# 13. MOLECULAR REPLACEMENT

## 13.1. Noncrystallographic symmetry

BY D. M. BLOW

### 13.1.1. Introduction

Excellent reviews of noncrystallographic symmetry exist. The subject is also discussed in Volume B of this series (Rossmann & Arnold, 2001). Other important reviews include those by Rossmann (1990), Lawrence (1991) and Rossmann (1995). A volume produced for a Daresbury Study Weekend (Dodson *et al.*, 1992) has many interesting chapters.

In this introductory chapter, effort has been made to cite some of the earliest work which initiated the methods which have now become familiar.

### 13.1.2. Definition of noncrystallographic symmetry

#### 13.1.2.1. *Standard noncrystallographic symmetry*

The standard cases of noncrystallographic symmetry arise when *there is more than one similar subunit in the crystallographic asymmetric unit.*

The phrase 'noncrystallographic symmetry' is used because the operation required to superimpose one subunit on another is similar to a symmetry operation, but it operates only over a local volume, and the symmetry is inexact because the subunits are in different environments.

The 'subunit' can be a molecular aggregate, a single molecule, a monomer unit of an oligomeric molecule, or a fragment of a molecule.

The word 'similar' is used because protein subunits in different environments are never identical. At the very least, surface side chains are differently ordered, and solvation is different because of different interactions with adjacent subunits.

If noncrystallographic symmetry exists, methods are available to define the operation required to superimpose one unit on another (Rossmann & Blow, 1962; Rossmann *et al.*, 1964). When this has been done, new information is available to improve the accuracy of structural results (Rossmann & Blow, 1963).

Table 13.1.2.1 presents different types of symmetry situations which may arise when noncrystallographic symmetry exists. It is frequently observed that the local symmetry corresponds or approximates to point-group symmetry. This arises very often because a natural molecular form is a symmetric oligomer whose symmetry is not fully expressed in the crystal symmetry [cases (1) and (3)]. Helical symmetry or pseudo-helical symmetry is also common [case (2)], especially in biological materials, but it cannot always be exploited crystallographically because the specimens are often noncrystalline fibres.

#### 13.1.2.2. *Generalized noncrystallographic symmetry*

Crystallographic methods similar to those which exploit standard noncrystallographic symmetry can often be applied to a more general situation, where *similar subunits exist in different crystals* (Scouloudi, 1969; Tollin, 1969) *or where the structure of a subunit is already predictable* (Hoppe, 1957; Lattman & Love, 1970). The types of relationship which may arise are summarized in the right-hand column of Table 13.1.2.1.

#### 13.1.2.3. *Exploitation of noncrystallographic symmetry*

In order to draw structural information from noncrystallographic symmetry, the different classes of subunit must provide different information. Cases of pseudo-crystallographic symmetry, where subunits are *almost* in an arrangement of higher crystallographic symmetry, are difficult to exploit by the techniques discussed in this chapter. Typically only weak reflections (those which would be forbidden by the higher symmetry if it were exact) provide extra information. This situation often arises in case (6), Table 13.1.2.1. Similarly in cases (7) and (8), comparison of crystals whose cell dimensions or contents are only slightly altered gives little new information.

### 13.1.3. Use of the Patterson function to interpret noncrystallographic symmetry

#### 13.1.3.1. *Rotation operations*

The first step towards identifying and exploiting noncrystallographic symmetry is to find the operation that is required to

Table 13.1.2.1. *Noncrystallographic symmetry in crystals*

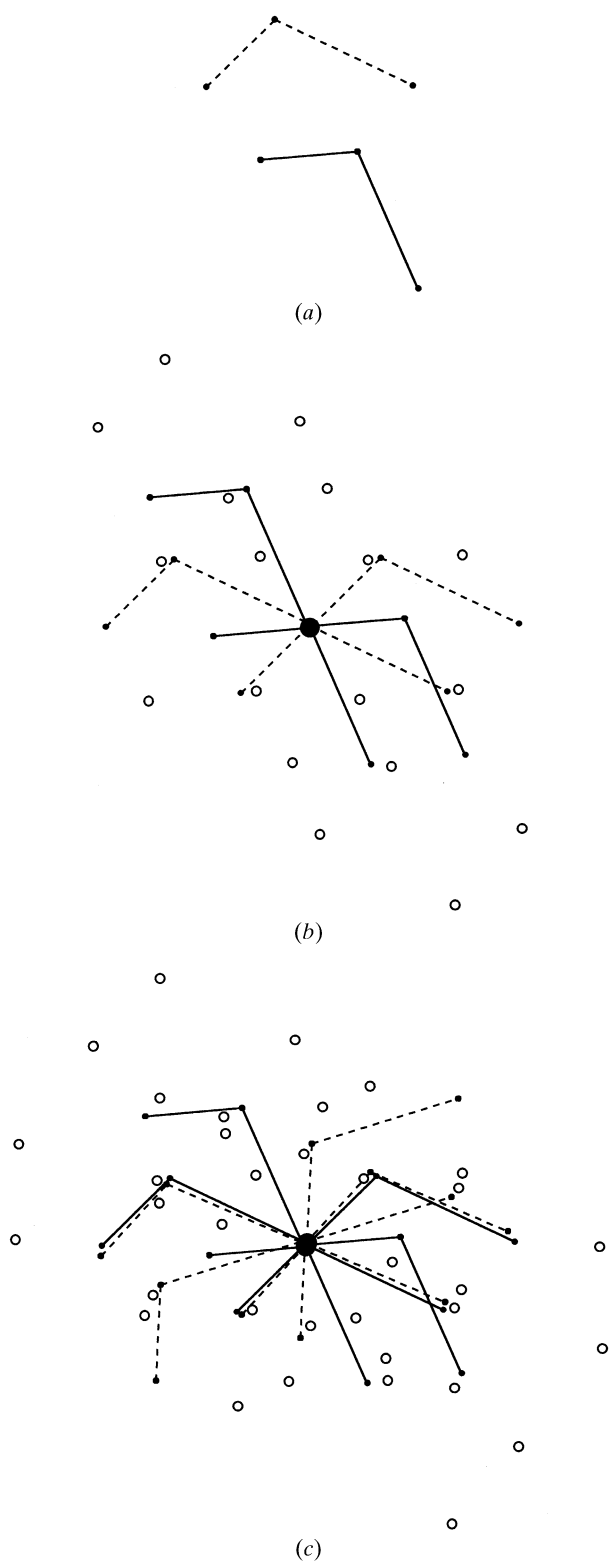|  | Relationships within the same crystal | Relationships between different crystals |
|---|---|---|
| Symmetry relations | (1) Symmetry of a noncrystallographic point group (*e.g.* 532) <br> (2) Infinite non-closed symmetry (helix) <br> (3) Crystallographic point-group symmetry, not incorporated into lattice | (6) Simple crystallographic relationship between two crystal forms |
| Relations not forming a group | (4) Similar subunits without systematic relationship | (7) Polymorphism (identical molecules crystallize differently) |
| Partial structural relationship | (5) Similar subunits account for only a part of unit-cell contents | (8) Crystals of different but similar molecules <br><br> (9) Crystals containing similar molecules, with other scattering material |

superimpose one subunit upon another. Superposition of one asymmetric rigid body upon a similar one requires in general that it be rotated and translated. A general rotation in three-dimensional space requires three variables to specify it: these can be the latitude and longitude of a rotation axis and the angle of rotation ($\kappa$, $\psi$, $\varphi$), or they can be the Euler angles. These rotational systems are presented in *IT* B (Rossmann & Arnold, 2001) and are discussed in detail in Chapter 13.2 by Navaza. Similarly, a general three-dimensional translation is specified by three variables. The operation to superimpose subunits therefore requires six variables to define it. Surveys or searches in six-dimensional functions are overwhelmingly laborious, though they have long been possible (Milledge, 1962; Kayushina & Vainshtein, 1965) and are now becoming easier (Kissinger *et al.*, 1999; Sheriff *et al.*, 1999).

The Patterson function, a function directly calculable from observed diffraction intensities, is a function which defines its own origin. Being a function in vector space, its origin is necessarily at a point representing a vector of zero length. It is this special property of the Patterson function which allows its use to factorize the six-dimensional problem into two three-dimensional ones. It is available without information about the phases of the diffraction data. Even when structural data are available, it is usually easier to make a three-dimensional rotation search based on the Patterson function than to carry out a full six-dimensional search using the known structure.

The Patterson function of a crystal may be considered to have two components: vectors between scattering centres in the same subunit, and those between different subunits. The intra-subunit vectors are necessarily shorter than the maximum subunit dimension. Inter-subunit vectors, though a few may be short, are clustered about the distances separating different subunits in the crystal, so they are mostly of the magnitude of the subunit dimensions or longer. By considering the region closer to the origin of the Patterson function, it is possible to include a high proportion of intra-subunit vectors.

Fig. 13.1.3.1 shows how the relation between the set of intra-subunit vectors of one subunit and the intra-subunit vectors of a similar subunit defines a rotation operation. This rotation is identical to the rotational part of the operation required to superimpose the subunits. If the whole Patterson function is rotated in this way, and then superimposed upon itself, one set of intermolecular vectors of the rotated Patterson function is superimposed upon a set of intermolecular vectors of the original (Fig. 13.1.3.1*c*). The self-rotation function (Rossmann & Blow, 1962) searches for correlations between a rotated Patterson function and the original.

By working with Patterson-function vectors, there is no dependence on the relative positions of the two subunits. Relative rotations can be determined in the standard case (Rossmann & Blow, 1962) or in the generalized case (Prothero & Rossmann, 1964; Lattman & Love, 1970). Methods for using the Patterson function to identify the rotation operation are detailed in Chapter 13.2.

### 13.1.3.2. *Translation operations*

When the rotation operation has been identified, the translation between the two differently oriented subunits needs to be determined (see Chapter 13.3 by Tong). The translation can only be defined in relation to an assigned origin of the subunit. It is possible to define the translation vector relative to the 'centres of gravity' (more precisely, the centres of scattering density) of the subunits. In this way, a translation may be defined between subunits of unknown structure. This approach, based only on information in the Patterson function, is the only available method if no phase information is available, but has several difficulties – it only applies

Fig. 13.1.3.1. (*a*) A dimer in which the two subunits are related by arbitrary rotation and translation. One member of the dimer is indicated by dashed lines joining its atoms. (*b*) The vector set representing the Patterson function of this dimer. The intra-subunit vectors of each subunit are indicated by filled circles linked by lines; the inter-subunit vectors are shown as open circles. (*c*) A second copy of the Patterson function has been rotated over the original. Intra-subunit vectors of the original Patterson function are indicated by full lines; the rotated intra-subunit vectors are distinguished by dashed lines. The rotation is almost the same as the rotation of one subunit to make it parallel to the other. When the rotation is exactly correct, half of the intra-subunit vectors of one Patterson will superimpose onto the other. None of the inter-subunit vectors superimpose.

accurately in certain cases, and the results are difficult to interpret and are imprecise (Blow *et al.*, 1964; Rossmann *et al.*, 1964). In practice, when dealing with a totally unknown structure, translational relationships are more frequently discovered by using specific markers on the subunit (intense scattering centres or 'heavy atoms', or anomalous scattering centres).

### 13.1.4. Interpretation of generalized noncrystallographic symmetry where the molecular structure is partially known

#### 13.1.4.1. *The cross-rotation function*

The rotation function can be used in the generalized case to compare the Patterson functions of different crystals. When used in this way, it is called the cross-rotation function. In the most usual case, information providing some kind of structural model is available for one of the crystals.

The power of the cross-rotation function may be greatly improved by removing all intermolecular vectors from the 'model' Patterson function. This may be done by constructing an imaginary crystal structure in which a single copy of the structural model is placed in a unit cell that is large enough for all intermolecular vectors to be longer than the longest intramolecular vector of the model. In this cell, the self-Patterson vectors may be completely isolated and used for comparison with the Patterson function of a crystal containing a molecule of unknown orientation.

#### 13.1.4.2. *The cross-translation function*

In searching for the position of a molecule in the generalized case of noncrystallographic symmetry, a molecular model defines an origin of coordinates in the model structure, and the corresponding position can be sought in an unknown structure (Nordman & Nakatsu, 1963; Tollin & Cochran, 1964; Huber, 1965; Crowther & Blow, 1967). The procedure is to calculate a three-dimensional function whose peaks should lie at the inter-subunit vectors. In

some procedures (Tollin & Cochran, 1964; Crowther & Blow, 1967), this function may be calculated as a Fourier series.

The translation functions will fail if the corresponding rotation is incorrect, or even if it is insufficiently accurate to give a good overlap between the structures. To avoid this danger, Brünger (1997) recommends computing translation functions using rotations corresponding to many (*e.g.* 200) high values of the rotation function. Though this is a huge increase in computing load, it still compares favourably with a full six-dimensional search.

These methods are considered further in Chapter 13.3.

#### 13.1.4.3. *Structure determination*

Table 13.1.4.1 distinguishes a number of different situations in which noncrystallographic symmetry can be used to aid structure determination. The most frequent application of molecular-replacement methods is to cases where a structure is partially known, but is not yet susceptible to refinement by standard techniques.

Two types of situation arise in the standard case, where noncrystallographically related subunits exist in the same crystal. Most frequently [type (2)], the noncrystallographic symmetry allows the electron density to be improved at the given resolution. Occasionally, high-order noncrystallographic symmetry may be used to extend the resolution to the point where conventional structural refinement becomes possible (Schevitz *et al.*, 1981; McKenna *et al.*, 1992). In the most favourable case, high-order noncrystallographic symmetry constraints may allow direct structure determination [type (1)], starting from the position of a symmetric particle in the asymmetric unit (Jack, 1973).

In the generalized case, most often, similarities with a known molecular structure can be employed to improve an unknown structure [types (5) and (6)]. Such techniques were first used by Tollin (1969) (before structural refinement was possible) and by Fehlhammer & Bode (1975).

It is also possible that a refinable structure could be generated from intensity data observed from several different crystal forms,

Table 13.1.4.1. *Structure determination using noncrystallographic symmetry*

| Starting structural information | Relationships within the same crystal (standard case) | Relationships between different crystals (generalized case) |
|---|---|---|
| None | (1) Subunit arrangement defined by relation between noncrystallographic and crystallographic symmetry. Resolution extended by noncrystallographic symmetry constraints | (3)* Subunit arrangement defined by relation between noncrystallographic and crystallographic symmetry in at least one crystal. Cross-rotation and translation functions applied to other crystals. Resolution extended by noncrystallographic symmetry constraints |
| Poorly resolved structure, unsuitable for refinement | (2) Electron density improved or resolution extended by noncrystallographic symmetry constraints | (4)* Resolution extended by noncrystallographic symmetry constraints |
| Similar structure known | | (5) Subunit orientation found by cross-rotation and translation functions. Phases derived from structural model and may be improved by noncrystallographic symmetry constraints |
| Part of unknown structure resembles a known structure | | (6) Subunit orientation found by cross-rotation and translation functions. Phases derived from structural model and may be improved by noncrystallographic symmetry constraints |

* Structure determinations of this kind have not been reported.

265