2.3. PATTERSON AND MOLECULAR-REPLACEMENT TECHNIQUES

$$\mathbf{F}_{\mathbf{h}} = A_{\mathbf{h}} + iB_{\mathbf{h}} = F_H \cos \varphi_{\mathbf{h}} + iF_H \sin \varphi_{\mathbf{h}},$$

where $A_{\mathbf{h}}$ and $B_{\mathbf{h}}$ are the real and imaginary components of the presumed heavy-atom sites. Therefore,

$$\Delta\rho(\mathbf{x}) = \frac{1}{V} \sum \frac{m|\Delta_{\mathbf{h}}|}{|\mathbf{F}_H|} (A_{\mathbf{h}} \cos 2\pi \mathbf{h} \cdot \mathbf{x} + B_{\mathbf{h}} \sin 2\pi \mathbf{h} \cdot \mathbf{x}).$$

If this SIR difference electron-density map shows significant peaks at sites related by noncrystallographic symmetry, then those sites will be at the position of a further set of heavy atoms. Hence, a suitable criterion for finding heavy-atom sites is

$$C_{\text{SIR}} = \sum_{j=1}^{n} \Delta\rho(\mathbf{x}_j),$$

or by substitution

$$C_{\text{SIR}} = \sum_{j=1}^{n} \frac{1}{V} \sum_{\mathbf{h}} \frac{m|\Delta_{\mathbf{h}}|}{|\mathbf{F}_H|} (A_{\mathbf{h}} \cos 2\pi \mathbf{h} \cdot \mathbf{x}_j + B_{\mathbf{h}} \sin 2\pi \mathbf{h} \cdot \mathbf{x}_j).$$

But

$$a_{\mathbf{h}} = \sum_{j=1}^{n} \cos 2\pi \mathbf{h} \cdot \mathbf{x}_j \quad \text{and} \quad b_{\mathbf{h}} = \sum_{j=1}^{n} \sin 2\pi \mathbf{h} \cdot \mathbf{x}_j.$$

Therefore,

$$C_{\text{SIR}} = \frac{1}{V} \sum_{\mathbf{h}} \frac{m|\Delta_{\mathbf{h}}|}{|\mathbf{F}_H|} (A_{\mathbf{h}} a_{\mathbf{h}} + B_{\mathbf{h}} b_{\mathbf{h}}). \qquad (2.3.3.6)$$

This expression is similar to (2.3.3.5) derived by consideration of a Patterson search. It differs from (2.3.3.5) in two respects: the Fourier coefficients are different and expression (2.3.3.6) is lacking a second term. Now the figure of merit m will be small whenever $|\mathbf{F}_H|$ is small as the SIR phase cannot be determined well under those conditions. Hence, effectively, the coefficients are a function of $|\Delta_{\mathbf{h}}|$, and the coefficients of the functions (2.3.3.5) and (2.3.3.6) are indeed rather similar. The second term in (2.3.3.5) relates to the use of the search atoms in phasing and could be included in (2.3.3.6), provided the actual feedback sites in each of the n electron-density functions tested by C_{SIR} are omitted in turn. Thus, a systematic Patterson search and an SIR difference Fourier search are very similar in character and power.

2.3.3.6. *Direct structure determination from difference Pattersons*

The difference Patterson computed with coefficients $F_{HN}^2 - F_N^2$ contains information on the heavy atoms (*HH* vectors) and the macromolecular structure (*HL* vectors) (Section 2.3.3.3). If the scaling between the $|\mathbf{F}_{HN}|$ and $|\mathbf{F}_N|$ data sets is not perfect there will also be noise. Rossmann (1961*b*) was partially successful in determining the low-resolution horse haemoglobin structure by using a series of superpositions based on the known heavy-atom sites. Nevertheless, Patterson superposition methods have not been used for the structure determination of proteins owing to the successful error treatment of the isomorphous replacement method in reciprocal space. However, it is of some interest here for it gives an alternative insight into SIR phasing.

The deconvolution of an arbitrary molecule, represented as '?', from an $(F_{HN}^2 - F_N^2)$ Patterson, is demonstrated in Fig. 2.3.3.5. The original structure is shown in Fig. 2.3.3.5(*a*) and the corresponding Patterson in Fig. 2.3.3.5(*b*). Superposition with respect to one of the heavy-atom sites is shown in Fig. 2.3.3.5(*c*) and the other in Fig. 2.3.3.5(*d*). Both Figs. 2.3.3.5(*c*) and (*d*) contain a centre of symmetry because the use of only a single *HH* vector implies a centre of symmetry half way between the two sites. The centre is

broken on combining information from all three sites (which together lack a centre of symmetry) by superimposing Figs. 2.3.3.5(*c*) and (*d*) to obtain either the original structure (Fig. 2.3.3.5*a*) or its enantiomorph. Thus it is clear, in principle, that there is sufficient information in a single isomorphous derivative data set, when used in conjunction with a native data set, to solve a structure completely. However, the procedure shown in Fig. 2.3.3.5 does not consider the accumulation of error in the selection of individual images when these intersect with another image. In this sense the reciprocal-space isomorphous replacement technique has greater elegance and provides more insight, whereas the alternative view given by the Patterson method was the original stimulus for the discovery of the SIR phasing technique (Blow & Rossmann, 1961).

Other Patterson functions for the deconvolution of SIR data have been proposed by Ramachandran & Raman (1959), as well as others. The principles are similar but the coefficients of the functions are optimized to emphasize various aspects of the signal representing the molecular structure.

2.3.3.7. *Isomorphism and size of the heavy-atom substitution*

It is insufficient to discuss Patterson techniques for locating heavy-atom substitutions without also considering errors of all kinds. First, it must be recognized that most heavy-atom labels are not a single atom but a small compound containing one or more heavy atoms. The compound itself will displace water or ions and locally alter the conformation of the protein or nucleic acid. Hence, a simple Gaussian approximation will suffice to represent individual heavy-atom scatterers responsible for the difference between native and heavy-atom derivatives. Furthermore, the heavy-atom compound often introduces small global structural changes which can be detected only at higher resolution. These problems were considered with some rigour by Crick & Magdoff (1956). In general, lack of isomorphism is exhibited by an increase in the size of the isomorphous differences with increasing resolution (Fig. 2.3.3.6).
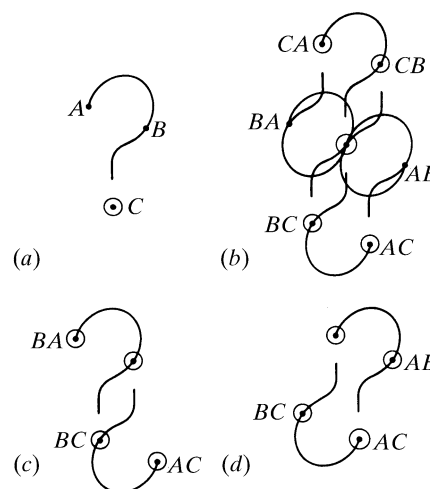

Fig. 2.3.3.5. Let (*a*) be the original structure which contains three heavy atoms *ABC* in a noncentrosymmetric configuration. Then a Fourier summation, with $(F_{NH}^2 - F_N^2)$ coefficients, will give the Patterson shown in (*b*). Displacement of the Patterson by the vector *BC* and selecting the common patterns yields (*c*). Similarly, displacement by *AC* gives (*d*). Finally, superposition of (*c*) on (*d*) gives the original figure or its enantiomorph. This series of steps demonstrates that, in principle, complete structural information is contained in an SIR derivative.

245