The MTF based on the double-weighted Cu–Cl vector gave the correct position of the Cu atom and the trivial peak at $(\mathbf{r}_{Cu}-\mathbf{r}_{Cl})/2$ as top peaks.

References

ALDEN, R. A., STOUT, G. H., KRAUT, J. & HIGH, D. F. (1964). *Acta Cryst.* **17**, 109–121.

EGERT, E. & SHELDRICK, G. M. (1985). *Acta Cryst.* A**41**, 262–268.

HIGH, D. F. & KRAUT, J. (1966). *Acta Cryst.* **21**, 88–96.

HUBBARD, C. R., BABICH, M. W. & JACOBSON, R. A. (1977). *A PL/1 Program System for Generalized Patterson Superposition.* Ames Laboratory-ERDA. Iowa State Univ., Ames, USA.

JACOBSON, R. A. & BECKMAN, D. E. (1979). *Acta Cryst.* A**35**, 339–340.

KETTMANN, V., PAVELČÍK, F. & RYBÁR, A. (1987). In preparation.

KIM, S. H., JEFFREY, G. A., ROSENSTEIN, R. D. & CORFIELD, P. W. R. (1967). *Acta Cryst.* **22**, 733–743.

LUGER, P. & FUCHS, J. (1986). *Acta Cryst.* A**42**, 380–386.

PAVELČÍK, F. (1986). *J. Appl. Cryst.* **19**, 488–491.

PAVELČÍK, F. & HAVETTA, K. (1988). In preparation.

PAVELČÍK, F., ŽEMLIČKA, M., KETTMANN, V. & KRÄTSMÁR-ŠMOGROVIČ, J. (1987). *Chem. Pap.*, **41**, 433–440.

RABINOWITZ, I. N. & KRAUT, J. (1964). *Acta Cryst.* **17**, 159–168.

SHELDRICK, G. M. (1985*a*). In *Crystallographic Computing* 3, edited by G. M. SHELDRICK, C. KRUGER & R. GODDARD, pp. 175–189. Oxford Univ. Press.

SHELDRICK, G. M. (1985*b*). *J. Mol. Struct.* **130**, 9–16.

SIMPSON, P. G., DOBROTT, R. D. & LIPSCOMB, W. N. (1965). *Acta Cryst.* **18**, 169–179.

TERWILLIGER, T. C., KIM, S.-H. & EISENBERG, D. (1987). *Acta Cryst.* A**43**, 1–5.

ULICKÁ, L., PAVELČÍK, F. & HUML, K. (1987). *Acta Cryst.* C**43**, 2266–2268.

VRÁBEL, V., PAVELČÍK, F., KELLÖ, E., MIERTUŠ, S., KONEČNÝ, S. & LOKAJ, J. (1987). *Collect. Czech. Chem. Commun.* **52**, 692–706.

Acta Cryst. (1988). A**44**, 729–735

A Priori Estimation of Scale and Overall Anisotropic Temperature Factors from the Patterson Origin Peak

BY ROBERT H. BLESSING AND DAVID A. LANGS

Medical Foundation of Buffalo, 73 High Street, Buffalo, New York 14203, *USA*

(*Received* 1 *October* 1987; *accepted* 26 *April* 1988)

Abstract

An idea due to D. Rogers [*Computing Methods in Crystallography* (1965), edited by J. S. Rollett, pp. 117–148. Oxford: Pergamon Press] has been developed and implemented. The method is an advantageous alternative to Wilson plot or K-curve scaling of intensity data. On the relative experimental scale the structure factor can be written in matrix notation as $F(\mathbf{h}) = k^{-1}\sum_j f_j(\mathbf{h}) \exp(2\pi i \mathbf{h}^T\mathbf{x}_j) \exp(-\mathbf{h}^T\mathbf{b}_j\mathbf{h})$; and the squared structure-factor magnitude can be written as $|F(\mathbf{h})|^2 = k^{-2}\exp(-2\mathbf{h}^T\mathbf{b}\mathbf{h})\{\sum_j f_j^2 + 2\sum_j\sum_{k>j} f_j f_k \exp[2\pi i \mathbf{h}^T(\mathbf{x}_j-\mathbf{x}_k)]\}$, if a common, or average, anisotropic temperature factor is factored out of the atomic summations. The f_j^2 summation corresponds to the Patterson origin peak, and the f_jf_k double summation to the off-origin Patterson peaks. A trivariate Gaussian density function, $P(\mathbf{u}) - P_{min} = p_0 \exp(-\mathbf{u}^T\mathbf{p}\mathbf{u})$, is fitted by least squares to the origin peak from a Patterson synthesis with coefficients $|F|^2_{meas}/\sum_j f_j^2$. Fourier inversion of the fitted Gaussian gives the scale and thermal parameters, $k^2 = (\det \mathbf{p})^{1/2}/(\pi^{3/2}V_{cell}P_0)$ and $\mathbf{b} = (\pi^2/2)\mathbf{p}^{-1}$. The fit of the parameter P_{min} is constrained by the condition that $P_{min} = -F(000)^2/(k^2 V_{cell}\sum_j Z_j^2)$, and thus only p_0 and the six coefficients p_{ij} $(i<j=1,2,3)$ are independent parameters.

Background

One of the first steps in an X-ray crystal structure analysis is the estimation of the scale and overall temperature factor for the diffraction intensities. Over the years, several methods for doing this have been devised (Rogers, 1965, 1980; Levy, Thiessen & Brown, 1970; Ladd, 1978; Giacovazzo, 1980; Subramanian & Hall, 1982; Hall & Subramanian, 1982).

The Wilson plot

The widely employed method of Wilson (1942) takes advantage of the cosine form of the product of the structure factor with its complex conjugate (Patterson, 1935). In matrix notation,

$$F(\mathbf{h}) = \sum_j f_j(s) \exp(-B_j s^2) \exp(2\pi i \mathbf{h}^T\mathbf{x}_j), \quad (1)$$

0108-7673/88/050729-07$03.00

where $s = (\sin\theta)/\lambda$ and $j = 1, 2, \ldots, n$ atoms per unit cell, and

$$|F(\mathbf{h})|^2 = \sum_j f_j^2 \exp(-2B_j s^2) + 2\sum_j \sum_{k>j} f_j f_k \exp[-(B_j + B_k)s^2] \times \cos[2\pi \mathbf{h}^T(\mathbf{x}_j - \mathbf{x}_k)]. \quad (2)$$

The form of (1) and (2) assumes that the atomic thermal vibrational displacements are isotropic. If it is further assumed that the individual atomic thermal parameters B_j can be fairly approximated by a common or average B, and if the scale factor k for the relative measurements of $|F(\mathbf{h})|^2$ is explicitly noted, then (2) becomes

$$k^2|F|^2_{\text{meas}} = \exp(-2Bs^2)\left\{\sum_j f_j^2 + 2\sum_j \sum_{k>j} f_j f_k \times \cos[2\pi \mathbf{h}^T(\mathbf{x}_j - \mathbf{x}_k)]\right\}.$$

If this is averaged in shells of s, the sum in cosine terms goes to zero, and one obtains

$$\left\langle |F|^2_{\text{meas}} / \varepsilon \sum_j f_j^2 \right\rangle_s \simeq k^{-2} \exp(-2B\langle s^2\rangle),$$
$$\ln\left\langle |F|^2_{\text{meas}} / \varepsilon \sum_j f_j^2 \right\rangle_s \simeq -2\ln k - 2B\langle s^2\rangle, \quad (3)$$

which provides estimates of k and B from the intercept and slope of a plot of the logarithms of the shell-averaged normalized intensities against the averaged $(\sin\theta)^2$.

The quantity $\varepsilon = \varepsilon(\mathbf{h})$ in (3) is a projection symmetry multiplier* that is introduced to account for the enhanced average intensity in certain zones and rows due to superposition of symmetry-equivalent atoms in projection. For general reflections $\varepsilon = 1$, and for special reflections $\varepsilon = 2, 3, 4, 6, 8$, or 12 (Iwasaki & Ito, 1977). For example, $\varepsilon = 2$ for $h0l$ and $0k0$ reflections in the monoclinic point group $2/m$.

The K curve

Karle & Hauptman (1953) devised a composite scale and temperature factor function,

$$K(s) = \left\langle \varepsilon \sum_j f_j^2(s) / |F|^2_{\text{meas}} \right\rangle_s \simeq k^2 \exp(2B\langle s^2\rangle),$$

for estimating normalized structure factor amplitudes, $|E(\mathbf{h})| = [|F(\mathbf{h})|^2_{\text{meas}}/K(s)]^{1/2}$. The extrapolated value of K at $s = 0$ gives the scale factor, but no explicit evaluation of the temperature factor is needed to estimate the $|E|$ values.

Anisotropic extension of the Wilson plot

Maslen (1967) devised an approximate method to allow for thermal anisotropy by means of a second series of Wilson plots against the Miller index products h^2, k^2, l^2, hk, kl, hl following a scaling of the intensities from a first Wilson plot against s^2.

Least-squares refinement method

Very recently, Sheriff & Hendrickson (1987) published a straightforward method for estimating the scale and overall anisotropic thermal parameters. In effect, they use

$$F^2_{\text{calc}}(\mathbf{h}) = k^{-2} \exp(-2\mathbf{h}^T\mathbf{b}\mathbf{h}) \sum_j f_j^2(s)$$

for a least-squares minimization of

$$\chi^2 = \sum_{\mathbf{h}} (|F|^2_{\text{meas}} - F^2_{\text{calc}})^2 / \sigma^2(|F|^2_{\text{meas}})$$

by fitting k and b^{ij} $(i<j = 1, 2, 3)$ for a 'composite atom' at $\mathbf{x} = (0, 0, 0)$ with 'atomic scattering factor' $\sum_j f_j^2(s)$. The authors report significant improvements in the refinement of macromolecular structure models when the overall thermal anisotropy is introduced. Essentially the same method was reported by Levy, Thiessen & Brown (1970), who were faced with the problem of estimating $|E|$ values for a structure for which the thermal attenuation of the diffraction intensities was severely anisotropic.*

Rogers analysis

Rogers (1965) emphasized the correspondence between the terms in the expression for the squared structure-factor magnitude and the peaks in the Patterson function. Allowing for Gaussian distributions of atomic thermal vibrational displacements (Johnson & Levy, 1974), we obtain

$$F(\mathbf{h}) = \sum_j f_j(\mathbf{h}) \exp(2\pi i \mathbf{h}^T\mathbf{x}_j - \mathbf{h}^T\mathbf{b}_j\mathbf{h}), \quad (4)$$

and, if

$$g_j(\mathbf{h}) = f_j(\mathbf{h}) \exp(-\mathbf{h}^T\mathbf{b}_j\mathbf{h}),$$
$$|F(\mathbf{h})|^2 = \sum_j g_j^2 + 2\sum_j \sum_{k>j} g_j g_k \exp[2\pi i \mathbf{h}^T(\mathbf{x}_j - \mathbf{x}_k)].$$

If this is substituted into the Patterson function,

$$P(\mathbf{u}) = V^{-1} \sum_{\mathbf{h}} |F(\mathbf{h})|^2 \exp(-2\pi i \mathbf{h}^T\mathbf{u}),$$

* Rogers (1965, p. 129; 1980) uses the symbol p and the Welsh word *pwys* (plural *pwysau*) meaning power, strength, weight, or importance.

* The work of Levy, Thiessen & Brown (1970) was called to our attention by one of the referees, who reports that the method was adopted in his laboratory and has been in routine use there for about 15 years with excellent results.

one obtains

$$P(\mathbf{u}) = V^{-1}\sum_{\mathbf{h}}\left(\sum_j g_j^2\right)\exp(-2\pi i\mathbf{h}^T\mathbf{u}) + V^{-1}\sum_{\mathbf{h}}\left\{2\sum_j\sum_{k>j} g_jg_k \times\exp[2\pi i\mathbf{h}^T(\mathbf{x}_j-\mathbf{x}_k-\mathbf{u})]\right\}.$$

The terms in g_j^2 give the contribution of $|F(\mathbf{h})|^2$ to the Patterson origin peak, and the g_jg_k terms give the contribution to the off-origin Patterson peaks. Fourier inversion of the origin peak,

$$P_0(\mathbf{u}) = V^{-1}\sum_{\mathbf{h}}\left(\sum_j g_j^2\right)\exp(-2\pi i\mathbf{h}^T\mathbf{u}),$$

gives the expectation of the squared structure-factor magnitude,

$$\langle|F(\mathbf{h})|^2\rangle = \sum_j g_j^2(\mathbf{h}) = V\int_v P_0(\mathbf{u})\exp(2\pi i\mathbf{h}^T\mathbf{u})\,\mathrm{d}^3\mathbf{u}.$$

Implementation of the Rogers analysis

We have written a program to compute a sharpened Patterson density,

$$P(\mathbf{u}) = \frac{1}{V_{\text{cell}}}\sum_{-\infty}^{-\infty}\sum_{\mathbf{h}}\frac{|F(\mathbf{h})|^2}{\sum_j f_j^2(\mathbf{h})}\exp(-2\pi i\mathbf{h}^T\mathbf{u}); \qquad (5)$$

fit a trivariate Gaussian density function,

$$P_0(\mathbf{u}) = p_0\exp(-\mathbf{u}^T\mathbf{p}\mathbf{u}), \qquad (6)$$

around $\mathbf{u}=0$; and calculate the Fourier inverse of this origin peak,

$$\left\langle\frac{|F(\mathbf{h})|^2}{\sum_j f_j^2(\mathbf{h})}\right\rangle = V_{\text{cell}}\int_{-\infty}^{+\infty} P_0(\mathbf{u})\exp(2\pi i\mathbf{h}^T\mathbf{u})\,\mathrm{d}^3\mathbf{u}.$$

The form of $P(\mathbf{u})$ given in (5) is not the same as an $|E|^2$ Patterson, which would have an additional factor $k^{-2}\exp[-2Bs^2(\mathbf{h})]$ in the denominator of the sharpened coefficients. The Gaussian form of $P_0(\mathbf{u})$ given in (6) follows by analogy with (3), assuming a common or average anisotropic thermal vibration tensor $\mathbf{b}$, *i.e.*, for $|F(\mathbf{h})|$ on the relative experimental scale,

$$\left\langle\frac{|F(\mathbf{h})|^2}{\sum_j f_j^2(\mathbf{h})}\right\rangle = k^{-2}\exp(-2\mathbf{h}^T\mathbf{b}\mathbf{h}). \qquad (7)$$

This is a Gaussian, and, therefore, so is its Fourier transform $P_0(\mathbf{u})$.

In practice, the cosine forms of the Patterson function and its inverse are used. That is, since $|F(-\mathbf{h})| = |F(\mathbf{h})|$,

$$P(\mathbf{u}) = \frac{2}{V_{\text{cell}}}\sum_{-\infty}^{+\infty}{}_h\sum_{-\infty}^{+\infty}{}_k\sum_{0}^{+\infty}{}_l\frac{|F(\mathbf{h})|^2}{\sum_j f_j^2(\mathbf{h})}\cos(2\pi\mathbf{h}^T\mathbf{u});$$

and since $P(-\mathbf{u}) = P(\mathbf{u})$,

$$\langle|F(\mathbf{h})|^2/\textstyle\sum_j f_j^2(\mathbf{h})\rangle = 2V_{\text{cell}}\int_0^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} P_0(\mathbf{u})\cos(2\pi\mathbf{h}^T\mathbf{u})\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w.$$

If we substitute the Gaussian [(6)] for $P_0(\mathbf{u})$, this becomes

$$\left\langle|F(\mathbf{h})|^2/\sum_j f_{j(\mathbf{h})}^2\right\rangle = 2V_{\text{cell}}p_0\int_0^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp(-\mathbf{u}^T\mathbf{p}\mathbf{u}) \times\cos(2\pi\mathbf{h}^T\mathbf{u})\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w.$$

The standard form

$$\int_0^{\infty}\exp(-a^2x^2)\cos mx\,\mathrm{d}x = [\pi^{1/2}/(2a)]\exp[-m^2/(2a)^2]$$

gives

$$\langle|F(\mathbf{h})|^2/\textstyle\sum_j f_j^2(\mathbf{h})\rangle = \pi^{3/2}V_{\text{cell}}[p_0/(\det\mathbf{p})^{1/2}] \times\exp(-\pi^2\mathbf{h}^T\mathbf{p}^{-1}\mathbf{h}), \qquad (8)$$

and it follows from (7) that

$$k^2 = (\det\mathbf{p})^{1/2}/(\pi^{3/2}V_{\text{cell}}p_0) \text{ and } \mathbf{b} = (\pi^2/2)\mathbf{p}^{-1}. \qquad (9)$$

The Patterson density is calculated on a hemispherical polar coordinate grid around the origin using coefficients $|F|^2_{\text{meas}}/\sum_j f_j^2$. The radius of the grid is a user-selected multiple, typically 2·5, of the half-height radius of the origin peak, determined as described below. With 25 r-divisions at 30° θ and φ intervals the grid contains $1+25+25\times12\times3=926$ grid points.

The Gaussian density function must be fitted to an adjusted origin peak (Fig. 1),

$$P_0(\mathbf{u}) - P_{\min} = p_0\exp(-\mathbf{u}^T\mathbf{p}\mathbf{u}),$$

where

$$p_{\min} = -F(000)^2\Big/\left(k^2V_{\text{cell}}\sum_j Z_j^2\right) \qquad (10)$$

accounts for the omitted $F(000)$ term in the Patterson–Fourier summation.

Starting values for $P_{\min}$, p_0 and the six p_{ij} ($i\le j=1,2,3$) are obtained from the minimum and maximum values and the half-height radii of the Patterson densities calculated along the a, b and c axes out to a user-selected radius, typically 2·0 Å, at intervals of 1/50 of the radius. The chances of including off-origin interatomic peaks within the radius of the origin peak

fit (Fig. 2) are somewhat reduced by the increasing sparseness of the polar coordinate grid with increasing radius. To deal with peaks that are included, any grid points with

$$P(\mathbf{u}) - P_{\min} < 0{\cdot}5\, p_0$$

and

$$P(\mathbf{u}) - P_{\min} - p_0 \exp(-\mathbf{u}^T\mathbf{p}\mathbf{u}) > 2\sigma,$$

where σ is the root-mean-square deviation of fit of the parameters $P_{\min}$, p_0 and p_{ij}, are down-weighted to a relative weight of 0·1 in the iterative least-squares refinement of the parameters. In this way, a minimum-profile origin peak fit is obtained (Rogers, 1965, p. 144), as indicated in Fig. 2. The fit of the parameter $P_{\min}$ is constrained by (10).

Test calculations

We have tested our program against a number of structures recently studied in our institute. The structures, space groups and unit-cell contents are listed in Table 1. The overall thermal vibration tensors in the form of mean-square displacement parameters,

$$U^{ij} = b^{ij}/(2\pi^2 a^{*i} a^{*j}),$$

are listed in Table 2, and the scale factors and thermal parameters contracted to equivalent isotropic scalars,

$$U_{\text{iso}} = (1/6\pi^2) \sum_i \sum_j \mathbf{a}_i \cdot \mathbf{a}_j b^{ij},$$

$$U_{\text{iso}} = \langle u^2 \rangle_{\text{iso}} = B_{\text{iso}}/(8\pi^2),$$

are listed in Table 3, along with the normalized root-mean-square deviation of fit for the anisotropic Gaussian origin peak.

All the data sets of Tables 1-3, except the monoclinic DL-biotin and the insulin data, were subjected to either a Wilson plot or K-curve scaling at the start of the structure analyses so that, ideally, the refined scale factors should be unity. When the CAB520, tetrahymanol, and pressinoic acid data were first scaled, the structures were not known to contain solvent of crystallization, so it is reasonable that their scale factors should refine to smaller values. Table 3 shows that in most cases the Rogers method gave better estimates of the scale factors than did the routinely applied customary methods. Presumably this is because the fit to the Patterson origin peak achieves a separation or smoothing of the non-random interatomic vectors (Rogers, 1965, pp. 140-148; 1980, pp. 83-85) that is equivalent to the effect of the Debye radial distribution curve as employed by Main (1976) to utilize known molecular structure information in the scaling process.

The triple entries in Tables 2 and 3 give, first, the *a priori* estimate from the Rogers analysis, then two differently weighted averages over the unit cell of the *a posteriori* atomic thermal parameters from the converged least-squares structure refinements. The second entry is weighted by the atomic masses, which seem reasonable for estimation of overall mean-square vibrational displacements. The third entry is weighted by the squared atomic numbers, which reflect the atomic scattering cross sections for X-rays. The results indicate that there is little to choose between the two kinds of averaging.

Table 2 shows that the Rogers method does quite well at getting the relative anisotropy of the overall thermal parameters about right. With the exception of the CAB610 structure, which turned out to have rotationally disordered trifluoromethyl groups, the agreement between the *a priori* and *a posteriori* anisotropic thermal parameters is best for the calcium channel blocker structures and the urea-phosphoric acid structure. These have smaller mole fractions of hydrogen than the other structures, and it appears that the equivalent anisotropic components of the refined isotropic thermal parameters for the hydrogen atoms noticeably bias the thermal-parameter averages. This effect is compounded in the cases of pressinoic acid and the valinomycin analogs, for which hydrogen atoms were included in the structure models at fixed calculated positions with fixed estimated thermal parameters, and for which anisotropic

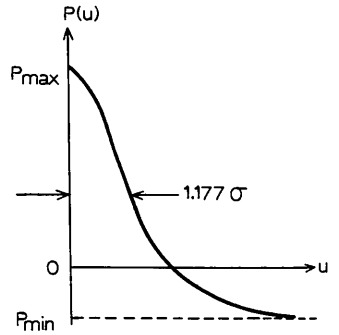

Fig. 1. Schematic illustration of a Patterson origin peak showing the half-height peak radius and the $P_{\min}$ value corresponding to equation (10). At $P = (P_{\min} + P_{\max})/2$, $u \simeq 1{\cdot}177\sigma$, where σ^2 is the variance of a Gaussian peak function centered at $u = 0$, $P(u) = P_{\min} + (P_{\max} - P_{\min}) \exp(-u^2/2\sigma^2)$.

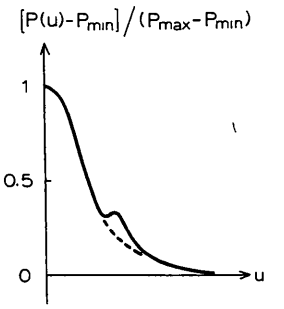

Fig. 2. Schematic illustration of the superposition of a Patterson origin peak and an off-origin interatomic peak.

Table 1. *Test data sets*

	Space group	Unit cell contents
Urea-phosphoric acid	*Pbca*	$8CH_7N_2O_5P$
DL-Biotin structures		
monoclinic	$P2_1/n$	$4C_{13}H_{22}N_2O_3S$
triclinic	$P1$	$2C_{13}H_{22}N_2O_3S.0{\cdot}5C_6H_{12}$
Calcium channel blockers		
CAB611	$C2/c$	$8C_{15}H_{14}N_3O_7$
CAB610	$P2_1/c$	$4C_{16}H_{15}F_3N_2O_4$
CAB520	$P2/c$	$8C_{17}H_{17}ClFNO_4.0{\cdot}25C_6H_{12}$
CAB438	$P2_12_12_1$	$4C_{19}H_{20}N_2O_6$
Tetrahymanol	$P2_1$	$4C_{30}H_{52}O.0{\cdot}5H_2O$
Pressinoic acid	$P2_1$	$2C_{33}H_{42}N_8O_{10}S_2.4{\cdot}5H_2O$
Valinomycin analogs		
ILED	$P2_12_12_1$	$4C_{60}H_{102}N_6O_{18}$
F4	$P2_12_12_1$	$4C_{60}H_{102}N_6O_{18}$
Insulins		
'4-zinc' porcine insulin		
$d_{min}=1{\cdot}50$Å	$R3$	$9C_{512}H_{762}Cl_{0\cdot22}N_{130}O_{152}S_{12}Zn_{0\cdot88}.\sim339H_2O$
'4-zinc' human insulin		
$d_{min}=1{\cdot}85$ Å	$R3$	$9C_{514}H_{766}Cl_{0\cdot22}N_{130}O_{154}S_{12}Zn_{0\cdot88}.(196+\sim143)H_2O$
monoclinic human insulin		
$d_{min}=2{\cdot}25$ Å	$P2_1$	$2C_{1542}H_{2298}N_{390}O_{462}S_{36}Zn_2.\sim1434H_2O$

Table 2. *Overall anisotropic temperature factor coefficients* (Å^2)

$F(\mathbf{h}) = F_0(\mathbf{h}) \exp(-2\pi^2 \sum_i \sum_j h_i h_j a^{*i} a^{*j} U^{ij})$.

The three values given are respectively the Rogers estimate, atomic-mass-weighted average of structure refined values, and squared-atomic-number-weighted average of structure refined values.

	U_{11}	U_{22}	U_{33}	U_{12}	U_{13}	U_{23}
Urea.H3PO4	0·026	0·028	0·022	0	0	0
	0·031	0·035	0·026	0	0	0
	0·028	0·032	0·024	0	0	0
Mon.DL-B	0·069	0·058	0·060	0	0·010	0
	0·057	0·045	0·048	0	0·012	0
	0·058	0·045	0·049	0	0·011	0
Tri.DL-B	0·071	0·076	0·053	0·021	−0·0037	0·0009
	0·063	0·065	0·055	0·008	0·0002	0·0037
	0·060	0·064	0·051	0·010	0·0001	0·0016
CAB611	0·071	0·084	0·093	0	0·018	0
	0·073	0·085	0·092	0	0·016	0
	0·075	0·088	0·093	0	0·017	0
CAB610	0·086	0·057	0·049	0	0·016	0
	0·116	0·090	0·080	0	0·009	0
	0·123	0·094	0·085	0	0·006	0
CAB520	0·080	0·074	0·091	0	0·036	0
	0·077	0·073	0·085	0	0·034	0
	0·078	0·073	0·086	0	0·035	0
CAB438	0·056	0·071	0·043	0	0	0
	0·059	0·070	0·045	0	0	0
	0·060	0·072	0·044	0	0	0
Thym.	0·052	0·062	0·049	0	−0·008	0
	0·046	0·055	0·040	0	−0·0006	0
	0·045	0·054	0·038	0	−0·002	0
Press.	0·068	0·063	0·093	0	0·025	0
	0·052	0·042	0·072	0	0·021	0
	0·053	0·042	0·074	0	0·023	0
Val.ILED	0·079	0·084	0·084	0	0	0
	0·088	0·092	0·092	0	0	0
	0·088	0·090	0·091	0	0	0
Val.F4	0·125	0·124	0·145	0	0	0
	0·150	0·149	0·161	0	0	0
	0·143	0·144	0·157	0	0	0
4Zn-P.Ins.	0·29	0·29	0·35	0·14	0	0
4Zn-H.Ins.	0·38	0·38	0·53	0·19	0	0
Mon.H.Ins.	0·58	0·53	0·60	0	0·26	0

thermal parameters could not be fitted for disordered water molecules of crystallization or disordered side-chain carbon atoms.

For the CAB610 structure, with the disordered trifluoromethyl groups, the Patterson origin peak gave scale and thermal parameters that were apparently biased towards the ordered part of the structure. This suggests that for a heavy-atom structure, such as an organometallic compound, the results would be biased towards the heavy atom to the extent that it

Table 3. *Scale factors, equivalent isotropic mean-square displacements* ($Å^2$) *and agreement-of-fit indices for the anisotropic Gaussian peaks*

$$R = \{\sum w[P(u) - P_0(u)]^2 / \sum wP(u)^2\}^{1/2}.$$

Values are listed in the order Rogers estimate/atomic-mass-weighted average of structure refined values/squared-atomic-number-weighted average of structure refined values.

	k	% error*	$\langle u^2 \rangle_{iso}$	R
Urea . H3PO4	1·11	−1	0·0252	0·020
	1·12	−11	0·0304	
			0·0280	
Mon.DL-B	4·51	−12	0·0625	0·018
	5·14		0·0503	
			0·0508	
Tri.DL-B	1·49	−13	0·0664	0·026
	1·73	−42	0·0609	
			0·0585	
CAB611	0·996	−9	0·0825	0·031
	1·090	−8	0·0834	
			0·0852	
CAB610	0·945	13	0·0639	0·026
	0·837	19	0·0955	
			0·1010	
CAB520	0·933	−6	0·0819	0·015
	0·990	1	0·0784	
			0·0789	
CAB438	0·773	−5	0·0567	0·031
	0·816	23	0·0583	
			0·0587	
Thym.	0·835	8	0·0544	0·042
	0·912	10	0·0472	
			0·0457	
Press.	0·733	−21	0·0744	0·028
	0·931	7	0·0556	
			0·0564	
Val.ILED	0·895	−0·1	0·0825	0·012
	0·896	12	0·0905	
			0·0895	
Val.F4	0·757	−7	0·132	0·031
	0·811	23	0·153	
			0·148	
4Zn-P. Ins.	1·30		0·308	0·008
4Zn-H.Ins.	0·370	12	0·434	0·014
	0·331		0·422	
Mon.H.Ins.	0·207	93	0·56	0·012
	0·107		0·35	

* The first value listed under '% error' for each entry is the percentage difference between the Rogers estimate of the scale factor and the structure refined value, and the second value is the percentage difference between unity and the structure-refined value.

dominates the scattering. The estimates would also be biased by uncorrected anisotropic absorption in strongly absorbing crystals.

Anisotropic thermal parameters are not yet available from the refinements of the insulin structures listed in the tables, but according to the experience of Sheriff & Hendrickson (1987), the introduction of thermal anisotropy estimates from Table 2 should improve the insulin structure models considerably. The overall thermal parameters for the insulin crystals (Tables 2 and 3) parallel the data-set resolution (Table 1), which, in turn, parallels the attenuation of the higher-angle X-ray scattering by the thermal motion and disorder in the crystals. The differences between the estimated and refined scale factors for the human insulin crystals (Table 3) reflect approximately the proportion of disordered side-chain and solvent structure not yet included in the structure models. When the very-low-angle data with $(\sin\theta)/\lambda < 0{\cdot}0625$ $Å^{-1}$ ($d > 8$ Å) were omitted from the Patterson–Fourier summation, and the water of crystallization was omitted from $\sum_j f_j^2$, the Rogers estimates for the monoclinic human insulin data were much closer to the refined values for the partial structure model. We have also seen that, not surprisingly, the analysis of the Patterson origin peak is unreliable when only a partial low-resolution data set is available. The data should extend to the scattering-angle radius at which the scattered intensity becomes experimentally insignificant owing to the thermal motion or disorder.

For all the examples listed in the tables, the Gaussian peak was fitted out to a radius of 2·5 half-height radii, with the points outside one half-height radius weighted as described above. The radius of fit was less than 1·25 Å for all but the Val.F4 and insulin crystals, for which the radii were 1·5, 2·4, 3·0, and 3·1 Å, respectively. In these latter cases, off-origin interatomic C–C, C–N and C–O peaks certainly fell within the radius of fit, and to a greater or lesser extent, depending on the effectiveness of the weighting, biased the results. The general effect should be to broaden and shorten the fitted Gaussian peak. This, in turn, would give an inverse Gaussian too narrow and tall, and would bias both the thermal parameters and the scale factor to values that are too large [see *e.g.* (7)]. To test for significant bias, the calculations for the Val.F4 and insulin data sets were repeated, restricting the radius of fit to one half-height radius. The results changed by at most a few percent, indicating that our simple scheme to down-weight the off-origin peaks works surprisingly well.

Two applications that we had in mind at the start of this work still remain to be pursued. The anisotropic Rogers expectation values for $|F|^2$ can be used to advantage in the Bayesian processing of weak high-angle diffraction data (French & Wilson, 1978), and in the estimation of experimental $|E|$ values for direct methods of structure determination. The overall anisotropic thermal parameters also provide the basis for an anisotropic thermal diffuse scattering correction that would not require that the elastic constants of the crystal be known (Blessing, 1987).

A listing of our Fortran program (~1000 lines of code) has been deposited as supplementary publication material.* A machine-readable copy of the program is available from the authors on request.

* This listing has been deposited with the British Library Document Supply Centre as Supplementary Publication No. SUP44936 (36 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

We thank our colleagues G. David Smith and Zdzisław Wawrzak for the use of their unpublished data. The porcine insulin data were courtesy of Professor Guy Dodson, University of York, England, *via* David Smith. We are grateful for the support of our work by USDHHS PHS NIH grants GM34073 and DK19856 (RHB) and HL32303 (DAL).

References

BLESSING, R. H. (1987). *Crystallogr. Rev.* **1**, 3-58.
FRENCH, S. & WILSON, K. (1978). *Acta Cryst.* A**34**, 517-525.
GIACOVAZZO, C. (1980). *Direct Methods in Crystallography*, pp. 5-56. London: Academic Press.
HALL, S. R. & SUBRAMANIAN, V. (1982). *Acta Cryst.* A**38**, 590-598; 598-608.
IWASAKI, H. & ITO, T. (1977). *Acta Cryst.* A**33**, 227-229.
JOHNSON, C. K. & LEVY, H. A. (1974). In *International Tables for X-ray Crystallography*, Vol. IV, edited by J. A. IBERS & W. C. HAMILTON, p. 314. Birmingham: Kynoch Press. (Present distributor Kluwer Academic Publishers, Dordrecht.)
KARLE, J. & HAUPTMAN, H. (1953). *Acta Cryst.* **6**, 473-476.
LADD, M. F. C. (1978). *Z. Kristallogr.* **147**, 279-296.
LEVY, H. A., THIESSEN, W. E. & (in part) BROWN, G. M. (1970). Am. Crystallogr. Assoc. Meeting, New Orleans, Louisiana, March 1970. Abstract No. B6.
MAIN, P. (1976). In *Crystallographic Computing Techniques*, edited by F. R. AHMED, pp. 97-105. Copenhagen: Munksgaard.
MASLEN, E. N. (1967). *Acta Cryst.* **22**, 945-946.
PATTERSON, A. L. (1935). *Z. Kristallogr.* **90**, 517-542.
ROGERS, D. (1965). In *Computing Methods in Crystallography*, edited by J. S. ROLLETT, pp. 117-148. Oxford: Pergamon Press.
ROGERS, D. (1980). In *Theory and Practice of Direct Methods in Crystallography*, edited by M. F. C. LADD & R. A. PALMER, pp. 83-92. London: Plenum Press.
SHERIFF, S. & HENDRICKSON, W. A. (1987). *Acta Cryst.* A**43**, 118-121.
SUBRAMANIAN, V. & HALL, S. R. (1982). *Acta Cryst.* A**38**, 577-590.
WILSON, A. J. C. (1942). *Nature (London)*, **150**, 152.

Acta Cryst. (1988). A**44**, 735-740

Groupes d'Espace à Quatre et Six Couleurs

PAR JEAN SIVARDIÈRE

Département de Recherche Fondamentale, Centre d'Etudes Nucléaires, 85X, 38041 Grenoble CEDEX, France

(*Reçu le 21 octobre* 1987, *accepté le 6 mai* 1988)

Abstract

A direct algebraic method is discussed which enumerates subgroups of index four and six of three-dimensional space groups, and the four- and six-coloured space groups. There are three types of four-coloured space groups, since the lattice can be one-, two- or four-coloured, and four types of six-coloured space groups, since the lattice can be one-, two-, three- or six-coloured.

Notations

G	Groupe ponctuel ordinaire.
H	Sous-groupe invariant de G.
$G_4(G_6)$	Groupe ponctuel à quatre (six) couleurs, isomorphe de G.
G_e	Groupe d'espace ordinaire, de classe G.
T	Réseau de G_e.
$T_4(T_6)$	Réseau à quatre (six) couleurs, isomorphe de T.
T_k	Réseau des translations monocolores de T_4 ou T_6.
H_e	Sous-groupe invariant de G_e, d'indice 4 ou 6.
$G_{e4}(G_{e6})$	Groupe d'espace à quatre ou six couleurs, isomorphe de G_e.

Introduction

Les groupes d'espace à deux couleurs, ou groupes magnétiques, et à trois couleurs sont bien connus (Opechowski & Guccione, 1965; Harker, 1981). Les groupes d'espace à quatre et six couleurs ont été moins étudiés (Jarratt & Schwarzenberger, 1980; Sénéchal, 1983; Roth, 1985). Un cas particulier intéressant est celui où les permutations des quatre ou six couleurs associées aux opérations géométriques sont cycliques: les groupes ponctuels colorés correspondants ont été énumérés à partir des représentations cycliques des groupes ponctuels ordinaires (Indenbom, Belov & Neronova, 1960; Niggli & Wondratschek, 1960); les réseaux colorés ont été énumérés par Zamorzaev (1969); la méthode des représentations cycliques a été étendue aux groupes d'espace (Koptsik & Kuzhukeev, 1973).

Nous discutons dans cet article une méthode algébrique directe de recherche des groupes d'espace à quatre et six couleurs, équivalente à la méthode des

0108-7673/88/050735-06$03.00