Supplementary material

The normalized structure factor of index **h** is defined as

$$E(\mathbf{h}) = \left(\sum_{j=1}^{N} f_j^2 \right)^{-1/2} \sum_{j=1}^{N} f_j \exp 2\pi i \mathbf{h} \cdot \mathbf{r}_j \quad , \tag{i}$$

where there are N atoms in the unit cell, the j^{th} of which has scattering factor f_j for the scattering angle appropriate to **h**. Similarly the unitary structure factor is defined as

$$U(\mathbf{h}) = \left(\sum_{j=1}^{N} f_j^2 \right)^{-1} \sum_{j=1}^{N} f_j \exp 2\pi i \mathbf{h} \cdot \mathbf{r}_j = \frac{\left(\sum_{j=1}^{N} f_j^2 \right)^{1/2}}{\sum_{j=1}^{N} f_j} E(\mathbf{h}) \quad . \tag{ii}$$

The density at a general point in an E-map is

$$\rho(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{h}} |E(\mathbf{h})| \cos(2\pi \mathbf{h} \cdot \mathbf{r} - \phi_{\mathbf{h}}) \tag{iii}$$

where $\phi_{\mathbf{h}}$ is the phase of the structure factor of index **h**. This map has an average value $E(\mathbf{0})/V$, which in general is very small compared with the density at an atomic centre. The standard deviation of the map density is

$$\sigma_\rho = \frac{1}{V} \sum_{\substack{\mathbf{h} \\ \mathbf{h} \neq 0}} |E(\mathbf{h})|^2 \quad . \tag{iv}$$

From (iii) the density at the centre of atom j is

$$\rho_j = \frac{1}{V} \sum_{\mathbf{h}} |E(\mathbf{h})| \cos(2\pi \mathbf{h} \cdot \mathbf{r}_j - \phi_{\mathbf{h}}) \tag{v}$$

and from Woolfson (1961) the expected value of the cosine term is

$$\left\langle \cos(2\pi \mathbf{h} \cdot \mathbf{r}_j - \phi_{\mathbf{h}}) \right\rangle = \frac{f_j \sum_{j=1}^{N} f_j}{\sum_{j=1}^{N} f_j^2} \left| U(\mathbf{h}) \right| = \frac{f_j}{\left(\sum_{j=1}^{N} f_j^2 \right)^{1/2}} \left| E(\mathbf{h}) \right| \ . \qquad \text{(vi)}$$

Hence the expected value of the density at the atomic centre is

$$\left\langle \rho_j \right\rangle = \frac{1}{V} \sum_{\mathbf{h}} \eta_j(\mathbf{h}) \left| E(\mathbf{h}) \right|^2 \quad , \qquad \text{(vii)}$$

where

$$\eta_j(\mathbf{h}) = \frac{f_j}{\left(\sum_{j=1}^{N} f_j^2 \right)^{1/2}} \quad . \qquad \text{(viii)}$$

The expected value of the density at the centre of atom *j* in units of the map standard deviation is thus

$$\frac{\left\langle \rho_j \right\rangle}{\sigma_\rho} = \frac{\sum_{\mathbf{h}} \eta_j(\mathbf{h}) \left| E(\mathbf{h}) \right|^2}{\left(\sum_{\substack{\mathbf{h} \\ \mathbf{h} \neq 0}} \left| E(\mathbf{h}) \right|^2 \right)^{1/2}} \quad . \qquad \text{(ix)}$$

In the case of an equal-atom structure $\eta_j(\mathbf{h}) = N^{-1/2}$ and since $\left\langle \left| E \right|^2 \right\rangle = 1$ we then find

$$\frac{\left\langle \rho_j \right\rangle}{\sigma_\rho} = \sqrt{\frac{M}{N}} \qquad \text{(x)}$$

where *M* is the total number of data. For 1.0 Å data the average protein gives *M*/*N* about 120 so that an atomic peak has density about $11\sigma_\rho$, which makes it very distinctive. For a resolution of 1.7 Å *M*/*N* falls to 24 giving a peak of density about 5 σ_ρ, a value that will still define the atom reasonably well. Thus with 1.7 Å data and phases from a refined structure a map shows the structure quite well although the spread of the peaks due to the resolution limit may not give sharply defined atomic positions.

For unequal atoms the analysis needs to take account of the distributions of $\eta_j(\mathbf{h})$. The general tendency is for the concentration of atomic density to increase with atomic number so that for higher atomic numbers the scattering factors fall off more slowly. The ratio of the peak density of two atoms will be

$$\frac{\rho_j}{\rho_k} = \frac{\overline{f_j}}{\overline{f_k}} \tag{xi}$$

in which the averages are for all the terms in the density summation. Thus the ratio of peak heights for oxygen and carbon will be higher than 4:3, the ratio of the atomic numbers.

When there are phase errors then the atomic peak density becomes

$$\rho'_j = \frac{1}{V}\sum_{\mathbf{h}} |E(\mathbf{h})| \cos(2\pi\mathbf{h}\cdot\mathbf{r}_j - \phi_{\mathbf{h}} + \Delta\phi_{\mathbf{h}}) \tag{xii}$$

Since the expectation values of $\sin(2\pi\mathbf{h}\cdot\mathbf{r} - \phi_{\mathbf{h}})$ and $\sin(\Delta\phi_{\mathbf{h}})$ are both zero it follows that

$$\langle \rho'_j \rangle = \frac{1}{V}\sum_{\mathbf{h}} \eta_j(\mathbf{h}) |E(\mathbf{h})|^2 \cos(\Delta\phi_{\mathbf{h}}) \tag{xiii}$$

and

$$\frac{\langle \rho_j' \rangle}{\sigma_\rho} = \frac{\sum_{\mathbf{h}} \eta_j(\mathbf{h}) |E(\mathbf{h})|^2 \cos(\Delta\phi_{\mathbf{h}})}{\left(\sum_{\substack{\mathbf{h} \\ \mathbf{h} \neq 0}} |E(\mathbf{h})|^2 \right)^{1/2}} \quad . \tag{xiv}$$

The phase errors make the expectation value of the peak density less by a factor $\langle \cos(\Delta\phi_{\mathbf{h}}) \rangle_{|E|^2}$, the $|E|^2$-weighted average value of $\cos(\Delta\phi_{\mathbf{h}})$. If this is 0.25, roughly corresponding to a mean phase error of 75° then, for 1.0 Å data and an equal-atom structure, the expectation height of an atomic peak would be about 2.75 σ, which is towards the low-probability end of a random density fluctuation. If the actual density at a point was as high as, say, 4 σ then it would be likely that this point is at, or close to, an atomic centre. On the other hand a density of 0.5 σ at a point makes it an unlikely candidate to be at or near an atomic location. This provides the rationale of the DDM process.

For 1.7 Å resolution data a mean phase error of 75° gives the expected density at an atomic peak of about 1.25 σ, well within the common range of random fluctuation. DDM could not be applied in such a case. However, an initial mean phase error of 60°, would give an expected atomic-peak height of 2.5 σ, again allowing DDM to be used successfully. This shows the need for better starting phases for lower-resolution structures.

References: Woolfson, M.M. (1961). Direct methods in Crystallography. Oxford University Press.