be applied to the observed structure-factor amplitudes, and the phased translation function, rather than the Patterson-correlation translation function, can be used in the search for additional molecules (Bentley & Houdusse, 1992; Driessen *et al.*, 1991; Read & Schierbeek, 1988). This could prove especially useful in locating the last few molecules in cases where there are several molecules in the asymmetric unit.

### 13.3.5. Packing check in translation functions

A correct molecular-replacement solution should lead to the placement of the search model at the correct orientation and position in the crystal unit cell. For this solution, there should be no or minimal steric clashes among the crystallographically related and noncrystallographically related molecules in the unit cell. Therefore, proper packing of the search model in the crystal unit cell is an important component of the molecular-replacement structure solution.

The packing of the search model in the unit cell can be estimated by determining the electron-density overlap among the molecules. This overlap can be calculated numerically, given the molecular envelope (Hendrickson & Ward, 1976). It can also be estimated by an analytical function (Harada *et al.*, 1981),

$$O(v_0) = \sum_h \left|\bar{F}_h^c(v_0)\right|^2 \Big/ N \sum_h \left|\bar{f}_{h,1}\right|^2, \qquad (13.3.5.1)$$

where $N$ is the number of crystallographic symmetry operators. This overlap function assumes a value of 1 when there is no overlap among the molecules, and higher values when there is overlap. This function has been used to replace the $\left(\sum_h \left|\bar{F}_h^c\right|^4\right)^{1/2}$ term in the denominator of equation (13.3.3.3) (Harada *et al.*, 1981). Consequently, those positions that lead to steric clashes among the molecules will be down-weighted, thereby increasing the signal for the correct solution.

The overlap functions provide an overall estimate for the packing of the search model in the unit cell. A more detailed packing analysis can be based on the checking of atomic contacts. For example, the number of $C_\alpha \cdots C_\alpha$ contacts below a pre-specified distance cutoff (normally between 2 to 3 Å) in a protein crystal can be determined. Too many such contacts would indicate significant overlap of the molecules. For nucleic acid structures, a set of representative atoms (for example, P, $N_1$, $C_4'$) can be selected from each nucleotide for this packing analysis.

### 13.3.6. The unique region of a translation function (the Cheshire group)

The region of the unit cell that should be covered during a translation search does not generally correspond to the asymmetric unit of the space group. Since the search model has a defined orientation, it can only reside in one of the asymmetric units in the unit cell. Lacking knowledge as to which asymmetric unit the model occupies, the entire unit cell would need to be searched. However, most space groups possess alternative origins, which means the position of a molecule in the unit cell can only be determined to within certain sets of translations. For example, in space group $P2_12_12_1$, there are eight alternative origins at $(0, 0, 0)$, $(\frac{1}{2}, 0, 0)$, $(0, \frac{1}{2}, 0)$, $(0, 0, \frac{1}{2})$, $(\frac{1}{2}, \frac{1}{2}, 0)$, $(\frac{1}{2}, 0, \frac{1}{2})$, $(0, \frac{1}{2}, \frac{1}{2})$ and $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. This implies that the region that should be searched to locate a molecule need only be $\frac{1}{8}$ of the volume of the unit cell [for example, $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$]. In addition, for polar space groups, the position of the molecule along the polar axis is arbitrary. The symmetry, as defined by these unique regions, is also known as the Cheshire group (Hirshfeld, 1968), and has been defined for all the 230 space groups.

Once the first molecule is positioned, the origin of the unit cell is fixed as well. The search for subsequent molecules will need to cover the entire unit cell.

### 13.3.7. Combined molecular replacement

The traditional division of the molecular-replacement problem into two steps is partly due to limited computer power. Such a division has placed more pressure on the rotation function, since generally, only a few rotation angles are examined by translation functions. The correct orientation, therefore, needs to be among the top few peaks either directly in the rotation functions or after Patterson-correlation refinement (Brünger, 1990).

With modern computers, it is no longer necessary to maintain the strict division between the rotational and the translational components. Even though a full six-dimensional search is still generally impractical, a limited six-dimensional search can certainly be performed. The Patterson-correlation translation function is preferred for this limited six-dimensional search since it can be evaluated quickly with the FFT technique. Using the $R$ factor or the correlation coefficient as the translation function would severely limit either the exploration of rotational space (Fujinaga & Read, 1987) or the reflection data that are used in the calculation (Rabinovich & Shakked, 1984).

Recently, an automated molecular-replacement protocol has been implemented which automatically examines the top peaks in the rotation function by translation functions (Navaza, 1994). This protocol has proven to be remarkably powerful. It assumes that the correct rotation solution is near the top peaks in the rotation function. A more general assumption is that the correct rotation for the search model should produce high values in the rotation function, even though they may not be near peaks in the rotation function. An error of 6° between the correct rotation angles and the peak in the rotation function, which often occurs, can make it impossible to obtain the correct translation-function solution (Fujinaga & Read, 1987).

The combined-molecular-replacement protocol (Tong, 1996*a*) therefore consists of examining all the grid points in the rotation function with heights greater than a defined cutoff value using the Patterson-correlation translation function. The top peaks (usually 10 to 20) in each translation function are all examined as possible solutions. The results from these translation functions are converted to the $R$ factor or the correlation coefficient, enabling comparisons among the various orientations. This protocol allows the automatic examination of not only the top peaks in the rotation functions, but also those angles that produce high rotation-function values. In addition, packing of the model in the crystal is examined automatically to eliminate those solutions that have severe clashes among the molecules (see Section 13.3.5). This generalized protocol has proven more powerful than conventional methods in a few structure determinations (Tong, 1996*a*; Wu *et al.*, 1997).

An alternative approach, examining the neighbourhood surrounding the rotation-function peaks, is also possible (Urzhumtsev & Podjarny, 1995).

### 13.3.8. The locked translation function

In the presence of noncrystallographic symmetry (NCS), locked self-rotation functions can be used to determine the orientation of the NCS elements in the crystal unit cell (Tong & Rossmann, 1990). Often, an atomic model for the monomer of the NCS assembly is available, but not the model of the entire assembly. This atomic model can be used in ordinary cross-rotation-function calculations.





references