Acta Cryst. (1996). D52, 447–452

Locating a Local Symmetry Axis from Patterson Map Cross Vectors: Application to Crystal Data from GroEL, GTP Cyclohydrolase I and the Proteosome

MILTON T. STUBBS,[a] HERBERT NAR,[a] JAN LÖWE,[a] ROBERT HUBER,[a] RUDOLF LADENSTEIN,[b] MICHAEL D. SPANGFORT[c] AND L. ANDERS SVENSSON[d]

[a]*Max-Planck Institut für Biochemie, D-82152 Martinsried bei München, Germany,* [b]*Centrum för Strukturbiokemi, Karolinska Institutet, NOVUM, S-141 57 Huddinge, Sweden,* [c]*Department of Biochemistry, Chemical Center, University of Lund, S-221 00 Lund, Sweden, and* [d]*Department of Molecular Biophysics, Chemical Center, University of Lund, S-221 00 Lund, Sweden. E-mail: stubbs@biochem.mpg.de*

(*Received 15 June 1995; accepted 11 October 1995*)

Abstract

The cross vectors of the native Patterson map are shown to exhibit non-crystallographic symmetry in the case of local axes parallel to one another. This information can be used to determine the translation component of such axes. A program is described to search for this cross vector, and is tested on low-resolution data from crystals of the tetradecameric GroEL molecule, the decameric GTP cyclohydrolase I and the tetradecameric proteosome. For GroEL, the function produces a packing arrangement optimal for sevenfold symmetry, and is in agreement with the dimensions of the molecule as given by electron microscopy data and the recently determined crystal structure. Positioning of local axes is confirmed by two high-resolution crystal structure analyses: the fivefold axis in cyclohydrolase I and the sevenfold axis in the proteosome. Implications for the location of heavy-atom positions are discussed for these two cases.

1. Introduction

The Patterson map, derived directly from the experimental native intensity data, can be thought of as consisting of self-vector (intramolecular) and cross-vector (intermolecular) components (Hoppe, 1957). The self-vector set can be used to determine the relative orientations of individual molecules in the crystal by way of the rotation function (Rossmann & Blow, 1962). The cross-vector set determines the relative positions of these molecules, and is the target for the translation function. These functions provide the basis for so-called Patterson search or molecular-replacement methods [for a collection of papers, see Rossmann (1972)]. Particular success has been found in the solution of unknown structures using that of a known model.

In the case of an unknown molecule, the rotation function is used to determine the orientation of local symmetry elements. The cross-vector set is rarely examined, other than in the case of a local twofold axis parallel to a crystallographic axis; this results in a set of vectors that overlap to produce a 'pseudo-origin peak' (Eagles, Johnson, Joynson, McMurray & Gutfreund, 1969; Epp, Steigemann, Formanek & Huber, 1971). Rossmann, Blow, Harding & Coller (1964) showed that it is possible to determine the translational component parallel to a rotation axis (the 'precise' parameter) relating two molecules within the asymmetric unit.

Knowledge of local symmetry elements can be of great importance in the determination of heavy-atom positions, particularly in the case of weak binding to large multimeric complexes. Moreover, the power of non-crystallographic averaging techniques provides the impetus to determine both the orientations and positions of regions of equivalent density as early and as accurately as possible.

In this paper, we shall show that the cross-vector set of a *native* Patterson map can provide information on the relative translational component of symmetry axes parallel to one another, with the pseudo origin described above being a special case. This is illustrated using 19 Å resolution data from crystals of GroEL (Spangfort *et al.*, 1993), in which a local sevenfold axis lies approximately parallel to a local twofold axis (Svensson, Surin, Dixon & Spangfort, 1994). The solution found provides a packing arrangement for the molecules in the unit cell, and confirms dimensions for the molecule obtained from electron microscopy. The technique is also applied to 7 Å resolution crystal data from GTP cyclohydrolase, where the position of the local fivefold axis has been confirmed by the subsequent crystal structure (Nar *et al.*, 1995). Analysis of 15 Å resolution data from the proteosome (Löwe *et al.*, 1995) reveals the position of the local sevenfold axis, the implications of which are discussed for the interpretation of heavy-atom-derivative data.

These results may be of use in the analysis of derivative Patterson maps, in setting up of averaging procedures, and possibly in *ab initio* phasing methods.

2. Theory

Consider a molecule A consisting of n subunits with proper symmetry $[N]$ *i.e.* such that $[N]^n = [1]$, the identity matrix. Then coordinates in the ith subunit can be written as

$$[N]^i\mathbf{x}_A,$$

where $\mathbf{x}_A$ corresponds to a position in the reference subunit.

If we now consider a second molecule B related by another (crystallographic) symmetry matrix $[C]$ and a translation vector $\mathbf{t}$, coordinates in the jth subunit can be written as

$$[C][N]^j\mathbf{x}_B + \mathbf{t}.$$

Cross vectors in the Patterson map (*i.e.* difference vectors between A and B) may then be written as

$$[N]^i\mathbf{x}_A - [C][N]^j\mathbf{x}_B - \mathbf{t}.$$

If $[C]$ and $[N]$ represent parallel rotation axes, then they commute, and these cross vectors can be written

$$[N]^i(\mathbf{x}_A - [N]^{j-i}[C]\mathbf{x}_B) - \mathbf{t}.$$

Thus, for any two real-space vectors $\mathbf{x}_A$, $\mathbf{x}_B$ within one subunit of the molecule, we get a set of vectors in the Patterson map related by an n-fold axis about the cross vector $-\mathbf{t}$ (and $\mathbf{t}$ by Patterson inversion symmetry). This is illustrated for a two-dimensional example in Fig. 1. The vector found is that perpendicular to the rotation axis, *i.e.* the 'imprecise' parameter described by Rossmann, Blow, Harding & Coller (1964).

3. Programming the method

The above has been incorporated into the direct-space search routines of *PROTEIN* (Steigemann, 1974). The conventional self-rotation function is used to determine the orientation of the local symmetry axes from the (origin-removed) Patterson map. The search for a cross-vector symmetry axis proceeds as follows:

(i) selection of a set of peaks above a given threshold within a sphere of given radius centred on the current gridpoint;

(ii) rotation of these peaks about the required symmetry axis passing through the current centre;

(iii) calculation of the correlation of the peak set with the Patterson map through multiplication, scaled to the correlation for unrotated peaks at that point;

(iv) incrementation of the centre through the grid;

(v) go to (i).

In general, this search can be restricted to two dimensions, usually on a Harker section where the cross peaks are expected. Searching of a section perpendicular to the given axis results in a single peak where the cross-vector axis intersects the plane; a section parallel to the axis would result in a 'streak', with highest density presumably at the cross vector between the centres of mass (although this is dependent on the mass distribution of the molecule).

4. Application to crystals of GroEL

The crystallization of the *Escherichia coli* molecular chaperone cpn60 (GroEL) has been reported (Spangfort *et al.*, 1993). The crystals, diffracting to 7 Å, contain two tetradecameric GroEL molecules in the triclinic $P1$ unit cell ($a = 143.3$, $b = 154.6$, $c = 265.0$ Å, $\alpha = 82$, $\beta = 95$, $\gamma = 107°$). Analysis of the rotation function (Svensson, Surin, Dixon & Spangfort, 1994) suggested two molecules of D_7 point-group symmetry, related by a local twofold axis. The orientation of each of the sevenfold axes and of the intermolecular twofold axis

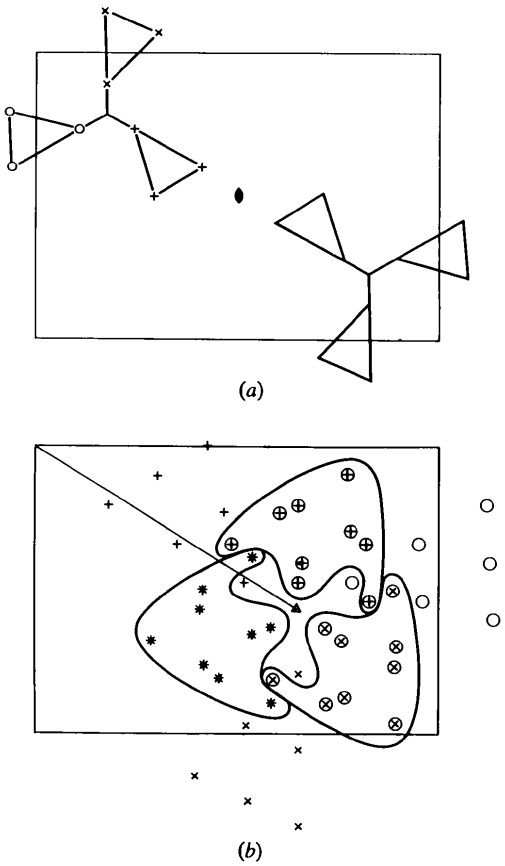


Fig. 1. (*a*) A threefold symmetric 'molecule' in a $p2$ planar unit cell. (*b*) The cross vectors arising from (*a*). + symbols show vectors from subunit $A1$ to B; × symbols show those from $A2$ to B; and o symbols show those from $A3$ to B. Despite the high degree of overlap (emphasized), a threefold axis is clear to see in the cross-vector set, at a vector corresponding to that between the centres of each rotation. [First presented by Stubbs *et al.* (1989).]

is approximately perpendicular to the *ac* plane. There should therefore be a local sevenfold axis at a cross vector corresponding to that between the two sevenfold axes.

An origin-removed Patterson map was calculated on a 3 Å grid (50 × 50 × 90) using data between 30.0 and 19.0 Å. An integration radius of 115 Å was chosen, corresponding to that giving the best results for the self-rotation function (Svensson, Surin, Dixon & Spangfort, 1994). Peaks greater than eight arbitrary units were picked from the Patterson map. The search for a local sevenfold axis, with an orientation in polar angles (ψ, φ, κ) = (5, 93, 51.4°) (ψ: inclination from y axis; φ: inclination from x axis; orthogonalization: x parallel a, z perpendicular to x in ac plane; y perpendicular to x and z) was made in the $v = 0$ plane. The results of this search are shown in Fig. 2.

Two large peaks were observed: one at grid coordinates (40, 0, 44) with a height 65% of the unrotated peaks, and one at (10, 0, 46) with a height of 40%. It will be noted that the second peak is equivalent to (–40, 0, –44). The peaks are of unequal height due to sampling on a coarse grid; subsequent scans on a 1 Å grid produced equivalent correlations for both peaks (data not shown).

The presence of such a high peak is in itself insufficient to prove that this is the sought cross vector. The behaviour of the correlation as a function of κ in the direction (ψ, φ, κ) = (5°, 93°, κ) provides clear evidence of a (slightly distorted) sevenfold axis (Fig. 3). This distortion is probably due to slight differences in the orientations of the two molecules, both from that observed in the self-rotation function and from each other. Indeed, further analysis of the cross-rotation function shows that the axis 'wobbles' in its orientation.

Analysis of the sevenfold cross correlation in different v sections revealed the highest correlation to be in the $v = 0$ layer, confirming that there is no difference in y coordinate of the two molecules (data not shown). There is also evidence for twofold axes perpendicular to the local sevenfold axis, but the peaks are considerably weaker than those of the self-rotation function.

Fig. 4 displays the packing arrangement of GroEL molecules concluded from these calculations in this crystal form. Maximal interactions occur in the direction of the crystallographic c axis, where the molecules tesselate, forming near-equivalent contacts for two faces of the heptagons. Less robust contacts are formed in the a direction. The crystal is then formed of stacks of these two-dimensional layers. From this packing arrangement, the maximal dimensions of the GroEL molecule are seen to be 70 Å radius by 150 Å in height, confirming measurements suggested by electron microscopy (Saibil & Wood, 1993) and in agreement with the recently determined X-ray structure (Braig *et al.*, 1994).

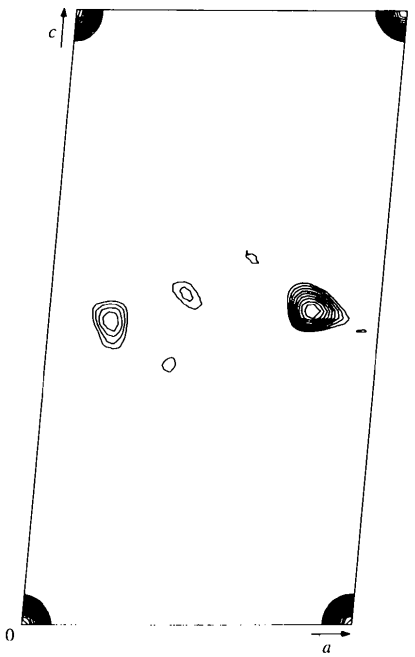


Fig. 2. Correlation function for $2\pi/7$ rotation in a search of the *ac* plane of GroEL crystals. At the four corners, a self-rotation-function correlation of 96% of the unrotated peaks is seen. The large peak to the right is at (40/50, 0/50, 44/90), with a height of 65%; the second largest peak at (10/50, 0/50, 46/90) [= (–40/50, 0/50, –44/90)] has a height of 40%. Average value of the correlation = 1.8%; σ = 10.5%. See text for further details.

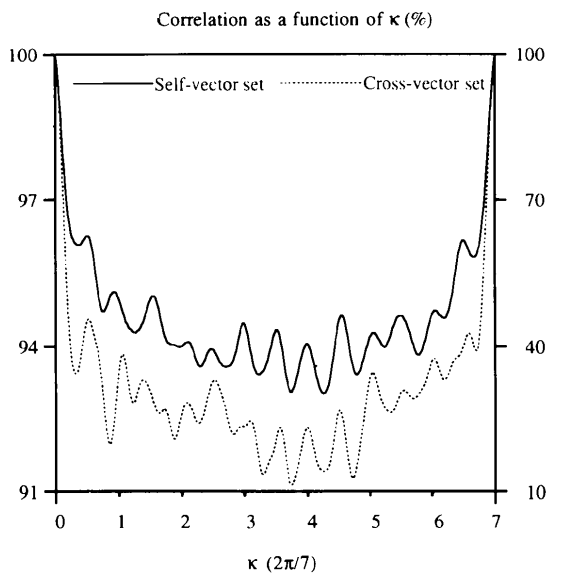


Fig. 3. Polar distribution about an axis of (5°, 94°, κ) for vectors centred at the origin (self-rotation function) and at the cross vector between the two sevenfold axes in the GroEL crystals shown in Fig. 2. The 14-fold nature of the self-rotation function arises from the overlap of the sevenfold particle symmetries related by a local twofold axis. The cross rotation shows the clear signature of a sevenfold axis; the presence of extra peaks suggests twinning of these crystals.

Surprisingly, Fig. 3 suggests a 14-fold axis rather than the expected sevenfold symmetry at the cross-vector position. This can be explained if one assumes a twinning in the crystals. The packing arrangement as determined in this study can be as shown in Fig. 4; an equivalent packing would be obtained if both molecules were rotated about their respective sevenfold axes by 180°. If both of these arrangements were found in the same crystal, then one would observe a 14-fold axis at the cross-vector position. This would also result in an overall translational symmetry, resulting in a pseudo-origin peak. Analysis of the native Patterson map indeed shows a small peak at the position of the cross vector (data not shown).

5. Application to crystals of GTP cyclohydrolase I

Monoclinic crystals of GTP cyclohydrolase I (space group $P2_1$, $a = 204.2$, $b = 210.4$, $c = 71.8$ Å, $\beta = 95.7°$) diffract to 3 Å resolution and contain two decameric enzyme complexes in the asymmetric unit (Schmid, Ladenstein, Luecke, Huber & Bacher, 1992). Electron microscopic studies and rotation-function calculations using 7 Å resolution X-ray data suggested D_5 point symmetry (Schmid *et al.*, 1993), and two distinct orientations of the independent decamers, with their fivefold axes perpendicular to the crystallographic b axis at an inclination of ($\psi = 90$, $\varphi = 84°$) and ($\psi = 90$, $\varphi = 96°$). A search was made for fivefold symmetry axes in the Patterson cross vectors in the uv plane ($w = 0$). Vectors of length 15–50 Å were picked from the 7 Å resolution Patterson map calculated on a 1 Å grid (200 × 200 × 72). A single large peak was observed for each orientation of the fivefold axis [($\psi = 90$, $\varphi = 84$, $\kappa = 72°$) and ($\psi = 90$, $\varphi = 96$, $\kappa = 72°$)], with peak heights 33 and 25% of the unrotated peaks, respectively. The two peaks coincided at grid coordinates (33, 100, 0); a scan in φ ($\psi = 90°$, $80 < \varphi < 100°$, $\kappa = 72°$) at this position confirmed the presence of maxima at the two expected orientations. A plot of the correlation as a function of κ (Fig. 5) clearly shows the fivefold symmetry around the cross-vector position. It was not possible to determine the cross vector between the two independent fivefold axes, indicating that the relative orientation of 12° is already too large to produce a local symmetry axis in the Patterson map at this resolution.

The positioning of the fivefold axes was confirmed by the subsequent X-ray structure determination at 3 Å resolution (Nar *et al.*, 1995), where the calculated cross-vectors coincide at grid coordinates (34, 100, 0). The structure was solved using a single Ta_6Br_{14} derivative. Inspection of the Harker sections revealed that the heavy-metal cluster was positioned on the local fivefold axes determined previously, thereby simplifying the interpretation of the difference Patterson to a search for four independent heavy atoms.

6. Application to crystals of the proteosome

Orthorhombic crystals of the tetradecameric proteosome diffracting to 3.2 Å resolution (space group $P2_12_12_1$, $a = 311.9$, $b = 209.0$, $c = 117.2$ Å) contain one proteosome in the asymmetric unit (Jap *et al.*, 1993), with a local sevenfold axis inclined 6.4° to the crystallographic a axis. A search for a local sevenfold axis in the Patterson cross vectors in the vw plane ($u = 1/2$) using 15 Å resolution data and a 2 Å grid (150 × 100 × 50) yielded an elongated peak centred at grid coordinates (88, 0, 75) (Fig. 6). This peak did not appear using higher-resolution data, presumably as a result of the 12.8° relative orientation of symmetry-related sevenfold

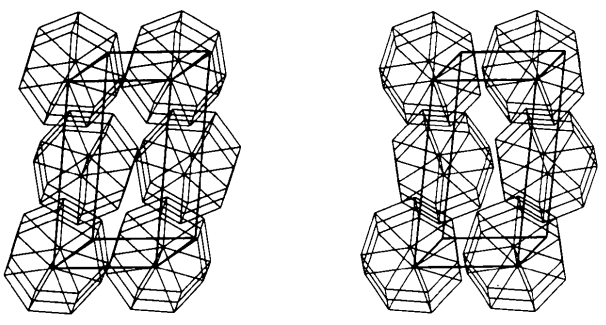


Fig. 4. Packing arrangement for GroEL derived from self- and cross-rotation-function results.

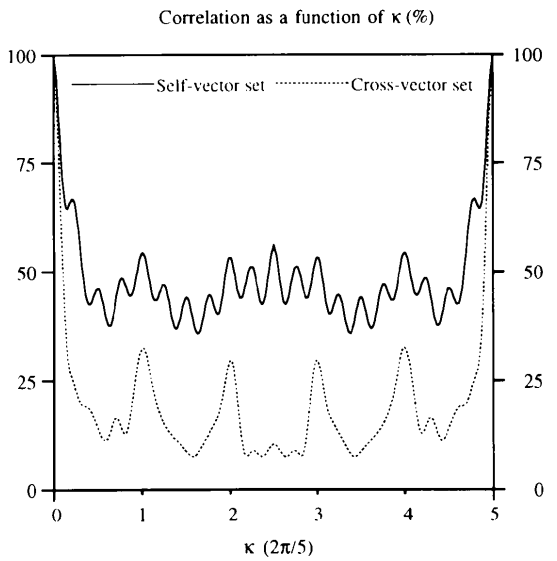
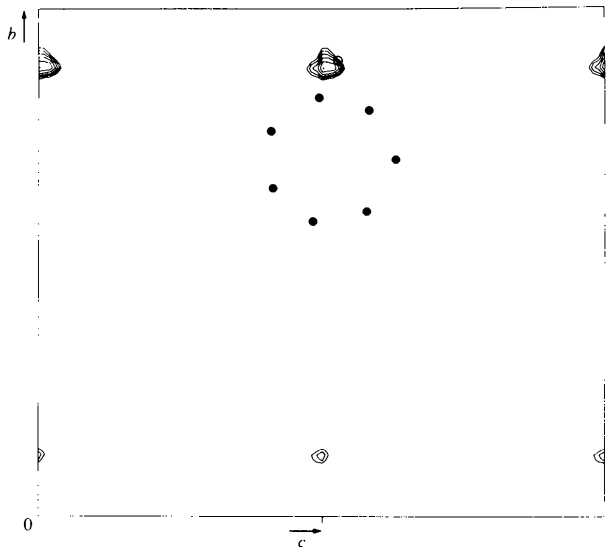
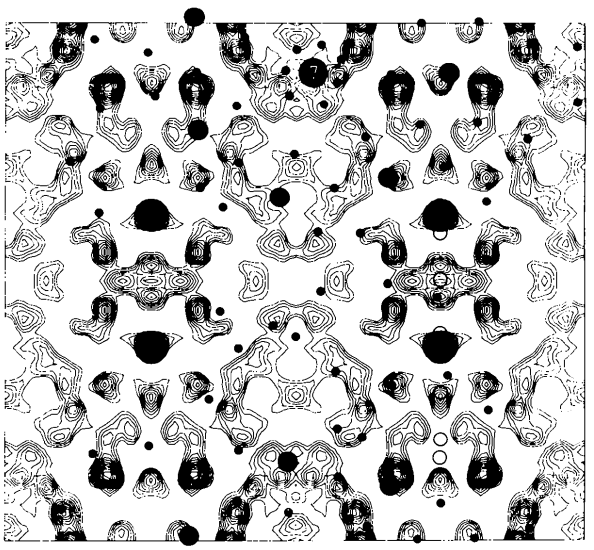


Fig. 5. Polar distribution about an axis of (90°, 84°, κ) for vectors centred at the origin (self-rotation function) and at the cross vector between the two fivefold axes in the monoclinic GTP cyclohydrolase I crystals. The 20-fold axis of the self-rotation function comes from the overlap of two parallel fivefold axes, with the perpendicular twofold axes oriented at approximately 90° to one another. This symmetry reduces to a fivefold axis at the cross vector between the two molecules.

axes (*cf.* the case for cyclohydrolase above), and was therefore not pursued further. The 3.4 Å resolution structure determination of the proteosome (Löwe *et al.*, 1995) revealed this peak to be in the correct position, however.



(*a*)



(*b*)

Fig. 6. (*a*) Correlation function for $2\pi/7$ rotation in a search of the *bc* plane ($a = 1/2$) of native proteosome crystals, using data to 15 Å resolution. An elongated peak centred at (75/150, 88/100, 50/50) is observed, with a height 52% of the unrotated peaks. Superimposed is a projection of the Ta_6Br_{14} positions used for phasing (filled circles), and the calculated position of the cross peak (open circle). (*b*) Harker section ($u = 1/2$) for difference Patterson (Ta_6Br_{14}–NATI) showing all positive contours. Superimposed are the position of the local sevenfold axis ('7'), the calculated Harker peaks (large filled circles), the local symmetry satellite peaks (small filled circles), and the peaks arising from the $w = 1/2$ Harker section (open circles). All of the Harker peaks are defined by density; some of the satellite peaks and $w = 1/2$ peaks do not appear as the local sevenfold axis is not absolutely parallel to the *a* axis.

Initial phases for the proteosome were obtained using a single Ta_6Br_{14} derivative at 7 Å resolution, the positions of which were determined using consecutive rotational and translational searches (Löwe *et al.*, 1995). It is of interest to know whether prior knowledge of the position of the local sevenfold axis would have allowed ready analysis of the heavy-atom positions. The $u = 1/2$ Harker section for the 7 Å resolution Ta_6Br_{14}–NATI difference Patterson is shown in Fig. 6. One would expect to see a sevenfold ring of Harker vectors in this section, although this is obscured for several reasons.

(*a*) The inherent *mmm* Patterson symmetry.

(*b*) The presence of 'satellite peaks'. As the heptagon plane of heavy atoms is almost perpendicular to the *a* axis, not only do the cross vectors between one atom and its symmetry mate (the 'Harker peaks') appear in the $u = 1/2$ layer, but also cross vectors between each atom and the symmetry mates of the other six atoms (the 'satellite peaks'). These appear to varying extents depending on the heptagon's inclination.

(*c*) The additional overlap of peaks from the $w = 1/2$ Harker section, resulting from the near-zero *x* coordinate of the heptagon ring.

The Harker section appears uninterpretable without knowledge of the local symmetry axis position. Having this information allows suitable cross vectors to be searched for. Inspection of Fig. 6 shows that all Harker peaks are defined, whereas many of the satellite peaks do not appear on this layer, allowing identification of the former through sevenfold averaging. The strong cross peaks arising from the $w = 1/2$ Harker section are the main problem in an *a priori* identification of Harker peaks in this case. Thus, the very special organization of Ta_6Br_{14} clusters in the proteosome, coupled with the inconsistency of the sevenfold axis position, mean that the heavy-atom positions probably could not have been determined using this method.

7. Discussion

This paper shows that, given the special condition that two local axes are parallel to one another, it is possible to determine the relative positions of such axes in the unit cell from the native Patterson map, and hence from the raw intensity data. A recent survey suggests that the preferred orientation of a local symmetry axis is approximately parallel to a crystallographic cell axis (Wang & Janin, 1993). Knowledge of the orientation and position of local symmetry elements can be invaluable in the interpretation of heavy-atom-derivative data. Thus, the ability to calculate these *a priori* is a major advantage in the search for suitable derivatives. Such information is also of use in preparation for averaging techniques; the immense power of this latter method even raises the possibility of *ab initio* phase calculation given this starting point, as has been suggested recently

for maltoporin (Schirmer, Keller, Wang & Rosenbusch, 1995). The limitation to cases where the local axes are parallel to each other represents a drawback to this approach; ways of overcoming this are currently under investigation.

References

Braig, K., Otkinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L. & Sigler, P. B. (1994). *Nature (London)*, **371**, 578–586.

Eagles, P. A. M., Johnson, L. N., Joynson, M. A., McMurray, C. H. & Gutfreund, H. (1969). *J. Mol. Biol.* **45**, 533–544.

Epp, O., Steigemann, W., Formanek, H. & Huber, R. (1971). *Eur. J. Biochem.* **20**, 432–437.

Hoppe, W. (1957). *Acta Cryst.* **10**, 750–751.

Jap, B., Pühler, G., Lücke, H., Typke, D., Löwe, J., Stock, D., Huber, R. & Baumeister, W. (1993). *J. Mol. Biol.* **234**, 881–884.

Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W. & Huber, R. (1995). *Science*, **268**, 533–539.

Nar, H., Huber, R., Meining, W., Schmid, C., Weinkauf, S. & Bacher, A. (1995). *Structure*, **3**, 459–466.

Rossmann, M.G. (1972). Editor. *The Molecular-Replacement Method. A Collection of Papers on the Use of Non-Crystallographic Symmetry*. New York: Gordon and Breach.

Rossmann, M. G. & Blow, D. M. (1962). *Acta Cryst.* **15**, 24–31.

Rossmann, M. G., Blow, D. M., Harding, M. M. & Coller, E. (1964). *Acta Cryst.* **17**, 338–342.

Saibil, H. & Wood, S. (1993). *Curr. Opin. Struct. Biol.* **3**, 207–213.

Schirmer, T., Keller, T. A., Wang, Y.-F. & Rosenbusch, J. P. (1995). *Science*, **267**, 512–514.

Schmid, C., Ladenstein, R., Luecke, H., Huber, R. & Bacher, A. (1992). *J. Mol. Biol.* **226**, 1279–1281.

Schmid, C., Meining, W., Weinkauf, S., Bachmann, L., Ritz, H., Eberhart, S., Gimkel, W., Werner, T., Lahm, H.-W., Nar, H. & Bacher, A. (1993). *Chemistry and Biology of Pteridines and Folates*, edited by J. E. Ayling, pp. 157–162. New York: Plenum Press.

Spangfort, M. D., Surin, B. P., Oppentocht, J. E., Weibull, C., Carlemalm, E., Dixon, N. E. & Svensson, L. A. (1993). *FEBS Lett.* **320**, 160–164.

Steigemann, W. (1974). PhD thesis, Technische Universität München, Germany.

Stubbs, M. T., Summers, L. J., Mayr, I., Schneider, M., Huber, R., Ries, A. & Kühn, K. (1989). *Proceedings of the 1st European Workshop on Crystallography of Biological Macromolecules*, Como, Italy.

Svensson, L. A., Surin, B. P., Dixon, N. E. & Spangfort, M. D. (1994). *J. Mol. Biol.* **235**, 47–52.

Wang, X. & Janin, J. (1993). *Acta Cryst.* D**49**, 505–512.