PART 14. ANOMALOUS DISPERSION

Chapter 14.1. Heavy-atom location and phase determination with single-wavelength diffraction data

B. W. MATTHEWS

14.1.1. Introduction

As is well known, the successful introduction of the method of isomorphous replacement by Green *et al.* (1954) was the turning point in the subsequent development of protein crystallography as we now know it.

The idea that the phases of X-ray reflections from a protein crystal could be obtained by the introduction of heavy atoms into the crystal was not new, having been suggested by J. D. Bernal in 1939 (Bernal, 1939). The isomorphous-replacement method was used as early as 1927 by Cork (1927) in studying the alums. Bokhoven *et al.* (1951) subsequently extended the method to the study of a noncentrosymmetric projection of strychnine sulfate, using what would now be termed the method of single isomorphous replacement. They also suggested that by using a double isomorphous replacement, a unique phase determination could be obtained, even for noncentrosymmetric reflections. The details of the double (or multiple) isomorphous-replacement method were worked out by Harker (1956), who introduced the very useful concept of phase circles. Another contribution which was of great practical value, and which will provide the basis for much of the subsequent discussion, is the method introduced by Blow & Crick (1959) for the treatment of errors in the isomorphous-replacement method. In addition to the determination of protein phases by the method of substitution with heavy atoms, it is now routine to supplement this information by utilizing the anomalous scattering of the substituted atoms. The underlying principles trace back to articles by Bijvoet (1954), Ramachandran & Raman (1956), and Okaya & Pepinsky (1960). The first application of the anomalous-scattering method to protein crystallography was by Blow (1958), who used the anomalous scattering of the iron atoms to determine phase information for a noncentrosymmetric projection of horse oxyhaemoglobin.

In the following discussion, we first review the classical method of phase determination by isomorphous replacement, then discuss the inclusion of single-wavelength anomalous-scattering data, and conclude by discussing the use of such data for heavy-atom location. Part of the review is based on Matthews (1970).

14.1.2. The isomorphous-replacement method

Consider a protein crystal with an isomorphous heavy-atom derivative, *i.e.* a modified crystal in which heavy atoms occupy specific sites throughout the crystal, but which is in all other respects identical to the unsubstituted 'parent' crystal. Let the structure factors of the protein crystal be $\mathbf{F}_P(\mathbf{h})$, of the isomorph be $\mathbf{F}_{PH}(\mathbf{h})$, and of the heavy atoms $\mathbf{F}_H(\mathbf{h})$. (Note: Structure *amplitudes* are indicated by italic type, *e.g.* F_P, and *vectors* by bold-face type, *e.g.* $\mathbf{F}_P$.) In practice, one can measure the structure amplitudes F_P and F_{PH}, and it is desired to obtain from these observable quantities the value of the phase angle of $\mathbf{F}_P(\mathbf{h})$ so that

a Fourier synthesis showing the electron density of the protein structure may be calculated. It will be assumed, for the moment, that the positions and occupancy of the sites of heavy-atom binding have been determined as accurately as possible.

(a)

(b)

Figure 14.1.2.1
(a) Harker construction for a single isomorphous replacement. φ_1 and φ_2 are the 'most probable' phases for $\mathbf{F}_P$. (b) Phase probability distribution for a single isomorphous replacement. This and subsequent probabilities are unnormalized. [All figures in this chapter are reproduced with permission from Matthews (1970).]

(a)

(b)

Figure 14.1.3.1
(*a*) Harker construction for a double isomorphous replacement. φ_M is the 'most probable' phase for $\mathbf{F}_P$. (*b*) Phase probability distribution corresponding to the double isomorphous replacement shown in part (*a*). The curve for derivative 1 is solid, that for derivative 2 is dashed, and that for the combined distribution is drawn as a dotted-and-dashed line.

From the heavy-atom parameters, the corresponding structure factor $\mathbf{F}_H(\mathbf{h})$ is calculated. To determine φ, the phase of $\mathbf{F}_P(\mathbf{h})$, we construct a set of phase circles, as proposed by Harker (1956). From a chosen origin O (Fig. 14.1.2.1*a*), the vector OA is drawn equal to $-\mathbf{F}_H$. Circles of radius F_P and F_{PH} are then drawn about O and A, respectively. The intersections of the phase circles at B and B' define two possible phase angles for F_P. Note that the angles are symmetrical about $\mathbf{F}_H$. This ambiguity may in principle be resolved in two ways: (*a*) by using a second heavy-atom isomorphous derivative or (*b*) by utilizing the anomalous-scattering effects for the first isomorph.

14.1.3. The method of multiple isomorphous replacement

The phase information provided by a second isomorph is illustrated in Fig. 14.1.3.1(*a*). In theory, the three phase circles will

intersect at a point and the phase ambiguity will be resolved. In practice, there will be errors in the observed amplitudes F_P and F_{PH} and in the heavy-atom parameters (and thus in $\mathbf{F}_H$). Also, the isomorphism may be imperfect. As a result, the intersections of the three phase circles may not coincide. Another complication arises from the fact that for reflections where $\mathbf{F}_H$ is small, the circles will be essentially concentric and will not have well defined points of intersection. In other words, the phase determination will become indeterminate. The method of Blow & Crick (1959) was introduced as a way to take all these factors into account. It has had an extraordinary impact, not only as a practical method for protein phase determination, but also in influencing all subsequent thinking in this area.

14.1.4. The method of Blow & Crick

Blow & Crick pointed out that in practice the phase angle φ can never be determined with complete certainty. Rather, there is a finite probability that any arbitrary phase angle may be the correct one. Consider the vector diagram shown in Fig. 14.1.4.1, in which $\mathbf{F}_H$ is known and we wish to determine the probability $P(\varphi)$ that the arbitrary phase angle φ is the correct phase of $\mathbf{F}_P$. Strictly, one should allow for the possibility of errors in $\mathbf{F}_H$, F_P and F_{PH}, and should consider the probability that the vector $\mathbf{F}_P$ occupies all possible positions in the Argand diagram. However, Blow & Crick suggested that the analysis might be considerably simplified by assuming that F_P and $\mathbf{F}_H$ are known accurately and that all the error lies in the observation of F_{PH}. In other words, it was assumed that the vector $\mathbf{F}_P$ must lie on the circle of radius F_P, and the probability distribution of F_P could be evaluated as a function of φ only.

For an arbitrary phase angle φ, the phase triangle (Fig. 14.1.4.1) will not close exactly. If we define F_C to be the vector sum of $\mathbf{F}_H$ and $F_P \exp(i\varphi)$, then the lack of closure of the phase triangle is given by

$$\varepsilon = F_C - F_{PH}. \qquad (14.1.4.1)$$

Following Blow & Crick, if E is the r.m.s. error associated with the measurements, and the distribution of error is assumed to be Gaussian, then the probability $P(\varphi)$ of the phase φ being the true phase is

$$P(\varphi) = N \exp(-\varepsilon^2/2E^2), \qquad (14.1.4.2)$$
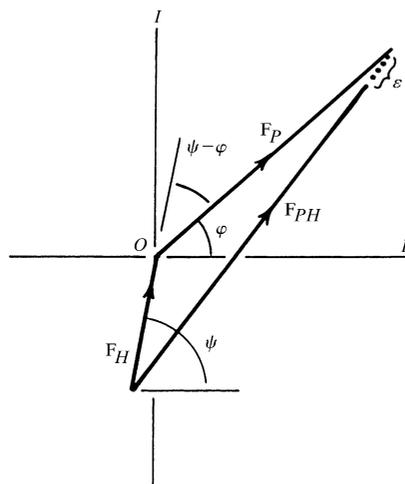
Figure 14.1.4.1
Vector diagram illustrating the lack of closure, ε, of an isomorphous-replacement phase triangle.

references