Acta Cryst. (1990). A46, 372–377

The Use of Sayre's Equation with Constraints for the Direct Determination of Phases

By Peter Main

Department of Physics, University of York, Heslington, York YO1 5DD, England

(*Received* 21 *July* 1989; *accepted* 20 *October* 1989)

Abstract

Sayre's equation and density modification may be combined into a single procedure for the direct determination of phases or for the determination of electron density. The constraints on either the phases or the density are expressed in terms of a system of non-linear simultaneous equations with twice as many equations as unknowns. A method of solving these equations is described which combines the conjugate gradient technique with the use of fast Fourier transforms (FFT's). This gives the possibility of obtaining a least-squares solution of the equations for tens of thousands of unknowns using only minutes of computer time. It is trivial to add different types of electron-density constraint to the system. The use of the complete diffraction pattern in the calculations makes this method more powerful than conventional direct methods of phase determination. Examples of phase determination and refinement are given, together with an indication of the high quality of the resulting maps.

Introduction

Direct methods of phase determination have made progress by increasing the constraints on the electron density or by imposing them more rigorously. For inequality relationships (Harker & Kasper 1948), the electron density is constrained never to go negative, but there is no constraint on the positive density. A more powerful constraint is therefore provided by Cochran's criterion (Cochran 1952) that $\int \rho^3(\mathbf{x})\,\mathrm{d}V$ is a maximum, which applies to all the density in the cell. This criterion is approximately satisfied by having discrete atomic peaks on a smooth background. It leads to the phase relationships used in symbolic addition (Karle & Karle 1963) and upon which the tangent formula is based (Karle & Hauptman, 1956). Individual phase relationships take no account of the complete probability distribution of their values, so this is done by including a weighting scheme in the tangent formula (Hull & Irwin, 1978). The tangent formula originally used only the strongest E's, but more magnitudes, especially the very weak ones, can also be used in the modified tangent formula described by Debaerdemaeker, Tate & Woolfson (1985). All of these have led to more powerful means of phase determination.

In order to make further progress, additional constraints must be applied. An obvious one is that the complete set of magnitudes could be used instead of only the very smallest and the very largest. This may be achieved by using Sayre's equation (Sayre, 1952) instead of the tangent formula. Other constraints are those which are already used successfully in density-modification techniques such as maximum and minimum density values, background smoothing, partially known structure or known solvent regions for macromolecules. A new constraint has recently been added to this list, that of density histogram matching described by Zhang & Main (1990*a*).

Sayre has already applied his equation to phase refinement and extension for a small protein (Sayre, 1974), but there were no other constraints on the electron density. This paper describes how Sayre's equation can be used simultaneously with histogram matching and other density-modification techniques to produce a powerful procedure for phase determination and refinement. It is suitable both for small molecular structures and for macromolecules.

The electron-density equations

Sayre's equation may be written as

$$F(\mathbf{h}) = [\theta(h)/V] \sum_{\mathbf{k}} F(\mathbf{k})F(\mathbf{h}-\mathbf{k}) \qquad (1)$$

where $\theta(h) = f(h)/g(h)$; $f(h)$ = atomic scattering factor; $g(h)$ = scattering factor of squared atom. It expresses the fact that the electron density consists solely of equal atoms whose shape conforms to the scattering factor $f(h)$ and also to $g(h)$ for the squared atoms. For real atoms, this means there is no negative density. In addition, the atoms can only be equal if they are spherically symmetric and they do not overlap, so atomic resolution is assumed.

Because electron-density constraints are more easily expressed in real space than in reciprocal space, it is convenient to express Sayre's equation in terms of the electron density. Also, the electron density is expressed as a discrete function, $\rho(\mathbf{n})$, evaluated on a three-dimensional grid of N points $\mathbf{n}$. Taking the Fourier transform of both sides of (1), Sayre's equation becomes

$$\rho(\mathbf{n}) = (V/N) \sum_{\mathbf{m}} \rho^2(\mathbf{m})\psi(\mathbf{n}-\mathbf{m}) \qquad (2)$$

0108-7673/90/050372-06$03.00

where

$$\psi(\mathbf{n}) = (1/V) \sum_{\mathbf{n}} \theta(h) \exp(-2\pi i \mathbf{h}\,.\,\mathbf{n}).$$

The scale factor N in (2) is a matter of definition and arises when FFT's are used to perform the convolutions. Equation (2) expresses the equality of the electron density with the squared density convoluted with the function $\psi(\mathbf{n})$. Now $\psi(\mathbf{n})$ is the Fourier transform of $\theta(h)$, which defines the atomic shape, so this also gives the shape of the peaks in $\rho(\mathbf{n})$.

The equations representing the density modification can be written as

$$\rho(\mathbf{n}) = H(\mathbf{n}) \tag{3}$$

where $H(\mathbf{n})$ is the electron-density map modified according the particular density-modification techniques employed. Specifically, the modifications used by the author are the solvent flattening and histogram matching as described by Zhang & Main (1990*a*,*b*) together with background smoothing. The latter is a common technique, but it is used here in conjunction with histogram matching. A threshold electron density level, $\rho_{\min}$, is obtained such that all density above this level can be considered to be atomic density and the rest regarded as background. The criterion for setting this level is that the volume of atomic density corresponds to the expected atomic volume. The peak density is modified by histogram matching and a shift of $\Delta\rho$ is applied to the background density where

$$\begin{aligned} \Delta\rho &= 0{\cdot}25\rho(\rho/\rho_{\min}-1) && \text{for } 0<\rho<\rho_{\min} \\ &= -0{\cdot}25\rho && \text{for } \rho<0. \end{aligned} \tag{4}$$

This has the effect of shifting all background density towards zero without introducing any discontinuities. Unless the structure-factor magnitudes are extrapolated to very high resolution, negative density should not be modified to zero. It can quite genuinely be negative given the particular magnitudes and resolution used. The weighting factor of 0·25 in (4) is arbitrary, but it is the value which gave the best results in trials of the method. Equations (3) are effectively non-linear because the actual density modification carried out is a complicated function of the density which is already present.

Equations (2) and (3) form a system of non-linear simultaneous equations with as many unknowns, $\rho(\mathbf{n})$, as grid points in the asymmetric unit of the map and twice as many equations as unknowns. An alternative formulation would be to take the Fourier transform of both sides of (3) and express it as

$$F(\mathbf{h}) = \mathscr{F}[H(\mathbf{n})]. \tag{5}$$

We could now solve the system of equations formed by (1) and (5) for the unknown phases. Again, there are twice as many equations as unknowns. Because of the method of solving the equations, described in the next section, the computer time is the same no matter which formulation of the equations is used, but the real-space formulation will be described.

Solution of the equations

A least-squares solution of the system of equations (2) and (3) is sought for the unknown electron-density values $\rho(\mathbf{n})$. The Newton–Raphson technique is used, the starting point for which is the approximate map which was modified to produce $H(\mathbf{n})$ in (3). At the very beginning, this could be given by random phases for small molecules or an MIR map for macromolecules.

The equations to be solved for the electron-density shifts, $\delta\rho(\mathbf{n})$, are obtained from the Jacobian of (2) and (3) as

$$\begin{aligned} 2(V/N)\sum_{\mathbf{m}} \rho(\mathbf{m})\psi(\mathbf{n}-\mathbf{m})\delta\rho(\mathbf{m}) - \delta\rho(\mathbf{n}) &= \Delta\rho(\mathbf{n}) \\ \delta\rho(\mathbf{n}) &= \Delta H(\mathbf{n}) \end{aligned} \tag{6}$$

where

$$\Delta\rho(\mathbf{n}) = \rho(\mathbf{n}) - V/N \sum_{\mathbf{m}} \rho^2(\mathbf{m})\psi(\mathbf{n}-\mathbf{m})$$

$$\Delta H(\mathbf{n}) = \rho(\mathbf{n}) - H(\mathbf{n}).$$

The equations (6) may easily contain tens of thousands of unknowns and setting up the normal matrix of least squares could take about 100 years of computer time at the rate of 10^6 multiplications per second. This can be completely avoided by using the conjugate-gradient technique for solving the equations which does not require the normal matrix at all:

To obtain the least-squares solution of

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{7a}$$

start from the trial solution $\mathbf{x}_0$, which may be the null vector, then form

$$\mathbf{r}_0 = \mathbf{A}^{\mathrm{T}}(\mathbf{b} - \mathbf{A}\mathbf{x}_0) \tag{7b}$$

$$\mathbf{p}_0 = \mathbf{r}_0. \tag{7c}$$

Now start with $k=0$ and improve the trial solution by the iteration

$$\mathbf{q}_k = \mathbf{A}\mathbf{p}_k \tag{8a}$$

$$\alpha_k = (\mathbf{r}_k^T\mathbf{p}_k/\mathbf{q}_k^T\mathbf{q}_k) \tag{8b}$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\mathbf{p}_k \tag{8c}$$

$$\mathbf{s}_k = \mathbf{A}^T\mathbf{q}_k \tag{8d}$$

$$\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k\mathbf{s}_k \tag{8e}$$

$$\beta_k = -(\mathbf{r}_{k+1}^T\mathbf{s}_k/\mathbf{q}_k^T\mathbf{q}_k) \tag{8f}$$

$$\mathbf{p}_{k+1} = \mathbf{r}_{k+1} + \beta_k\mathbf{p}_k. \tag{8g}$$

The number of iterations required for an exact solution is equal to the number of unknowns. However, acceptable solutions are normally obtained after very few iterations, thus saving an enormous amount of time. Note that the normal matrix never appears explicitly, although it is implicit in (7*b*) and (8*d*). Most of the work of calculation comes in forming the matrix-vector products in (7*b*), (8*a*) and (8*d*). These can be expressed as convolutions and performed using FFT's, thus saving considerably more time.

With $\mathbf{x}_0 = \mathbf{0}$, (7*b*) becomes $\mathbf{r}_0 = \mathbf{A}^T\mathbf{b}$ where $\mathbf{A}$ is the l.h.s. matrix and $\mathbf{b}$ the r.h.s. vector of (6). The solution of (7*a*) then represents the shifts $\delta\rho(\mathbf{n})$ to be applied to the electron density. The residual vector $\mathbf{r}_0$ is calculated as follows:

let

$$G(\mathbf{h}) = (V/N)\sum_{\mathbf{m}} \rho^2(\mathbf{m}) \exp(2\pi i\mathbf{h}\,.\,\mathbf{m}) \qquad (9a)$$

then

$$\Delta F(\mathbf{h}) = F(\mathbf{k}) - \theta(h)G(\mathbf{h}) \qquad (9b)$$

and

$$\Delta\rho(\mathbf{n}) = (1/V)\sum_{\mathbf{n}} \Delta F(\mathbf{h}) \exp(-2\pi i\mathbf{h}\,.\,\mathbf{n}) \qquad (9c)$$

giving the $\mathbf{n}$th component of $\mathbf{r}_0$ as

$$r_0(\mathbf{n}) = (2/V)\rho(\mathbf{n})\sum_{\mathbf{h}} \theta(h)\Delta F(\mathbf{h}) \exp(-2\pi i\mathbf{h}\,.\,\mathbf{n}) - \Delta\rho(\mathbf{n}) + \Delta H(\mathbf{n}). \qquad (9d)$$

Thus, three FFT's in (9*a*), (9*c*) and (9*d*) are required to calculate $\mathbf{r}_0$. On the way, (9*b*) gives the residual of Sayre's equation.

The calculation of $\mathbf{q}_k$ in (8*a*) is achieved in a similar manner:

let

$$a(\mathbf{h}) = (V/N)\sum_{\mathbf{m}} \rho(\mathbf{m})p_k(\mathbf{m}) \exp(2\pi i\mathbf{h}\,.\,\mathbf{m}) \qquad (10a)$$

and

$$b(\mathbf{h}) = (V/N)\sum_{\mathbf{m}} p_k(\mathbf{m}) \exp(2\pi i\mathbf{h}\,.\,\mathbf{m}) \qquad (10b)$$

where $p_k(\mathbf{m})$ is the $\mathbf{m}$th component of the vector $\mathbf{p}_k$. Then

$$\mathbf{q}_k = \left\{\begin{array}{c} \dfrac{1}{V}\sum_{\mathbf{h}} [2a(\mathbf{h})\theta(h) - b(\mathbf{h})] \exp(-2\pi i\mathbf{h}\,.\,\mathbf{m}) \\ \hline p_k(\mathbf{m}) \end{array}\right\} = \left[\begin{array}{c} Q_k(\mathbf{m}) \\ \hline p_k(\mathbf{m}) \end{array}\right] \qquad (10c)$$

where the vector $\mathbf{q}_k$ is partitioned as shown.

Similarly, $\mathbf{s}_k$ in (8*d*) is obtained from

$$s_k(\mathbf{n}) = (2/V)\rho(\mathbf{n})\sum_{\mathbf{h}} \theta(h)[2a(\mathbf{h})\theta(h) - b(\mathbf{h})] \times \exp(-2\pi i\mathbf{h}\,.\,\mathbf{n}) - Q_k(\mathbf{n}) + p_k(\mathbf{n}) \qquad (11)$$

where $Q_k(\mathbf{n})$ is defined in (10*c*).

The remaining calculations in (8*b*), (8*c*), (8*e*), (8*f*) and (8*g*) require either the inner product of a pair of vectors or a linear combination of vectors, both of which are very quick to calculate. Each iteration of the conjugate-gradient technique is therefore seen to require the four FFT's described in (10) and (11).

The criterion for stopping the iterations is that the length of the residual vector $\mathbf{r}_k$ should have decreased by a factor of 100 from the original $\mathbf{r}_0$. This normally takes less than ten iterations even when there are tens of thousands of unknowns. One complete iteration of the full-matrix least-squares process therefore requires between 30 and 40 FFT's. This can be achieved in minutes, as opposed to the 100 years of computer time hinted at earlier. The electron-density shifts, $\delta\rho(\mathbf{n})$, obtained in this way are applied to the electron density to produce an improved map.

The final stage in the cycle is to calculate structure factors from the improved map and calculate their Sim weights. When MIR phases are available, they are combined with the new phases, taking their weights into account. Finally, a weighted electron-density map is calculated, using the observed magnitudes, to start the next cycle of map improvement.

Diagonal approximation

One of the most obvious ways of speeding up the calculation of the electron-density shifts is to use the diagonal approximation to the normal matrix. As with the full-matrix calculation, it can be done entirely by FFT's and linear combinations of vectors:

given the overdetermined system of equations

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \qquad (12a)$$

the least-squares estimate of $\mathbf{x}$ is given by solving

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}. \qquad (12b)$$

If $\mathbf{A}$ is the l.h.s. matrix and $\mathbf{b}$ the r.h.s. vector of (6), then $\mathbf{x}$ represents the electron-density shifts $\delta\rho(\mathbf{n})$.

The right-hand side of (12*b*) is identical to the residual vector $\mathbf{r}_0$ which was calculated in (9). It only remains to calculate the diagonal element of $\mathbf{A}^T\mathbf{A}$. The required expression can be shown to be

$$a_0(\mathbf{n}) = (4/N)\rho(\mathbf{n})\left[\rho(\mathbf{n})\sum_{\mathbf{h}}|\theta(h)|^2 - \sum_{\mathbf{h}}\theta(h)\right] + 2. \qquad (13)$$

So the electron-density shifts are calculated from

$$\delta\rho(\mathbf{n}) = r_0(\mathbf{n})/a_0(\mathbf{n}). \qquad (14)$$

The calculation of $\mathbf{r}_0$ takes three FFT's. Two more FFT's are required to impose the observed magnitudes on the next map as described at the end of the last section. Thus five FFT's are needed for each iteration of the diagonal approximation. This makes it about seven or eight times faster than the full-matrix calculation and this difference in speed is actually observed.

The scale factor $\theta(h)$

The calculation of $\Delta F(\mathbf{h})$ in (9b) is critically dependent upon the value of the scale factor $\theta(h)$, which is defined in (1). It was found that this expression gave the correct shape for the function $\theta(h)$, but that the absolute value is a sensitive function of resolution. It is necessary therefore to calculate and apply a linear overall scale factor to $\theta(h)$ to match the magnitudes of $F(\mathbf{h})$ and $\theta(h)G(\mathbf{h})$. As a further convenience, it was assumed that the atoms were Gaussian in shape so a formula could be obtained for $\theta(h)$:

for an atomic scattering factor of

$$f(S) = A\exp(-aS^2) \tag{15a}$$

it can be shown that

$$g(S) = A^2[\pi/(2a)]^{3/2}\exp(-aS^2/2) \tag{15b}$$

so that

$$\theta(S) = [f(S)/g(S)] = (1/A)(2a/\pi)^{3/2}\exp(-aS^2/2) \tag{15c}$$

where

$$S = (\sin\vartheta)/\lambda = |\mathbf{h}|/2. \tag{15d}$$

Fitting a Gaussian of the form (15a) to the scattering factor of nitrogen gave $A = 6{\cdot}53$ and $a = 1{\cdot}87$ for $0{\cdot}0 < S^2 < 0{\cdot}5$, hence $\theta(S)$ from (15c). Refinement of these parameters using Sayre's equation with correct phases for an equal-atom structure showed that A was sensitive to resolution and that the value of a was close to that determined theoretically.

An alternative way of obtaining $\theta(h)$ is by setting up Sayre's equation for a similar but known structure at the same resolution. The scale factor can be determined as a function of $(\sin\vartheta)/\lambda$ by spherical averaging. This method has been employed for macromolecules.

Since $\theta(h)$ depends upon atomic shape, (15c) is only accurate at high resolution. At very low resolution, $\theta(h)$ becomes highly structure dependent and cannot be predicted. However, for a resolution higher than about 3·0 Å, (15c) gives acceptable results.

Method

To test the ability of the proposed technique to determine and refine phases, the following scheme was used:

(a) Start with exact phases for a selected number of the strongest reflexions, *i.e.* all those with magnitudes greater than F_{min}.

(b) Calculate the electron density using all known phases and weighted observed magnitudes.

Table 1. *Details of enniatin B used as the test structure*

Molecular formula: $C_{33}H_{57}N_3O_9$
$a = 29{\cdot}28$, $b = 28{\cdot}29$, $c = 10{\cdot}84$ Å, $\gamma = 121{\cdot}1°$
space group = $P2_1$, c-axis unique, $Z = 8$
180 non-H atoms per asymmetric unit

(c) Modify the electron density by density histogram matching (Zhang & Main, 1990a) and background smoothing (4). This produces the map $H(\mathbf{n})$ in (3).

(d) Solve (2) and (3) for the unknown electron density, starting from the approximate density calculated in (b). All reflexions not included in the density calculated in (b) are given a magnitude of zero.

(e) Calculate structure factors from the latest electron density together with their Sim weights.

(f) For phase determination, lower F_{min} and accept all phases $\varphi(\mathbf{h})$ for which $|F(\mathbf{h})| > F_{\text{min}}$. For phase refinement, keep F_{min} the same.

(g) Repeat from (b) until all phases have been determined and refined.

The rate at which new reflexions are accepted at step (e) is governed by the increase of variance they bring to the electron-density map. The values of F_{min} are chosen to increase $\sum |F(\mathbf{h})|^2$ by constant amounts, where the F's included in the summation are those used to calculate the map.

Results

The power of phase determination given by Sayre's equation with electron-density constraints has not yet been fully explored. Most of the tests done to date have been to confirm the validity of the mathematical techniques. However, promising results have been obtained, some of which are presented here. The work on phase refinement and extension for macromolecules is presented in the following paper by Zhang & Main (1990b).

Calculations were carried out on the known structure of enniatin B, the details of which are presented in Table 1. Observed F's, sharpened by removing the overall temperature factor, were used and all calculations were performed at 1·1 Å resolution. It was found that the full-matrix calculation gave only marginally better results than the diagonal approximation. Since the latter is about eight times faster, the diagonal approximation was used almost exclusively. For a comparison between the two types of calculation, see Zhang & Main (1990b).

Two examples of phase determination and refinement are shown in Table 2. The weights used in the calculation of the weighted mean phase error are the $|F(\mathbf{h})|$ and the mean is taken over all reflexions in the data. The mean phase error for the strongest reflexions is less than half that shown in the Table. The R factor is between the observed magnitudes and

Table 2. *Results of phase determination and refinement on the test structure, starting from exact phase values*

Number of starting reflexions	Number of stages of phase expansion	Number of cycles of phase refinement	Number of phases determined	Weighted mean phase error	Final R factor
200	7	5	6500	19·9°	13·2%
100	17	20	6500	19·2°	15·4%

those calculated from the map. Even better results can normally be obtained by performing the phase determination in smaller stages and using more of them. It is clear that the phases are very good and this was confirmed by looking at the final electron-density map for the first entry in Table 2. With the peaks numbered in decreasing order of height, the highest spurious peak was 178, the lowest atomic peak was 182 and only one atom was missing from the map out of the 180 expected.

Another measure of the quality of the map was obtained almost by accident. Only after the structure had been solved was it realised that some strong low-angle reflexions were completely missing from the data, including the strongest reflexion in the whole diffraction pattern. To deal with these during phase determination and refinement, their magnitudes and phases were calculated along with all the other reflexions and these values were used in the calculation of the next map. However, no weights could be calculated and no direct restrictions placed on the magnitudes so obtained. The magnitudes and phases obtained from the final map are shown in Table 3. The agreement with the structure factors calculated from the atomic coordinates is astonishing. The R factor for the magnitudes shown is 4·2% and the weighted mean phase error is 3·4°.

Discussion

One reason why the phases are so accurate in the present method when compared with those obtained from more conventional direct methods is that the complete diffraction pattern is used in the calculations. This gives smoother electron-density maps with better peak shapes which, in turn, leads to better density modification. At the beginning of the calculation, however, not all phases are used. It was found that significantly better results were obtained when the missing reflexions were entered into the calculations with zero magnitude than when they were left out completely. Leaving them out places no restriction on their magnitudes. Putting them in as zero forces them at least to be smaller than those already present and a value of zero is a good approximation to the correct magnitude for most of the reflexions.

Equations (2) and (3) have been given equal weight in the formulation of their solution in (9), (10) and (11). This need not be the case and different weights can be given to Sayre's equation and the various types of density modification. From the small number of trials already carried out, it appears that different weights are either unnecessary or have little effect.

Table 3. *Low-resolution reflexions missing from the original data but calculated from the final map as F_{map} and φ_{map}; they are compared with the structure factors calculated from the known structure*

h	k	l	F_{map}	φ_{map}	F_c	φ_c	ΔF	$\Delta\varphi$
0	0	2	282	194	280	203	2	9
0	1	0	11	360	29	180	18	180
0	2	0	276	360	270	360	6	0
0	2	1	238	194	249	191	11	3
0	4	1	293	358	287	2	6	4
1	−1	0	12	180	19	180	7	0
1	−1	2	520	113	522	113	2	0
1	0	2	296	197	279	204	17	7
2	−2	0	159	360	165	360	6	0
2	−2	1	351	14	356	11	5	3
2	0	0	438	360	385	360	53	0
2	1	2	265	110	268	105	3	5
2	2	0	274	180	277	280	3	0
4	−4	0	305	360	325	360	20	0
4	−2	1	229	187	229	188	0	1
4	0	0	264	180	248	180	16	0

In an effort to speed up the computer calculations, the grid of the electron-density map was made slightly more coarse than the maximum reflexion indices required. Even though the reflexions which were thus cut off were never used, this immediately led to very poor phase determination at high angles. The errors fed back into the rest of the reflexions and substantially degraded the quality of all the phases. This is a well known phenomenon in signal processing, called aliasing. It should be carefully avoided in these calculations because the phases are particularly sensitive to it.

Since the phases determined by this method are very accurate, given a good starting point, the question arises whether good phases could be determined from very inaccurate starting values. In other words, could this phase-determination procedure be successful in a Monte Carlo (multi-solution) technique? This has yet to be determined, but trials of this approach will soon be carried out, both for small molecular crystals and for small proteins.

The method of phase determination and refinement described here may be regarded as a development of Sayre's own work on the phase refinement of rubredoxin (Sayre, 1974). However, additional constraints have been placed on the electron density and the technique used for solving the equations gives acceptably short computing times. For example, the times for the calculations described in Table 2 were 14 and 35 min respectively on an Inmos T800 transputer controlled by an IBM PC-compatible host.

Because of the completely general way in which the electron-density constraints can be applied, it is very easy to incorporate almost any other kind of constraint into the method. An obvious example of this is when a structure is partially known, *e.g.* the main chain of a macromolecule. The known density can be enforced on the map as part of the density modification and the remaining density determined along with the known density. An alternative approach would be to subtract the known density from the map and to determine the remaining density as a smaller structure. It is intended to try both of these developments of the method.

The most efficient way to propagate knowledge of phases throughout the reciprocal lattice is to use normalized structure factors, E's, rather than observed structure factors, F's. However, if E's are used, the map is subject to large series-termination errors and a considerable amount of negative density is produced. The background smoothing will get rid of this and, in so doing, extrapolate the E's to higher resolution. The extrapolated E's should be used in the calculation of the next map or the negative density will reappear. There are therefore two ways of proceeding with the calculations. One is to use F's, which produce no series-termination errors, and perform the calculations at the observed resolution. The other way is to use E's and perform the calculations at a suitably high resolution with extrapolated data, as is normally done in maximum-entropy calculations. This will increase the computing time considerably. The author has chosen to use F's to keep the computing time short, although the possibility of using E's needs to be investigated properly.

References

COCHRAN, W. (1952). *Acta Cryst.* **5**, 65-67.
DEBAERDEMAEKER, T., TATE, C. & WOOLFSON, M. M. (1985). *Acta Cryst.* **A41**, 286-290.
HARKER, D. & KASPER, J. S. (1948). *Acta Cryst.* **1**, 70-75.
HULL, S. E. & IRWIN, M. J. (1978). *Acta Cryst.* **A34**, 863-870.
KARLE, I. L. & KARLE, J. (1963). *Acta Cryst.* **16**, 969-975.
KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635-651.
SAYRE, D. (1952). *Acta Cryst.* **5**, 60-65.
SAYRE, D. (1974). *Acta Cryst.* **A30**, 180-184.
ZHANG, K. Y. J. & MAIN, P. (1990*a*). *Acta Cryst.* **A46**, 41-46.
ZHANG, K. Y. J. & MAIN, P. (1990*b*). *Acta Cryst.* **A46**, 377-381.

Acta Cryst. (1990). **A46**, 377-381

The Use of Sayre's Equation with Solvent Flattening and Histogram Matching for Phase Extension and Refinement of Protein Structures

BY KAM Y. J. ZHANG AND PETER MAIN

Department of Physics, University of York, Heslington, York YO1 5DD, England

(*Received* 21 *July* 1989; *accepted* 24 *October* 1989)

Abstract

A new method for phase refinement and extension, which combines Sayre's equation with solvent flattening and histogram matching, has been developed. Equations which express electron-density constraints are solved using discrete Fourier transforms to give a new approximation to the electron density. The formulation of the equations is in real space, which allows any set of constraints to be entered directly into the calculation. An application to the known structure of 2Zn insulin refined the 3 Å MIR phases from a mean phase error of 46 to 39° and extended the phases to 2 Å resolution with a mean overall phase error of 57°. Analysis of the phase errors shows that, for the strong reflexions, the new method determines phases with half the mean error of MIR phases.

Introduction

The dominant method in the determination of macromolecular structures is that of multiple isomorphous replacement (MIR). The phases obtained by MIR suffer from inaccuracies due to experimental error and lack of isomorphism and they are not always determined to the full resolution of the native data. All of this detracts from the quality of the electron-density map and may lead to difficulties in its interpretation. Thus, the ability to improve the quality of the MIR phases and to extend them to the full resolution of the native data would be a valuable contribution to protein crystallography.

The most successful technique of phase refinement and extension uses density modification. In its various forms it applies constraints to the electron density such as positivity, atomicity, boundedness, solvent flatness, connectivity and non-crystallographic symmetry. For a review, see Podjarny, Bhat & Zwick (1987). A recent addition to density modification is the histogram matching of Zhang & Main (1990) which imposes the correct electron-density histogram on the map. When combined with solvent flattening (Wang, 1985), it successfully refined the 1·9 Å MIR

0108-7673/90/050377-05$03.00