Acta Cryst. (1976). A32, 125

The Combination of Independent Phase Information Obtained from Separate Protein Structure Determinations of Yeast Hexokinase

BY ROBERT J. FLETTERICK* AND THOMAS A. STEITZ

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520, *U.S.A.*

(*Received* 14 *April* 1975; *accepted* 3 *July* 1975)

A procedure is presented for combining in reciprocal space the information contained in protein structures which have been independently determined by multiple isomorphous replacement (MIR). Application of this method to the two crystal forms of yeast hexokinase B which have been solved at 3·5 Å resolution results in an electron density map that is a striking improvement over either of the two MIR maps or the map obtained by simply averaging the MIR maps in direct space. The background is lower, the apparent resolution is higher and most importantly, side chains are visible in this new hybrid map where none exist in either of the MIR maps. The method consists of determining a linear transformation matrix and vector relating the subunits of the two crystals, orienting the electron density map of one crystal form in the unit cell of the other and transforming this map to produce a set of calculated phases and structure-factor amplitudes. These calculated phases from the first crystal are then combined with the MIR phases of the second to produce a hybrid phase set which is used to calculate a new electron density map.

Introduction

Very often it is necessary to improve the electron density map of a protein which has been produced by the multiple isomorphous replacement (MIR) method (Green, Ingram & Perutz, 1954; Blow & Crick, 1959). In the absence of additional suitable heavy-atom derivatives, several methods for improving the Fourier map have been proposed and used. These include: (1) Averaging of the electron density of identical subunits related by non-crystallographic symmetry (Matthews, Sigler, Henderson & Blow, 1967; Muirhead, Cox, Mazzarella & Perutz, 1967; Buehner, Ford, Moras, Olsen & Rossmann, 1974); (2) Application of the tangent formula or related direct methods to refine the MIR phases (Reeke & Lipscomb, 1969; Weinzierel, Eisenberg & Dickerson, 1969; Hendrickson, 1973; Sayre, 1972); (3) Difference-Fourier and least-squares methods of refinement of protein coordinates (Watenpaugh, Sieker, Herriott & Jensen, 1973; Freer, Alden, Carter & Kraut, 1975); and (4) Electron density modification procedures (Barrett & Zwick, 1971; Collins, 1975). The first is useful at any resolution while the latter methods are successful at very high resolution or in the final stages of the structure analysis.

Rossmann & Blow (1963) have pointed out that phase information can be obtained if two or more copies of a subunit (or molecule) are contained in the asymmetric unit of a crystal or if the molecule is found in more than one crystal form. Method (1) above, the averaging of electron density maps, is in fact one way of utilizing this information. We show here that significantly greater improvement in the resultant averaged electron density map can be achieved by combining the phases in reciprocal space and making use of the estimation of error in the phases being combined. This is particularly true if the electron density maps of a protein obtained in different crystal forms are of unequal quality.

We can consider three ways of combining the information from two different crystal forms. The simplest method is to average appropriately the electron density maps. Then if N subunits are averaged, the signal-to-noise ratio will increase by at least $\sqrt{N}$ if the maps are of equal quality. The second method is to perform the averaging operation as above and then, by taking the Fourier transform of the 'average' subunit, calculate a new set of phases. The electron density map which is obtained using these phases and one set of observed structure-factor amplitudes is superior to the simple averaged map. Bricogne (1974) has shown that this second method is in fact equivalent to a least-squares solution of the general 'molecular-replacement' phase equations in reciprocal space.

The third method, which is the one we describe below, is a variation of the second. Instead of calculating the Fourier transform of the 'average' subunit, we can calculate the Fourier transform of one of the subunits (transformed into the Cartesian space of the other) to produce a set of calculated phases. Then these phases can be combined with the experimental phases and a map calculated with these hybrid phases.

These three methods can be illustrated schematically for two subunits in two crystal forms. Let ϱ_1 represent the electron density function for subunit 1 and ϱ_2' the electron density function for subunit 2 transformed into the Cartesian space of ϱ_1. Define T to be the

* Current Address: Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7.

Fourier transform operator, then for method I:

I. $\langle\varrho\rangle = \frac{1}{2}(\varrho_1 + \varrho_2')$.

The second method then is

II. $\langle\varrho\rangle \xrightarrow{T^{-1}} \alpha(\text{calc}), F(\text{calc})$

and $\alpha(\text{calc}), F(1,\text{obs}) \xrightarrow{T} \varrho_{1+2}$,

where α(calc) is the set of calculated phases. Method III is then represented schematically as

III. $\varrho_2' \xrightarrow{T^{-1}} \alpha(2,\text{calc}), F(2,\text{calc})$

$f[\alpha(1,\text{obs}), \alpha(2,\text{calc})], F(1,\text{obs}) \xrightarrow{T} \varrho_{\text{hybrid}}$,

where here the phases are combined using some function f. Clearly with this method the function f must be chosen to take advantage of the information regarding the errors in the observed and calculated phases. One simple function that does this is described below. The advantages or disadvantages of method III in comparison to method II depend on the choice of this function. In this paper we demonstrate clearly that method III is better than I in improving the map. No comparison is made between methods II and III which we feel to be essentially similar except insofar as one has, by virtue of defining the function f, somewhat more flexibility in manipulating the phase information with method III.

Electron density maps have been obtained by the MIR method for two different crystal forms of yeast hexokinase B at high resolution (Fletterick, Bates & Steitz, 1975; Fletterick, Anderson, Anderson & Steitz, 1976). One crystal form, BIII, contains one monomer of molecular weight 51000 per asymmetric unit in space group $P2_12_12_1$ while the other crystal form, BII, contains one dimer per asymmetric unit in the same space group. To 3·5 Å resolution the phases for the BIII map show an average figure of merit of 0·78 while those for the BII map show only a 0·65 average figure of merit. Suitable averaging of electron density maps (method I), of these two crystal forms produces an 'average' map which is better than the BII MIR map but worse than the BIII MIR map. We show here that combining the structural information by a properly weighted averaging of MIR phases in reciprocal space results in an average electron density map which is superior to either the BIII or BII MIR map.

Summary of the method

Before this phase combination method and its application to hexokinase are discussed in detail, a brief outline of the procedure is given. For the general application of the method consider the case in which two crystal forms exist, each with one subunit in the crystallographic asymmetric unit. After the matrix and vector relating the subunits in the two crystals have been determined, the electron density of the first subunit ($S1$) is oriented into the space of the second ($S2$). Phases $\alpha(\text{calc})_{S1}$ can be calculated from $S1$ oriented in the space of $S2$ with the fast Fourier transform (FFT) and combined using suitable weighting with the MIR phases $\alpha(\text{MIR})_{S2}$ to produce hybrid phases and a corresponding electron density map. The essential stages of this method are:

(1) The linear transformation relating subunits $S1$ and $S2$ is calculated. This can be achieved by using the rotation and translation functions (Rossmann & Blow, 1962), by an analysis of the corresponding heavy-atom sites found for $S1$ and $S2$, or by direct visual inspection of the two Fourier maps.

(2) This preliminary linear transform is then refined in real space to maximize the overlap of the $S2$ electron density map with that of $S1$.

(3) The $S1$ electron density map is modified so that the electron densities of the solvent regions are set to a uniform background level. The neighboring molecule fragments are also removed so that the $S1$ electron density map contains only one complete subunit.

(4) The refined matrix and vector relating $S1$ and $S2$ and the symmetry operators of the $S2$ crystal form are used to construct an electron density map in the space of $S2$ but with the subunit density distribution of $S1$.

(5) The FFT algorithm is used to transform this map in order to produce a new set of phases, $\alpha(\text{calc})_{S1}$, and structure factors $F(\text{calc})_{S1}$.

(6) The calculated phases, $\alpha(\text{calc})_{S1}$, are combined with the observed phases, $\alpha(\text{MIR})_{S2}$ to give hybrid phases making use of the figure of merit as an estimate of the error in the $\alpha(\text{MIR})_{S2}$ phases and the agreement between $F(\text{obs})_{S2}$ and $F(\text{calc})_{S1}$ as an estimate of the error in $\alpha(\text{calc})_{S1}$ phases.

(7) An electron density map is calculated using these hybrid phases and the observed structure factors $F(\text{obs})_{S2}$.

The method in detail

Relation of subunits

We shall now describe in more detail for each of the above steps the mathematical algorithms used to carry out this phase-combination procedure. The initial step, which is the determination of the spatial relationship between the two subunits, can be done in a variety of ways, as mentioned above. Only one, the utilization of the heavy-atom positions will be described here. After the heavy-atom positions have been assigned to one subunit in each of the crystal forms, the intrasubunit distances between these heavy-atom positions are calculated. If a number of heavy atoms bind to the same location in both crystal forms, then the distances between these common sites will be the same. In this way, the heavy-atom sites in crystal one which have counterparts in crystal two can be identified. In practice while four or five heavy-atom sites in common between the two crystal forms will suffice, eight to ten common sites will generate a more reliable linear transformation matrix and translation vector.

If $\mathbf{x}_{i1}$ represents the unit-cell fractional coordinates of the ith heavy-atom position in $S1$ and $\mathbf{x}_{i2}$ represents the coordinates of the corresponding mate in $S2$, then the relationship of the heavy-atom positions in the two subunits $S1$ and $S2$ is given by:

$$\mathbf{x}_{i2} = A\mathbf{x}_{i1} + \mathbf{T} . \tag{1}$$

Here A is not a rotation matrix but rather is a 3×3 linear transformation matrix of elements a_{pq}. $\mathbf{T}$ is a translation vector with elements t_p. The linear least-squares fitting of equation (1) for the nine elements of A and the three of $\mathbf{T}$, for the N observational equations, will give starting values for the parameters which relate the two subunits.

Stage 2 of the procedure is to refine the matrix A and the vector $\mathbf{T}$ in real space using the electron density maps for each subunit. A similar scheme for this operation but using an Eulerian matrix has been given by Muirhead *et al.* (1967).

The procedure presented here differs in that rotational constraints are not applied to elements of the transformation matrix A. By allowing for shears and expansions this modification makes the construction of the normal equations considerably simpler than for an Eulerian matrix and avoids possible degeneracies in the solution vector. Our experience with this method as outlined below shows that elastic deformations of the electron density map are not observed.

Here we wish to shift the elements a_{pq} and t_p in such a way as to minimize the difference in the electron density between corresponding points in the two electron density maps of $S1$ and $S2$, that is, we wish to minimize the function

$$\delta = \varrho_1(\mathbf{x}_1) - \varrho_2(\mathbf{x}_2) \tag{2}$$

for a suitable large number of grid points, $\mathbf{x}_1$, where ϱ represents the electron density distribution function. Define $\mathbf{C}$ to be the 12-element vector of shifts in the elements of A and $\mathbf{T}$ so that the density at the new point, $\mathbf{x}_2'$, is given by

$$\varrho_2(\mathbf{x}_2') = \varrho_2(\mathbf{x}_2) + B\mathbf{C} , \tag{3}$$

where for m observations the matrix B is of order $m\times12$.

Muirhead *et al.* (1967) have shown that the solution for the parameter-shift vector is given by

$$\mathbf{C} = (\tilde{B}B)^{-1}\tilde{B}\delta .$$

The calculation of the elements of B from the electron density map for $S2$ is straightforward. For example:

$$B_{11} = \frac{\partial\varrho_2}{\partial x_2}\frac{\partial x_2}{\partial a_{11}} + \frac{\partial\varrho_2}{\partial y_2}\frac{\partial y_2}{\partial a_{11}} + \frac{\partial\varrho_2}{\partial z_2}\frac{\partial z_2}{\partial a_{11}} = \frac{\partial\varrho_2}{\partial x_2}x_1 ,$$

since $\dfrac{\partial x_2}{\partial a_{11}} = x_1$ and the other terms vanish.

So that the rows of the matrix B are given by

$$\frac{\partial\varrho_2}{\partial x_2}x_1 \quad \frac{\partial\varrho_2}{\partial x_2}y_1 \quad \frac{\partial\varrho_2}{\partial x_2}z_1 \quad \frac{\partial\varrho_2}{\partial y_2}x_1 \quad \frac{\partial\varrho_2}{\partial y_2}y_1 \quad \frac{\partial\varrho_2}{\partial y_2}z_1$$

$$\frac{\partial\varrho_2}{\partial z_2}x_1 \quad \frac{\partial\varrho_2}{\partial z_2}y_1 \quad \frac{\partial\varrho_2}{\partial z_2}z_1 \quad \frac{\partial\varrho_1}{\partial x_2} \quad \frac{\partial\varrho_2}{\partial y_2} \quad \frac{\partial\varrho_2}{\partial z_2} .$$

Maximizing the overlap in electron density between $S1$ and $S2$ produces a suitably accurate linear transformation matrix and vector.

Electron density map modification

In order to place the electron density distribution for $S1$ into the space of $S2$, the map of $S1$ must be modified to contain only the electron density for the subunit. That is, the electron density of solvent regions and the molecular fragments from other symmetry-related subunits must be set to a uniform background level. This can be done by a visual inspection of the map and constructing a circumscribing polygon which coarsely outlines the subunit boundary. The set of polygons (one for each map section) is then used as a boundary such that all the electron density at points outside this boundary can be set to the average background value of the solvent region.

The polygon must be chosen so as to exclude all neighboring molecular fragments and not remove density from the subunit of interest. An exact outline is not critical because, provided this requirement has been met, all unwanted structural information has been excluded. It is of no consequence that low-density solvent regions might be included inside the bounding polygon. This was the case in our application to hexokinase where we found considerable decrease in background for relatively extensive regions of solvent regions which were included within the bounds of the subunit. In practice a ten-sided irregular polygon is more than adequate to exactly isolate the molecule. Because there are generally very large intersubunit solvent regions in protein crystals and relatively few intimate contacts between the subunits this isolation procedure is not difficult for a map of even moderate to poor quality.

The final map modification is to construct an electron density map of a full unit cell in $S2$ space containing the $S1$ density distribution. This is readily accomplished using the refined linear transformation and the space-group-symmetry operators in the crystal form of $S2$. The FFT algorithm is then used to transform this map and produce a set of calculated phases, $\alpha(\text{calc})_{S1}$, and calculated structure-factor amplitudes, $F(\text{calc})_{S1}$.

Phase combination

The optimal method for combining these calculated phases with the observed MIR phases would be to multiply the total observed phase probability profile for each of the experimental MIR phases by a cal-

culated Gaussian phase probability profile determined for each of the calculated phases (Rossmann & Blow, 1961; Hendrickson & Lattman, 1970; Hendrickson, 1973). The centroid phase from this joint probability profile could then be appropriately weighted and used as the hybrid phase.

We have used a computationally simpler method for the phase combination which takes advantage of the estimated phase errors from the two sources of MIR phase information. The figure of merit, m, is a sensitive indicator of the relative error for the $S2$ MIR phases, α(MIR). The quantity $r=|F(\text{obs})_{S2}-F(\text{calc})_{S1}|/|F(\text{obs})_{S2}+F(\text{calc})_{S1}|$ was chosen to estimate the relative error in the calculated phases, $\alpha(\text{calc})_{S1}$. The details of the method in which the phases were combined and the justification are described below.

Application to hexokinase

This method of combining the information from independent determinations of the same structure has been successfully used to improve the electron density maps of two different crystal forms of yeast hexokinase B. Table 1 presents a summary of some crystallographic parameters for two orthorhombic crystal forms of yeast hexokinase B. Crystal form BII contains two subunits in the asymmetric unit (Anderson *et al.*, 1974) and form BIII contains one subunit in the asymmetric unit (Fletterick *et al.*, 1975) so that there are three independently phased structure determinations of the hexokinase subunit. The following gives the description of the methodology used to isolate the BIII subunit, $S3$, and to place the $S3$ electron density in the positions of the two subunits ($S1$ and $S2$) in crystal form BII, and the subsequent application of the fast Fourier transform and phase combination.

Table 1. *Hexokinase crystallographic data*

Crystal form	Cell constants	Space group	Refinement parameters
BII	$a=144{\cdot}2$ Å	$P2_12_12_1$	Five heavy-atom
	$b=87{\cdot}4$	Two 51000 m.w.	derivatives
	$c=99{\cdot}4$	subunits/asym.	$\langle m\rangle=0{\cdot}67$
	$V=1{\cdot}25\times10^6$ Å^3	unit	3·5 Å resolution
BIII	$a=166{\cdot}5$ Å	$P2_12_12_1$	Seven heavy-
	$b=59{\cdot}2$	One 51000 m.w.	atom derivatives
	$c=58{\cdot}5$	subunit/asym.	$\langle m\rangle=0{\cdot}78$
	$V=5{\cdot}99\times10^5$ Å^3	unit	3·5 Å resolution

In the crystal form BII the trial matrix and vector relating the two subunits of the dimer ($S1$ and $S2$) were determined using equation (1) to relate eight heavy-atom positions in one subunit to the equivalent eight heavy-atom positions in the second subunit. The root-mean-square distance between the observed positions and those calculated using the matrix and vector for these eight sites was 1·09 Å. The relationship between $S1$ and $S3$ was determined in the same way but with six heavy-atom positions that were found to occupy similar positions of the protein subunit in the two crystal forms BII and BIII. The root-mean-square distance between the six observed and calculated sites in this case was 1·0 Å.

The matrix that relates the two subunits in the asymmetric unit of form BII was then refined using 1000 randomly selected grid points of moderate to high electron density within the $S1$ subunit of the 3·5 Å electron density map. The gradient calculation was limited to 1000 points instead of the entire subunit array because this calculation proved to be computationally the most time consuming. The electron density gradient was calculated at each of these grid points by evaluating the electron density at 1 Å intervals in the three perpendicular coordinate directions. The electron density was estimated with the use of an eight-point linear interpolation between $S1$ grid points in three dimensions. Three cycles showed that the refinement had converged to give negligible shifts in the elements of the A matrix and **T** vector. The convergence of this refinement was further checked by careful visual comparison by superposition of each of the rotated density-map sections for $S2$ with the unaltered map of $S1$. No systematic errors in the transformation parameters or distortions of the electron density were found. The two subunits of the dimer, $S1$ and $S2$, were found to be related by a 156° rotation and a 13·8 Å translation along the rotation axis. The subunit electron density map of $S2$ was next rotated and translated so as to superpose with the $S1$ density and the average density was taken. Electron density at non-integral lattice points was estimated using an eight-point linear interpolation. This step produced a more noise-free $S1$ density map for the next step in the refinement.

The matrix relating $S3$ and $S1$ was refined in the same way with the averaged map in the $S1$ position. Four refinement cycles were necessary for convergence. Once again a visual comparison of the rotated and unaltered map sections superposed showed no errors or distortions. In this case a rather more substantial change in the transformation parameters was observed than for the case relating $S1$ to $S2$.

The electron density of the monomer ($S3$) was isolated from the rest of the BIII electron density map by tracing a decagon around the molecule on each of the 48 Fourier map sections. The boundary in most cases was placed more than 5 Å from density features assigned to the subunit. The separation between subunits in the crystal is obvious in all areas of the map. All density outside this boundary was set to zero, the average electron density of the unit-cell. The map of $S3$ was then placed in the two positions occupied by $S1$ and $S2$ subunits in crystal form BII and an entire unit cell was generated with the BII space-group-symmetry operators.

Before the three subunits in real space or reciprocal space were averaged, it was necessary to show that their tertiary structures are the same. This was checked by visual comparison of the structures of the MIR-phased electron density distributions of each of the

three subunits rotated into the space of BIII (Figs. 3 and 4) and by subtracting the subunit density maps pairwise and looking for any extensive regions of difference density. These difference maps showed that the tertiary structures of the subunits are identical at 3·5 Å resolution in most regions, although there are indeed some small tertiary structural differences between the three subunits, particularly at the dimer interface.

Since the electron density map has only been interpreted in terms of a polypeptide backbone, analysis of the difference electron density maps between the various subunits has not been complete or quantitative. A few side chains at the subunit interface move by several Ångstroms. Variation in the positions of some of the polypeptide backbone on the surface or at the subunit interface may be as large as an Ångstrom, but most differences in structure would appear to be substantially less. We did not attempt to analyze possible alternative positions for side chains in the subunits. The structural differences will be described elsewhere (Fletterick *et al.*, 1976) since they are sufficiently small that they will not affect the results being presented here.

The Fourier transform of $S3$ density in the BII cell was calculated to produce $F(\text{calc})_{S3}$ and $\alpha(\text{calc})_{S3}$. The grid intervals used in calculating the fast Fourier transform satisfy the sampling criterion proposed by Collins (1975). The number of grid points along the cell axes in crystal form BII was 128, 64 and 64 for x, y and z, respectively.

The F(obs) for BII were scaled to these calculated structure-factor amplitudes with a temperature factor of $B=30{\cdot}1$ Å^2 applied to F(calc). This value of B corrects for the different rates of fall-off with $\sin\theta/\lambda$ of the observed structure amplitudes for the two crystal forms and the opposite effect of the figure-of-merit weighting used to compute the BIII electron density map. The residual, $R=2\sum|F(\text{obs})-F(\text{calc})|/\sum|F(\text{obs})+F(\text{calc}|$, was 0·57 at 10 Å resolution and smoothly varied to 0·34 at 3·5 Å resolution. The R value for all 8600 reflections was 0·38. It is interesting to note for comparison that the R value between BIII F(obs) and BIII F(calc) was 0·33. In this case F(calc) was computed by applying the fast Fourier transform to the BIII crystal form with the electron density of the solvent regions set to zero. We conclude from the calculated R values then that the refined linear transformation relating $S1$ to $S3$ was sufficiently accurate and that the subunit tertiary structure is homologous in the two crystal forms.

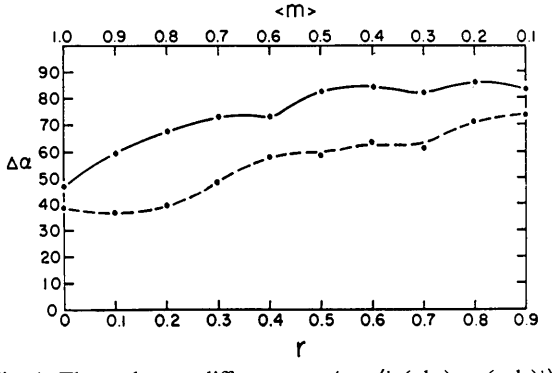

Fig. 1. The phase difference, $\Delta\alpha=\langle|\alpha(\text{obs})-\alpha(\text{calc})|\rangle$ in degrees between the observed MIR phases for BII and those calculated from the BIII electron density. The dashed line shows the phase differences as a function of the quantity $r=|F(\text{obs})-F(\text{calc})|/|F(\text{obs})+F(\text{calc})|$ for those BII MIR reflections having a figure of merit greater than 0·9. As expected, the agreement between observed and calculated structure factors is correlated with the agreement between the observed and calculated phases. The solid line shows the phase difference between observed and calculated phases as a function of figure of merit for all reflections. Again, the figure of merit is correlated with phase error, as anticipated.

Phase combination

In order to combine the observed MIR phases with the calculated phases, it is necessary to make use of the estimated errors in these phases. The figure of merit, m, indicates the relative error in the BII MIR phases, α(MIR). The quality r, which measures the agreement between the observed and calculated structure factors, was chosen to measure the relative error in the calculated phases, $\alpha(\text{calc})_{S1}$. Fig. 1 shows a plot (dashed line) of r *versus* the phase difference ($\Delta\alpha$) between the calculated phases and those observed phases having a figure of merit larger than 0·90. The phase difference, $\Delta\alpha=\alpha(\text{obs})-\alpha(\text{calc})$ generally increases for increasing values of r. Assuming the observed phases of $m\geq 0{\cdot}9$ to be relatively error free, it can be seen that the calculated phases for r greater than about 0·4 to 0·5 have a 60° to 75° phase error. Fig. 1 also shows the variation between the observed and calculated phases as a function of the figure of merit of the MIR-phased reflections. This demonstrates that while the phase change is quite large, the smallest phase changes are associated with high figure of merit reflections.

Based on the data in Fig. 1 the following four cases were defined for purposes of combining the phases. Case 1: for $r>0{\cdot}5$ and $m>0{\cdot}6$ the reflection was not used (530 reflections). Case 2: for $r<0{\cdot}5$ and $m<0{\cdot}7$ the α(calc) (3500 reflections) phases were used with a figure of merit of 1·0. Case 3: $r>0{\cdot}5$ and $m>0{\cdot}6$ the figure of merit weighted α(obs) from MIR were used (920 reflections). For the remaining data (4650 reflections) with $r<0{\cdot}5$, and $m>0{\cdot}7$, a weighted average of the α(obs) and α(calc) was used. The combination scheme in this case allowed the α(obs) phase to rotate in the direction of α(calc) by an amount $f\Delta\alpha$ where f was constrained to fall in the range (0·25–0·75). The magnitude of this fractional rotation, f, was determined from the function $f=1{\cdot}33-0{\cdot}5r-0{\cdot}83m$.

Using this functional form for f, with $m=1{\cdot}0$ and $r=0{\cdot}5$, the observed phase (presumably more reliable), α(obs), rotates in the direction of α(calc) by an amount $0{\cdot}25\ \Delta\alpha$. With $m=0{\cdot}7$ and $r=0{\cdot}0$ (the calculated phase presumably more reliable), α(obs) rotates in the direction of α(calc) by an amount $0{\cdot}75\ \Delta\alpha$. At other values

of r and m for this case f will take on an intermediate value between 0·25 and 0·75. For $r=0·5$ and $m=0·7$, for example, $f=0·5$ and the resulting phase is midway between α(obs) and α(calc). The phase change relating these hybrid phases to the original MIR phases is shown in Fig. 2 by the dotted line. The average phase change for the 8600 reflections was 50° which is very close to the average error in the MIR phases (48°) predicted from the average figure of merit (0·67).

Any errors in the linear transforms used to relate the three subunits or differences in the tertiary structure of the three subunits would result in increasing phase differences between the BII MIR phases and the BIII calculated phases with increasing resolution. In order to evaluate these phase differences the BII MIR phases with high figure of merit were compared with the BIII calculated phases after the phase error estimated by the figure of merit had been subtracted from these latter phases. The dashed line in Fig. 2 is calculated by subtracting for each resolution shell the arccosine of the average figure of merit for the BIII phases from the observed phase change (smooth line) between the BII MIR phases with $\langle m \rangle > 0·9$ and the BIII calculated phases. Since the phase difference is relatively flat as a function of resolution, there is probably little systematic error introduced into the calculated phase by errors in the linear transformation matrix or differences in tertiary structure.

Results and discussion

Our most striking result is that the electron density map obtained using these phase-combination procedures is a substantial improvement over either of the two MIR maps or the map obtained by averaging the two MIR electron density maps. The background is lower, the resolution is higher and most importantly, side chains are present in the map calculated with hybrid phases where none exist in either of the MIR maps. While a simple quantitative analysis of the degree of phase improvement achieved by these procedures is not possible, a qualitative assessment can be obtained by simply comparing each of the initial electron density maps with the resulting electron density maps computed with the hybrid phases.

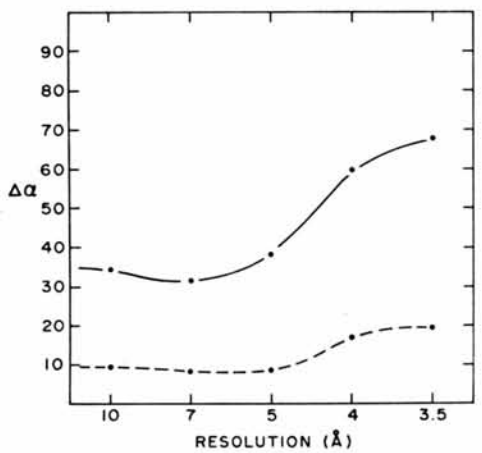

Fig. 2. The solid line shows the phase difference between the observed MIR phases for BII with figure of merit greater than 0·9, and the phases calculated from the BIII MIR electron density plotted as a function of resolution. The dashed line was obtained by subtracting for each resolution shell the arccosine of the BIII average figure of merit from the points on the solid line.

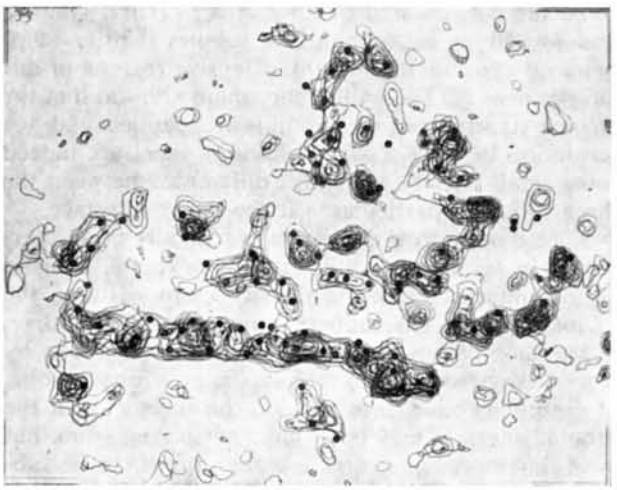

Fig. 3. A 6 Å slice of the BII MIR electron density map with the two subunits averaged in real space. The first contour above zero has been omitted.

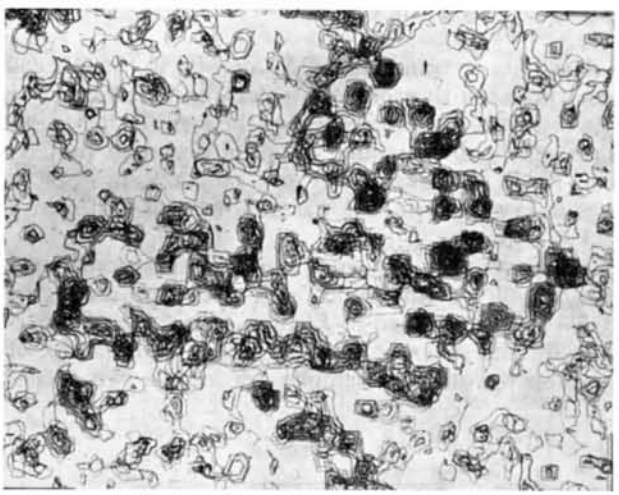

Fig. 4. A 6 Å slice of the BIII MIR electron density map.

In the following discussion we shall need to refer to five different electron density maps: the BIII MIR, BII MIR; the BII hybrid and BIII hybrid; and the BII–BIII average. The method used to calculate each of these maps was as follows: (1) the BIII MIR map was computed using the observed figure-of-merit-weighted structure-factor amplitudes for BIII and the associated MIR phases. (2) The BII MIR map was computed using the observed figure-of-merit-weighted

structure-factor amplitudes for BII and the MIR phases. After electron density maps had been calculated for each subunit of the dimer, the electron density distributions were averaged. The electron density of the averaged subunit was rotated into the same orientation as observed for the BIII MIR map. (3) The BII hybrid map was computed using the observed structure amplitudes for BII and the hybrid phases determined as described above. The electron density of the two subunits was then averaged and again rotated into the same orientation as the BIII MIR map. Thus, the only differences in computation of the two BII maps are in the phases used. (4) The BIII hybrid map was computed using the observed structure amplitudes for the BIII crystal form and hybrid phases. These phases were calculated from the Fourier inversion of the BII hybrid map oriented in the space of BIII. These calculated phases were not combined with the BIII MIR phases. (5) The BII–BIII average map is the direct-space average of the BII and BIII MIR maps.

In order to compare the quality of these maps, slices about 6 Å thick from each of these five electron density maps are shown in Figs. 3 through 7. The contour intervals were equally spaced for each map with the zero contour level omitted; in the BII MIR map the first contour level is also absent. A prominent α helix, 50 Å long, can be seen in each case in the lower third of the map. The dots on the two BII density maps represent α carbon positions determined from interpretation of the 2·7 Å resolution BIII MIR electron density map (Fletterick *et al.*, 1975). These α carbon positions were plotted on the BII electron density maps in order to check the accuracy of the linear transformations relating the subunits and the degree of similarity of the tertiary structures.

A comparison of the BII MIR map and the BIII MIR map (Figs. 3 and 4) shows that the two are of very different quality. Examination of the 50 Å helix in particular shows that the BII MIR map is of much lower resolution than the BIII MIR map even though both are calculated from data extending to 3·5 Å resolution. This is due to the much higher overall temperature factor exhibited by the BII structure factors, and the much poorer MIR phasing of the BII amplitudes between 3·5 and 4·5 Å resolution (Fletterick *et al.*, 1976). The BII MIR map appears to have a lower background because the densities of two subunits of the dimer have been averaged and because the first contour has been omitted. However, the α helix more nearly resembles a low-resolution rod of electron density with no side chains, instead of the helix of backbone density with periodic side chains seen in the BIII MIR map.

The object of the phase-combination procedures described here has been to reduce the obvious noise in the BIII MIR map without losing the higher-resolu-

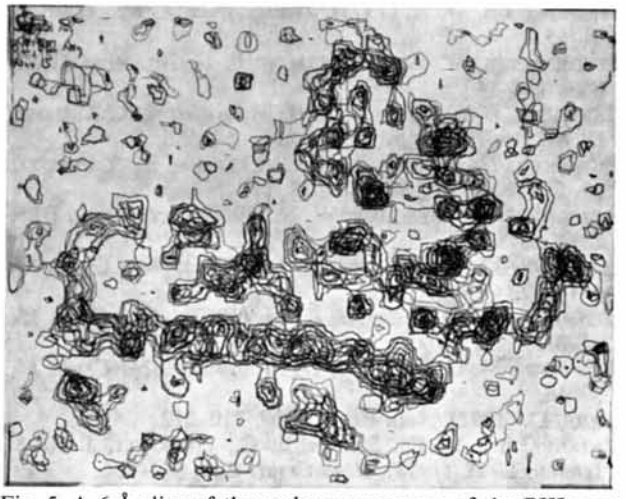

Fig. 5. A 6 Å slice of the real-space average of the BIII map with the BII MIR map.

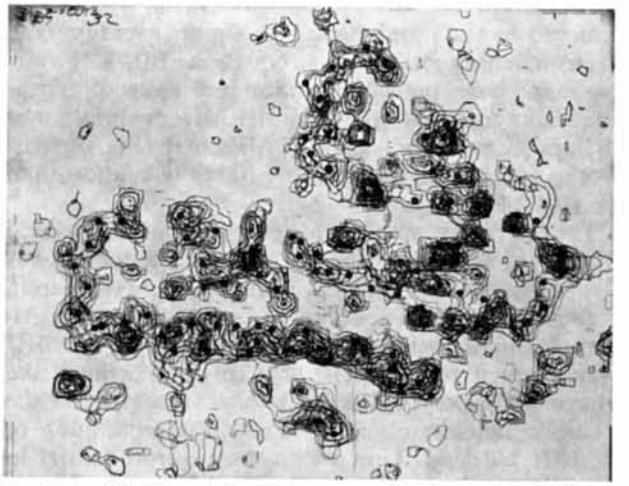

Fig. 6. A 6 Å slice of the BII electron density map calculated with hybrid phases.

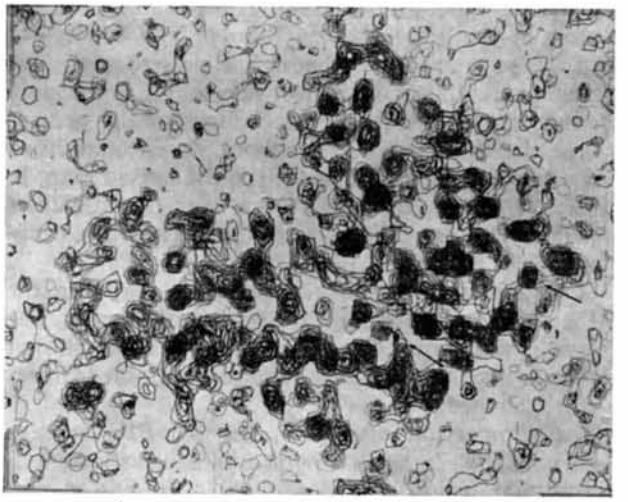

Fig. 7. A 6 Å slice of the BIII electron density map calculated with hybrid phases. Several side chains (arrowed) which are visible in this map do not appear in any of the previous maps. Also, the long α helix is better resolved.

tion information which it contains. This goal is not achieved by the usual method of simply averaging the electron density maps. For although the BII–BIII average map shows higher resolution and better polypeptide chain continuity than the BII MIR map, it is not an improvement upon the BIII MIR map (Fig. 4). This average map is of considerably lower resolution than the BIII MIR map.

The BII hybrid map in which the combined or hybrid phases are applied to the BII observed amplitudes is a considerable improvement over the BII MIR map in both resolution and continuity of electron density. Although this map is not as well resolved as the BIII MIR map, it was in fact easier to trace through the course of the polypeptide backbone. In comparing the α carbon markers obtained from an interpretation of the BIII 2·7 Å MIR map with the electron density in the BII hybrid map, they were found to generally lie inside an area of non-zero density. In a few cases, however, it was found that an error had been made in tracing through a loop of chain near the surface of the protein in the BIII 2·7 Å MIR map. The BII hybrid map was so much better that it was clear that in a few places we had interpreted a few side chains as the main chain in the 2·7 Å map. No changes in the general course of the polypeptide backbone were found, however, after careful inspection of the BII hybrid map. In contrast to the MIR maps, there were no difficulties in tracing through the course of the polypeptide backbone in the BII hybrid map.

Comparison of the BII–BIII average map and the BII hybrid map shows rather subtle differences. The hybrid phase map, however, appears to be somewhat better resolved and the density of recognizable features is better shaped.

The comparison of Fig. 5, the BII–BIII average map, with Fig. 7, the BIII hybrid map, clearly demonstrates that reciprocal-space phase averaging is preferred to a simple real-space average. The background is lower, the apparent resolution is higher and the map is far easier to interpret.

The comparison of the BIII hybrid map with the BII hybrid map in Figs. 6 and 7 shows the BIII hybrid map to be of significantly higher apparent resolution. Since these two maps were calculated with essentially the same hybrid phases, the rapid fall-off of the F(obs) with increasing resolution for the BII crystal form is responsible for this phenomenon. It should be noted that a temperature factor of $B=30{\cdot}1$ Å was required to scale the F(calc) from the BIII crystal form to the observed structure amplitudes for the BII crystal form. Not only is the BIII hybrid map superior in resolution to both the BII hybrid and the BII–BIII average maps, but it is a striking improvement over the BIII MIR map. The background is lower, the resolution of the α helix is higher and, most importantly, side chains are visible in this map where none exist in the MIR maps.

In summary, we have been able to use the information contained in the two independent crystal structure determinations to produce a map which is superior to either of the starting maps.

The computational times on an IBM 370/158 computer involved in the various steps of the above analysis were as follows: The CPU time required for one cycle of matrix refinement by the gradient technique was 2·5 min. The CPU time needed to isolate the subunit density, rotate the density distribution function, reconstruct the unit cell and calculate a set of hybrid phases was 12 min.

Note added in proof: G. Bricogne has recently applied a similar but more rigorous method of phase improvement to the solution of TMV coat protein and glyceraldehyde 3-phosphate dehydrogenase (Bricogne, 1975*a*, *b*).

We thank Charles Anderson and Wayne Anderson for discussions and help with this work.

References

ANDERSON, W. F., FLETTERICK, R. J. & STEITZ, T. A. (1974). *J. Mol. Biol.* **86**, 261–269.
BARRETT, A. N. & ZWICK, M. (1971). *Acta Cryst.* A**27**, 6–10.
BLOW, D. M. & CRICK, F. H. C. (1959). *Acta Cryst.* **12**, 794–802.
BRICOGNE, G. (1974). *Acta Cryst.* A**30**, 395–405.
BRICOGNE, G. (1975*a*). Presented at the Summer School on Crystallographic Computing, Prague.
BRICOGNE, G. (1976). Personal communication.
BUEHNER, M., FORD, G. C., MORAS, D., OLSEN, K. W. & ROSSMANN, M. G. (1974). *J. Mol. Biol.* **82**, 563–585.
COLLINS, D. M. (1975). *Acta Cryst.* A**31**, 388–389.
FLETTERICK, R. J., ANDERSON, W. F., ANDERSON, C. & STEITZ, T. A. (1975*b*). *J. Mol. Biol.* Manuscript submitted.
FLETTERICK, R. J., BATES, D. J. & STEITZ, T. A. (1975). *Proc. Natl. Acad. Sci. US.* **72**, 38–42.
FREER, S. T., ALDEN, R. A., CARTER, C. W. & KRAUT, J. J. (1975). *Biol. Chem.* **250**, 46–54.
GREEN, D. W., INGRAM, V. M. & PERUTZ, M. F. (1954). *Proc. Roy. Soc.* A**225**, 287–296.
HENDRICKSON, W. A. (1973). *Trans. Amer. Cryst. Assoc.* **9**, 61–65.
HENDRICKSON, W. A. & LATTMAN, E. (1970). *Acta Cryst.* B**26**, 136–143.
MATTHEWS, B. W., SIEGLER, P. B., HENDERSON, R. & BLOW, D. M. (1967). *Nature, Lond.* **214**, 652–656.
MUIRHEAD, H., COX, J. M., MAZZARELLA, L. & PERUTZ, M. F. (1967). *J. Mol. Biol.* **28**, 117–130.
REEKE, G. N. & LIPSCOMB, W. N. (1969). *Acta Cryst.* B**25**, 2614–2623.
ROSSMANN, M. G. & BLOW, D. M. (1961). *Acta Cryst.* **14**, 641–647.
ROSSMANN, M. G. & BLOW, D. M. (1962). *Acta Cryst.* **15**, 24–31.
ROSSMANN, M. G. & BLOW, D. M. (1963). *Acta Cryst.* **16**, 39–45.
SAYRE, D. (1972). *Acta Cryst.* A**28**, 210–212.
WATENPAUGH, K. D., SIEKER, L. C., HERRIOTT, J. R. & JENSEN, L. H. (1973). *Acta Cryst.* B**29**, 943–956.
WEINZIEREL, J. E., EISENBERG, D. & DICKERSON, R. E. (1969). *Acta Cryst.* B**25**, 380–387.