

Conditional ambiguity of one-dimensional crystal structures determined from a minimum of diffraction intensity data

Irina A. Shkel,^a Ho Seung Lee^{b,c} and Oleg V. Tsodikov^{b,c*}

^aDepartment of Biomedical Engineering, University of Wisconsin-Madison, Madison, WI 53706, USA, ^bDepartment of Medicinal Chemistry, University of Michigan, Ann Arbor, MI 48109, USA, and ^cUndergraduate Research Opportunities Program, University of Michigan, Ann Arbor, MI 48109, USA. Correspondence e-mail: olegt@umich.edu

When the number of intensities greatly exceeds the number of unknown atomic coordinates, the problem of obtaining a crystal structure from the intensities is overdetermined and, for a sufficiently small structure, a chemically meaningful solution can be found by direct methods. A difficulty in determining a structure has been historically attributed to the non-uniqueness of such a structure owing to multiple, or homometric, structures that yield the same set of intensities. The number of homometric structures has not been rigorously analyzed owing to the complexity of this problem. By using the method of elementary symmetric polynomials with a new origin definition, one-dimensional crystal structures of a small number of identical atoms ($N < 5$), determined from a minimum ($N - 1$) of the lowest-resolution intensities, are enumerated. It is demonstrated that such a structure is unique for $N \leq 3$. Interestingly, for $N = 4$, the structure can be determined either uniquely or twofold ambiguously, depending on the intensity values. These results suggest that, even for larger structures, a minimum set of (or not many more) accurately measured intensities can yield a unique structure.

© 2011 International Union of Crystallography
Printed in Singapore – all rights reserved

1. Introduction

In his pioneering study, Ott noted that in a typical diffraction experiment the number of measured intensities I_{hkl} exceeds the number of unknown atomic coordinates, $3N$ (Ott, 1927). Therefore, the problem of obtaining atomic coordinates from well measured diffraction intensities is overdetermined if the data at a sufficiently high resolution are collected. Ott (1927) and later Avrami (1939) demonstrated that the atomic coordinates can, in principle, be obtained from the intensities as a solution of a system of polynomial equations relating diffraction intensities to atomic coordinates. Their approach was deemed impractical at that time, owing to a cumbersome structure of the polynomial system and the lack of computing power needed to handle it. Since then, the algebraic approach has been revisited, which resulted in a number of important advances, including new methods of calculating the structure from a subset of structure factors and evaluating the role of experimental errors in structure determination (Cervellino & Ciccariello, 1999, 2005; Pilz & Fischer, 2000, 1998; Fischer & Pilz, 1997). However, a practical method of structure determination has not yet emerged. Nevertheless, the original ideas put forth by Ott served as an impetus for development of the direct methods (Sheldrick, 1984; Miller *et al.*, 1993, 1994; Karle & Hauptman, 1956), reviewed by Usón & Sheldrick (1999) and Hauptman (1997). The direct methods are now the

preferred and time-tested technique for determining crystal structures of small molecules and are gaining popularity in determining heavy-atom substructures in macromolecular crystallography (Dauter *et al.*, 1999; Schneider & Sheldrick, 2002; Weeks *et al.*, 2003). Owing to their probabilistic character, as a way of determining full structures, the direct methods remain applicable to structures of at most a hundred or so atoms in the asymmetric unit for which exceptionally high resolution data ($< 1.2 \text{ \AA}$) are available. This limitation, at least in principle, can be overcome by the original deterministic algebraic approach provided that (a) the number of structure solutions of the above polynomial system is reasonably small (*e.g.* not exponentially increasing with the increasing number of atoms), (b) the solution(s) are stable with respect to small changes in diffraction intensities (as these are determined with some experimental uncertainty) and, finally, as long as (a) and (b) hold, that (c) a practically useful method of solving the above-mentioned high-order polynomial system is found.

Because of an intrinsic limitation of a diffraction intensity data set to yield interatomic distances in the structure of interest, but not the atomic coordinates, even an ideal set of intensities can, at best, produce multiple structures. Patterson coined the term *homometric* to describe such structures, *i.e.* structures that contain the same set of interatomic distances (Patterson, 1939). This homometric ambiguity goes beyond

congruent structures that are related by operations of reflection, rotation and translation, which preserve the interatomic distances. A report of the existence of homometric structures that are not congruent dates back to 1930 (Pauling & Shappell, 1930). In 1944, Patterson presented an elegant method of generating such homometric one-dimensional structures made up of identical atoms located at subsets of the vertices of an equilateral polygon, or cyclotomic sets (Patterson, 1944). Many examples of homometric structures and even methods of constructing them have been reported since then (Grimm & Baake, 2008; Rau *et al.*, 1979; Zobetz, 1993; Bullough, 1964, 1961). Hauptman & Karle (1951) found an example where two lowest resolution intensities (I_1 and I_2) yielded an ambiguous one-dimensional crystal structure of three non-identical atoms, but, when a higher-resolution intensity (I_3) was considered, this ambiguity was eliminated. Today, the problem of reconstructing structures from interatomic distances is an unsolved problem in several areas of physics and mathematics, known to be notoriously difficult. Because a method for enumerating homometric structures has not been developed, some important questions that remain unanswered are: (i) what is the number of structures that can be determined from the minimum number of intensities, and (ii) can a unique structure or a reasonably small number of structures be determined in some cases? Here we use the method of symmetric polynomials to enumerate one-dimensional crystal structures of a small number ($N < 5$) of identical atoms determined from $N - 1$ lowest-resolution diffraction intensities.

2. Preliminaries

2.1. Defining unique solution in one-dimensional crystal structure determination

We will consider a general one-dimensional crystal structure of N point atoms in the unit cell of size $a = 1$. Structure factors F_h for this structure are

$$F_h \equiv |F_h| \exp(2\pi i \varphi_h) = \sum_{j=1}^N f_j \exp(2\pi i h x_j),$$

$$h = -\infty, \dots, -1, 0, 1, \dots, +\infty, \quad (1)$$

where f_j , x_j are individual atomic scattering factors and their unknown coordinates, respectively. Phases φ_h are unknown, which constitutes the phase problem. For a structure of equal light atoms (without sulfur or heavier atoms and neglecting H atoms), we will normalize $f_j = 1$. Suppose a solution (x_1, x_2, \dots, x_N) has been found based on a set of observed and properly scaled reflection intensities $I_h = I_{-h}$ (Friedel's law),

$$I_h = |F_h|^2 = \sum_{j,p=1}^N \exp[2\pi i h (x_j - x_p)]$$

$$= N + \sum_{\substack{j=1 \\ j < p}}^N 2 \cos[2\pi h (x_j - x_p)], \quad (2)$$

then the reflection (or enantiomer) solution $(-x_1, -x_2, \dots, -x_N)$ relative to the same origin yields identical I_h . Distinguishing between these enantiomers is straightforward if the chirality of the molecule of interest is known *a priori*, which, for example, is the case with proteins, or experimentally by using anomalous dispersion (resonant scattering). We will call these two structures *origin enantiomers*. We will define then that the system of equation (2) has a *unique solution* if it has two and only two structure solutions (relative to a fixed origin) that are origin enantiomers.

2.2. Defining an origin

Assuming that the total number of atoms in the unit cell, N , is known, we will introduce $\xi_j = \exp(-2\pi i x_j)$. A set of ξ_j ($j = 1, \dots, N$) thus may be considered a solution of the structure. Then

$$F_h = \sum_{j=1}^N \xi_j^h \text{ and } I_h = \sum_{j,p=1}^N \xi_j^h \xi_p^{-h}. \quad (3)$$

In this notation the origin enantiomer of solution $(\xi_1, \xi_2, \dots, \xi_N)$ is $(\xi_1^{-1}, \xi_2^{-1}, \dots, \xi_N^{-1})$. Because $I_h = I_{-h}$ (Friedel's law) and $h = 0$ is a trivial case, we will use only I_h . The origin can be chosen arbitrarily; therefore, we will set it to coincide with the center of gravity of all atoms in the unit cell so that $\sum_{j=1}^N x_j = 0$, *i.e.*

$$\prod_{q=1}^N \xi_q^h = 1. \quad (4)$$

Note that, relative to this new origin, x_j no longer lies between 0 and 1.

As we discussed above, if the origin is fixed, then a set of x_j and its enantiomer $(-x_j)$ define a unique structure. However, the above choice of origin, $\sum_{j=1}^N x_j = 0$, does not uniquely fix the origin. Owing to crystal periodicity, for a structure of N atoms there are N origins defined this way, since the ordered atom numbering can be frame-shifted in N different ways and each of such N frames defines its own origin. It can be readily demonstrated that these N origins, each defined as the center of gravity of the structure, are separated from each other by $1/N$. The origin shift by $1/N$ is equivalent to a phase shift by $2\pi/N$ for atomic coordinates x_j . Therefore, a unique crystal structure is given by $2N$ (congruent) structures: N structures defined relative to different origins and their respective enantiomers. An analogous definition of uniqueness neglecting the congruence was introduced by Patterson (1944) and it was most recently employed in the one-dimensional case (Zimmermann & Fischer, 2009). Then, the number of solutions is, generally, a multiple of $2N$ ($2Nn$) and the problem is to establish n . In other words, for $n = 1$, the structure obtained from the minimum of intensity data is unique, and for $n > 1$ the minimum set yields n non-congruent structures.

3. Results

3.1. Application of elementary symmetric polynomials to the crystal structure determination problem

We define elementary symmetric polynomials e_h ($h = 1, 2, \dots, N$) as a sum of all possible different products of j distinct ξ_p , *i.e.*

$$e_1 = \sum_{j=1}^N \xi_j, \tag{5}$$

$$e_2 = \sum_{\substack{j_1, j_2=1 \\ j_1 > j_2}}^N \xi_{j_1} \xi_{j_2}, \tag{6}$$

$$e_N = \prod_{j=1}^N \xi_j. \tag{7}$$

We will also define the elementary symmetric polynomials with negative subscripts as follows,

$$e_{-1} = \sum_{j=1}^N \xi_j^{-1}, \tag{8}$$

$$e_{-2} = \sum_{\substack{j_1, j_2=1 \\ j_1 > j_2}}^N (\xi_{j_1} \xi_{j_2})^{-1}, \tag{9}$$

$$e_{-N} = \left(\prod_{j=1}^N \xi_j \right)^{-1}. \tag{10}$$

Notice that for our choice of the coordinate origin [equation (3)], $e_N = e_{-N} = 1$. Also, we define $e_0 = 1$. Then we obtain

$$e_{-h} = e_{N-h}, \quad h = 0, 1, \dots, N. \tag{11}$$

In algebra, the relationships between the elementary symmetric polynomials and structure factors (power sums) are known as Newton's identities, which can be formulated in terms of matrix determinants (Littlewood, 1950),

$$F_h = \begin{vmatrix} e_1 & 1 & 0 & \dots \\ 2e_2 & e_1 & 1 & 0 & \dots \\ 3e_3 & e_2 & e_1 & 1 & \dots \\ \vdots & & & & \ddots \\ he_h & e_{h-1} & \dots & & e_1 \end{vmatrix}, \quad h = 1, \dots, N-1. \tag{12}$$

By using (11) we also obtain

$$F_{-h} = \begin{vmatrix} e_{N-1} & 1 & 0 & \dots \\ 2e_{N-2} & e_{N-1} & 1 & 0 & \dots \\ 3e_{N-3} & e_{N-2} & e_{N-1} & 1 & \dots \\ \vdots & & & & \ddots \\ he_{N-h} & e_{N-h+1} & \dots & & e_{N-1} \end{vmatrix}, \quad h = 1, \dots, N-1. \tag{13}$$

Now, intensity I_h ($h = 1, \dots, N-1$) can be related to symmetric polynomials as a product of the two determinants,

$$I_h = F_h F_{-h} = \begin{vmatrix} e_1 & 1 & 0 & \dots \\ 2e_2 & e_1 & 1 & 0 & \dots \\ 3e_3 & e_2 & e_1 & 1 & \dots \\ \vdots & & & & \ddots \\ he_h & e_{h-1} & \dots & & e_1 \end{vmatrix} \begin{vmatrix} e_{N-1} & 1 & 0 & \dots \\ 2e_{N-2} & e_{N-1} & 1 & 0 & \dots \\ 3e_{N-3} & e_{N-2} & e_{N-1} & 1 & \dots \\ \vdots & & & & \ddots \\ he_{N-h} & e_{N-h+1} & \dots & & e_{N-1} \end{vmatrix}. \tag{14}$$

These equations form a system of $N-1$ polynomial equations for $h = 1, \dots, N-1$, whose unknowns are e_h . In this formalism Newton's identities dictate that the crystal structure of N atoms ($\xi_j, j = 1, \dots, N$) is determined from the $N-1$ elementary symmetric polynomials with the smallest positive indices by solving a univariate polynomial equation (Macdonald, 1979),

$$\sum_{j=0}^N (-1)^j e_j \xi^{N-j} = 0, \tag{15}$$

where $e_0 = 1$ by definition, and $e_N = 1$ by the choice of origin.

A corollary of the fundamental theorem of algebra is that for a given set of e_h ($h = 1, 2, \dots, N-1$), the solution of equation (15), *i.e.* the crystal structure, is unique. Therefore, solving system (14) is, in essence, equivalent to determining a crystal structure. We will then check whether different sets of e_h obtained from the intensities by solving the polynomial system of equations (14) for $h = 1, 2, \dots, N-1$ yield the same or different structures after their substitution into equation (15) as coefficients.

Note that all coefficients of system (14) are real; therefore, for each e_h there is a complex conjugate solution \bar{e}_h . This, in turn, means that for a structure solution x_j obtained from a set of e_h by using (15), there is a respective complex conjugate solution \bar{x}_j . Because $|\xi_j| = 1, \bar{\xi}_j = 1/\xi_j$, *i.e.* \bar{e}_h yield the enantiomer structure.

3.2. Uniqueness of the structure for $N < 5$

In the case $N = 1$, the solution is trivial and unique, $x_1 = 0$. The enantiomer coincides with the structure.

For $N = 2$, because a two-atom structure is always centrosymmetric owing to the choice of the origin, an enantiomer coincides with the solution. Two choices of the origin are possible; therefore a unique structure corresponds to the only two solutions of system (14), which in this case is just one equation: $e_1^2 = I_1$. The two solutions are $e_1 = \pm(I_1)^{1/2}$; therefore, the structure is unique. As expected, these solutions yield the two sets of coordinates x_j phase-shifted relative to each other by π , which is the origin shift for $N = 2$, as follows directly from (15) for this case:

$$\xi^2 - e_1 \xi + 1 = 0.$$

For $N \geq 3$, non-centrosymmetric structures are possible; therefore, as described in the above section, a unique structure would correspond to $2N = 6$ solutions of system (14),

$$\begin{aligned} I_1 &= e_1 e_2, \\ I_2 &= (e_1^2 - 2e_2)(e_2^2 - 2e_1). \end{aligned}$$

This system yields

$$e_1^6 - \frac{1}{2}(I_1^2 + 4I_1 - I_2)e_1^3 + I_1^3 = 0. \quad (16)$$

Indeed, (16) has at most six distinct solutions [sets of (e_1, e_2)]; therefore, the structure obtained from system (14) is unique. The solutions can be readily obtained explicitly as equation (16) is a quadratic in terms of e_1^3 . This result also implies that, for $N = 3$, all intensities I_h , for $h > 2$, can be calculated from the intensities from the minimum set, through polynomial relationships.

For $N = 4$, system (14) is

$$\begin{aligned} I_1 &= e_1 e_3, \\ I_2 &= (e_1^2 - 2e_2)(e_3^2 - 2e_2), \\ I_3 &= (e_1^3 - 3e_1 e_2 + 3e_3)(e_3^3 - 3e_3 e_2 + 3e_1). \end{aligned} \quad (17)$$

Eliminating e_1 and e_3 yields a biquadratic equation in terms of e_2 ,

$$\begin{aligned} e_2^4(3I_1 - 6) + e_2^2 \left[\frac{3}{2}(I_1^2 - I_2)(1 - I_1) + I_1(I_1 - 3)^2 - I_3 \right] \\ + \frac{3}{4}(I_1^2 - I_2)^2 = 0, \end{aligned} \quad (18)$$

and a polynomial equation of the 16th degree for e_1 (not shown). Alternatively, one can simply note that for each of the four e_2 solutions of (18) the first two equations of system (17) yield a quartic equation in terms of e_1 . This results in 16 pairs (e_1, e_2) . e_3 is determined uniquely from e_1 [e.g. by using the first equation in system (17)]. Therefore, system (17) can have at most 16 distinct solutions [sets (e_1, e_2, e_3)]. As shown in the above section, $2N = 8$ solutions correspond to the same structure. Therefore, system 17 can yield at most two non-congruent structures. We will investigate the presence of this ambiguity for different values of the intensities I_1, I_2 and I_3 .

Note that for $N = 4$ the choice of the origin given by (4) yields $\text{Im}(e_2) = 0$. Because $|x_j| = 1$, the choice of the origin yields that $\xi_j \xi_k$ is a complex conjugate of $\xi_p \xi_q$, where all j, k, p, q are distinct. This directly results in $\text{Im}(e_2) = 0$ from the definition of e_2 . For $I_1 = 2$, equation (18) becomes

$$e_2^2(4 - \frac{3}{2}I_2 + I_3) = \frac{3}{4}(4 - I_2)^2. \quad (19)$$

Because e_2 is real, $8 - 3I_2 + 2I_3 \geq 0$ [or $I_2 \leq (8 + 2I_3)/3$] and, at most, two values of e_2 can be obtained in this case. Each value of e_2 yields at most four pairs of (e_1, e_3) , i.e. for $I_1 = 2$, at most eight solutions are possible. This means that for $I_1 = 2$, the structure obtained from the three lowest-resolution intensities is unique.

For $I_1 < 2$, the first coefficient of (18) is negative and, because the third coefficient is always positive, the product of the two solutions e_2^2 is negative, by Vieta's theorem. The negative e_2^2 does not yield a physically meaningful structure since e_2 is real. Therefore, for $I_1 < 2$, similarly to $I_1 = 2$, there is at most one possible structure, i.e. there is no ambiguity.

In the last case, $I_1 > 2$, the two solutions e_2^2 are either both positive or both negative. The latter case yields no physically

meaningful structures and the former case yields four values of e_2 and two distinct structures. By Vieta's theorem, two structures are possible only for $(3/2)(I_1^2 - I_2)(1 - I_1) + I_1(I_1 - 3)^2 - I_3 < 0$ and $I_1 > 2$. This result is also important because it states that no more than two homometric (excluding congruent) structures can be obtained for $N = 4$ in one dimension even when one uses more than a minimum (or even a complete) set of intensities. Examples of pairs of homometric structures for $N = 4$ exist (Patterson, 1944). Therefore, we have proven that the minimum set of intensities is sufficient for determining this pair of homometric structures. If the two structures determined this way are indeed homometric, all intensities beyond the minimum set can be determined from those in the minimum set, as polynomial functions. In addition, we have proven that no more than two homometric structures are possible for $N = 4$ as using more intensities cannot increase the number of solutions. It remains to be investigated, however, whether the two non-congruent structures obtained from the minimum intensity set for $N = 4$ are always homometric.

These results indicate that a region of the intensity space exists where a structure can be uniquely determined from the minimal number of intensities even as N increases. Therefore, the number of structures that can be obtained generally depends on the structure itself, which yields these intensities.

4. Discussion

Determining a crystal structure from diffraction intensity data has been a central problem in crystallography for almost 80 years. It has been appreciated that even a complete intensity data set is generally insufficient for determining a structure of N atoms unambiguously (Patterson, 1944, 1939). Direct methods rely on a large intensity data set greatly exceeding the number of unknown atomic coordinates, to yield a chemically meaningful solution. Nevertheless, structures of as many as several hundreds of atoms have been determined by the direct methods. These successes suggest that the number of structures that yield the same set of intensities (minimum or not) cannot increase too fast with the increasing structure size (N). If such number of structures always increased very rapidly, e.g. exponentially with increasing N , then no practically obtainable set of intensities could be sufficient by $N = 100$ or so. Therefore this multiplicity of a crystal structure determined from intensities merits investigation. In the ideal case of a sufficiently large number of intensities, this task is equivalent to the difficult problem of enumeration of homometric structures.

In order to be able to obtain a structure larger than 100 atoms directly from the intensities, one would need to use a deterministic algebraic approach as the existing direct methods are no longer applicable. In addition, the phase problem generally becomes less overdetermined as the structure size N increases, since macromolecular crystals generally do not diffract at high resolutions normally observed for small-molecule crystals. Nevertheless, with bright synchrotron X-ray beam sources available today, intensity

data sets collected on macromolecular crystals at a relatively modest resolution (near 2 Å) are sufficiently overdetermined so that even the data at the highest resolution of the minimum data set are measured with little uncertainty, important for an algebraic approach. To illustrate this point by an example from protein crystallography, we consider a recently reported structure of the N-terminal domain of *Mycobacterium tuberculosis* DnaB helicase determined by our group (Biswas & Tsodikov, 2008). The structure contains $N = 6697$ non-H atoms in the asymmetric unit, *i.e.* $3N = 20091$ unknown coordinates. The minimum set of 20090 intensities is contained within the resolution of approximately 3 Å, considerably lower than the refinement cutoff of 1.9 Å. At the resolution of 3 Å, the relative uncertainty of the measured intensities is only 5%; therefore, most of the data of the minimal intensity data set are very well measured. Therefore the accuracy of a minimum data set should not be a serious issue in the implementation of an algebraic approach at resolutions that are far too low for the direct methods. Nevertheless, the error of the intensity measurements is expected to lead to an increase in the ambiguity (or quasi-homometry) (Fischer *et al.*, 2005). Detailed studies of the stability of the structure solution obtained from the minimum intensity data set with respect to the experimental error in such data are yet to be carried out. Another potential complication of an algebraic method is that a small number of lowest-resolution intensities are not available owing to a blocked direct beam. We envision that if a deterministic approach is developed, the missing intensities could be used as fitting parameters of an optimization routine, in which more than a minimum of intensities are used or the lack of these intensity data is compensated by a few chemical constraints, *i.e.* known covalent bond lengths and bond angles.

Therefore, it appears that the question of homometry is still one of the most critical ones in addressing the feasibility of an algebraic approach in macromolecular crystallography. We set out to address this question starting with the simplest system of a one-dimensional crystal of a small number of identical atoms and a minimum of intensity data. We obtain that the smallest structures of identical atoms, $N = 2$ and $N = 3$, are determined uniquely (discounting the congruency) from the minimum set of $N - 1$ lowest-resolution intensities. The multiplicity begins at $N = 4$, where at most two solutions are possible. This leads to important corollaries that (i) at most two homometric structures can be obtained for $N = 4$ (from a complete set of intensities!) and (ii) the minimum set of intensities is sufficient for their determination. Interestingly, a four-atom structure can still be uniquely determined even from a minimum set of intensities in a certain region of the intensity space (*i.e.* a structure coordinate space). This result is important as it indicates that a unique structure can be obtained even for larger N and for higher dimensions and that the ambiguity need not always be overwhelming. Nevertheless, we demonstrate that the ambiguity is fundamental to the structure determination directly from the intensity data

and should be a serious consideration in developing a practical method of structure determination algebraically directly from the intensities. The structure multiplicity for more complex cases of increasing N and space dimensionality is a subject of ongoing investigation in this group. In all the cases considered, we reduced the structure determination problem to a univariate polynomial that can be solved either by methods of elementary algebra or by routine numerical techniques. The method of symmetric polynomials systematically developed in this study significantly simplifies algebraic manipulations and holds promise for further analysis in cases when more than a minimum set of intensities is considered and in higher dimensions.

We are indebted to Dr Karl Fischer for his critical reading of the manuscript and for his numerous insightful (and in many instances eye-opening) comments and suggestions.

References

- Avrami, M. (1939). *Z. Kristallogr.* **100**, 381–393.
 Biswas, T. & Tsodikov, O. V. (2008). *FEBS J.* **275**, 3064–3071.
 Bullough, R. K. (1961). *Acta Cryst.* **14**, 257–268.
 Bullough, R. K. (1964). *Acta Cryst.* **17**, 295–308.
 Cervellino, A. & Ciccariello, S. (1999). *Z. Kristallogr.* **214**, 739–750.
 Cervellino, A. & Ciccariello, S. (2005). *Acta Cryst.* **A61**, 494–500.
 Dauter, Z., Dauter, M., de La Fortelle, E., Bricogne, G. & Sheldrick, G. M. (1999). *J. Mol. Biol.* **289**, 83–92.
 Fischer, K. F., Kirfel, A. & Zimmermann, H. (2005). *Z. Kristallogr.* **220**, 643–656.
 Fischer, K. F. & Pilz, K. (1997). *Acta Cryst.* **A53**, 475–483.
 Grimm, U. & Baake, M. (2008). *Z. Kristallogr.* **223**, 777–781.
 Hauptman, H. (1997). *Curr. Opin. Struct. Biol.* **7**, 672–680.
 Hauptman, H. & Karle, J. (1951). *Acta Cryst.* **4**, 383.
 Karle, J. & Hauptman, H. (1956). *Acta Cryst.* **9**, 635–651.
 Littlewood, D. E. (1950). *The Theory of Group Characters and Matrix Representations of Groups*, 2nd ed. Oxford: Clarendon Press.
 Macdonald, I. G. (1979). *Symmetric Functions and Hall Polynomials*. Oxford University Press, Clarendon Press.
 Miller, R., DeTitta, G. T., Jones, R., Langs, D. A., Weeks, C. M. & Hauptman, H. A. (1993). *Science*, **259**, 1430–1433.
 Miller, R., Gallo, S. M., Khalak, H. G. & Weeks, C. M. (1994). *J. Appl. Cryst.* **27**, 613–621.
 Ott, H. (1927). *Z. Kristallogr.* **66**, 136–153.
 Patterson, A. L. (1939). *Nature (London)*, **143**, 939–940.
 Patterson, A. L. (1944). *Phys. Rev.* **65**, 195–201.
 Pauling, L. & Shappell, M. D. (1930). *Z. Kristallogr.* **75**, 128–142.
 Pilz, K. & Fischer, K. F. (1998). *Acta Cryst.* **A54**, 273–282.
 Pilz, K. & Fischer, K. F. (2000). *Z. Kristallogr.* **215**, 640–649.
 Rau, V. G., Iliukhin, V. V. & Belov, N. V. (1979). *Dokl. Akad. Nauk SSSR+*, **248**, 363–365.
 Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* **D58**, 1772–1779.
 Sheldrick, G. (1984). *Acta Cryst.* **A40**, C440.
 Usón, I. & Sheldrick, G. M. (1999). *Curr. Opin. Struct. Biol.* **9**, 643–648.
 Weeks, C. M., Adams, P. D., Berendzen, J., Brunger, A. T., Dodson, E. J., Grosse-Kunstleve, R. W., Schneider, T. R., Sheldrick, G. M., Terwilliger, T. C., Turkenburg, M. G. W. & Uson, I. (2003). *Macromol. Crystallogr. D*, **37A**, 37–83.
 Zimmermann, H. & Fischer, K. F. (2009). *Acta Cryst.* **A65**, 443–455.
 Zobetz, E. (1993). *Z. Kristallogr.* **207**, 209–222.