

13. MOLECULAR REPLACEMENT

longer dependent on the number of atoms present in the unit cell [equation (13.3.2.7)]. The starting position is usually chosen such that the centre of the search model is at (0, 0, 0). Then the vector that is determined from the translation searches will define the centre of the model in the unit cell.

Equation (13.3.2.7) can be generalized to allow for the presence of other molecules that are to remain stationary during the translation search:

$$\bar{F}_h^c = \bar{A}_h + \sum_n \bar{f}_{h,n} \exp(2\pi i h [T_n] v_0), \quad (13.3.2.10)$$

where \bar{A}_h is the contribution from the stationary molecules. This formulation is useful if there is more than one molecule in the asymmetric unit. The position of one of the molecules can be determined first, and the model is then included as a stationary molecule for the position search of the next molecule.

Evaluation of the R factor and the correlation coefficient [equations (13.3.2.1) and (13.3.2.3)] in a translation search is generally rather slow. A method has been developed to calculate the correlation coefficient by the fast Fourier transform (FFT) technique, which involves reciprocal vectors up to four times the resolution of the reflection data (Navaza & Vernoslova, 1995).

Equation (13.3.2.7) can also be generalized to allow for the presence of two (or more) search models that are to move independently of each other during the translation search. However, this will generally lead to a six- (or more) dimensional problem and is extremely expensive in computation time. With recent improvements in computer technology (especially parallel processing), it might be feasible to carry out such searches in special cases. However, this aspect of translation functions will not be discussed further here.

13.3.3. Patterson-correlation translation function

The most commonly used translation-search indicator is based on the correlation between the observed and the calculated Patterson maps (Crowther & Blow, 1967). Rotation functions are based on the overlap of only a subset of the interatomic vectors in the Patterson map, *i.e.* only those near the origin of the unit cell, which generally contain the self vectors within each crystallographically unique molecule. The correct orientation and position of a search model in the crystal unit cell should lead to the maximal overlap of both the self and the cross vectors, *i.e.* maximal overlap between the observed and the calculated Patterson maps throughout the entire unit cell (Tong, 1993),

$$PC = \int_{\Omega} P^o(u) P^c(u) du \quad (13.3.3.1)$$

$$= \sum_h (F_h^o)^2 |\bar{F}_h^c|^2.$$

The calculated structure factor is a function of the translation vector v_0 [equation (13.3.2.10)]. Combining equations (13.3.2.10) and (13.3.3.1) gives

$$PCTF(v_0) = \sum_h (F_h^o)^2 |A_h|^2 + \sum_h \sum_n (F_h^o)^2 |\bar{f}_{h,n}|^2$$

$$+ 2 \sum_h \sum_n (F_h^o)^2 \bar{A}_h \bar{f}_{h,n}^* \exp(-2\pi i h [T_n] v_0)$$

$$+ \sum_h \sum_n \sum_{m \neq n} (F_h^o)^2 \bar{f}_{h,n} \bar{f}_{h,m}^* \exp\{-2\pi i h ([T_m] - [T_n]) v_0\}. \quad (13.3.3.2)$$

The first two terms in equation (13.3.3.2) contribute a constant to the correlation and are generally ignored in this calculation (but see Section 13.3.8). The Patterson-correlation translation function is therefore a Fourier transform, with $h[T_n]$ (from the third term) and

$h([T_m] - [T_n])$ (from the fourth term) as the indices. This function can be evaluated quickly by the FFT technique. The term $h([T_m] - [T_n])$ often leads to the doubling of the original reflection indices. For example, $h([T_1] - [T_2])$ will give rise to $(2h, 0, 2l)$ for monoclinic space groups (b -unique setting). Therefore, the Patterson-correlation Fourier transform should normally be sampled with a grid size about 1/6 of the maximum resolution of the reflection data used in the calculation.

A disadvantage of the Patterson-correlation translation function as formulated in equation (13.3.3.2) is that the results from the calculation are on an arbitrary scale. This makes it difficult to compare results from different calculations. The R factor or the correlation coefficient can be calculated for the top peaks in the Patterson-correlation translation function [equation (13.3.3.2)] to place the results on an 'absolute' scale. Alternatively, the correlation, as defined by equation (13.3.3.1), can be normalized to become a Patterson-correlation coefficient (Harada *et al.*, 1981),

$$PC = \frac{\sum_h (F_h^o)^2 |\bar{F}_h^c|^2}{\left[\sum_h (F_h^o)^4 \sum_h |\bar{F}_h^c|^4 \right]^{1/2}}. \quad (13.3.3.3)$$

This correlation coefficient is equivalent to that defined by equation (13.3.2.3) for reflection intensities (Fujinaga & Read, 1987). The FFT technique can be used to evaluate the $\left(\sum_h |\bar{F}_h^c|^4 \right)^{1/2}$ term in the denominator of equation (13.3.3.3), which involves reciprocal vectors up to four times the data resolution (Navaza & Vernoslova, 1995). Alternatively, this term can be approximated by an expression which also measures the packing of the search molecules in the crystal (see Section 13.3.5).

The intramolecular vectors can be removed (Crowther & Blow, 1967) from the Patterson maps by subtracting, after appropriate scaling, the structure factors calculated from individual molecules [equation (13.3.2.8)].

The translation functions, as defined above, are based on the structure-factor amplitudes. Normalized structure factors (E 's) may provide better results under certain circumstances, since they increase the weight of the high-resolution reflections in the translation function (Harada *et al.*, 1981; Tickle, 1985; Tollin, 1966). Since the Patterson-correlation translation function is also based on reflection intensities, the 'large term' approach can be used to accelerate the calculation (Tollin & Rossmann, 1966).

13.3.4. Phased translation function

If an atomic model needs to be placed in an electron-density map that has been obtained through other methods (*e.g.* the multiple-isomorphous-replacement method or partial model phases), the phased translation function can be used (Bentley & Houdusse, 1992; Colman *et al.*, 1976; Read & Schierbeek, 1988; Tong, 1993). It is essentially based on the correlation between observed and calculated electron-density values throughout the unit cell:

$$PTF(v_0) = \sum_h \sum_n \bar{F}_h^o \bar{f}_{h,n}^* \exp(-2\pi i h [T_n] v_0). \quad (13.3.4.1)$$

Therefore, the phased translation function can also be evaluated by the FFT technique. As with the Patterson-correlation translation function, the phased translation function can be placed on an 'absolute' scale by introducing appropriate normalizing factors or by converting the results to R factors or correlation coefficients. It should be noted that the prior phase could be in the wrong hand, so both \bar{F}_h^o and $(\bar{F}_h^o)^*$ may need to be tried in the phased translation function.

The stationary molecules contribute a constant to the phased translation function and are not shown in equation (13.3.4.1). However, the phase information from the stationary molecules can