaid many times over at the indexing, structure solution and structure refinement stages.

References

Boultif, A. and Louër, D., Indexing of powder diffraction patterns for low-symmetry lattices by the successive dichotomy method. *J. Appl. Cryst.*, 24 (1991) 987–993.

Cernik, R.J, Cheetham, A.K., Prout, C.K., Watkin, D.J., Wilkinson, A.P. and Willis, B.T.M., The structure of cimetidine ($C_{10}H_{16}N_6S$) solved from synchrotron-radiation X-ray powder diffraction data. *J. Appl. Cryst.*, 24 (1991) 222–226.

Cheetham, A.K., Ab initio structure solution. In Young, R.A. (Ed.), *The Rietveld Method*, Oxford University Press, Oxford, 1995, pp. 276–292.

Cullity, B.D., *Elements of X-Ray Diffraction*, 2nd edn., Addison-Wesley, Reading, MA, 1978.

Gadret, P.M., Goursolle, M., Leger, J.M., Colleter, J.C. and Carpy, A., Structure cristalline du chlorhydrate de sotalol, chlorhydrate de p-(hydroxy-1 isopropylamine-2 éthyl méthanesulfonanilide. *Acta Cryst.*, B32 (1976) 2757–2761.

Klug, H.P. and Alexander, L.E., *X-Ray Diffraction Procedures for Polycrystalline and Amorphous Materials*, 2nd edn., Wiley Interscience, New York, 1973.

Rietveld, H.M., A profile refinement method for nuclear and magnetic structures. *J. Appl. Cryst.*, 2 (1969) 65–71.

Sivia, D.S., David, W.I.F, Knight, K.S. and Gull, S.F., An introduction to Bayesian model selection. *Physica D*, 66 (1993) 234–242.

Tremayne, M., Lightfoot, P., Mehta, M.A., Bruce, P.G., Harris, K.D.M., Shankland, K., Gilmore, C.J. and Bricogne, G., An ab initio determination of $LiCF_3SO_3$ from X-ray powder diffraction data using entropy maximisation and likelihood ranking. *J. Solid State Chem.*, 100 (1992) 191–196.