

## 2. RECIPROCAL SPACE IN CRYSTAL-STRUCTURE DETERMINATION

from the known portion. In such single-site search procedures, single atoms are placed at all possible positions in a crystal, using a search grid of the same fineness as for the stored Patterson function, preferably about one-third of the resolution of the Patterson map. Solutions are gauged to be acceptable if all expected vectors due to symmetry-related atoms are observed above a specified threshold in the Patterson.

Systematic computerized Patterson search procedures for orienting and positioning known molecular fragments were also developed in the early 1960s. These hierarchical procedures rely on first using the 'self'-vectors which depend only on the orientation of a molecular fragment. A search for the position of the fragment relative to the crystal symmetry elements then uses the cross-vectors between molecules (see Sections 2.3.6 and 2.3.7). Nordman constructed a weighted point representation of the predicted vector set for a fragment (Nordman & Nakatsu, 1963; Nordman, 1966) and successfully solved the structure of a number of complex alkaloids. Huber (1965) used the convolution molecule method of Hoppe (1957a) in three dimensions to solve a number of natural-product structures, including steroids. Various program systems have used Patterson search methods operating in real space to solve complex structures (Braun *et al.*, 1969; Egert, 1983).

Others have used reciprocal-space procedures for locating known fragments. Tollin & Cochran (1964) developed a procedure for determining the orientation of planar groups by searching for origin-containing planes of high density in the Patterson. General procedures using reciprocal-space representations for determining rotation and translation parameters have been developed and will be described in Sections 2.3.6 and 2.3.7, respectively.

Although as many functions have been used to detect solutions in these Patterson search procedures as there are programs, most rely on some variation of the sum, product and minimum functions (Section 2.3.2.4). The quality of the stored Patterson density representation also varies widely, but it is now common to use 16 or more bits for single density values. Treatment of vector overlap is handled differently by different investigators and the choice will depend on the degree of overlapping (Nordman & Schilling, 1970; Nordman, 1972). General Gaussian multiplicity corrections can be employed to treat coincidental overlap of independent vectors in general positions and overlap which occurs for symmetric peaks in the vicinity of a special position or mirror plane in the Patterson (G. Kamer, S. Ramakumar & P. Argos, unpublished results; Rossmann *et al.*, 1972).

## 2.3.3. Isomorphous replacement difference Pattersons

## 2.3.3.1. Introduction

One of the initial stages in the application of the isomorphous replacement method is the determination of heavy-atom positions. Indeed, this step of a structure determination can often be the most challenging. Not only may the number of heavy-atom sites be unknown, and have incomplete substitution, but the various isomorphous compounds may also lack isomorphism. To compound these problems, the error in the measurement of the isomorphous difference in structure amplitudes is often comparable to the differences themselves. Clearly, therefore, the ease with which a particular problem can be solved is closely correlated with the quality of the data-measuring procedure.

The isomorphous replacement method was used incidentally by Bragg in the solution of NaCl and KCl. It was later formalized by J. M. Robertson in the analysis of phthalocyanine where the coordination centre could be Pt, Ni and other metals (Robertson, 1935, 1936; Robertson & Woodward, 1937). In this and similar cases, there was no difficulty in finding the heavy-atom positions. Not only were the heavy atoms frequently in special positions, but

they dominated the total scattering effect. It was not until Perutz and his colleagues (Green *et al.*, 1954; Bragg & Perutz, 1954) applied the technique to the solution of haemoglobin, a protein of 68 000 Da, that it was necessary to consider methods for detecting heavy atoms. The effect of a single heavy atom, even uranium, can only have a very marginal effect on the structure amplitudes of a crystalline macromolecule. Hence, techniques had to be developed which were dependent on the difference of the isomorphous structure amplitudes rather than on the solution of the Patterson of the heavy-atom-derivative compound on its own.

## 2.3.3.2. Finding heavy atoms with centrosymmetric projections

Phases in a centrosymmetric projection will be 0 or  $\pi$  if the origin is chosen at the centre of symmetry. Hence, the native structure factor,  $\mathbf{F}_N$ , and the heavy-atom-derivative structure factor,  $\mathbf{F}_{NH}$ , will be collinear. It follows that the structure amplitude,  $|\mathbf{F}_H|$ , of the heavy atoms alone in the cell will be given by

$$|\mathbf{F}_H| = |(|\mathbf{F}_{NH}| \pm |\mathbf{F}_N|)| + \varepsilon,$$

where  $\varepsilon$  is the error on the parenthetic sum or difference. Three different cases may arise (Fig. 2.3.3.1). Since the situation shown in Fig. 2.3.3.1(c) is rare, in general

$$|\mathbf{F}_H|^2 \simeq (|\mathbf{F}_{NH}| - |\mathbf{F}_N|)^2. \quad (2.3.3.1)$$

Thus, a Patterson computed with the square of the differences between the native and derivative structure amplitudes of a centrosymmetric projection will approximate to a Patterson of the heavy atoms alone.

The approximation (2.3.3.1) is valid if the heavy-atom substitution is small enough to make  $|\mathbf{F}_H| \ll |\mathbf{F}_{NH}|$  for most reflections, but sufficiently large to make  $\varepsilon \ll (|\mathbf{F}_{NH}| - |\mathbf{F}_N|)^2$ . It is also assumed that the native and heavy-atom-derivative data have been placed on the same relative scale. Hence, the relation (2.3.3.1) should be re-written as

$$|\mathbf{F}_H|^2 \simeq (|\mathbf{F}_{NH}| - k|\mathbf{F}_N|)^2,$$

where  $k$  is an experimentally determined scale factor (see Section 2.3.3.7). Uncertainty in the determination of  $k$  will contribute further to  $\varepsilon$ , albeit in a systematic manner.

Centrosymmetric projections were used extensively for the determination of heavy-atom sites in early work on proteins such as haemoglobin (Green *et al.*, 1954), myoglobin (Bluhm *et al.*,

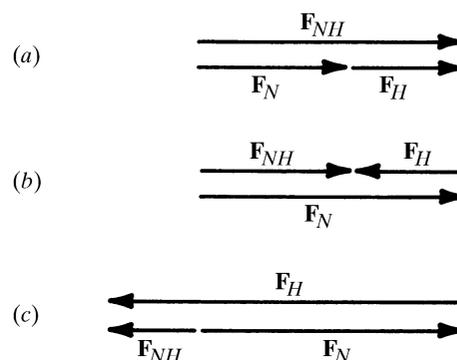


Fig. 2.3.3.1. Three different cases which can occur in the relation of the native,  $\mathbf{F}_N$ , and heavy-atom derivative,  $\mathbf{F}_{NH}$ , structure factors for centrosymmetric reflections.  $\mathbf{F}_N$  is assumed to have a phase of 0, although analogous diagrams could be drawn when  $\mathbf{F}_N$  has a phase of  $\pi$ . The crossover situation in (c) is clearly rare if the heavy-atom substitution is small compared to the native molecule, and can in general be neglected.