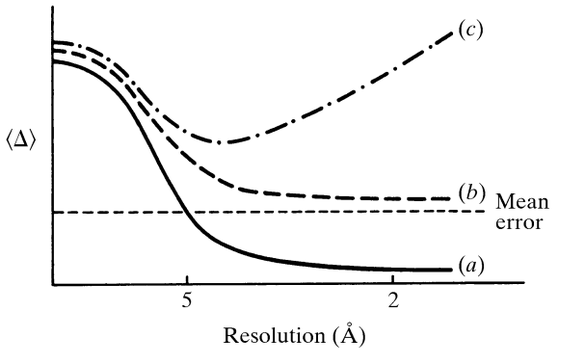

Fig. 2.3.3.6. A plot of mean isomorphous differences as a function of resolution. (*a*) The theoretical size of mean differences following roughly a Gaussian distribution. (*b*) The observed size of differences for a good isomorphous derivative where the smaller higher-order differences have been largely masked by the error of measurement. (*c*) Observed differences where 'lack of isomorphism' dominates beyond approximately 5 Å resolution.

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).

Crick & Magdoff (1956) also derived the approximate expression

$$\sqrt{\frac{2N_{\rm H}}{N_{\rm P}}}\cdot\frac{f_{\rm H}}{f_{\rm P}}$$

to estimate the r.m.s. fractional change in intensity as a function of heavy-atom substitution. Here, $N_{\rm H}$ represents the number of heavy atoms attached to a protein (or other large molecule) which contains $N_{\rm P}$ light atoms. $f_{\rm H}$ and $f_{\rm P}$ are the scattering powers of the average heavy and protein atom, respectively. This function was tabulated by Eisenberg (1970) as a function of molecular weight (proportional to $N_{\rm P}$). For instance, for a single, fully substituted, Hg atom the formula predicts an r.m.s. intensity change of around 25% in a molecule of 100 000 Da. However, the error of measurement of a reflection intensity is likely to be arround 10% of I, implying perhaps an error of around 14% of I on a difference measurement. Thus, the isomorphous replacement difference measurement for almost half the reflections will be buried in error for this case.

Scaling of the different heavy-atom-derivative data sets onto a common relative scale is clearly important if error is to be reduced. Blundell & Johnson (1976, pp. 333–336) give a careful discussion of this subject. Suffice it to say here only that a linear scale factor is seldom acceptable as the heavy-atom-derivative crystals frequently suffer from greater disorder than the native crystals. The heavy-atom derivative should, in general, have a slightly larger mean value for the structure factors on account of the additional heavy atoms (Green *et al.*, 1954). The usual effect is to make $\sum|\mathbf{F}_{\rm NH}|^2/\sum|\mathbf{F}_{\rm N}|^2 \simeq 1.05$ (Phillips, 1966).

As the amount of heavy atom is usually unknown in a yet unsolved heavy-atom derivative, it is usual practice either to apply a scale factor of the form $k\exp[-B(\sin\theta/\lambda)^2]$ or, more generally, to use local scaling (Matthews & Czerwinski, 1975). The latter has the advantage of not making any assumption about the physical nature of the relative intensity decay with resolution.

2.3.4. Anomalous dispersion

2.3.4.1. Introduction

The physical basis for anomalous dispersion has been well reviewed by Ramaseshan & Abrahams (1975), James (1965), Cromer (1974) and Bijvoet (1954). As the wavelength of radiation approaches the absorption edge of a particular element, then an atom will disperse X-rays in a manner that can be defined by the complex scattering factor

$$f_0 + \Delta f' + i\Delta f'',$$

where f_0 is the scattering factor of the atom without the anomalous absorption and rescattering effect, $\Delta f'$ is the real correction term (usually negative), and $\Delta f''$ is the imaginary component. The real term $f_0 + \Delta f'$ is often written as f', so that the total scattering factor will be $f' + i\Delta f''$. Values of $\Delta f'$ and $\Delta f''$ are tabulated in *IT* IV (Cromer, 1974), although their precise values are dependent on the environment of the anomalous scatterer. Unlike f_0, $\Delta f'$ and $\Delta f''$ are almost independent of scattering angle as they are caused by absorption of energy in the innermost electron shells. Thus, the anomalous effect resembles scattering from a point atom.

The structure factor of index $\mathbf{h}$ can now be written as

$$\mathbf{F}_{\mathbf{h}} = \sum_{j=1}^{N} f_j' \exp(2\pi i\mathbf{h}\cdot\mathbf{x}_j) + i\sum_{j=1}^{N}\Delta f_j'' \exp(2\pi i\mathbf{h}\cdot\mathbf{x}_j). \qquad (2.3.4.1)$$

(Note that the only significant contributions to the second term are from those atoms that have a measurable anomalous effect at the chosen wavelength.)

Let us now write the first term as $A + iB$ and the second as $a + ib$. Then, from (2.3.4.1),

$$\mathbf{F} = (A + iB) + i(a + ib) = (A - b) + i(B + a). \qquad (2.3.4.2)$$

Therefore,

$$|\mathbf{F}_{\mathbf{h}}|^2 = (A - b)^2 + (B + a)^2$$

and similarly

$$|\mathbf{F}_{\bar{\mathbf{h}}}|^2 = (A + b)^2 + (-B + a)^2,$$

demonstrating that Friedel's law breaks down in the presence of anomalous dispersion. However, it is only for noncentrosymmetric reflections that $|\mathbf{F}_{\mathbf{h}}| \neq |\mathbf{F}_{\bar{\mathbf{h}}}|$.

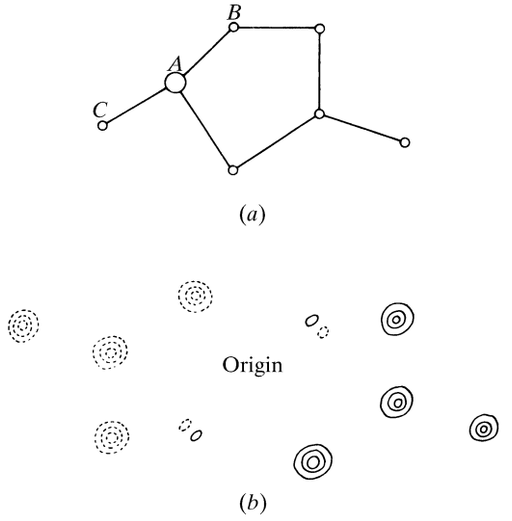

Fig. 2.3.4.1. (*a*) A model structure with an anomalous scatterer at *A*. (*b*) The corresponding $P_s(\mathbf{u})$ function showing positive peaks (full lines) and negative peaks (dashed lines). [Reprinted with permission from Woolfson (1970, p. 293).]

Now,

$$\rho(\mathbf{x}) = \frac{1}{V}\sum_{\mathbf{h}}^{\text{sphere}} \mathbf{F}_{\mathbf{h}} \exp(2\pi i \mathbf{h}\cdot\mathbf{x}).$$

Hence, by using (2.3.4.2) and simplifying,

$$\rho(\mathbf{x}) = \frac{2}{V}\sum_{\mathbf{h}}^{\text{hemisphere}} [(A\cos 2\pi\mathbf{h}\cdot\mathbf{x} - B\sin 2\pi\mathbf{h}\cdot\mathbf{x}) + i(a\cos 2\pi\mathbf{h}\cdot\mathbf{x} - b\sin 2\pi\mathbf{h}\cdot\mathbf{x})]. \quad (2.3.4.3)$$

The first term in (2.3.4.3) is the usual real Fourier expression for electron density, while the second term is an imaginary component due to the anomalous scattering of a few atoms in the cell.

2.3.4.2. The $P_s(\mathbf{u})$ function

Expression (2.3.4.3) gives the complex electron density expression in the presence of anomalous scatterers. A variety of Patterson-type functions can be derived from (2.3.4.3) for the determination of a structure. For instance, the $P_s(\mathbf{u})$ function gives vectors between the anomalous atoms and the 'normal' atoms.

From (2.3.4.1) it is easy to show that

$$\begin{aligned}\mathbf{F}_{\mathbf{h}}\mathbf{F}_{\mathbf{h}}^* &= |\mathbf{F}_{\mathbf{h}}|^2 \\ &= \sum_{i,j}(f_i' f_j' + f_i'' f_j'')\cos 2\pi\mathbf{h}\cdot(\mathbf{x}_i - \mathbf{x}_j) \\ &\quad + \sum_{i,j}(f_i' f_j'' - f_i'' f_j')\sin 2\pi\mathbf{h}\cdot(\mathbf{x}_i - \mathbf{x}_j).\end{aligned}$$

Therefore,

$$|\mathbf{F}_{\mathbf{h}}|^2 + |\mathbf{F}_{\bar{\mathbf{h}}}|^2 = 2\sum_{i,j}(f_i' f_j' + f_i'' f_j'')\cos 2\pi\mathbf{h}\cdot(\mathbf{x}_i - \mathbf{x}_j)$$

and

$$|\mathbf{F}_{\mathbf{h}}|^2 - |\mathbf{F}_{\bar{\mathbf{h}}}|^2 = 2\sum_{i,j}(f_i' f_j'' - f_i'' f_j')\sin 2\pi\mathbf{h}\cdot(\mathbf{x}_i - \mathbf{x}_j).$$

Let us now consider the significance of a Patterson in the presence of anomalous dispersion. The normal Patterson definition is given by

$$\begin{aligned}P(\mathbf{u}) &= \int_V \rho^*(\mathbf{x})\rho(\mathbf{x}+\mathbf{u})\,\mathrm{d}\mathbf{x} \\ &= \frac{1}{V^2}\sum_{\mathbf{h}}^{\text{sphere}} |\mathbf{F}_{\mathbf{h}}|^2 \exp(-2\pi i\mathbf{h}\cdot\mathbf{u}) \\ &\equiv P_c(\mathbf{u}) - iP_s(\mathbf{u}),\end{aligned}$$

where

$$P_c(\mathbf{u}) = \frac{2}{V}\sum_{\mathbf{h}}^{\text{hemisphere}} (|\mathbf{F}_{\mathbf{h}}|^2 + |\mathbf{F}_{\bar{\mathbf{h}}}|^2)\cos 2\pi\mathbf{h}\cdot\mathbf{u}$$

and

$$P_s(\mathbf{u}) = \frac{2}{V}\sum_{\mathbf{h}}^{\text{hemisphere}} (|\mathbf{F}_{\mathbf{h}}|^2 - |\mathbf{F}_{\bar{\mathbf{h}}}|^2)\sin 2\pi\mathbf{h}\cdot\mathbf{u}.$$

The $P_c(\mathbf{u})$ component is essentially the normal Patterson, in which the peak heights have been very slightly modified by the anomalous-scattering effect. That is, the peaks of $P_c(\mathbf{u})$ are proportional in height to $(f_i' f_j' + f_i'' f_j'')$.

The $P_s(\mathbf{u})$ component is more interesting. It represents vectors between the normal atoms in the unit cell and the anomalous scatterers, proportional in height to $(f_i' f_j'' - f_i'' f_j')$ (Okaya *et al.*, 1955). This function is antisymmetric with respect to the change of the direction of the diffraction vector. An illustration of the function is given in Fig. 2.3.4.1. In a unit cell containing N atoms, n of which are anomalous scatterers, the $P_s(\mathbf{u})$ function contains only $n(N-n)$ positive peaks and an equal number of negative peaks related to the former by anticentrosymmetry. The analysis of a $P_s(\mathbf{u})$ synthesis presents problems somewhat similar to those posed by a normal Patterson. The procedure has been used only rarely [*cf.* Moncrief & Lipscomb (1966) and Pepinsky *et al.* (1957)], probably because alternative procedures are available for small compounds and the solution of $P_s(\mathbf{u})$ is too complex for large biological molecules.

2.3.4.3. The position of anomalous scatterers

Anomalous scatterers can be used as an aid to phasing, when their positions are known. But the anomalous-dispersion differences (Bijvoet differences) can also be used to determine or confirm the heavy atoms which scatter anomalously (Rossmann, 1961*a*). Furthermore, the use of anomalous-dispersion information obviates the lack of isomorphism but, on the other hand, the differences are normally far smaller than those produced by a heavy-atom isomorphous replacement.

Consider a structure of many light atoms giving rise to the structure factor $\mathbf{F}_{\mathbf{h}}(N)$. In addition, it contains a few heavy atoms which have a significant anomalous-scattering effect. The non-anomalous component will be $\mathbf{F}_{\mathbf{h}}(H)$ and the anomalous component is $\mathbf{F}_{\mathbf{h}}''(H) = i(\Delta f''/f')\mathbf{F}_{\mathbf{h}}(H)$ (Fig. 2.3.4.2*a*). The total structure factor will be $\mathbf{F}_{\mathbf{h}}$. The Friedel opposite is constructed appropriately (Fig. 2.3.4.2*a*). Now reflect the Friedel opposite construction across the real axis of the Argand diagram (Fig.

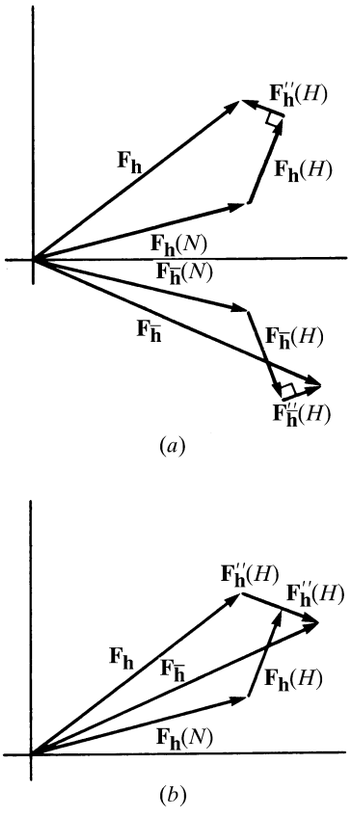

Fig. 2.3.4.2. Anomalous-dispersion effect for a molecule whose light atoms contribute $\mathbf{F}_{\mathbf{h}}(N)$ and heavy atom $\mathbf{F}_{\mathbf{h}}(H)$ with a small anomalous component of $\mathbf{F}_{\mathbf{h}}''(H)$, 90° ahead of the non-anomalous $\mathbf{F}_{\mathbf{h}}(H)$ component. In (*a*) is seen the construction for $\mathbf{F}_{\mathbf{h}}$ and $\mathbf{F}_{\bar{\mathbf{h}}}$. In (*b*) $\mathbf{F}_{\bar{\mathbf{h}}}$ has been reflected across the real axis.

2.3.4.2*b*). Let the difference in phase between $\mathbf{F}_{\mathbf{h}}$ and $\mathbf{F}_{\bar{\mathbf{h}}}$ be φ. Thus

$$4|\mathbf{F}_{\mathbf{h}}''(H)|^2 = |\mathbf{F}_{\mathbf{h}}|^2 + |\mathbf{F}_{\bar{\mathbf{h}}}|^2 - 2|\mathbf{F}_{\mathbf{h}}||\mathbf{F}_{\bar{\mathbf{h}}}|\cos\varphi,$$

but since φ is very small

$$|\mathbf{F}_{\mathbf{h}}''(H)|^2 \simeq \tfrac{1}{4}(|\mathbf{F}_{\mathbf{h}}| - |\mathbf{F}_{\bar{\mathbf{h}}}|)^2.$$

Hence, a Patterson with coefficients $(|\mathbf{F}_{\mathbf{h}}| - |\mathbf{F}_{\bar{\mathbf{h}}}|)^2$ will be equivalent to a Patterson with coefficients $|\mathbf{F}_{\mathbf{h}}''(H)|^2$ which is proportional to $|\mathbf{F}_{\mathbf{h}}(H)|^2$. Such a Patterson (Rossmann, 1961*a*) will have vectors between all anomalous scatterers with heights proportional to the number of anomalous electrons $\Delta f''$. This 'anomalous dispersion' Patterson function has been used to find anomalous scatterers such as iron (Smith *et al.*, 1983; Strahs & Kraut, 1968) and sulfur atoms (Hendrickson & Teeter, 1981). The anomalous signal from Se atoms in selenomethionine-substituted proteins has been found to be extremely powerful and is now routinely used for protein structure determinations (Hendrickson, 1991). Anomalous signals from halide ions or xenon atoms have also been used to solve protein structures (Dauter *et al.*, 2000; Nagem *et al.*, 2003; Schiltz *et al.*, 2003). The anomalous signal from sulfur atoms, though very small (Hendrickson & Teeter, 1981), has recently been applied successfully to solve several protein structures (Debreczeni *et al.*, 2003; Ramagopal *et al.*, 2003; Yang *et al.*, 2003).

It is then apparent that a Patterson with coefficients

$$\Delta F_{\mathrm{ANO}}^2 = (|\mathbf{F}_{\mathbf{h}}| - |\mathbf{F}_{\bar{\mathbf{h}}}|)^2$$

(Rossmann, 1961*a*), as well as a Patterson with coefficients

$$\Delta F_{\mathrm{ISO}}^2 = (|\mathbf{F}_{\mathrm{NH}}| - |\mathbf{F}_{\mathrm{H}}|)^2$$

(Rossmann, 1960; Blow, 1958), represent Pattersons of the heavy atoms. The $\Delta F_{\mathrm{ANO}}^2$ Patterson suffers from errors which may be larger than the size of the Bijvoet differences, while the $\Delta F_{\mathrm{ISO}}^2$ Patterson may suffer from partial lack of isomorphism. Hence, Kartha & Parthasarathy (1965) have suggested the use of the sum of these two Pattersons, which would then have coefficients $(\Delta F_{\mathrm{ANO}}^2 + \Delta F_{\mathrm{ISO}}^2)$.

However, given both SIR and anomalous-dispersion data, it is possible to make an accurate estimate of the $|\mathbf{F}_{\mathrm{H}}|^2$ magnitudes for use in a Patterson calculation [Blundell & Johnson (1976, p. 340), Matthews (1966), Singh & Ramaseshan (1966)]. In essence, the Harker phase diagram can be constructed out of three circles: the native amplitude and each of the two isomorphous Bijvoet differences, giving three circles in all (Blow & Rossmann, 1961) which should intersect at a single point thus resolving the ambiguity in the SIR data and the anomalous-dispersion data. Furthermore, the phase ambiguities are orthogonal; thus the two data sets are cooperative. It can be shown (Matthews, 1966; North, 1965) that

$$F_{\mathrm{N}}^2 = F_{\mathrm{NH}}^2 + F_{\mathrm{N}}^2 \mp \frac{2}{k}(16k^2F_{\mathrm{P}}^2F_{\mathrm{H}}^2 - \Delta I^2)^{1/2},$$

where $\Delta I = F_{\mathrm{NH}}^{+\,2} - F_{\mathrm{NH}}^{-\,2}$ and $k = \Delta f''/f'$. The sign in the third-term expression is $-$ when $|(\alpha_{\mathrm{NH}} - \alpha_{\mathrm{H}})| < \pi/2$ or $+$ otherwise. Since, in general, $|\mathbf{F}_{\mathrm{H}}|$ is small compared to $|\mathbf{F}_{\mathrm{N}}|$, it is reasonable to assume that the sign above is usually negative. Hence, the heavy-atom lower estimate (HLE) is usually written as

$$F_{\mathrm{HLE}}^2 = F_{\mathrm{NH}}^2 + F_{\mathrm{H}}^2 - \frac{2}{k}(16k^2F_{\mathrm{P}}^2F_{\mathrm{H}}^2 - \Delta I^2)^{1/2},$$

which is an expression frequently used to derive Patterson coefficients useful in the determination of heavy-atom positions when both SIR and anomalous-dispersion data are available.

2.3.4.4. Computer programs for automated location of atomic positions from Patterson maps

Several programs are currently used for automated systematic interpretation of (difference) Patterson maps to locate the positions of heavy atoms and/or anomalous scatterers from isomorphous replacement and anomalous-dispersion data (Weeks *et al.*, 2003). These include *Solve* (Terwilliger & Berendzen, 1999), *CNS* (Brünger *et al.*, 1998), *CCP4* (Collaborative Computational Project, Number 4, 1994) and *Patsol* (Tong & Rossmann, 1993). In these programs, sets of trial atomic positions (seeds) are produced based on one- and two-atom solutions to the Patterson map (see Section 2.3.2.5) (Grosse-Kunstleve & Brunger, 1999; Terwilliger *et al.*, 1987; Tong & Rossmann, 1993). Information from a translation search with a single atom can also be used in this process (Grosse-Kunstleve & Brunger, 1999). Scoring functions have been devised to identify the likely correct solutions, based on agreements with the Patterson map or the observed isomorphous or anomalous differences, as well as the quality of the resulting electron-density map (Terwilliger, 2003*b*; Terwilliger & Berendzen, 1999). The power of modern computers allows the rapid screening of a large collection of trial structures, and the correct solution is found automatically in many cases, even when there is a large number of atomic positions (Weeks *et al.*, 2003).