2.3. PATTERSON AND MOLECULAR-REPLACEMENT TECHNIQUES

1958) and lysozyme (Poljak, 1963). However, with the advent of faster data-collecting techniques, low-resolution (e.g. a 5 Å limit) three-dimensional data are to be preferred for calculating difference Pattersons. For noncentrosymmetric reflections, the approximation (2.3.3.1) is still valid but less exact (Section 2.3.3.3). However, the larger number of three-dimensional differences compared to projection differences will enhance the signal of the real Patterson peaks relative to the noise. If there are $N$ terms in the Patterson synthesis, then the peak-to-noise ratio will be proportionally $\sqrt{N}$ and $1/\varepsilon_3$. With the subscripts 2 and 3 representing two- and three-dimensional syntheses, respectively, the latter will be more powerful than the former whenever

$$\frac{\sqrt{N_2}}{\varepsilon_3} > \frac{\sqrt{N_3}}{\varepsilon_2}.$$ 

Now, as $\varepsilon_3 \approx \sqrt{2}\varepsilon_2$, it follows that $N_2$ must be greater than $2N_3$ if the three-dimensional noncentrosymmetric computation is to be more powerful. This condition must almost invariably be true.

2.3.3.3. Finding heavy atoms with three-dimensional methods

A Patterson of a native bio-macromolecular structure (coefficients $F^2$) can be considered as being, at least approximately, a vector map of all the light atoms (carbons, nitrogens, oxygens, some sulfurs, and also phosphorus for nucleic acids) other than hydrogen atoms. These interactions will be designated as $LL$. Similarly, a Patterson of the heavy-atom derivative will contain $HH + HL + LL$ interactions, where $H$ represents the heavy atoms. Thus, a true difference Patterson, with coefficients $F^2_{HH} - F^2_{HL}$, will contain only the interactions $HH + HL$. In general, the carpet of $HL$ vectors completely dominates the $HH$ vectors except for very small proteins such as insulin (Adams et al., 1969). Therefore, it would be preferable to compute a Patterson containing only $HH$ interactions in order to interpret the map in terms of specific heavy-atom sites.

Blow (1958) and Rossmann (1960) showed that a Patterson with $(|F_{HH} - |F_N|)^2$ coefficients approximated to a Patterson containing only $HH$ vectors. If the phase angle between $F_N$ and $F_{HH}$ is $\varphi$ (Fig. 2.3.3.2), then

$$|F_H|^2 = |F_N|^2 + |F_{HH}|^2 - 2|F_N||F_{HH}|\cos\varphi.$$ 

In general, however, $|F_H| \ll |F_N|$. Hence, $\varphi$ is small and

$$|F_H|^2 \approx (|F_{HH} - |F_N|)^2,$$

which is the same relation as (2.3.3.1) for centrosymmetric approximations. Since the direction of $F^2_H$ is random compared to $F^2_N$, the root-mean-square projected length of $F^2_H$ onto $F^2_N$ will be $F^2_H / \sqrt{2}$. Thus it follows that a better approximation is $|F_{HH}|^2 \approx \sqrt{2}(|F_{HH} - |F_N|)^2$.

$$|F_{HH}|^2 \approx \sqrt{2}(|F_{HH} - |F_N|)^2,$$ (2.3.3.2)

which accounts for the assumption (Section 2.3.3.2) that $\varepsilon_3 = \sqrt{2}\varepsilon_2$. The almost universal method for the initial determination of major heavy-atom sites in an isomorphous derivative utilizes a Patterson with $(|F_{HH} - |F_N|)^2$ coefficients. Approximation (2.3.3.2) is also the basis for the refinement of heavy-atom parameters in a single isomorphous replacement pair (Rossmann, 1960; Cullis et al., 1962; Terwilliger & Eisenberg, 1983).

2.3.3.4. Correlation functions

In the most general case of a triclinic space group, it will be necessary to select an origin arbitrarily, usually coincident with a heavy atom. All other heavy atoms (and subsequently also the macromolecular atoms) will be referred to this reference atom. However, the choice of an origin will be independent in the interpretation of each derivative’s difference Patterson. It will then be necessary to correlate the various, arbitrarily chosen, origins. The same problem occurs in space groups lacking symmetry axes perpendicular to the primary rotation axis (e.g. $P2_1$, $P6$ etc.), although only one coordinate, namely parallel to the unique rotation axis, will require correlation. This problem gave rise to some concern in the 1950s. Bragg (1958), Blow (1958), Perutz (1956), Hoppe (1959) and Bodo et al. (1959) developed a variety of techniques, none of which were entirely satisfactory. Rossmann (1960) proposed the $(F_{HH} - F_{HL})^2$ synthesis and applied it successfully to the heavy-atom determination of horse haemoglobin. This function gives positive peaks ($H1 \cdot H1$) at the end of vectors between the heavy-atom sites in the first compound, positive peaks ($H2 \cdot H2$) between the sites in the second compound, and negative peaks between sites in the first and second compound (Fig. 2.3.3.3). It is thus the negative peaks which provide the necessary correlation. The function is unique in that it is a Patterson containing significant information in both positive and negative peaks. Steinrauf (1963) suggested using the coefficients $(|F_{HH1} - |F_N|) \cdot (|F_{HH2} - |F_N|)$ in order to eliminate the positive $H1 \cdot H1$ and $H2 \cdot H2$ vectors. Although the problem of correlation was a serious concern in the early structural determination of proteins during the late 1950s and early 1960s, the problem has now been by-passed. Blow &

![Fig. 2.3.3.3. A Patterson with coefficients $(F_{HH1} - F_{HH2})^2$ will be equivalent to a Patterson whose coefficients are $(AB)^2$. However, $AB = -F_{HH1} + F_{HH2}$. Thus, a Patterson with $(AB)^2$ coefficients is equivalent to having negative atomic substitutions in compound 1 and positive substitutions in compound 2, or vice versa. Therefore, the Patterson will contain positive peaks for vectors of the type $H1 \cdot H1$ and $H2 \cdot H2$, but negative vector peaks for vectors of type $H1 \cdot H2$.](image-url)