which implies

$$\rho(\mathbf{x}) = \rho(\mathbf{y}_1) = \rho(\mathbf{y}_2) = \ldots. \qquad (2.3.8.6)$$

Furthermore, in the **h** cell

$$\rho(\mathbf{x}) = \frac{1}{V_\mathbf{h}} \sum_\mathbf{h} \mathbf{F}_\mathbf{h} \exp(-2\pi i \mathbf{h} \cdot \mathbf{x}), \qquad (2.3.8.7)$$

and thus, by combining with (2.3.8.5), (2.3.8.6) and (2.3.8.7),

$$\rho(\mathbf{y}_1) = \frac{1}{V_\mathbf{h}} \sum_\mathbf{h} \mathbf{F}_\mathbf{h} \exp[-2\pi i(\mathbf{h}[\boldsymbol{C}] \cdot \mathbf{y}_1 + \mathbf{h} \cdot \mathbf{d})]. \qquad (2.3.8.8)$$

Now using (2.3.8.4) and (2.3.8.8) it can be shown that

$$\mathbf{F}_\mathbf{p} = \frac{U}{V_\mathbf{h}} \sum_\mathbf{h} \mathbf{F}_\mathbf{h} \sum_{n=1}^{N} G_{\mathbf{h}\mathbf{p}_n} \exp[2\pi i(\mathbf{p} \cdot \mathbf{S}_n - \mathbf{h} \cdot \mathbf{S})], \qquad (2.3.8.9)$$

where

$$UG_{\mathbf{h}\mathbf{p}_n} = \int_U \exp[2\pi i(p[\boldsymbol{C}_n] - \mathbf{h}[\boldsymbol{C}]) \cdot \mathbf{u}] \, \mathrm{d}\mathbf{u}. \qquad (2.3.8.10)$$

S is a chosen molecular origin in the **h** crystal and $\mathbf{S}_n$ is the corresponding molecular position in the nth asymmetric unit of the **p** crystal.

2.3.8.2. *Using noncrystallographic symmetry for phase improvement*

The use of noncrystallographic symmetry for phase determination was proposed by Rossmann & Blow (1962, 1963) and subsequently explored by Crowther (1967, 1969) and Main & Rossmann (1966). These methods were developed in reciprocal space and were primarily concerned with *ab initio* phase determination. Real-space averaging of electron density between noncrystallographically related molecules was used in the structure determination of deoxyhaemoglobin (Muirhead *et al*., 1967) and of α-chymotrypsin (Matthews *et al*., 1967). The improvement derived from the averaging between the two noncrystallographic units was, however, not clear in either case. The first obviously successful application was in the structure determination of lobster glyceraldehyde-3-phosphate dehydrogenase (Buehner *et al*., 1974; Argos *et al*., 1975), where the tetrameric molecule of symmetry 222 occupied one crystallographic asymmetric unit. The improvement in the essentially SIR electron-density map was considerable and the results changed from uninterpretable to interpretable. The uniqueness and validity of the solution lay in the obvious chemical correctness of the polypeptide fold and its agreement with known amino-acid-sequence data. In contrast to the earlier reciprocal-space methods, noncrystallographic symmetry was used as a method to improve poor phases rather than to determine phases *ab initio*.

Many other applications followed rapidly, aided greatly by the versatile techniques developed by Bricogne (1976). Of particular interest is the application to the structure determination of hexokinase (Fletterick & Steitz, 1976), where the averaging occurred both between different crystal forms and within the same crystal.

The most widely used procedure for real-space averaging is the 'double sorting' technique developed by Bricogne (1976) and also by Johnson (1978). An alternative method is to maintain the complete map stored in the computer (Nordman, 1980*b*). This avoids the sorting operation, but is only possible given a very large computer or a low-resolution map containing relatively few grid points.

Bricogne's double sorting technique involves generating real-space non-integral points (D_i) which are related to integral grid points (I_i) in the cell asymmetric unit by the noncrystallographic symmetry operators. The elements of the set D_i are then brought back to their equivalent points in the cell asymmetric unit (D_i') and sorted by their proximity to two adjacent real-space sections. The set I_i', calculated on a finer grid than I_i and stored in the computer memory two sections at a time, is then used for linear interpolation to determine the density values at D_i' which are successively stored and summed in the related array I_i. A count is kept of the number of densities received at each I_i, resulting in a final averaged aggregate, when all real-space sections have been utilized. The density to be assigned outside the molecular envelope (defined with respect to the set I_i) is determined by averaging the density of all unused points in I_i. The grid interval for the set I_i' should be about one-sixth of the resolution to avoid serious errors from interpolation (Bricogne, 1976). The grid point separation in the set I_i need only be sufficient for representation of electron density, or about one-third of the resolution.

Molecular replacement in real space consists of the following steps (Table 2.3.8.1): (*a*) calculation of electron density based on a starting phase set and observed amplitudes; (*b*) averaging of this density among the noncrystallographic asymmetric units or molecular copies in several crystal forms, a process which defines a molecular envelope as the averaging is only valid within the range of the noncrystallographic symmetry; (*c*) reconstructing the unit cell based on averaged density in every noncrystallographic asymmetric unit; (*d*) calculating structure factors from the reconstructed cell; (*e*) combining the new phases with others to obtain a weighted best-phase set; and (*f*) returning to step (*a*) at the previous or an extended resolution. Decisions made in steps (*b*) and (*e*) determine the rate of convergence (see Table 2.3.8.1) to a solution (Arnold *et al*., 1987).

The power of the molecular-replacement procedure for either phase improvement or phase extension depends on the number of

Table 2.3.8.1. *Molecular replacement: phase refinement as an iterative process*

(*A*)	$\mathbf{F}_\text{obs}, \alpha_n', m_n' \rightarrow \rho_n$
(*B*)	$\rho_n \rightarrow \rho_n$(modified) (i) Use of noncrystallographic symmetry operators (ii) Definition of envelope limiting volume within which noncrystallographic symmetry is valid (iii) Adjustment of solvent density* (iv) Use of crystallographic operators to reconstruct modified density into a complete cell
(*C*)	ρ_n(modified) $\rightarrow \mathbf{F}_{\text{calc},\,n+1}$; $\alpha_{\text{calc},\,n+1}$
(*D*)	$(\mathbf{F}_{\text{calc},\,n+1}, \alpha_{\text{calc},\,n+1}) + (\mathbf{F}_\text{obs}, \alpha_0) \rightarrow \mathbf{F}_\text{obs}, \alpha_{n+1}', m_{n+1}'$ (i) Assessment of reliability of new phasing set α_{n+1} in relation to original phasing set $\alpha_0(w)$ (ii) Use of figures of merit m_0, m_{n+1} and reliability w to determine modified phasing set α_{n+1}', m_{n+1}'† (iii) Consideration of α_{n+1} and m_{n+1} where there was no prior knowledge of (*a*) $\mathbf{F}_\text{obs}$ (*e.g.* very low order reflections or uncollected data) (*b*) α_0 (*e.g.* no isomorphous information or phase extension)
(*E*)	Return to step (*A*) with α_{n+1}', m_{n+1}' and a possibly augmented set of $\mathbf{F}_\text{obs}$.

* Wang (1985); Bhat & Blow (1982); Collins (1975); Schevitz *et al*. (1981); Hoppe & Gassmann (1968).
† Rossmann & Blow (1961); Hendrickson & Lattman (1970).

references