

The Effect of Atoms in Fixed Positions on the Statistical Tests for Symmetry Centers*

By ROBERT L. COLLIN

Cancer Research Institute, New England Deaconess Hospital, Boston 15, Massachusetts, U.S.A.

(Received 11 February 1955 and in revised form 2 March 1955)

The probability distribution of X-ray intensities is altered by the presence of atoms in fixed positions. If these atoms are heavy enough, they will in turn affect the statistical tests that are used to distinguish centric from acentric structures. However, if the positions of these fixed atoms are known, distribution curves can be calculated for both centric and acentric structures. A comparison of these curves with the observed intensity distribution will allow one to infer the presence or absence of a symmetry center.

Introduction

Wilson (1949) has derived the probability distribution of X-ray intensities† for crystals that contain only atoms in random positions. This distribution (which we will refer to as a random-position distribution) has been tested and put into a practical form for distinguishing centric from acentric crystals by Howells, Phillips & Rogers (1950).

If a crystal contains some atoms that occupy fixed positions‡, and if these atoms are heavy enough, they will alter the intensity distribution of the X-ray reflections. That this effect is to be expected, has been pointed out previously by both Wilson (1949) and Howells *et al.* (1950). A more exact derivation of the probability distribution has been carried out for both the random-position case and the case where some atoms are in fixed positions by Karle & Hauptman (1953) and Hauptman & Karle (1953). They have derived the probability distribution of the structure-factor magnitudes in terms of infinite series from which the probability function can be computed to any desired accuracy. In this paper we will make use of approximate distributions, similar to those of Wilson, which are invalid at very low values of $\sin \theta$ but sufficiently accurate at higher values to allow a study to be made of the effect of the fixed-position atoms on the test used to distinguish centric from acentric crystals.

Intensity distribution

When a crystal contains atoms in fixed positions, these atoms will contribute a constant term to the tri-

* This work was done under United States Atomic Energy Commission Contract AT(30-1)-901 with the New England Deaconess Hospital.

† The word *intensity* is used here to denote the observed intensity corrected for geometric and polarization effects, and thus is the quantity that is compared with the square of the structure factor.

‡ An atom in a fixed position is one that has no variable position parameter. When one is considering only a zone of reflections then an atom will be thought of as in a fixed position if it has no variable position parameter with respect to that zone.

gonometric part of the structure factor expression for a large fraction, if not all, of the reflections. For example, in cysteylglycine sodium iodide (Dyer, 1951) the iodine and sodium atoms are in two sets of special positions of the space group $A2$ and the $(h0l)$ projection is required to have a center. The space group extinguishes reflections with l even and it is convenient, for this projection, to choose a pseudo-projected cell with sides a and $\frac{1}{2}c$. In this way we do away with the need to include the systematically absent reflections in the statistical treatment. For the reflections that are allowed, the sodium plus iodine contribution is $f_I - f_{Na}$ for reflections with h odd, and $f_I + f_{Na}$ for reflections with h even.

In the analysis to follow the reflections will be divided up into groups so that each group will have a constant contribution from the atoms in fixed positions. The expressions that are derived for the intensity distribution will then apply only within this one group of reflections. In the example above the reflections would be divided into two groups, one with reflections where h is even and the other with reflections where h is odd. The intensity distribution would have to be worked out separately for each group and then averaged, since the distributions are different in each case.

Centric crystal

For a centric structure with N atoms per unit cell in general positions and M atoms per unit cell in fixed positions, if the reflections are grouped in the manner described above, we can write the structure factor for a reflection, hkl , within any group as

$$F = 2 \sum_{j=1}^{N/2} f_j \cos 2\pi(hu_j + kv_j + lw_j) + F_s, \quad (1)$$

where $F_s = \sum_{j=1}^{M/2} f_j A_j$, the contribution to the structure factor of the atoms in fixed positions, and where A_j is independent of h , k and l . u_j , v_j , w_j are the position parameters and f_j is the atomic scattering factor of

the j th atom. Following Wilson's (1949) argument in his use of the central limit theorem, we arrive at the relation

$$P(|F|)d|F| = \frac{1}{\sqrt{(2\pi\Sigma')}} \left[\exp \left\{ -\frac{(|F|-F_s)^2}{2\Sigma'} \right\} + \exp \left\{ -\frac{(-|F|-F_s)^2}{2\Sigma'} \right\} \right] d|F|, \quad (2)$$

where $\Sigma' = 2 \sum_{j=1}^{N/2} f_j^2$, and where we have written $|F|$

to denote the magnitude of the structure factor and $P(|F|)$ the probability that F lies between $|F|$ and $|F|+d|F|$. This is the first term in the probability distribution of the structure factor magnitude derived by Karle & Hauptman (1953). We can now rearrange (2) and at the same time replace $|F|$ by \sqrt{I} . This will give the desired expression for the probability distribution of the intensity:

$$P(I)dI = \frac{1}{\sqrt{(2\pi\Sigma'I)}} \left[\exp \left\{ -\frac{F_s\sqrt{I}}{\Sigma'} - \frac{I}{2\Sigma'} \right\} + \exp \left\{ \frac{F_s\sqrt{I}}{\Sigma'} - \frac{I}{2\Sigma'} \right\} \right] \exp \left\{ -\frac{F_s^2}{2\Sigma'} \right\} dI. \quad (3)$$

Acentric crystal

In an acentric structure the fixed-position atoms will be related to each other either in a centric or acentric manner. With the fixed-position atoms in a centric array the mathematics is much less complex, and, since most structures will have the fixed-position atoms in this arrangement, we will work out the relations only for this case. The results will also apply to an acentric structure where the fixed-position atoms can be placed at the eight corners of the unit cell, that is, at a position with all parameters zero.

The complex structure factor will have a real and an imaginary part. The real part of the structure factor can be written as

$$X = x + x_s,$$

and the imaginary part as

$$Y = y,$$

where x and y are the real and imaginary contributions to the structure factor of the atoms in random positions and x_s is the contribution of the atoms in fixed positions. Following Wilson (1949), the probability distributions for X and Y can be written as

$$P(X)dX = \frac{1}{\sqrt{(\pi\Sigma')}} \exp \left[-\frac{(X-x_s)^2}{\Sigma'} \right] dx, \quad (4)$$

and

$$P(Y)dY = \frac{1}{\sqrt{(\pi\Sigma')}} \exp \left[-\frac{Y^2}{\Sigma'} \right] dY, \quad (5)$$

where Σ' is Σf_j^2 with the sum taken over all the random-position atoms. If we denote by $P(I)_X$ the

probability distribution of the intensity for a fixed value of X we can write

$$P(I)_X dI = \pi P(X, Y) dX dY \\ = \frac{1}{\Sigma'} \exp \left\{ -\frac{(I+x_s^2)}{\Sigma'} \right\} \exp \left\{ \frac{2x_s X}{\Sigma'} \right\} dI, \quad (6)$$

where we have let $X^2 + Y^2 = I$.

In order to find the probability distribution of the intensity as X varies over its allowable range, $-\sqrt{I}$ to $+\sqrt{I}$, we must evaluate the expression

$$P(I)dI = \frac{\int_{-\sqrt{I}}^{\sqrt{I}} P(I)_X P(X) dX}{\int_{-\sqrt{I}}^{\sqrt{I}} P(X) dX} dI. \quad (7)$$

If we substitute (4) and (6) into (7) and integrate we obtain the rather complex expression

$$P(I)dI = \frac{1}{\Sigma'} \exp \left\{ -\frac{(I-2x_s^2)}{\Sigma'} \right\} \times \\ \frac{\left\{ \operatorname{erf} \left(\frac{\sqrt{I+2x_s}}{\sqrt{\Sigma'}} \right) \pm \operatorname{erf} \left(\frac{\sqrt{I-2x_s}}{\sqrt{\Sigma'}} \right) \right\}}{\left\{ \operatorname{erf} \left(\frac{\sqrt{I+x_s}}{\sqrt{\Sigma'}} \right) \pm \operatorname{erf} \left(\frac{\sqrt{I-x_s}}{\sqrt{\Sigma'}} \right) \right\}} dI \quad \begin{cases} + & \text{for } \sqrt{I} \geq 2x_s \\ - & \text{for } \sqrt{I} < 2x_s \\ + & \text{for } \sqrt{I} \geq x_s \\ - & \text{for } \sqrt{I} < x_s \end{cases} \quad (8)$$

The sign used between the error functions (erf) depends on the relative magnitudes of \sqrt{I} and x_s , as shown at the right of the equation. Tables of the error function may be found in Jahnke & Emde (1938).

$N(z)$ distributions

In order to apply (3) and (8) to practical situations it is convenient to transform the expressions somewhat. Howells *et al.* (1950) have introduced the function $N(z)$, where $z = I/\langle I \rangle$, and where $\langle I \rangle$ denotes the average value of the intensity, for the case where all atoms are in random positions. (From (2) it can be shown that $\langle I \rangle = \Sigma' + F_s^2$.) $N(z)$ is the percent reflections with intensity less than or equal to $z\langle I \rangle$. A plot of $N(z)$ versus z for the centric and acentric cases gives distinctly different curves. We will also evaluate $N(z)$ when atoms occur at fixed positions and compare the results with those for structures with random-position atoms only.

First, we write the probability distributions of the intensity in terms of z . For the centric and acentric cases we have,

$$\text{Centric: } P(z)dz = \frac{1}{2} \sqrt{(a/2\pi)} \exp \left\{ -\frac{1}{2}w \right\} \\ \times \left[\exp \left\{ -\sqrt{(awz)} - \frac{1}{2}az \right\} + \exp \left\{ \sqrt{(awz)} - \frac{1}{2}az \right\} \right] dz/\sqrt{z}, \quad (9)$$

where $a = (\Sigma' + F_s^2)/\Sigma'$ and $w = F_s^2/\Sigma'$.

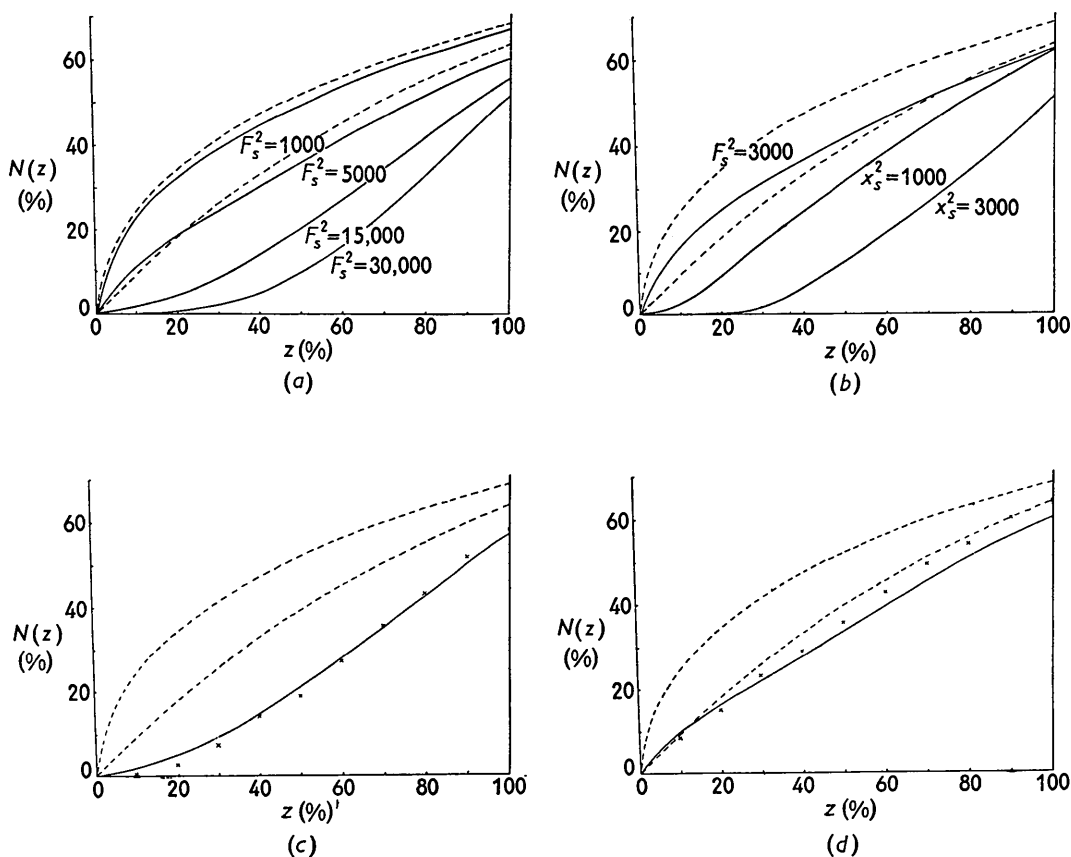


Fig. 1. (a) $N(z)$ curves for centric structures with $\Sigma' = 1500$ and four different values of F_s^2 (solid lines). The random-position curves for centric and acentric structures are shown broken. (b) $N(z)$ curves for acentric structures with $\Sigma' = 1500$ and $x_s^2 = 1000$ and 3000 . The curve for a centric structure with $\Sigma' = 1500$ and $F_s^2 = 3000$ as well as the centric and acentric curves (broken) for the random-position structures are included for comparison. (c) Calculated $N(z)$ curve for platinum phthalocyanine (solid line) and the calculated curves for the random-position structures (broken lines). The points computed from the intensity data for platinum phthalocyanine are shown as \times . (d) Calculated $N(z)$ curve for cysteylglycine sodium iodide (heavy line) and the calculated curves for the random-position structures (broken lines). The points computed from the intensity data for cysteylglycine sodium iodide are shown as \times .

Acentric: $P(z)dz = a \exp \{-(az-2w)\}$

$$\times \frac{\operatorname{erf}(\sqrt{az} + 2\sqrt{w}) \pm \operatorname{erf}(\sqrt{az} - 2\sqrt{w})}{\operatorname{erf}(\sqrt{az} + \sqrt{w}) \pm \operatorname{erf}(\sqrt{az} - \sqrt{w})} dz \begin{cases} + & \text{for } \sqrt{az} \geq 2\sqrt{w} \\ - & \text{for } \sqrt{az} < 2\sqrt{w} \\ + & \text{for } \sqrt{az} \geq \sqrt{w} \\ - & \text{for } \sqrt{az} < \sqrt{w} \end{cases} \quad (10)$$

where $a = (\Sigma' + x_s^2)/\Sigma'$ and $w = x_s^2 \Sigma'$.

The function we are interested in, $N(z)$, is given by

$$N(z) = \int_0^z P(z) dz \Big/ \int_0^\infty P(z) dz. \quad (11)$$

For the centric case the integrals above can be evaluated exactly and lead to

$$N(z) = \frac{1}{2} \left\{ \operatorname{erf} \left[\sqrt{\frac{1}{2}az} + \sqrt{\frac{1}{2}w} \right] \pm \operatorname{erf} \left[\sqrt{\frac{1}{2}az} - \sqrt{\frac{1}{2}w} \right] \right\} \begin{cases} - & \text{for } \sqrt{\frac{1}{2}az} < \sqrt{\frac{1}{2}w} \\ + & \text{for } \sqrt{\frac{1}{2}az} \geq \sqrt{\frac{1}{2}w} \end{cases}. \quad (12)$$

The integrals for the acentric case are much more

complicated but can be evaluated by graphical integration for specific cases.

As an illustration of the general form of the centric function, curves have been calculated for the case where $\Sigma' = 1500$ and F_s^2 varies from 1000 to 30,000. These are plotted in Fig. 1(a) along with the centric and acentric curves for the random-position case. The important point to notice is that under certain conditions a centric structure with atoms in fixed positions may give a distribution curve indistinguishable from the acentric curve calculated for a structure with all atoms in random positions.

The acentric $N(z)$ curve with fixed-position atoms is illustrated in Fig. 1(b) where a curve for $\Sigma' = 1500$ and $x_s^2 = 3000$ has been calculated by numerical integration of (11) with values of $P(z)$ given by (10). In addition, the centric curve for $\Sigma' = 1500$ and $F_s^2 = 3000$ is reproduced for comparison. The two curves differ significantly and both are very different from the random-position curves for centric and acentric structures.

Application to actual structures

The above equations and curves show that when the atoms that are in fixed positions are not outstandingly heavy compared with the other atoms in the structure, random-position statistics apply well enough to distinguish centric from acentric crystals. It is only when the atom at the fixed position is 'heavy' that the modifications introduced in this paper will be of practical importance.

Practical tests of the relation for $N(z)$ derived here are not as easy to carry out as in the case where all atoms are in random positions. The quantities a and w are not strictly constant but depend on $\sin \theta$. For a rough calculation one may evaluate a and w by using atomic numbers in place of atomic scattering factors, but, when more accurate results are needed, the best approach seems to lie in the calculation of separate $N(z)$ distributions for each small region of $\sin \theta$, weighting these with the number of reciprocal-lattice points in each region, and then taking an average of the weighted $N(z)$ values. Data for two structures with heavy atoms in fixed positions, platinum phthalocyanine (Robertson & Woodward, 1940) and cysteylglycine sodium iodide (Dyer, 1951), have been taken from the literature for a test of the applicability of the proposed $N(z)$ distributions.

Platinum phthalocyanine crystallizes with the space group $C_{2h}^5-P2_1/a$ and has two molecules in a unit cell. The [010] projection is required by the space group to be centric. Because of the systematic absences introduced by the space group it is convenient to transform to a pseudo-projected cell which contains one molecule and has an a axis one-half the length of the true a axis. The platinum atom may be placed at the origin and then $F_s = f_{Pt}$ for all reflections. The observed $N(z)$ distribution and the curve calculated from (12) along with the random-position curves for the centric and acentric cases are shown in Fig. 1(c). In evaluating (12) the change of a and w with $\sin \theta$ was taken into account by the method described above.

The comparison of the observed and calculated distributions for cysteylglycine sodium iodide is somewhat more difficult and the procedure outlined in the first part of the second section was followed, allowing for variations of a and w with $\sin \theta$ in the usual way. The $N(z)$ values for both groups of reflections are averaged to give the $N(z)$ distribution plotted in Fig. 1(d). This is a good example of a centric projection from a crystal with heavy atoms in fixed positions that gives an $N(z)$ distribution which would be mistaken for an acentric distribution.

After this article was submitted for publication an article appeared (Hargreaves, 1955) in which the effect of a heavy atom in a general position on the intensity distribution was considered. The relation between the present work and that of Hargreaves is difficult to

point out because in the present case the reflections are divided into groups which receive constant contributions from the fixed-position atoms and the $N(z)$ function is worked out within a group only, while in Hargreaves' work all reflections are considered together. The sum of the $N(z)$ functions of the individual groups is not the $N(z)$ function of all the groups considered together.

For a structure where the fixed atom contributes a constant term to all reflections the limiting expression for $N(z)$ can be formed by taking the limit in (9) as $\Sigma' \rightarrow 0$. The resulting function is the same as $\max N(z)$ of Hargreaves. If we now consider structures similar to the other two that Hargreaves discusses, and if we put the heavy atoms at fixed positions, a look at the numerical values of the structure factors will show that as the fixed atom is moved into positions where the number of reflections with constant contribution decreases and approaches a general position, the $N(z)$ curves for all reflections grouped together approach those of Hargreaves.

Concluding remarks

When applying the statistical test to determine whether an unknown structure is centric or acentric care must be exercised if atoms occur in fixed positions. If such a situation exists, it is usually apparent as soon as the unit-cell dimensions and space-group possibilities are known. $N(z)$ curves for both centric and acentric cases can then be calculated for the particular case of interest and compared with the observed $N(z)$ obtained from the intensity data.

The average value of the intensity, which is important for establishing an absolute scale, is not affected abnormally by fixed-position atoms and can be calculated in the same manner as for random structures.

In certain cases the methods outlined here may be of valuable aid in deciding between various fixed-position possibilities for heavy atoms, although it would seem that an examination of a Patterson projection would be more enlightening and would not involve too much more effort.

References

- DYER, H. B. (1951). *Acta Cryst.* **4**, 42.
 HARGREAVES, A. (1955). *Acta Cryst.* **8**, 12.
 HAUPTMAN, H. & KARLE, J. (1953). *Acta Cryst.* **6**, 136.
 HOWELLS, E. R., PHILLIPS, D. C. & ROGERS, D. (1950). *Acta Cryst.* **3**, 210.
 JAHNKE, E. & EMDE, F. (1938). *Tables of Functions*. Leipzig: Teubner.
 KARLE, J. & HAUPTMAN, H. (1953). *Acta Cryst.* **6**, 131.
 ROBERTSON, J. M. & WOODWARD, I. (1940). *J. Chem. Soc.* p. 36.
 WILSON, A. J. C. (1949). *Acta Cryst.* **2**, 318.