Acta Cryst. (1963). **16**, 1243

Computer Determination of Unit-Cell from Powder-Diffraction Data

BY HELMUT M. HAENDLER AND WILLIAM A. COONEY

Department of Chemistry, University of New Hampshire, Durham, N.H., U.S.A.

(*Received* 29 *January* 1963 *and in revised form* 4 *April* 1963)

A series of programs for the IBM 1620 computer has been prepared for the calculation of unit cell parameters from powder diffraction data. It has been tested with triclinic, monoclinic, and orthorhombic cells.

Any unit cell, whatever its symmetry, can be represented as a triclinic cell, corresponding to three noncoplanar vectors in reciprocal space. Each reflection of a powder pattern is related to a vector in reciprocal space. Selection of six appropriate vectors, or six appropriate lines, will permit determination of a unit cell and assignment of suitable indices. This method was suggested by Ito (1950), who discussed the technique of line selection and reduction of the unit cell to one of higher symmetry. We have written a series of programs for the IBM 1620 computer which carry out both of these operations and have tested the process with four representative examples. The complete procedure is indicated by the flow diagram in Appendix I, and the programs are listed with suitable descriptions and identification.

Selection of vector sets

The calculations are based on the reciprocal cell relationship

$$1/d^2_{hkl}=Q_{hkl}=h^2a^{*2}+k^2b^{*2}+l^2c^{*2}+2hka^*b^*\cos\gamma^*$$
$$+2hla^*c^*\cos\beta^*+2klb^*c^*\cos\alpha^*.$$

Two lines can be designated as corresponding to two vectors of the unit cell, with

$$Q_{100}=a^{*2}\quad\text{and}\quad Q_{010}=b^{*2}.$$

A search is made for two other lines evenly spaced about the sum $Q_{100}+Q_{010}$. The inherent inaccuracy of powder data makes it necessary to assign certain arbitrary limits to define 'evenly spaced'. If such an evenly spaced pair, within these limits, is found in the list of observed Q values, the indices 110 are assigned to the line of greater Q, $1\bar{1}0$ to the line of smaller Q. The reciprocal-cell angle can be determined by

$$\cos\gamma^*=\frac{(Q_{110}-Q_{1\bar{1}0})}{4a^*b^*}.$$

Another instance must be found in which a pair of observed Q values is evenly spaced around the sum of two other Q values, one of the latter Q values corresponding to either Q_{100} or Q_{010} of the previous calculation. The second component of the sum is now designated as Q_{001}, and the indices 101 assigned to the line of greater Q, $10\bar{1}$ to the line of smaller Q (or respectively 011 and $01\bar{1}$ if Q_{010} is the previously used component). The corresponding reciprocal cell angle can be calculated. The search for evenly spaced pairs is carried out with the sum $Q_{010}+Q_{001}$, and the remaining angle calculated.

The set of nine Q values, corresponding to nine lines of the powder pattern, will determine a reciprocal unit cell and will be consistent with respect to the criteria used in forming the set. The criteria, however, can be met by chance, and further checking is necessary to determine if the unit cell is correct. A large number of such sets can usually be formed from the powder pattern Q values. To guarantee that one of the sets so formed will define a correct cell, it is necessary to form every such set. A program (A) has been written for the IBM 1620 computer which finds every instance in which two Q values are evenly spaced around the sum of two other Q values.

The observed Q values are read, in ascending order, into the computer, which assigns to each a subscript number $(1, 2, 3, \ldots, M)$. The computer then adds $Q(1)+Q(1)$, subtracts from this sum each of the observed Q values, and compares the results. The process is repeated for the sums $Q(1)+Q(1), \ldots, Q(1)+Q(N)$; $Q(2)+Q(2), \ldots, Q(2)+Q(N)$; $Q(N)+Q(N)$, where N is a prespecified integer. The output lists only the subscripts of the two Q values which form the sum, the subscripts of the two Q values which are evenly spaced around this sum, and the amount by which the pair fails of being truly evenly spaced.

The operator then forms from the output data every possible internally consistent set of nine indexed Q values which will lead to a reciprocal cell. Although it would be possible to program this tabulation we feel that a major advantage of the IBM 1620 lies in the ease with which the operator can participate and interfere with the process of computation, perhaps by intuition. Although each of the sets can be tested, a judicious selection of a few at this point can sometimes speed the remaining calculations.

Testing of vector sets

The second program of the series (B) performs a preliminary test on each set. The observed Q values are again read, in ascending order, and subscripts assigned. The subscripts of Q_{100}, Q_{010}, Q_{001}, Q_{110}, $Q_{1\bar{1}0}$, Q_{101}, $Q_{10\bar{1}}$, Q_{011}, $Q_{01\bar{1}}$, are read in, in that order. The computer calculates the reciprocal cell corresponding to each set, and prints out the cell constants and the values for Q_{111}, $Q_{11\bar{1}}$, $Q_{1\bar{1}1}$, and $Q_{\bar{1}11}$. The reciprocal cell angles are changed by the computer from all acute to all obtuse and the corresponding Q values listed. If three or four of the calculated Q values match observed Q values, the cell is considered for further testing.

Those reciprocal cells which have not been eliminated at this point are entered, one by one, into a triclinic indexing program (C or C′) in an attempt to calculate the complete pattern. The values of h, k, and l are allowed to range from $-N$ to $+N$, where N is an integer usually less than 5. If the triclinic cell is a correct one, the parameters can be refined by selection of lines of suitable index for improved agreement. If the original vectors selected were not those of a primitive cell no indexing will be complete. In these cases a selection of certain submultiples of the initial Q values must be made. This again is a matter for the judgement of the operator.

Reduction of the triclinic cell

Although the triclinic cell is a correct one, the pattern can often be indexed on the basis of a unit cell of higher symmetry. This conversion uses a series of programs, each of which can also be used independently, and follows the scheme outlined by Azároff & Buerger (1958). The reciprocal cell is transformed to the corresponding direct cell and a search is made for the reduced cell, that cell with the three shortest non-coplanar vectors and with all angles acute or obtuse. Any vector can be expressed as the translation

$$\mathbf{t}=u\mathbf{a}+v\mathbf{b}+w\mathbf{c}$$

and the computer calculates the magnitude (t^2) of all vectors for values of u, v, and w ranging from -2 to $+2$. The program output lists u, v, w and the magnitude of $\mathbf{t}$ for all combinations of u, v, and w for which the magnitude is equal to or less than that of the longest direct cell vector. The three shortest non-coplanar vectors can readily be detected by examination by the operator. This method was selected in preference to the familiar Delaunay reduction because of the ease with which it could be programmed.

The three sets of u, v, and w, one for each original vector, form the matrix for transforming the old direct cell to the new one. This calculation can be done with a program which uses standard vector evaluation equations. If the reduced cell is not all acute or all obtuse, the next most likely matrix can easily be selected.

The reduced cell scalar matrix can be listed, and the transformation matrix for conversion to the correct unit cell determined by reference to the Niggli matrix representations (1953). These representations, which are based on the scalar products of the cell, are also discussed and listed by Azároff & Buerger (1958) and by Buerger (1957). The final cell can be computed, converted to the reciprocal cell if desired, and the powder pattern indexed by suitable calculation for the crystal system concerned.

Reliability of unit cells

Facility in computing a unit cell as described does not insure its inherent correctness, but only its correct conformation to the original data. The purity of the material must be unquestionable, and the powder data must be of high reliability. Failure to discern weak reflections may lead to unit cells which are submultiples of larger true cells or which bear a not immediately obvious relation to the true cell. Caution must be used in evaluation of powder data, although in those instances in which single crystals are unobtainable there is often no choice. The inclusion of a refinement program in the overall procedure would be useful, for the results would assist appreciably with the final indexing.

An indication of the reliability of the unit cell can be obtained by use of the methods discussed by de Wolff (1961). A comparison is made of the number of lines observed with those expected, and the discrepancy in Q (as $10^4/d^2$) is compared with that expected. This calculation has also been programmed and has been applied to the unit cells listed below.

Unit cell of $Ca(H_2PO_4)_2$-triclinic

The powder data used were those of $Ca(H_2PO_4)_2$, number 9-390 in the ASTM X-ray Powder Data File. Only the first twenty lines were used in this calculation. The procedure is sufficiently rapid that a preliminary survey with few lines may lead quickly to a possible solution. The set of twenty lines was surveyed with all combinations of the first five lines as the three cell vectors. There were 41 pairs of lines evenly spaced (within $\pm 0{\cdot}0010$ in Q) about all possible sums of pairs of the first five lines. Fifteen sets of 9 lines were selected for input to the first testing program. Selection was based upon minimum differences of the absolute values for the various pairs of 'evenly spaced' lines. Of these fifteen sets, six groups of reciprocal cell parameters gave agreement for the $\{111\}$ Q-values.

These six cells were used to compute the complete powder pattern, which was compared with the observed pattern. Only one of the six failed to index the 20 lines of the pattern. The five reciprocal cells were converted to direct cells and entered in the reduction program. The three shortest non-coplanar vectors were selected from the output of each cell calculation and the corresponding transformation matrices selected. Four of the five led to a common cell. The parameters of these four were averaged, and the pattern re-indexed. From a comparison of these values, some of the parameters were refined by inspection and a revised cell adopted. Reference to the Niggli matrix showed that this cell was the direct cell.

The cell constants, compared with the ASTM values, were

	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)
Computed	5·54	7·56	8·27	109·6	109·1	93·4
ASTM	5·549	7·566	8·255	109·82	109·15	93·31

A calculation was made for reference to de Wolff's criteria. There should be 7 lines of 10^4Q below 500 with a chance expected discrepancy of 23; 20 lines below 10^4Q of 1000, with a discrepancy of 16. The pattern showed 6 lines below 500, with an average discrepancy of less than 1; and 16 lines below 1000, with an average discrepancy of 3.

Unit cell of SnF_2-monoclinic

At the time this work was done, the structure or unit cell of tin(II) fluoride was not known. A powder pattern had been reported by Nebergall, Muhler & Day (1952). We had begun single-crystal work, but decided to test this program on the powder pattern obtained on a sample of freshly prepared SnF_2, using a Philips 114·6 mm camera. The pattern used consisted of 36 lines, with a maximum Q of about 0·7500 ($d_{hkl} = 1{\cdot}153$).

A survey of the Q values indicated that each of the first six lines might be the first in a series of the form $Q(A), 4Q(A), 9Q(A), \ldots$ *etc.* All six were thus considered as possible cell vectors, and 20 of the lines were used for the evenly spaced pairs calculation. Forty-one pairs of lines were found, leading to 365 possible cells. Each of these cells was used to calculate $\{111\}$ Q values, and 33 of these gave all four Q values in agreement by $\pm 0{\cdot}003$ or better.

These 33 cells were used to calculate the complete pattern and to compare it with the observed data. A limit of 0·003 was again used. Six of the reciprocal cells indexed the entire pattern. Six others indexed all but one or two lines of the pattern. The six cells were transformed to direct cells for conversion to the reduced cell. In some cases the axes and angles were renamed to insure that **c** had the largest magnitude. The six reduced cells were equivalent, and the results were averaged to give the cell

$$\begin{aligned} a &= 4{\cdot}041 \pm 0{\cdot}003\ \text{Å} & \alpha &= 103{\cdot}061 \pm 0{\cdot}032° \\ b &= 4{\cdot}043 \pm 0{\cdot}002\ \text{Å} & \beta &= 118{\cdot}589 \pm 0{\cdot}030° \\ c &= 4{\cdot}224 \pm 0{\cdot}001\ \text{Å} & \gamma &= 105{\cdot}191 \pm 0{\cdot}060° \end{aligned}$$

It remained to be seen if this cell could be converted to one of high symmetry. The Niggli matrix was found to be

$$\begin{matrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{matrix}$$

leading to a body-centered monoclinic cell, which was partially refined to

$$a = 4{\cdot}226,\ b = 4{\cdot}902,\ c = 5{\cdot}143\ \text{Å};\ \beta = 85{\cdot}96°.$$

The density of SnF_2 was found to be 4·69 g.cm^{-3}, giving $Z = 2$.

The complete pattern and the index assignments are given in Table 1.

Table 1. *Powder pattern of* SnF_2

d_{hkl}(obs)	d_{hkl}(calc)	I/I_0	hkl
3·544	3·544	100	100
3·368	3·376	43	101
3·196	3·196	79	011
3·147	3·150	42	$\bar{1}01$
2·567	2·565	21	200
2·450	2·451	17	020
2·106	2·108	17	002
2·055	2·055	43	211
1·983	1·983	32	121
1·952	1·950	28	$\bar{2}11$
1·934	1·936	29	$\bar{1}21$
1·851	1·852	22	112
1·772	1·772	42	220
—	1·774	—	$\bar{1}12$
1·689	1·688	7	202
1·628	1·625	11	301
1·616	1·615	19	310
1·598	1·598	13	022
1·576	1·575	7	$\bar{2}02$
1·556	1·557	14	130
1·524	1·524	12	031
1·391	1·390	10	222
1·379	1·380	4	103
1·352	1·351	7	013
1·326	1·325	18	231
—	1·325	—	312
—	1·325	—	$\bar{2}22$
1·309	1·308	7	$\bar{3}21$
1·297	1·296	5	$\bar{2}31$
1·283	1·283	2	400
1·265	1·265	6	132
1·242	1·243	7	$\bar{3}12$
—	1·240	—	$\bar{1}32$
1·229	1·226	7	040
1·214	1·213	6	411
1·203	1·203	4	123
1·182	1·181	5	330
1·171	1·169	—	$\bar{4}11$
—	1·170	—	$\bar{1}23$
1·165	1·163	—	$\bar{2}13$
1·153	1·152	—	141

Marignac (1857) reported the compound as monoclinic and assigned a morphological cell with $a:b:c = 2{\cdot}7204:1:1{\cdot}4078$ and $\beta = 109{\cdot}40°$. The cell reported here can be transformed to Marignac's cell by the matrix

$$\begin{matrix}\bar{3} & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 1\end{matrix}$$

Bergerhoff (1962) has reported the structure of SnF_2 based on single-crystal studies, and gives a unit cell with 16 formula units, twice the size of the Marignac cell (the *c* axis being doubled) and eight times the size of the one reported here. He reports that the pattern of Nebergall *et al.* shows only the strong reflections, and such would be the case with ours. His constants are

$$a=13{\cdot}46,\ b=4{\cdot}92,\ c=13{\cdot}86\ \text{Å};\ \beta=109{\cdot}5°,$$

and the matrix for the conversion of our cell to this one is

$$\begin{matrix}\bar{3} & 0 & 1\\ 0 & 1 & 0\\ 2 & 0 & 2\end{matrix}$$

H. E. Swanson, of the National Bureau of Standards, has determined the powder pattern of SnF_2 for inclusion in the powder data monograph. The pattern was refined, after some single-crystal work, by Evans (1963), and indexed on the basis of the Bergerhoff cell. The refinement gave

$$a=4{\cdot}228,\ b=4{\cdot}909,\ c=5{\cdot}156\ \text{Å};\ \beta=85{\cdot}96°$$

for the monoclinic cell corresponding to the one reported here. There were several faint lines in the Bureau of Standards diffractometer pattern which were not detected in our photographic pattern. It does not appear likely that these lines could have been indexed on the basis of the triclinic cell originally assigned. The ($\bar{1}11$) reflection of the Bergerhoff cell, for example, corresponds to (381) of the eightfold monoclinic cell. It would be necessary to search for a cell based in part on some of these faint reflections, a calculation that might be too lengthy for practical use of the 1620 computer. If a sufficient number of these faint reflections are present, the solution should be possible.

The de Wolff comparison of the powder data is shown in Table 2.

Table 2. *De Wolff comparison for* SnF_2

10^4Q	N_c	N_o	Δ	Max. Δ_o*
1000	4	3	95	5
2000	10	6	67	2
4000	28	17	47	14
7550	73	36	34	32

* The largest single Δ in each successive interval is given.

Unit cell of dickite-monoclinic

The powder data for dickite, a basic aluminum silicate, was taken from the ASTM file (file number 10-446). This was selected since the unit cell contained four formula units, and selection of a primitive triclinic cell would be unlikely. This proved to be the case. Again with the first five lines for cell vectors and 20 lines for pairs, 65 combinations were tested for {111} indexing. Agreement was poor, and no cell gave completely satisfactory indexing of the powder pattern.

The initial search for evenly spaced pairs was repeated with one-half the Q value of the first line taken as $Q_{100}(a^*)$. Not all combinations were recalculated, but a comparison was made of two vector combinations, with and without halving one reciprocal axes. The halved set led directly to complete indexing of the pattern and could readily be reduced to a triclinic cell which gave the correct monoclinic cell upon transformation. No refinement was attempted.

	a	*b*	*c*	*β*
Computed	5·15 Å	8·96 Å	14·43 Å	97·54°
ASTM	5·149	8·949	14·419	96·80

Unit cell of $LaBO_3$-orthorhombic

The evenly spaced pairs program was written so that if the sum of two vector lines corresponded to the Q value of a single line this was noted in the output. This would be the case if the first two lines were perpendicular to each other. It should be possible, in an orthorhombic pattern, to detect these relationships. The data for lanthanum borate were taken from the National Bureau of Standards (1962) Monograph 25.

Twenty lines were used, with the first eight used as cell vectors, and three sets resulted with possible mutually perpendicular axes. Each of these was tested directly with an orthorhombic indexing program as were the corresponding cells with all three direct cell lengths doubled. One of the eightfold cells indexed the pattern completely.

	a	*b*	*c*
Computed	5·87 Å	8·26 Å	5·11 Å
NBS	5·872	8·257	5·107

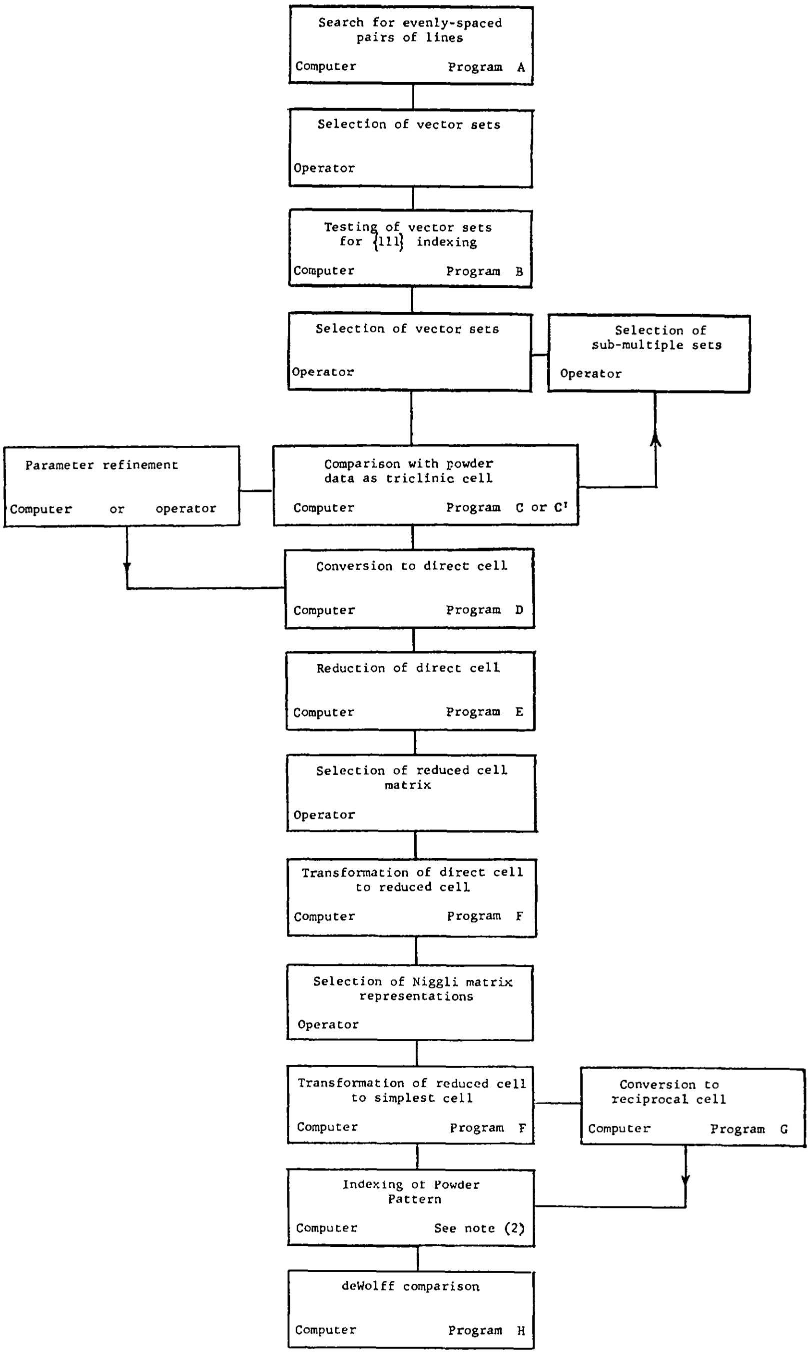
Comments

It is difficult to estimate the time required for the calculations discussed. A fortunate selection can save much time. The lanthanum borate computation required one day, including all machine work. A triclinic cell, such as the calcium phosphate, might require three hours of IBM 1620 time.

It should not be inferred that the proposed program automatically permits determination of all unit cells from powder patterns. It is subject to the limitations inherent in the Ito procedure and dependent upon the validity and precision of the powder pattern. If the cell is one containing many formula units it may be difficult to find sufficient weak reflections to permit unambiguous and direct determination of cell angles. Thus, many trial runs may be necessary. Adaptation to larger computers may be feasible in

APPENDIX I

Outline of calculation procedure



Notes: (1) Titles and reference numbers of the programs are listed in Appendix II.
(2) The indexing programs for different crystal systems are listed in the World List and the American Crystallographic Association list.

these cases. The programs outlined here, however, should be readily applicable to many substances of less complex nature and to the detection of portions of the unit cell, as in the case of SnF_2, from which other details can be deduced.

We are grateful for the assistance of Stanley Prince, who tested the programs by calculating two of the cells, without being aware of their nature, and of the University of New Hampshire Computer Center, supported by the National Science Foundation, which also supported one of us (W. A. C.) through an undergraduate research grant. The work was also supported in part by the U.S. Atomic Energy Commission. Print-outs of the FORTRAN programs (see Appendix II) are available from the authors.

APPENDIX II

Reference list of programs (FORTRAN)

Program	Title and Object	Authors' number	World List number
A	Evenly-spaced-pairs-finding program To determine line pairs for calculating possible triclinic primitive cell angles	17	—
B	Vector-set-testing program To determine which cell triplets give correct {111} Q values	18	—
C	Triclinic cell comparison program To calculate Q values for triclinic cell and to compare these to powder pattern, listing hkl values	23	—
C′	Triclinic cell indexing program To calculate Q, $\sin^2\theta$, d for all hkl values for triclinic unit cell	14	287
D	Direct lattice calculation program To calculate direct lattice parameters from reciprocal lattice parameters	3	278
E	Direct cell reduction program To determine from triclinic primitive cell if cell exists with shorter non-coplanar vectors	21	—
F	Axis transformation program To change axes and angles of a 3-coordinate system to those of a second system	19	289
G	Reciprocal lattice calculation program To calculate reciprocal lattice parameters from direct cell parameters	16	288
H	De Wolff comparison program To calculate expected discrepancy between observed and calculated Q values	22	—

Each of the following programs computes Q_{hkl}, $\sin^2\theta_{hkl}$, and d_{hkl} for all hkl values up to a preselected limit of hkl and θ.

	Simple cubic interplanar spacings	8	*
	Face-centered cubic	9	283
	Body-centered cubic	10	*
	Tetragonal	12	*
	Hexagonal	11	284
	Orthorhombic	2	277
	Monoclinic	4	279
	Triclinic	14	287

* Have been submitted.

References

Azároff, L. V. & Buerger, M. J. (1958). *The Powder Method in X-Ray Crystallography*. Chap. 10 and 11. New York: McGraw-Hill. (Some of the Niggli matrix representations of Table 6, p. 146, are incorrect.)

Bergerhoff, G. (1962). *Acta Cryst.* **15**, 509.

Buerger, M. J. (1957). *Z. Kristallogr.* **109**, 42.

Evans, H. T., Appleman, D. E. & Handwerker, D. S. (1963). American Crystallographic Association, Annual Meeting, March 28–March 30, 1963. Paper E-10.

Ito, T. (1950). *X-ray Studies on Polymorphism*, p. 187. Tokyo: Maruzen Co.

Marignac, C. (1857). *Ann. Min., Paris*, [5], **12**, 1.

National Bureau of Standards (1962). *Standard X-ray Diffraction Powder Patterns*. Monograph 25, Section 1, p. 20.

Nebergall, W. H., Muhler, J. C. & Day, H. G. (1952). *J. Amer. Chem. Soc.* **74**, 1604.

Niggli, A. (1953). *Schweiz. Min. Petrogr. Mitt.* **33**, 21.

Wolff, P. M. de (1961). *Acta Cryst.* **14**, 579.