

and $Q_{\alpha\beta}$ can be written as

$$\begin{aligned}
 Q_{xx} &= D_x - (1/\pi^2) \sum_{h=1}^{\infty} [A_i(h, 0, 0)/h^2] \\
 Q_{yy} &= D_y - (1/\pi^2) \sum_{k=1}^{\infty} [A_i(0, k, 0)/k^2] \\
 Q_{zz} &= D_z - (1/\pi^2) \sum_{l=1}^{\infty} [A_i(0, 0, l)/l^2] \\
 Q_{xy} &= (1/2)(D_x + D_y) \\
 &\quad - (1/4\pi^2) \sum_{h,k \neq 0} [F_i(h, k, 0)/hk] \\
 Q_{yz} &= (1/2)(D_y + D_z) \\
 &\quad - (1/4\pi^2) \sum_{k,l \neq 0} [F_i(0, k, l)/kl] \\
 Q_{zx} &= (1/2)(D_z + D_x) \\
 &\quad - (1/4\pi^2) \sum_{l,h \neq 0} [F_i(h, 0, l)/hl].
 \end{aligned} \tag{48}$$

Equation (48) is the sum rule relating the microscopic information to bulk quadrupolar tensor.

Concluding remarks

A careful discussion of the Coulomb potential has shown the necessity to consider the $V(0)$ term in its Fourier expansion. This term is important when comparative studies are to be done. It can be calculated either in direct space, using a multipolar expansion

of the density, or in reciprocal space, from total structure factors. The reciprocal-space expansion leads to sum rules which must be fulfilled by the structure factors, and which connect the microscopic charge density with bulk multipolar moments that can be measured by different techniques.

This work was supported in part by the US National Science Foundation (CHE8711736), the National Institute of Health (5R01HL2388408) and a NATO Research Contract No. 587/83.

References

- AVERY, J. (1987). In *Density Matrices and Density Functionals*, edited by R. ERDAHL & V. H. SMITH JR, p. 693. Dordrecht: Reidel.
 BECKER, P. & COPPENS, P. (1991). *International Tables for Crystallography*, Vol. C. Dordrecht: Kluwer. In the press.
 BERTAUT, E. F. (1952). *J. Phys. Radium*, **13**, 499-505.
 BERTAUT, E. F. (1978). *J. Phys. Chem. Solids*, **39**, 97-102.
 COPPENS, P. (1991). *International Tables for Crystallography*, Vol. B. Dordrecht: Kluwer. In the press.
 COPPENS, P. & HALL, M. B. (1982). Editors. *Electron Distributions and the Chemical Bond*. New York: Plenum.
 DAHL, J. P. & AVERY, J. (1984). Editors. *Local Density Approximations in Quantum Chemistry and Solid State Physics*. New York: Plenum.
 ERDAHL, R. & SMITH, V. H. JR (1987). Editors. *Density Matrices and Density Functionals*. Dordrecht: Reidel.
 HANSEN, N. K. & COPPENS, P. (1978). *Acta Cryst.* **A34**, 909-921.
 JACKSON, J. D. (1975). *Classical Electrodynamics*. New York: John Wiley & Sons.
 POLITZER, P. & TRUHLAR, D. G. (1981). Editors. *Chemical Applications of Atomic and Molecular Electrostatic Potentials*. New York: Plenum.
 STEWART, R. F. (1979). *Chem. Phys. Lett.* **65**, 335-342.

Acta Cryst. (1990). **A46**, 258-262

Crystal Structure Determination from Low-Resolution X-ray Powder Diffraction Data

BY C. C. WILSON AND J. W. WADSWORTH

Neutron Science Division, Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, England

(Received 16 May 1989; accepted 2 November 1989)

Abstract

The previously known structures of the modified nucleic acid bases 3-deazauracil (4-hydroxy-2-pyridone) and 6-azauracil (2*H*,4*H*-1,2,4-triazine-3,5-dione) have been shown to be determined by model-based reciprocal-space Patterson methods from low-resolution X-ray powder diffraction data. The data sets used in structural solution consisted of 26 and 23 reflections respectively, with $|F_h|$ values extracted by a simple fitting procedure. The structural

parameters found are of limited accuracy and only severely constrained refinements (on F) are possible from these data. However, these determinations indicate the strength of model-based Patterson methods to reveal structural information even from poor data.

Introduction

The solution of crystal structures from laboratory X-ray powder data using Patterson methods has

recently been demonstrated (Rius & Miravittles, 1988). In this work, a Patterson search method was used on $|F_n|$ values derived from the powder pattern and the solution of two known structures achieved from both simulated and experimental data. In addition, the direct-methods solution of heavy-atom structures from X-ray powder data is also becoming well established (e.g. Le Bail, Ferey, Amoros, Beltran-Porter & Villeneuve, 1989). However, the use of a standard laboratory diffractometer for structural solution of light-atom structures is very much in its infancy. The small number of reflections resolvable from such organic structures, often of poor crystal quality, using routine laboratory sources suggests the use of Patterson rather than direct methods (Wilson, 1989), and, with no heavy atom present, model-based Patterson methods are the obvious recourse. The reciprocal-space Patterson-methods program *PATMET* (Wilson & Tollin, 1986) has been shown to be able to solve structures from very small data sets (Wilson, 1989) and so hopefully will be applicable to the method.

The structures chosen for examination in this work were the simple planar modified nucleic acid base materials 3-deazauracil ($C_5H_5NO_2$; Fig. 1a) and 6-azauracil ($C_5H_3N_3O_2$; Fig. 1b). These materials, whose structures are known (Low & Wilson, 1983; Singh & Hodgson, 1974), were chosen as representative of a class of possible intercalative drugs for which single-crystal growth is difficult and hence structural information is often unavailable for even very simple materials.

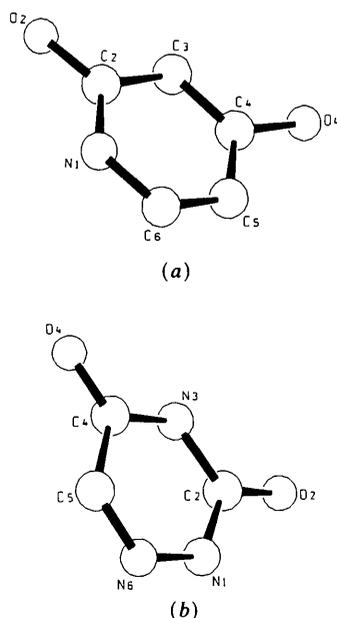


Fig. 1. (a) Perspective view of the 3-deazauracil molecule. H atoms are omitted for clarity. Drawn using *SCHAKAL88* (Keller, 1988). (b) Perspective view of the 6-azauracil molecule. H atoms are omitted for clarity. Drawn using *SCHAKAL88* (Keller, 1988).

Data collection and intensity extraction

For the determination of structures from powder data, one of the main problems is often cell and space-group determination. There are several autoindexing programs available for this, but none of these are completely reliable, especially for the low-symmetry space groups commonly adopted by small-molecule organic materials. For this reason, in the present study this information was assumed to be unknown and must initially be extracted from the X-ray powder trace before proceeding to solution. Data were collected for both materials from a Philips PW1710 powder diffractometer, a standard moderate-resolution instrument for routine work, using Cu $K\alpha$ radiation ($\lambda = 1.542 \text{ \AA}$), from samples purchased from Sigma Chemicals. The data used were in the 2θ range $15\text{--}40^\circ$, beyond which diffraction from these samples was too weak to observe. The step sizes used in the scans were 0.025° for 3-deazauracil and 0.05° for 6-azauracil, with collection times of 3 s step^{-1} . Data were transferred *via* the Philips microprocessor through an Atari ST-512 microcomputer used for data capture, to a VAX 8650 on which all further analysis was carried out.

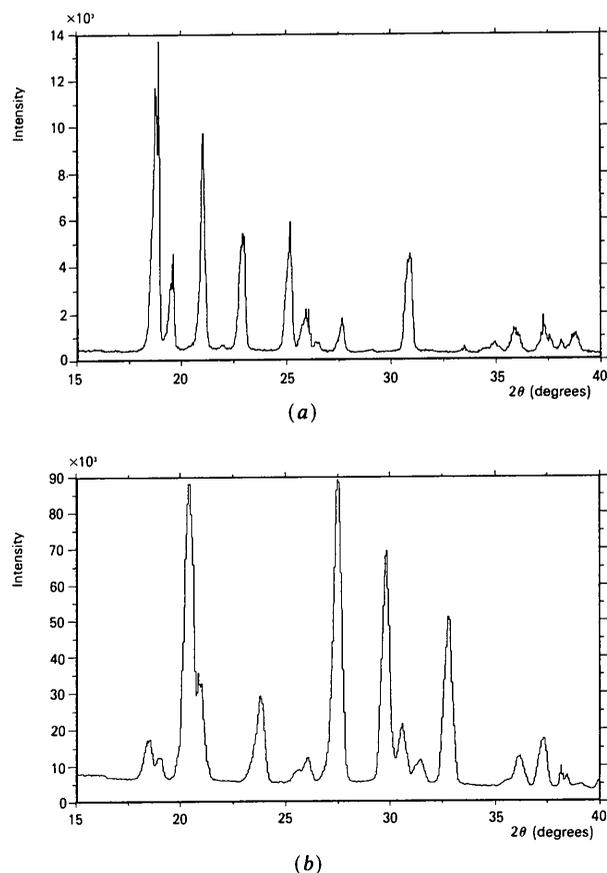


Fig. 2. (a) Powder diffraction trace in the region for which data were extracted for 3-deazauracil. (b) Powder diffraction trace in the region for which data were extracted for 6-azauracil.

The powder diffraction data collected for each material are shown in Figs. 2(a) and (b). The plotting program used was the local ISIS data-display program *GENIE* (David *et al.*, 1986). From these plots the positions of a series of peaks were extracted for use in autoindexing on each material.

Cell determination of 3-deazauracil

The positions of 19 reflections were extracted from the 3-deazauracil pattern and autoindexing attempted in the program *TREOR* (Werner, Eriksson & Westdahl, 1985). The first solution produced from *TREOR* was the orthorhombic cell $a = 11.28$, $b = 10.67$, $c = 8.67$ Å, which can be seen to compare reasonably with the published cell (Low & Wilson, 1983) $a = 8.638$, $b = 5.279$, $c = 11.220$ Å, given that a and c are switched and b is doubled in the cell found by *TREOR*. For convenience, the a - c transformation was carried out at this stage. If one tags the positions of the reflections predicted by the *TREOR* cell, it is clear that there is strong evidence for the true b dimension being half that determined by the autoindexing procedure, since all k -odd reflections appear to be absent. This cell doubling is common in such autoindexing procedures and should always be checked for. For intensity extraction in this case, therefore, the unit cell used was assumed to be $a = 8.67$, $b = 5.34$, $c = 11.28$ Å. From the pattern indexed on this cell, there is evidence for systematic absences consistent with the space group $P2_12_12_1$ ($h00$, $0k0$, $00l$ absent for h odd, k odd, l odd). This is the space group found in the single-crystal study.

Intensity extraction

With the pattern indexed as above, reflection intensities were extracted in a very simple way. For resolved reflections a Gaussian was fitted to the peak, the area of this giving I_h , whereas for overlapped reflections the total intensity under the envelope was extracted by simple background extraction and integration, with the apportioning of relative contributions estimated manually. This very primitive, and for X-ray data by no means the most correct, way of intensity extraction is in keeping with the object of demonstrating the relative robustness of Patterson methods for structure solution from poor data. There are more sophisticated procedures available (*e.g.* Pawley, 1981; Rius & Miravittles, 1988) which in general should be exploited if available. From the pattern in Fig. 2(a), it was found possible to extract the intensities of some 26 reflections. The derived $|F_h|$ values are shown in Table 1(a).

Cell determination and intensity extraction for 6-azauracil

The procedures described above for 3-deazauracil were closely followed for 6-azauracil (Fig. 2b), with

Table 1. $|F_h|$ data extracted for 3-deazauracil and 6-azauracil

(a) 3-Deazauracil

h	k	l	F	h	k	l	F
1	0	2	3.75	1	0	4	0.54
1	1	0	2.22	0	2	0	0.77
1	1	1	2.18	0	2	1	0.87
2	0	1	0.37	3	0	2	0.64
0	1	2	2.96	2	1	3	1.43
1	1	2	2.41	1	2	1	0.72
1	0	3	1.40	0	1	4	1.02
2	0	2	1.40	3	1	1	0.72
2	1	0	0.95	0	2	2	2.44
2	1	1	1.23	1	1	4	1.07
0	1	3	0.20	2	0	4	1.25
1	1	3	2.00	1	2	2	1.14
2	1	2	2.00	3	1	2	1.14

(b) 6-Azauracil

h	k	l	F	h	k	l	F
1	1	0	3.87	1	5	0	2.43
0	1	1	2.81	0	5	1	4.25
1	2	0	12.49	1	4	1	10.79
0	2	1	7.72	1	6	0	2.71
1	3	0	2.81	2	1	0	5.54
0	3	1	7.85	1	5	1	3.93
1	0	1	2.76	0	0	2	6.59
1	1	1	2.82	2	2	0	5.76
0	4	1	12.26	0	1	2	5.76
1	2	1	8.70	0	2	2	4.69
1	3	1	11.59	2	3	0	2.56
0	6	0	10.65				

similar results. Once again *TREOR* produced a cell with one dimension (a) doubled, which could be identified from the absence of all h -odd reflections in the pattern. There was again some evidence for absences corresponding to $P2_12_12_1$, which is the space group found by Singh & Hodgson (1974) in the single-crystal study. The cell used for intensity extraction was $a = 5.008$, $b = 17.60$, $c = 4.83$ Å, which compares reasonably well with the published cell $a = 5.022$, $b = 17.611$, $c = 4.875$ Å. Intensity extraction based on this cell and carried out using the simple method described above led to 23 I_h values, giving the $|F_h|$ values shown in Table 1(b).

Structural solution using Patterson methods

The fragment used as model in both structure determinations was a six-membered ring (C—C distance 1.4 Å, C—C—C angles 120°) with two exocyclic, *meta*, O atoms (C=O distances 1.2 Å, C—C—O angles 120°). Thus the model contains only the basic idealized molecular-shape information.

Structural solution of 3-deazauracil

The 26 $|F_h|$ data shown in Table 1(a) were used for structural solution in the Patterson-methods program *PATMET* (Wilson & Tollin, 1986; Wilson, 1988). The orientation of the planar ring was determined using the $I(\theta, \varphi)$ function, and its orientation within the plane indicated by the highest peak in

Table 2. *Published, PATMET solution and refined (with e.s.d.s) fractional coordinates for 3-deazauracil and 6-azauracil*

(a) 3-Deazauracil

	Published			PATMET			Refined (RBLs)		
	x	y	z	x	y	z	x	y	z
N1	0.1745	0.3429	0.2102	0.180	0.375	0.230	0.226 (16)	0.332 (17)	0.243 (9)
C2	0.1755	0.5441	0.1344	0.180	0.568	0.146	0.226 (17)	0.533 (24)	0.164 (13)
O2	0.0583	0.6865	0.1315	0.069	0.726	0.143	0.11 (2)	0.69 (3)	0.16 (2)
C3	0.3101	0.5811	0.0650	0.301	0.590	0.066	0.350 (13)	0.567 (21)	0.088 (9)
C4	0.4329	0.4147	0.0735	0.422	0.418	0.070	0.474 (16)	0.401 (20)	0.091 (11)
O4	0.5638	0.4432	0.0143	0.537	0.439	-0.006	0.592 (19)	0.434 (23)	0.019 (14)
C5	0.4239	0.2074	0.1526	0.422	0.225	0.153	0.474 (17)	0.201 (27)	0.171 (14)
C6	0.2955	0.1774	0.2184	0.301	0.204	0.233	0.350 (13)	0.166 (24)	0.247 (11)

(b) 6-Azauracil; the two atoms O2 and O4 are obviously not correctly located in the PATMET solution

	Published			PATMET			Refined (RBLs)		
	x	y	z	x	y	z	x	y	z
N1	0.4859	-0.4404	-0.0611	0.514	-0.431	-0.070	0.49 (9)	-0.433 (21)	0.00 (7)
C2	0.5640	-0.4004	-0.2814	0.606	-0.392	-0.290	0.56 (7)	-0.390 (17)	-0.22 (6)
O2	0.7558	-0.4192	-0.4240	0.550	-0.294	-0.71 (6)	0.71 (6)	-0.360 (12)	-0.44 (5)
N3	0.4127	-0.3357	-0.3276	0.464	-0.329	-0.381	0.42 (5)	-0.325 (11)	-0.28 (4)
C4	0.1949	-0.3128	-0.1789	0.230	-0.305	-0.250	0.20 (9)	-0.301 (21)	-0.12 (7)
O4	0.0635	-0.2554	-0.2294	0.648	-0.390	-0.341	0.15 (5)	-0.259 (16)	-0.33 (3)
C5	0.1343	-0.3622	0.0459	0.138	-0.344	-0.028	0.13 (7)	-0.344 (17)	0.10 (6)
N6	0.2731	-0.4227	0.0998	0.280	-0.407	0.062	0.27 (5)	-0.409 (11)	0.16 (4)

the $I(\theta, \varphi)$ ($\theta, \varphi = 50, 54^\circ$) was found using the $\sigma(\theta_1, \theta_2, \theta_3)$ function. The highest peak in the rotation function ($\theta_1 = 122^\circ$) was used to define the final orientation of the fragment. Location of the oriented fragment within the cell was accomplished using the Q functions $Q(Y_0, Z_0)$, $Q(X_0, Z_0)$, $Q(X_0, Y_0)$ to position the group with respect to the three screw axes. The highest peak in the combined Q map (Wilson & Tollin, 1988) indicated shifts of $(X_0, Y_0, Z_0) = (0.18, 0.375, 0.23)$. One concession made to the quality of data in this example was that Q -function peaks separated by up to 0.4 \AA (*i.e.* $\pm 0.2 \text{ \AA}$ from the 'true' shifts) were assumed to represent the same translation. With good-quality data this increment tends to be set at around 0.3 \AA . Apart from this one relaxing of the normal procedures, the PATMET solution of the structure was fairly routine. As can be seen from the coordinates shown in Table 2(a), the located fragment is in reasonable agreement with the published molecule.

The correctness of the solution obtained from the Patterson-methods solution was checked in two ways, by Fourier recycling and rigid-body least-squares refinement (on F), using the programs FFT/SEARCH (Fourier) and RBLs (refinement) in the GX system (Mallinson & Muir, 1985). The O atoms were removed from the located model, and Fourier recycling calculations using just the located six-membered ring as phasing model revealed the two O atoms as the highest two peaks in the Fourier map, confirming, among other things, the orientation of the fragment within the plane of the group. Rigid-body least-squares refinement on the resulting model was carried out. One should bear in mind that, with such poor data available, any refinement must be

performed cautiously, so the most basic refinement of the whole molecule as a rigid body was carried out, together with scale and thermal parameters. The resulting R factor of some 13% (unweighted) is not unreasonable given the poor data available. The final refined coordinates are shown in Table 2(a). The average bond length for the ring is $1.38 (19) \text{ \AA}$, and average angle $120 (12)^\circ$, and the broad packing features of the structure, including the hydrogen-bonding scheme, are maintained.

Refinements with O-atom positions freed from the rigid group led to slight alterations ($\sim 0.1 \text{ \AA}$) of $C=O$ bond lengths, but the O atoms remained close (bonding) to their parent C atoms.

Structural solution of 6-azauracil

With the 23 $|F_n|$ data shown in Table 1(b), the same calculations were employed for the 6-azauracil structure. The $I(\theta, \varphi)$ peak occurs at $(54, 47^\circ)$ and the rotation-function peak at $\theta_1 = 24^\circ$. In this case the Q functions are somewhat more difficult to interpret. There is good agreement between the shifts indicated in the $Q(X_0, Y_0)$ and $Q(Y_0, Z_0)$ maps, with the fourth-highest peaks in each map giving a combined shift of $(0.23, -0.305, -0.25)$. There is a problem with the $Q(X_0, Z_0)$ map, however, probably caused by the lack of data. This map is at extremely low resolution, showing one very broad maximum over most of the Q -function-calculation region. This map can certainly be interpreted in such a way as to agree with the above set of shifts from the $Q(X_0, Y_0)$ and $Q(Y_0, Z_0)$ maps, but not conclusively so. Nor, however, does the map conclusively disagree with the above findings and so these shifts were used for fragment location.

After application of the indicated shifts to the oriented model, a set of coordinates is found which proves to be the correct solution but rotated by 60° from the published model. This means that two atoms are in fact incorrectly located (O2 and O4) which may help to account for the relatively poor Q -function results. The re-interpreted coordinate lists are shown in Table 2(b).

In this case, with two O-atom positions determined incorrectly by the Patterson-method procedure, the solution of the structure is less transparent. Refinement of the (incorrect) located model in *R*BLS was relatively unsuccessful, with the lowest R factor obtained being 30%. In addition, refining the O atoms independently of the six-membered ring led to the movement of these atoms away from bonding positions. These O atoms were thus removed from the model and Fourier recycling performed. The first F map revealed O4 as the highest peak, with no other atoms in bonding positions. A subsequent difference Fourier map revealed O2 as the highest difference peak, but at a somewhat long distance from C2. Refinement in *R*BLS of the rigid six-membered ring + O4 along with the positional coordinates of O2, followed by rigid-body refinement of the whole group, led to an R factor of $\sim 18\%$. The precision of the solution in this case is substantially lower than for 3-deazauracil, perhaps reflecting a very poor initial data set and pointing to a reason for the difficulties encountered in the Patterson-method calculations.

It is clear again in this case that Fourier recycling coupled with rigid-body least-squares refinement has, with care, developed the Patterson-method-solution coordinates into an acceptable structural solution. The final refined coordinates are shown in Table 2(b), with the average bond length in this case being $1.4(4)$ Å and average angle $120(30)^\circ$.

Concluding remarks

It can be seen from these examples that it is certainly possible to solve crystal structures of simple organic materials from very poor-quality X-ray powder diffraction data. The use of Patterson methods where the deconvolution of the intensity pattern is aided by the provision of an input model corresponding to the known stereochemistry of at least part of the unknown molecule is the key to such determinations in cases

where very few reflections can be measured. As has been shown in the above two examples, the results are often not unambiguous in such cases and it is important to appreciate the limitations of very low-resolution data sets (e.g. in giving very broad Fourier and Fourier-type maps). The information obtained from such basic structure determinations is also limited – without accurate least-squares refinement one can deduce only the basic framework of the structure. It should be noted that in addition to the rigid-body refinements on F carried out here, constrained Rietveld (profile) refinement of the pattern could also be used in the refinement process, perhaps leading to greater accuracy and reliability. In spite of these limitations, however, the determination of crystal structures from data such as these is a further demonstration of the ability of Patterson methods to extract information from even very paltry and inaccurate reflection data (Wilson, 1989). There are still many simple molecules which are resistant to crystallization and which may not even give reasonable microcrystals for use in single-crystal synchrotron or high-resolution powder diffractometers. The fact that Patterson methods give one the opportunity to extract structural information from such materials using moderate-resolution X-ray powder diffractometers should be borne in mind when such problems are encountered.

References

- DAVID, W. I. F., JOHNSON, M. W., KNOWLES, K. J., MORETON-SMITH, C. M., CROSBIE, G. D., CAMPBELL, E. P., GRAHAM, S. P. & LYALL, J. S. (1986). Report RAL-86-102. Rutherford Appleton Laboratory, Didcot, Oxon, England.
- KELLER, E. (1988). *SCHAKAL88. A Fortran Program for the Graphical Representation of Molecular and Crystallographic Models*. Albert-Ludwigs-Univ., Freiburg, Federal Republic of Germany.
- LE BAIL, A., FERÉY, G., AMOROS, P., BELTRAN-PORTER, D. & VILLENEUVE, G. (1989). *J. Solid State Chem.* **79**, 169–176.
- LOW, J. N. & WILSON, C. C. (1983). *Acta Cryst.* **C39**, 1688–1690.
- MALLINSON, P. R. & MUIR, K. W. (1985). *J. Appl. Cryst.* **18**, 51–53.
- PAWLEY, G. S. (1981). *J. Appl. Cryst.* **14**, 357–361.
- RIUS, J. & MIRAVITTLES, C. (1988). *J. Appl. Cryst.* **21**, 224–227.
- SINGH, P. & HODGSON, D. J. (1974). *Acta Cryst.* **B30**, 1430–1435.
- WERNER, P.-E., ERIKSSON, L. & WESTDAHL, M. (1985). *J. Appl. Cryst.* **18**, 367–370.
- WILSON, C. C. (1988). Report RAL-88-087. Rutherford Appleton Laboratory, Didcot, Oxon, England.
- WILSON, C. C. (1989). *Acta Cryst.* **A45**, 833–839.
- WILSON, C. C. & TOLLIN, P. (1986). *J. Appl. Cryst.* **19**, 411–412.
- WILSON, C. C. & TOLLIN, P. (1988). *Acta Cryst.* **A44**, 226–230.