

## 2.2. QUALITY INDICATORS

between the  $E$  values (analogous to normalized structure factors as above) derived from the observed isomorphous or anomalous differences and those calculated from the substructure model. In contrast to CC(all), however, CC(weak) is calculated for the weak reflections only. As above,  $x$  and  $y$  are the  $E$  values derived from the observed and calculated differences. The  $E$ -value cutoff for defining a reflection as weak can be chosen by the user, but a typical threshold value is 1.5, although lower values may be required for low-resolution data (Schneider & Sheldrick, 2002). A value of CC(weak)  $\geq 0.15$  often indicates that the substructure has been correctly identified (Sheldrick, 2010).

**The minimal function,  $R(\varphi)$ .** The minimal function  $R(\varphi)$  is a measure of the mean-square difference between the values of the triplets calculated using a particular set of phases and the expected values of the same triplets as given by the ratio of modified Bessel functions. The minimal function is expected to have a constrained global minimum when the phases are equal to their correct values for some choice of origin and enantiomorph (the minimal principle).

$$R(\varphi) = \sum_{H,K} A_{HK} \left\{ \cos \varphi_{HK} - [I_1(A_{HK})/I_0(A_{HK})] \right\}^2 / \sum_{H,K} A_{HK}, \quad (2.2.4.1)$$

where  $A_{HK} = (2/N^{1/2})|E_H E_K E_{H+K}|$  with  $N$  being the number of atoms in the corresponding primitive unit cell. The minimal function is the function minimized in the program *Shake&Bake*, abbreviated as *SnB* (Weeks *et al.*, 1993, 1994).

## 2.2.5. Quality indicators for phase determination

Once the isomorphous or anomalous substructure has been successfully determined, it can be used as reference point for the calculation of phases. The quality of the resulting phases is dependent on the strength of the isomorphous or anomalous signal and the completeness and correctness of the isomorphous or anomalous substructure.

**Cullis  $R$  factor,  $R_{\text{Cullis}}$ .** The Cullis  $R$  factor (Cullis *et al.*, 1961) for phase determination by isomorphous replacement is defined as the ratio between the lack-of-closure error  $\varepsilon(\varphi_P)$  [equation (2.2.5.1a) below] and the isomorphous difference  $|F_{\text{PH}} - F_P|$ . It is the most useful signal for a usable heavy-atom derivative. Values  $< 0.6$  for centrosymmetric data are excellent, while values  $< 0.9$  are still usable.

$$\varepsilon(\varphi_P) = |F_{\text{PH}} - |F_P + F_H||, \quad (2.2.5.1a)$$

$$R_{\text{Cullis}} = \sum_{hkl} |F_{\text{PH}} - |F_P + F_H|| / \sum_{hkl} |F_{\text{PH}} - F_P|. \quad (2.2.5.1b)$$

**Anomalous Cullis  $R$  factor,  $R_{\text{Cullis,ano}}$ .** The Cullis  $R$  factor for phase determination by anomalous dispersion is defined as the ratio between the lack-of-closure error and the observed anomalous difference  $|F_{\text{PH}}(hkl) - F_{\text{PH}}(\bar{h}\bar{k}\bar{l})|$ . The lack-of-closure error in the anomalous-dispersion case is the difference between the observed anomalous difference and the calculated anomalous difference  $2F_H \sin \alpha_P$ , where  $\alpha_P$  is the protein phase. A value of  $R_{\text{Cullis,ano}} < 1.0$  suggests that a contribution to the phasing from the anomalous data is likely (*MLPHARE* program documentation; Collaborative Computational Project, No. 4, 1994).

$$R_{\text{Cullis,ano}} = \frac{\sum_{hkl} ||F_{\text{PH}}(hkl) - F_{\text{PH}}(\bar{h}\bar{k}\bar{l})| - |2F_H \sin \alpha_P|}{\sum_{hkl} |F_{\text{PH}}(hkl) - F_{\text{PH}}(\bar{h}\bar{k}\bar{l})|}. \quad (2.2.5.2)$$

**Phasing power,  $\text{PP}_{\text{iso}}$ .** The isomorphous phasing power  $\text{PP}_{\text{iso}}$  for phase determination by isomorphous replacement is defined

for a particular pair of native and heavy-atom-derivative data sets as the ratio of  $|F_H|$  and  $\varepsilon(\varphi_P)$ , where  $|F_H|$  is the calculated amplitude of the heavy-atom structure factor and  $\varepsilon(\varphi_P)$  is the lack-of-closure error [equation (2.2.5.1a)].  $F_P + F_H$  is a vector sum of the calculated structure factor  $F_H$  and the structure factor  $F_P$ .

$$\text{PP}_{\text{iso}} = \sum_{hkl} |F_H| / \sum_{hkl} |F_{\text{PH}} - |F_P + F_H||. \quad (2.2.5.3)$$

There is another, slightly different, definition of  $\text{PP}_{\text{iso}}$ , which is implemented in the program *SOLVE*. Here,  $\text{PP}_{\text{iso}}$  is defined as the ratio of the r.m.s. of the  $|F_H|$  values and the r.m.s. of the lack-of-closure errors  $\varepsilon(\varphi_P)$ . For each reflection  $hkl$ , a weighted average of  $\varepsilon(\varphi_P)$  is calculated by integrating  $\varepsilon(\varphi_P)$  over the whole phase circle.

$$\text{PP}_{\text{iso}} = (\sum_{hkl} |F_H|^2)^{1/2} / (\sum_{hkl} \langle |F_{\text{PH}} - |F_P + F_H|| \rangle^2)^{1/2}. \quad (2.2.5.4)$$

*Note.* Owing to the cancelling out of the factor  $N^{1/2}$  in the numerator and denominator, the latter  $\text{PP}_{\text{iso}}$  formula does not appear as a ratio of r.m.s. values at first sight.

**Anomalous phasing power,  $\text{PP}_{\text{ano}}$ .** The anomalous phasing power  $\text{PP}_{\text{ano}}$  for phase determination by anomalous-dispersion methods is defined as the ratio of the sum of calculated anomalous differences  $d''_{\text{calc}}$  and the sum of estimated standard uncertainties  $\sigma(d''_{\text{obs}})$  in the measurement of these anomalous differences:

$$\text{PP}_{\text{ano}} = \sum_{hkl} d''_{\text{calc}} / \sum_{hkl} \sigma(d''_{\text{obs}}). \quad (2.2.5.5)$$

As with  $\text{PP}_{\text{iso}}$  (see above), the program *SOLVE* uses a slightly different definition of  $\text{PP}_{\text{ano}}$ . Here, the anomalous phasing power is defined as the ratio of the r.m.s. of the  $d''_{\text{calc}}$  values and the r.m.s. of  $\sigma(d''_{\text{obs}})$ . For this, a weighted average of  $d''_{\text{calc}}$  is computed by integrating over the whole phase circle for each reflection.

$$\text{PP}_{\text{ano}} = (\sum_{hkl} \langle d''_{\text{calc}} \rangle^2)^{1/2} / [\sum_{hkl} \sigma(d''_{\text{obs}})^2]^{1/2}. \quad (2.2.5.6)$$

*Note.* As above in the  $\text{PP}_{\text{iso}}$  formula, the factors  $N^{1/2}$  in the numerator and denominator cancel out.

**Figure of merit (f.o.m.),  $m$ .** The figure of merit  $m$  is a term used in a number of contexts in X-ray crystallography. In its most common use, it is defined as the weight applied to an individual structure-factor amplitude that, in conjunction with its best phase, gives rise, in a Fourier synthesis, to the electron-density map with the minimum level of noise (Blow & Crick, 1959). Typically,  $m$  is given as an average value over all reflections in the data set or in a given resolution shell.

$$m = \int P(\alpha) \exp(i\alpha) d\alpha / \int P(\alpha) d\alpha = \langle \cos(\Delta\alpha) \rangle, \quad (2.2.5.7)$$

where  $P(\alpha)$  is the probability of the phase  $\alpha$ , initial or refined, being the best phase and  $\Delta\alpha = \alpha_{\text{best}} - \alpha$  is the error in the phase angle at  $\alpha$ . The integration is from 0 to  $2\pi$  and values for  $m$  range from 0 to 1.

## 2.2.6. Quality indicators for density modification and phase improvement

After determination of initial phases, a first electron-density map can be computed. It is expected that this map will contain significant errors and improbable features. Additional information, such as the flatness of the electron density in the solvent region or the similarity of electron-density regions of two or more identical molecules in the asymmetric unit, can be exploited to modify the electron density and hence improve the phases.