

2. RECIPROCAL SPACE IN CRYSTAL-STRUCTURE DETERMINATION

noncrystallographic asymmetric units, the size of the excluded volume expressed in terms of the ratio $(V - UN)/V$ and the magnitude of the measurement error on the structure amplitudes. Crowther (1967, 1969) and Bricogne (1974) have investigated the dependence on the number of noncrystallographic asymmetric units and conclude that three or more copies are sufficient to ensure convergence of an iterative phase improvement procedure in the absence of errors on the structure amplitudes. As with the analogous case of isomorphous replacement in which three data sets ensure reasonable phase determination, additional copies will enhance the power of the method, although their usefulness is subject to the law of diminishing returns. Another example of this principle is the sign determination of the $h0l$ reflections of horse haemoglobin (Perutz, 1954) in which seven shrinkage stages constituted the sampling of the transform of a single copy.

Procedures for real-space averaging have been used extensively with great success. The interesting work of Wilson *et al.* (1981) is noteworthy for the continuous adjustment of molecular envelope with increased map definition. Furthermore, the analysis of complete virus structures has only been possible as a consequence of this technique (Bloomer *et al.*, 1978; Harrison *et al.*, 1978; Abad-Zapatero *et al.*, 1980; Liljas *et al.*, 1982). Although the procedure has been used primarily for phase improvement, apparently successful attempts have been made at phase extension (Nordman, 1980b; Gaykema *et al.*, 1984; Rossmann *et al.*, 1985). *Ab initio* phasing of glyceraldehyde-3-phosphate dehydrogenase (Argos *et al.*, 1975) was successfully attempted by initially filling the known envelope with uniform density to determine the phases of the innermost reflections and then gradually extending phases to 6.3 Å resolution. Johnson *et al.* (1976) used the same procedure to determine the structure of southern bean mosaic virus to 22.5 Å resolution. Particularly impressive was the work on polyoma virus (Rayment *et al.*, 1982; Rayment, 1983; Rayment *et al.*, 1983) where crude initial models led to an entirely unexpected breakdown of the Caspar & Klug (1962) concept of quasi-symmetry. *Ab initio* phasing has also been used by combining the electron-diffraction projection data of two different crystal forms of bacterial rhodopsin (Rossmann & Henderson, 1982).

2.3.8.3. Equivalence of real- and reciprocal-space molecular replacement

Let us proceed in reciprocal space doing exactly the same as is done in real-space averaging. Thus

$$\rho_{AV}(\mathbf{x}) = \frac{1}{N} \sum_{n=1}^N \rho(\mathbf{x}_n),$$

where

$$\mathbf{x}_n = [\mathbf{C}_n]\mathbf{x} + \mathbf{d}_n.$$

Therefore,

$$\rho_{AV}(\mathbf{x}) = \frac{1}{N} \sum_N \frac{1}{V} \left[\sum_{\mathbf{h}} \mathbf{F}_{\mathbf{h}} \exp(2\pi i \mathbf{h} \cdot \mathbf{x}_n) \right].$$

The next step is to perform the back-transform of the averaged electron density. Hence,

$$\mathbf{F}_{\mathbf{p}} = \int_U \rho_{AV}(\mathbf{x}) \exp(-2\pi i \mathbf{p} \cdot \mathbf{x}) \, d\mathbf{x},$$

where U is the volume within the averaged part of the cell. Hence, substituting for ρ_{AV} ,

$$\mathbf{F}_{\mathbf{p}} = \int_U \left[\frac{1}{NV} \sum_N \sum_{\mathbf{h}} \mathbf{F}_{\mathbf{h}} \exp(2\pi i \mathbf{h} \cdot \mathbf{x}_n) \right] \exp(-2\pi i \mathbf{p} \cdot \mathbf{x}) \, d\mathbf{x},$$

which is readily simplified to

$$\mathbf{F}_{\mathbf{p}} = \frac{U}{NV} \sum_{\mathbf{h}} \mathbf{F}_{\mathbf{h}} \sum_N G_{\mathbf{h}\mathbf{p}n} \exp(2\pi i \mathbf{h} \cdot \mathbf{d}_n).$$

Setting

$$\mathbf{B}_{\mathbf{h}\mathbf{p}} = \frac{U}{NV} \sum_N G_{\mathbf{h}\mathbf{p}n} \exp(2\pi i \mathbf{h} \cdot \mathbf{d}_n),$$

the molecular-replacement equations can be written as

$$\mathbf{F}_{\mathbf{p}} = \sum_{\mathbf{h}} \mathbf{B}_{\mathbf{h}\mathbf{p}} \mathbf{F}_{\mathbf{h}} \quad (2.3.8.11)$$

(Main & Rossmann, 1966), or in matrix form

$$\mathbf{F} = [\mathbf{B}]\mathbf{F},$$

which is the form of the equations used by Main (1967) and by Crowther (1967). Colman (1974) arrived at the same conclusions by an application of Shannon's sampling theorem. It should be noted that the elements of $[\mathbf{B}]$ are dependent only on knowledge of the noncrystallographic symmetry and the volume within which it is valid. Substitution of approximate phases into the right-hand side of (2.3.8.11) produces a set of calculated structure factors exactly analogous to those produced by back-transforming the averaged electron density in real space. The new phases can then be used in a renewed cycle of molecular replacement.

Computationally, it has been found more convenient and faster to work in real space. This may, however, change with the advent of vector processing in 'supercomputers'. Obtaining improved phases by substitution of current phases on the right-hand side of the molecular-replacement equations (2.3.8.1) seems less cumbersome than the repeated forward and backward Fourier transformation, intermediate sorting, and averaging required in the real-space procedure.

2.3.9. Conclusions

Complete interpretation of Patterson maps is no longer used frequently in structure analysis, although most determinations of heavy-atom positions of isomorphous pairs are based on Patterson analyses. Incorporation of the Patterson concept is crucial in many sophisticated techniques essential for the solution of complex problems, particularly in the application to biological macromolecular structures. Patterson techniques provide important physical insights in a link between real- and reciprocal-space formulation of crystal structures and diffraction data.

2.3.9.1. Update

This article was originally completed in January 1986. Since then, some advances have occurred. In particular, the use of real-space averaging between noncrystallographically related electron density within the crystallographic asymmetric unit has become an accepted way of extending phase information to higher resolution, particularly for complex structures such as viruses (Gaykema *et al.*, 1984; Rossmann *et al.*, 1985; Hogle *et al.*, 1985; Arnold *et al.*, 1987; Hosur *et al.*, 1987; Luo *et al.*, 1987; Acharya *et al.*, 1989). The power of this procedure has been examined theoretically by Arnold & Rossmann (1986).

The availability of fast computers with large random access memories and even larger disk storage also makes many of the techniques considered here commonplace and no longer subject to