International Tables for Crystallography (2006). Vol. F, Chapter 13.1, pp. 263–268.

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) (2) Infinite non-closed symmetry (helix) (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 (9) Crystals containing similar molecules, with other scattering material

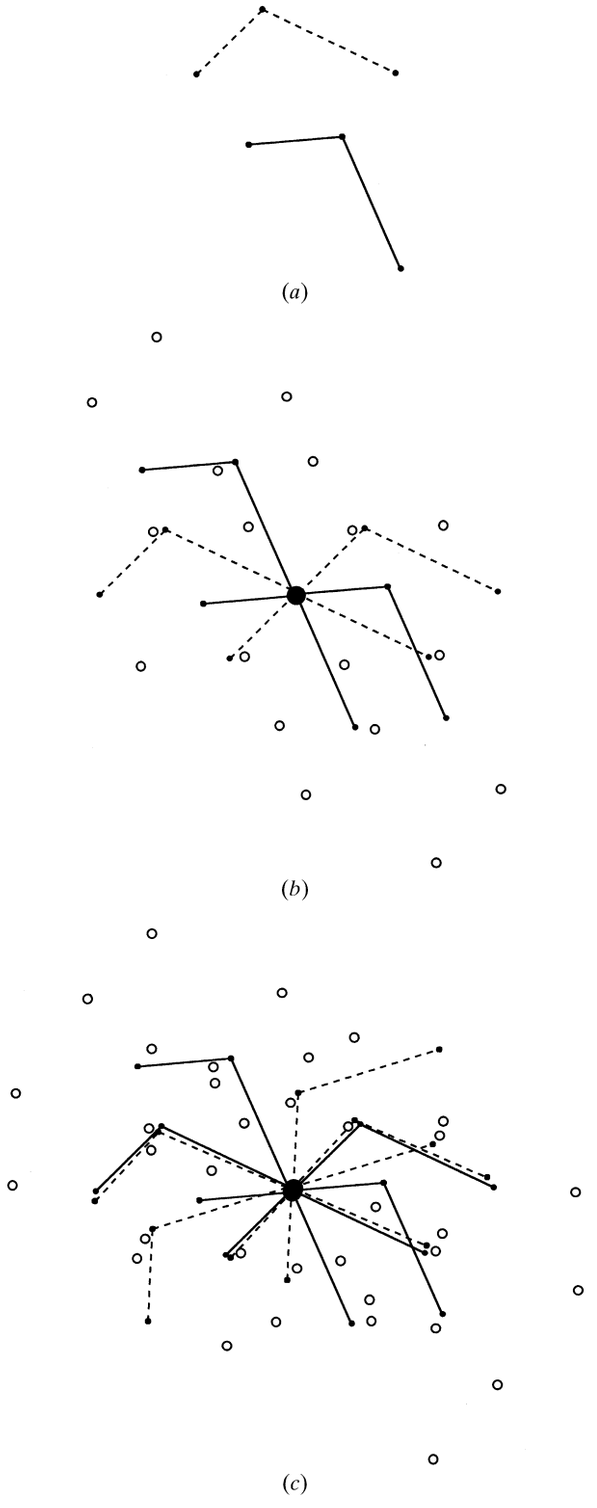

Fig. 13.1.3.1. (*a*) A dimer in which the two subunits are related by arbitrary rotation and translation. One member of the dimer is indicated by dashed lines joining its atoms. (*b*) The vector set representing the Patterson function of this dimer. The intra-subunit vectors of each subunit are indicated 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.

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 (κ, ψ, φ), 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

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.

using noncrystallographic symmetry constraints, but this is not known to have been done in practice [types (3) and (4)].

13.1.5. The power of noncrystallographic symmetry in structure analysis

13.1.5.1. *Relevant parameters: standard case*

A poorly determined structure, if known at sufficient resolution and accuracy, can be improved by structural refinement of an atomic model to fit the observations. These methods often use existing structural knowledge (of bond lengths and angles, for example) to improve the convergence of the refinement process.

The most important contributions of noncrystallographic symmetry arise before this point of structure determination is reached. In this stage of structural analysis, the distribution of scattering density may be constrained by the requirements of noncrystallographic symmetry. The density may be improved by imposing noncrystallographic symmetry on poorly defined scattering density, and, in favourable cases, cyclical improvement leads to a unique corrected structure. To avoid any confusion with refinement of the atomic structure, this process will be referred to as 'symmetry correction' or 'correction' (Hoppe & Gassmann, 1968).

Noncrystallographic symmetry can also be used to improve the accuracy and convergence of atomic structural refinement by increasing the number of observations to a given resolution (Section 13.1.5.5).

The power of correction methods in improving an unknown structure in the standard case (Section 13.1.2.1) depends on:

(1) the resolution of the analysis, d;

(2) the number, N, of subunits per asymmetric unit;

(3) the volume fraction, NU/V_a, of the asymmetric unit over which the noncrystallographic symmetry operation applies;

(4) whether the density between subunit volumes is constant;

(5) the degree of similarity of the subunits being matched; and

(6) the extent to which the noncrystallographic symmetry operations differ from the crystal symmetry operations.

The first three parameters are expressed in quantitative terms; parameter (4) might be true or false, but more often lies between these; and parameters (5) and (6) are not easily expressed in measurable form.

The resolution d should ideally be matched to the level of similarity of the subunits. The root-mean-square displacement between an atom in one subunit and the rotated and translated position of the corresponding atom from another subunit (or model subunit) provides an order of magnitude for the resolution d which can be used effectively. In many cases, the resolution is worse than this for practical reasons of crystal disorder and data collection.

This limit was encountered by Huber *et al.* (1974) working at 1.9 Å resolution. They found that a model structure with a mean coordinate difference of 1.9 Å was not usable for molecular replacement, while another model agreeing to 0.75 Å gave results which allowed the structure to be refined. This suggests that agreement significantly better than the resolution is required.

13.1.5.2. *Information gain from ideal noncrystallographic symmetry*

Rossmann & Blow (1962) wrote, "The effect of noncrystallographic symmetry . . . results in decreasing the size of the structure to be determined, while the number of observable intensities remains the same. This 'redundancy' in information might be used to help solve a structure." This idea is developed below.

First, it will be shown that (in the absence of noncrystallographic symmetry) there is a constant ratio between the number of independent measurements required to specify the scattering density at a chosen resolution and the volume of the asymmetric unit. Then the effect of noncrystallographic symmetry on this ratio is discussed. The importance of the ratio (volume of symmetry-constrained unit/volume of asymmetric unit) $(= U/V_a)$ is stressed. Another ratio is developed – available no. of measurements/ideally required no. of measurements – and this is referred to as the overdetermination ratio.

Consider a noncentrosymmetric crystal whose asymmetric unit volume is V_a and whose diffraction data have been measured to a resolution d. If the multiplicity of the space group (number of asymmetric units in the primitive unit cell) is Z, the volume of reciprocal space, V^*, per point of the primitive reciprocal lattice is given by

$$V^* = 1/V = 1/ZV_a,$$

where V is the volume of the primitive unit cell.

The number of independent orders of diffraction (the number of independent intensities) within the resolution sphere of radius $1/d$ is given by

$$N_{\rm ref} = \left(\frac{4\pi}{3d^3}\right)\left(\frac{1}{V^*}\right)\left(\frac{1}{2Z}\right) = \frac{2\pi V_a}{3d^3}. \qquad (13.1.5.1)$$

In this formula, Friedel's law is supposed to apply. A set of $2Z$ reflections have identical intensity due to the combined effects of Friedel's law and crystal symmetry. When $N_{\rm ref}$ is expressed in terms of V_a, the multiplicity factor disappears.

To calculate the scattering density over the volume V_a at resolution d, $2N_{\rm ref}$ independent quantities need to be specified (say, the real and imaginary parts of each of the $N_{\rm ref}$ independent structure factors). The required number of measurements is

$$R = 2N_{\rm ref} = (4\pi/3d^3)V_a. \qquad (13.1.5.2)$$

If only the diffracted intensities can be measured, they provide exactly half the $2N_{\rm ref}$ measurements required to calculate the density at resolution d.

In what follows, it is assumed that the required number of measurements, R, to specify the scattering density at the chosen resolution is proportional to the volume over which the density must be specified. This is true when the volume is a crystallographic asymmetric unit [equation (13.1.5.2)], and it agrees with another analysis discussed below. Following this argument, the overdetermination ratio

$$\frac{\text{available no. of measurements}}{\text{ideally required no. of measurements}} = \frac{N_{\rm ref}}{R} = \frac{V_a}{2X}, \qquad (13.1.5.3)$$

where X is the volume whose density is unknown.

Next, consider that ideal noncrystallographic symmetry applies. The crystal asymmetric unit contains N identical subunits and no other scattering matter. Since the symmetry is noncrystallographic, it is never possible to fit the subunit volumes together so as to fill the unit cell exactly. The volume assigned to each subunit, U, has to be less than V_a/N, leaving some parts of the unit cell not assigned to any subunit. In the case of ideal noncrystallographic symmetry, these regions are necessarily empty. In this case $X = U$, which is less than V_a/N, so from equation (13.1.5.3)

$$\text{overdetermination ratio} = (V_a/2U) \; > \; N/2.$$

Even where N is only 2, more intensity data are available than the number of measurements ideally required to specify the electron density at resolution d.

A more sophisticated analysis of the number of variables required to define a structure with noncrystallographic symmetry has been made in terms of sets of orthogonal 'eigendensity

functions', which satisfy the noncrystallographic symmetry (Crowther, 1967). Any structure satisfying the symmetry requirements can be constructed from the appropriate set of eigendensities. Crowther (1969) demonstrated that the number of eigendensities m is approximately $(2N_{\rm ref}U/V_a)$.

The structure is specified by m weights, which are applied to the m allowed eigendensities (which depend only on the symmetry constraints), so the overdetermination ratio

$$\frac{\text{available no. of measurements}}{\text{ideally required no. of measurements}} = \frac{N_{\rm ref}}{m} = \frac{V_a}{2U},$$

the same result as before, showing that the two methods of analysis approximately agree.

13.1.5.3. *Information gain in the non-ideal case*

In the non-ideal case, the definition of volume U assigned to each subunit assumes an important role. It is particularly important that the volumes should not overlap, since this may set up a chain of unrealizable constraints. In imposing noncrystallographic symmetry, the volumes between subunits are often unconstrained, allowing for differences in solvent structure and surface side chains following from their different environments. In addition to the volume U of one subunit, whose structure is to be defined, an additional volume, $V_a - NU$, is left unconstrained. The volume X of unknown electron density is $U + (V_a - NU)$, and, using equation (13.1.5.3), the overdetermination ratio

$$\frac{\text{available no. of measurements}}{\text{ideally required no. of measurements}} = \frac{V_a}{2[V_a - (N-1)U]}. \quad (13.1.5.4)$$

If $N = 2$, U must be less than $V_a/2$, and the overdetermination ratio in equation (13.1.5.4) must be less than 1, so in the non-ideal case there is no chance of convergent correction. This confirms the practical observation that although averaging electron density with $N = 2$ can improve the structure (Matthews *et al.*, 1967), it does not lead to convergent correction (B. W. Matthews, unpublished results). Slowly convergent *ab initio* structure correction was reported at 6.3 Å resolution for $N = 4$ (Argos *et al.*, 1975). In this case, the volume $4U$ of the constrained tetramer was reported to be only about $V_a/2$. Substituting $N = 4$, $U = V_a/8$ in the above expression gives an overdetermination ratio of only 1.6, which was sufficient to allow convergent correction.

An alternative possibility is to constrain the density between subunits to a constant value, even when this may not be precisely correct, in order to improve the convergence of symmetry correction. There is a close analogy to solvent-flattening techniques used in density modification and atomic structural refinement (Schevitz *et al.*, 1981; Wang, 1985). The volume constrained to a constant value is now $V_a - NU$. The volume whose structure is to be determined is only U, and in place of equation (13.1.5.4),

$$\frac{\text{available no. of measurements}}{\text{ideally required no. of measurements}} = \frac{V_a}{2U},$$

as in the ideal case. Such a constraint, while only approximately valid, may allow structure correction to proceed convergently, as found by Rossmann *et al.* (1992). The constraint may be released at a later stage.

This analysis also emphasizes the importance of specifying the size and shape of the subunit volume U as closely as possible (Wilson *et al.*, 1981). Methods of automatic refinement of the chosen volume are available (Rossmann *et al.*, 1992; Abrahams & Leslie, 1996).

13.1.5.4. *Relevant parameters: generalized case*

In the generalized case it is obvious that noncrystallographic symmetry makes more measurements available, since the data from more than one type of crystal are being used. The volume of the subunit U must be less than V_a in each type of crystal. Making a simple assumption of two crystals, each with one subunit in each asymmetric unit, the available number of measurements per volume of the subunit is, in the ideal case,

$$\frac{N_{\rm ref1} + N_{\rm ref2}}{U} = \left(\frac{2\pi}{3d^3}\right)\left(\frac{V_{a1} + V_{a2}}{U}\right) > 2\left(\frac{2\pi}{3d^3}\right).$$

Thus, the overdetermination ratio $= (N_{\rm ref1} + N_{\rm ref2})/(4\pi U/3d^3) > 1$, so even in this case the structure is theoretically overdetermined.

This type of reasoning can be applied to analysis of a crystal which includes a unit of known structure and also another unit whose structure is unknown [type (6), Table 13.1.4.1]. A complex between an enzyme of known structure with an unknown inhibitor provides a familiar example. Note first that the envelope defining the volume U of known structure must be tightly defined, since otherwise unwanted features will be taken over into the unknown structure.

If this known structure of volume U appears once in the asymmetric unit V_a of the partly unknown structure, can noncrystallographic symmetry correction be used to define a unique structure at resolution d?

The unconstrained volume is $V_a - U$. The number of measurements required to define this density at the given resolution is $4\pi(V_a - U)/3d^3$. The partially unknown structure provides $2\pi V_a/3d^3$ measured intensities to this resolution, and specification of the contents of U at resolution d is equivalent to $4\pi U/3d^3$ measurements. Thus the overdetermination ratio is

$$\frac{(2\pi V_a/3d^3) + (4\pi U/3d^3)}{4\pi(V_a - U)/3d^3} = \frac{V_a + 2U}{2(V_a - U)}.$$

The overdetermination ratio is greater than 1 if $U > V_a/4$, that is, if only a quarter of the asymmetric unit represents known structure. Although this relationship applies to an 'ideal' case and is therefore certainly too optimistic, it indicates the remarkable power of the molecular-replacement method. If, for example, half the unit cell is devoted to unknown structure $(U = V_a/2)$, the overdetermination ratio is ideally 2.

13.1.5.5. *Noncrystallographic symmetry in atomic coordinate refinement*

In atomic coordinate refinement, noncrystallographic symmetry again provides a useful increase in the ratio of the number of observed quantities to the number of atomic parameters to be refined. As is discussed by Cruickshank in Chapter 18.5, the application of restraints in refinement (on quantities like bond lengths, bond angles and the elimination of short contacts) is formally equivalent to an increase in the number of observational equations. However, if these restraints are tightly applied, they act more like constraints, and their effect is more like a reduction in the number of parameters to be determined. Meaningful refinement is not possible unless the number of observations exceeds the number of parameters, and in practice it usually needs to do so by a factor of 2 or so. If noncrystallographic symmetry is imposed, the number of observations required to define the structure is reduced, because the volume of unknown structure is reduced.

Noncrystallographic symmetry can thus provide a crucial advantage in leading to unambiguous interpretation of structure at relatively poor resolution (say, 3.0 to 3.8 Å), where the ratio of

refined parameters to the number of observations is marginal. Consider two crystals of the same material, one of which has one subunit per asymmetric unit and the asymmetric volume is V_1. The other has N subunits in an asymmetric unit of volume V_N. To the same resolution, the available number of observed reflections is increased in the ratio V_N/V_1, in order to obtain the same number of parameters if noncrystallographic symmetry is imposed.

What effect do these N subunits have on the precision of the final coordinates? The crystal allows the determination of N sets of atomic coordinates. If the errors were independent of each other, the precision of the mean value of each coordinate could be improved in the ratio $N^{-1/2}$ (compared to a well refined V_1 structure).

This improvement will be lost when constraints are applied to the mean coordinates (to make them conform to given bond lengths and angles, for example). If this is done, the errors are no longer independent, and the increase of precision will be less.

Cruickshank (1999 and Chapter 18.5) shows that at high resolution (examples at 0.94 and 1.0 Å) and for atoms of low B factor (less than say 10 Å^2), restraints make little difference to the precision of refinement. Under these conditions, N independent subunits in the asymmetric unit might improve the precision of the mean coordinates by a factor approaching $N^{-1/2}$. But at such good resolution, it is very possible that the differences between the calculated subunit conformations are not due to error, but reflect real structural differences. If so, the precision of the mean coordinates is less significant.

At less high resolution (example given at 1.7 Å), Cruickshank has shown that the precision of unrestrained refinement is significantly worse than the precision of the restraints. In this case, imposing noncrystallographic symmetry on the structure should provide some improvement. But because the coordinate errors then cease to be independent, the improvement in the mean coordinates would be less.

buffer and salt conditions. These variations can be exploited in a systematic fashion for phasing by electron-density averaging, so long as (1) the shrinkage relationships among the different crystals are not merely isotropic and (2) the boundaries and NCS parameters among related segments can be determined. Perutz (Perutz, 1946; Bragg & Perutz, 1952) recognized the potential utility of such shrinkage stages for crystallographic phasing in studies of haemoglobin crystals with varying degrees of hydration.

Recent examples of structure solutions involving multidomain and multiple-crystal-form averaging include studies of HIV reverse transcriptase (RT) (Ren *et al.*, 1995; Ding *et al.*, 1995). Studies of HIV RT by Stuart and coworkers involved multidomain and multiple-crystal-form averaging using different soaking solutions (Esnouf *et al.*, 1995; Ren *et al.*, 1995), in some cases with dramatically improved diffraction resolution. Arnold and co-workers have applied multidomain and multiple-crystal-form averaging to studies of HIV RT, including a systematic application of averaging electron density between 'frozen' and 'unfrozen' crystal forms (Ding *et al.*, 1995; Das *et al.*, 1996). Tong *et al.* (1997) recently described electron-density averaging among multiple closely related crystal forms of the human cytomegalovirus protease that were obtained by treatment of the crystals with different soaking buffers containing differing levels of precipitants, such as salt and polyethylene glycol.

13.4.14. Programs

This review hopefully covers most aspects encountered when employing electron-density averaging, yet the authors have drawn liberally from their own experience. There are now a large number of averaging programs and procedures available, some more suitable for structure determinations of proteins with low NCS redundancy and improper relationships (Jones, 1992) and others particularly suitable for high NCS redundancy, such as is encountered in the study of icosahedral viruses. For large structures, phase determination can be a very time-consuming computer operation. Therefore, attempts have been made to parallelize some programs (Cornea-Hasegan *et al.*, 1995), although this may lead to difficulties in exporting the programs to new and different computers.

Recently described program packages for symmetry averaging have been successfully applied to a number of cases. General program systems for averaging that are well suited to cases with high NCS include *ENVelope* (Rossmann *et al.*, 1992) and *GAP* (Jonathan Grimes and David Stuart, unpublished results); these same packages have also been used for multiple-crystal-form averaging and problems with low symmetry. A number of the program packages have been conveniently integrated with interactive computer-graphics programs such as *O* (Jones *et al.*, 1991) and most permit molecular-envelope definition by a number of possible approaches. *RAVE* and *MAVE* (Kleywegt & Jones, 1994), programs for graphics-assisted averaging within and between crystal forms, also come with an array of tools for flexible map handling and envelope definition (Kleywegt & Jones, 1996). The program systems *DMMULTI* (Cowtan & Main, 1993) and *MAGICSQUASH* (Schuller, 1996), which both derive from the program *SQUASH* (Zhang, 1993), can simultaneously apply real-space (symmetry averaging and solvent levelling with or without histogram matching) and reciprocal-space (phase refinement by the Sayre equation) constraints for phase improvement and extension. The advantage of adding phasing by the Sayre equation is greater at higher resolution, but appears to be significant in some cases, even at relatively low resolution (Cowtan & Main, 1993). *MAGICSQUASH* has been used to determine a number of structures which required multiple-domain and multiple-crystal-form averaging (Schuller, 1996). The *DEMON/ANGEL* package allows noncrystallographic averaging among multiple crystal forms together with solvent flattening and histogram matching (Vellieux *et al.*, 1995). Other versatile programs for electron-density averaging include *AVGSYS* (Bolin *et al.*, 1993) and *PHASES* (Furey & Swaminathan, 1990, 1997), both of which have features for facilitating definition and refinement of NCS parameters.

Acknowledgements

We are most grateful to Sharon Wilder and Cheryl Towell for extensive help in creating this manuscript. We are also grateful for decades of financial support by the National Science Foundation and the National Institutes of Health during the development of the techniques reported here.

References

13.1

Abrahams, J. P. & Leslie, A. G. W. (1996). *Methods used in the structure determination of bovine mitochondrial F_1 ATPase. Acta Cryst.* D**52**, 30–42.

Argos, P., Ford, G. C. & Rossmann, M. G. (1975). *An application of the molecular replacement technique in direct space to a known protein structure. Acta Cryst.* A**31**, 499–506.

Blow, D. M., Rossmann, M. G. & Jeffery, B. A. (1964). *The arrangement of α-chymotrypsin molecules in the monoclinic crystal form. J. Mol. Biol.* **8**, 65–78.

Brünger, A. T. (1997). *Patterson correlation searches and refinement. Methods Enzymol.* **276**, 558–580.

Crowther, R. A. (1967). *A linear analysis of the non-crystallographic symmetry problem. Acta Cryst.* **22**, 758–764.

Crowther, R. A. (1969). *The use of non-crystallographic symmetry for phase determination. Acta Cryst.* B**25**, 2571–2580.

Crowther, R. A. & Blow, D. M. (1967). *A method of positioning a known molecule in an unknown crystal structure. Acta Cryst.* **23**, 544–548.

Cruickshank, D. W. J. (1999). *Remarks about protein structure precision. Acta Cryst.* D**55**, 583–601; erratum D**55**, 1108.

Dodson, E., Gover, S. & Wolf, W. (1992). Editors. *Proceedings of the CCP4 study weekend. Molecular replacement.* Warrington: Daresbury Laboratory.

Fehlhammer, H. & Bode, W. (1975). *The refined crystal structure of bovine β-trypsin at 1.8 Å resolution. I. Crystallization, data collection and application of Patterson search techniques. J. Mol. Biol.* **98**, 683–692.

Hoppe, W. (1957). *Die 'Faltmolekülmethode' – eine neue Methode zur Bestimmung der Kristallstruktur bei ganz oder teilweise bekannter Molekülstruktur. Acta Cryst.* **10**, 750–751.

Hoppe, W. & Gassmann, J. (1968). *Phase correction, a new method to solve partially known structures. Acta Cryst.* B**24**, 97–107.

Huber, R. (1965). *Die automatisierte Faltmolekülmethode. Acta Cryst.* **19**, 353–356.

Huber, R., Kukla, D., Bode, W., Schwager, P., Bartels, K., Deisenhofer, J. & Steigemann, W. (1974). *Structure of the complex formed by bovine trypsin and bovine pancreatic trypsin inhibitor. II. Crystallographic refinement at 1.9 Å resolution. J. Mol. Biol.* **89**, 73–101.

13.1 (*cont.*)

Jack, A. (1973). *Direct determination of X-ray phases for tobacco mosaic virus protein using non-crystallographic symmetry. Acta Cryst.* A**29**, 545–554.

Kayushina, R. L. & Vainshtein, B. K. (1965). *Rentgenografiye opredepenie strukturi L-prolina. Kristallografiya*, **10**, 833–844.

Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). *Rapid automated molecular replacement by evolutionary search. Acta Cryst.* D**55**, 484–491.

Lattman, E. E. & Love, W. E. (1970). *A rotational search procedure for detecting a known molecule in a crystal. Acta Cryst.* B**26**, 1854–1857.

Lawrence, M. C. (1991). *The application of the molecular replacement method to the de novo determination of protein structure. Q. Rev. Biophys.* **24**, 399–424.

McKenna, R., Xia, D., Willingmann, P., Ilag, L. L. & Rossmann, M. G. (1992). *Structure determination of the bacteriophage φX174. Acta Cryst.* B**48**, 499–511.

Matthews, B. W., Sigler, P. B., Henderson, R. & Blow, D. M. (1967). *Three-dimensional structure of tosyl-α-chymotrypsin. Nature (London)*, **214**, 652–656.

Milledge, H. J. (1962). *The automatic selection of molecular-crystal structure by combining stereochemical criteria and high-speed computing. Proc. R. Soc. London Ser. A*, **267**, 566–589.

Nordman, C. E. & Nakatsu, K. (1963). *Interpretation of the Patterson function of crystals containing a known molecular fragment. The structure of an Alstonia alkaloid. J. Am. Chem. Soc.* **85**, 353.

Prothero, J. W. & Rossmann, M. G. (1964). *The relative orientation of molecules of crystallized human and horse oxyhaemoglobin. Acta Cryst.* **17**, 768–769.

Rossmann, M. G. (1990). *The molecular replacement method. Acta Cryst.* A**46**, 73–82.

Rossmann, M. G. (1995). *Ab initio phase determination and phase extension using non-crystallographic symmetry. Curr. Opin. Struct. Biol.* **5**, 650–655.

Rossmann, M. G. & Arnold, E. (2001). *Patterson and molecular-replacement techniques.* In *International tables for crystallography*, Vol. B. *Reciprocal space*, edited by U. Shmueli, pp. 235–263. Dordrecht: Kluwer Academic Publishers.

Rossmann, M. G. & Blow, D. M. (1962). *The detection of sub-units within the crystallographic asymmetric unit. Acta Cryst.* **15**, 24–31.

Rossmann, M. G. & Blow, D. M. (1963). *Determination of phases by the conditions of non-crystallographic symmetry. Acta Cryst.* **16**, 39–45.

Rossmann, M. G., Blow, D. M., Harding, M. M. & Coller, E. (1964). *The relative positions of independent molecules within the same asymmetric unit. Acta Cryst.* **17**, 338–342.

Rossmann, M. G., McKenna, R., Tong, L., Xia, D., Dai, J.-B., Wu, H., Choi, H.-K. & Lynch, R. E. (1992). *Molecular replacement real-space averaging. J. Appl. Cryst.* **25**, 166–180.

Schevitz, R. W., Podjarny, A. D., Zwick, M., Hughes, J. J. & Sigler, P. B. (1981). *Improving and extending the phases of medium- and low-resolution macromolecular structure factors by density modification. Acta Cryst.* A**37**, 669–677.

Scouloudi, H. (1969). *X-ray crystallographic studies of seal myoglobin at 6-Å and 5-Å resolution. J. Mol. Biol.* **40**, 353–377.

Sheriff, S., Klei, H. E. & Davis, M. E. (1999). *Implementation of a six-dimensional search using the AMoRe translation function for difficult molecular-replacement problems. J. Appl. Cryst.* **32**, 98–101.

Tollin, P. (1969). *Determination of the orientation and position of the myoglobin molecule in the crystal of seal myoglobin. J. Mol. Biol.* **45**, 481–490.

Tollin, P. & Cochran, W. (1964). *Patterson function interpretation for molecules containing planar groups. Acta Cryst.* **17**, 1322–1324.

Wang, B. C. (1985). *Resolution of phase ambiguity in macromolecular crystallography. Methods Enzymol.* **115**, 90–92.

Wilson, I. A., Skehel, J. J. & Wiley, D. C. (1981). *Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. Nature (London)*, **289**, 366–373.

13.2

Brink, D. M. & Satchler, G. R. (1968). *Angular momentum*, 2nd ed. Oxford University Press.

Burdina, V. I. (1971). *Symmetry of rotation function. Sov. Phys. Crystallogr.* **15**, 545–550.

Carter, C. W. & Sweet, R. M. (1997). *Molecular replacement. Methods Enzymol.* **276**, 558–619.

Crowther, R. A. (1972). In *The molecular replacement method*, edited by M. G. Rossmann, pp. 173–178. New York: Gordon and Breach.

DeLano, W. L. & Brünger, A. T. (1995). *The direct rotation function: rotational Patterson correlation search applied to molecular replacement. Acta Cryst.* D**51**, 740–748.

Dodson, E. J. (1985). In *Proceedings of the Daresbury study weekend. Molecular replacement*, edited by P. A. Machin, pp. 33–45. Warrington: Daresbury Laboratory.

Dodson, E. J., Gover, S. & Wolf, W. (1992). Editors. *Proceedings of the Daresbury study weekend. Molecular replacement*. Warrington: Daresbury Laboratory.

Landau, L. D. & Lifschitz, E. M. (1972). *Théorie quantique relativiste*, pp. 109–196. Moscow: Editions MIR.

Lattman, E. E. (1972). *Optimal sampling of the rotation function. Acta Cryst.* B**28**, 1065–1068.

Machin, P. A. (1985). Editor. *Proceedings of the Daresbury study weekend. Molecular replacement*. Warrington: Daresbury Laboratory.

Moss, D. S. (1985). *The symmetry of the rotation function. Acta Cryst.* A**41**, 470–475.

Navaza, J. (1993). *On the computation of the fast rotation function. Acta Cryst.* D**49**, 588–591.

Nordman, C. E. (1966). *Vector space search and refinement procedures. Trans. Am. Crystallogr. Assoc.* **2**, 29–38.

Rossmann, M. G. (1972). Editor. *The molecular replacement method*. New York: Gordon and Breach.

Rossmann, M. G. & Blow, D. M. (1962). *The detection of sub-units within the crystallographic asymmetric unit. Acta Cryst.* **15**, 24–31.

Rossmann, M. G., Ford, G. C., Watson, H. C. & Banaszak, L. J. (1972). *Molecular symmetry of glyceraldehyde-3-phosphate dehydrogenase. J. Mol. Biol.* **64**, 237–245.

Steigemann, W. (1974). *Die Entwicklung und Anwendung von Rechenverfahren und Regenprogrammen zur Strukturanalyse von Proteinen am beispiel des Trypsin–Trypsininhibitor Komplexes, des Freien und der L-Asparaginase*. PhD thesis, Technische Universität, Munich, Germany.

Tollin, P., Main, P. & Rossmann, M. G. (1966). *The symmetry of the rotation function. Acta Cryst.* **20**, 404–417.

Tong, L. & Rossmann, M. G. (1990). *The locked rotation function. Acta Cryst.* A**46**, 783–792.

Watson, G. N. (1958). *A treatise on the theory of Bessel functions*, 2nd ed. Cambridge University Press.

13.3

Bentley, G. A. & Houdusse, A. (1992). *Some applications of the phased translation function in macromolecular structure determination. Acta Cryst.* A**48**, 312–322.

Blow, D. M., Rossmann, M. G. & Jeffery, B. A. (1964). *The arrangement of α-chymotrypsin molecules in the monoclinic crystal form. J. Mol. Biol.* **8**, 65–78.

Brünger, A. T. (1990). *Extension of molecular replacement: a new search strategy based on Patterson correlation refinement. Acta Cryst.* A**46**, 46–57.