

## 9. BASIC STRUCTURAL FEATURES

Zorky (1995) show that about half the possibilities are realized in symmorphic space groups with centred cells. The situation is set out in Table 9.7.4.1.

Although about half the point groups are not represented in symmorphic space groups with one molecule in the appropriate special position, it is interesting to look for molecules of these symmetries in space groups of higher symmetry. A few are in fact to be found in non-symmorphic space groups, but seven point groups have no established examples.

## 9.7.5. Structural classes

As developed so far (Belsky, Zorkaya & Zorky, 1995), structural classes relate primarily to homomolecular structures – structures in which all molecules are the same. Nevertheless, they are important in a study of the space-group distribution of molecular organic structures, as structures belonging to the same space-group type but to different structural classes are found to have very different frequencies. The germ of the idea is implicit in Kitajgorodskij's subdivision of his four categories by molecular symmetry.

In its general form, the symbol of a structural class has the form

$$SG, Z = n[(x)^a, (y)^b, \dots],$$

where  $SG$  is the standard space-group symbol,  $n$  is the number of molecules in the unit cell,  $x, y, \dots$  are the symbols of the point-group symmetries of the Wyckoff positions occupied, and  $a, b, \dots$  are the numbers of occupied Wyckoff positions of those symmetries. An example will make this clearer. There is a structural class

$$P\bar{4}m2, Z = 32[(mm)^4, m^4, 1].$$

This indicates that the space group is  $P\bar{4}m2$  and that there are 32 molecules in the unit cell occupying four positions of symmetry  $mm$ , four of symmetry  $m$ , and one general position. On consulting the multiplicity of the special positions for this space group in Volume A of *International tables for crystallography* (Hahn, 1995), one finds that the 32 molecule total is accounted for as  $(4 \times 2) + (4 \times 4) + (1 \times 8)$ . If (as is usually the case) the square brackets are unnecessary, they are omitted, as in

$$P\bar{4}n2, Z = 2(222) \text{ and } P4/nnc, Z = 4(4).$$

Occasionally, two distinguishable structural classes will lead to the same symbol. As yet, Belsky, Zorkaya & Zorky prefer to deal with this problem on an *ad hoc* basis, rather than by attempting to devise any general rules.

Belsky, Zorkaya & Zorky divide the structural classes into six groups, in accordance with the number of examples found. The groups are 'anomalous' (up to five examples, 199 structural classes); 'rare' (up to 19 examples, 55 classes); 'small' (up to 49 examples, 24 classes); 'big' (13 classes); 'giant' (eight classes); and 'supergiant' (six classes). The last three are not explicitly defined, but examination of the tables shows that the dividing line between 'big' and 'giant' is about 250, and between 'giant' and 'supergiant' is about 750. All these statistics, of course, are subject to modification as the number of known structures increases.

## 9.7.6. A statistical model

Wilson (1988, 1990, 1991) has discussed the factors that, on statistical analysis, appear to govern the relative frequency of occurrence of the space groups of molecular organic crystals. In

its developed form (Wilson, 1990), the statistical model postulated that, within an arithmetic crystal class (see Chapter 1.4), the number of examples,  $N_{sg}$ , of a space-group type would depend exponentially on the numbers of symmetry elements within the unit cell, thus:

$$N_{sg} = A \exp \left\{ - \sum_j B_j [e_j]_{sg} \right\}. \quad (9.7.6.1)$$

In this equation,  $A$  is a normalizing constant depending on the arithmetic crystal class,  $[e_j]_{sg}$  is the number of symmetry elements of type  $e_j$  within the unit cell in the space group, and  $B_j$  is a parameter depending on the arithmetic crystal class and the symmetry element  $e_j$ ;  $A$  and  $B_j$  are independent of the space group. Empirically,  $B_j$  has a positive sign for the syntropic symmetry elements ( $k$  and  $\bar{k}$ , where  $n = 2, 3, 4, 6$ ) and a negative sign for  $1$  and the antitropic symmetry elements (glide planes and screw axes). Often, however, laws of 'conservation of symmetry elements', of the type

$$[2] + [2_1] = c, \quad (9.7.6.2)$$

or in general

$$\sum_j [e_j] = c, \quad (9.7.6.3)$$

where  $c$  is a constant for the crystal class, eliminate a separate dependence on one or more of the  $e_j$ 's (Wilson, 1990). Often a cohort larger than an arithmetic crystal class can be used, with arithmetic crystal class included as an additional factor.\* With this adjustment, the model can be used for such cohorts as geometric crystal class or even crystal system, giving fits within the usual crystallographic range ( $R_2 \lesssim 0.05$ ), but, for some classes (in particular  $mmm$ ), statistical tests based on the scaled deviance indicated residual systematic error (Wilson, 1980). This was traced (Wilson, 1991) to the failure of an explicit postulate: 'The second possibility is that the distribution is seriously affected by molecular symmetry. Some molecules possess inherent symmetry . . . , and this symmetry could coincide with the corresponding crystallographic symmetry element, again increasing the variance and/or bias of the number of examples per space group. . . The comparative rarity of utilization of molecular symmetry suggests that it can be ignored in an exploratory statistical survey . . .' (abbreviated from Wilson, 1988). This procedure was perhaps reasonable in a first 'exploratory' survey, and in fact agreement in the usual crystallographic range ( $R_2 \lesssim 0.05$ ) was achieved (Wilson, 1988, 1990). However, the scaled deviance indicated that systematic errors still remained (Wilson, 1980), and the discrepancy was traced to the use of molecular symmetry in many structures (Wilson, 1990, 1992). If such structures are eliminated,  $R_2$  falls to trivial values, and the agreement between observed and calculated frequencies becomes too good to be interesting.

## 9.7.7. Molecular packing

## 9.7.7.1. Relation to sphere packing

The effect of molecular symmetry cannot be ignored in overall statistical surveys as well as in structure prediction. However, in most structures, the molecular symmetry is low or it is not used

\* Statistical modelling programs distinguish between variates and factors. The values of variates are ordinary numbers;  $[2]$ ,  $[m]$ , . . . are variates. Factors are qualitative. In the immediate context, 'arithmetic crystal class' is a factor, but other categories, such as metal-organic compound, polypeptide, structural class (Belsky & Zorky, 1977), . . . , could be included if desired. The programs allow appropriately for both variates and factors; see Baker & Nelder (1978, Sections 1.2.1, 8.5.2, 22.1 and 22.2.1).

## 9.7. THE SPACE-GROUP DISTRIBUTION OF MOLECULAR ORGANIC STRUCTURES

Table 9.7.4.1. Occurrence of molecules with specified point group in centred symmorphic and other space groups, based on the statistics by Belsky, Zorkaya &amp; Zorky (1995)

There is no entry in the 'other space group' column if examples are found in the centred symmorphic group.

Point group	Symmorphic space group	Other space group	Frequency
2	<i>C</i> 2	...	18
<i>m</i>	<i>C</i> <i>m</i>	...	6
<i>2/m</i>	<i>C</i> 2/ <i>m</i>	...	20
222	None	<i>C</i> <i>c</i> <i>c</i> <i>a</i>	4
	...	<i>F</i> <i>d</i> <i>d</i>	2
	...	<i>P</i> $\bar{4}$ <i>n</i> 2	2
	...	<i>P</i> 4/ <i>n</i> <i>c</i> <i>c</i>	1
<i>mm</i> 2	<i>F</i> <i>m</i> <i>m</i> 2	<i>I</i> 4 <sub>1</sub> / <i>a</i> <i>c</i> <i>d</i>	3
<i>mmm</i>	None	...	2
	...	<i>P</i> 4 <sub>2</sub> / <i>m</i> <i>m</i> <i>m</i>	6
	...	<i>I</i> <i>m</i> $\bar{3}$	1
4	<i>I</i> 4	...	1
$\bar{4}$	None	<i>P</i> 4/ <i>n</i>	1
	...	<i>P</i> 4 <sub>2</sub> / <i>n</i>	3
	...	<i>I</i> 4 <sub>1</sub> / <i>a</i>	12
	...	<i>P</i> 4 <sub>2</sub> 1/ <i>c</i>	17
	...	$\bar{I}$ 4 <sub>2</sub> <i>d</i>	1
	...	<i>I</i> 4 <sub>1</sub> / <i>a</i> <i>c</i> <i>d</i>	1
4/ <i>m</i>	<i>I</i> 4/ <i>m</i>	...	1
422	None	<i>P</i> 4/ <i>n</i> <i>m</i> <i>c</i>	1
4 <i>mm</i>	None	None	None
$\bar{4}$ 2 <i>m</i>	$\bar{I}$ 4 <sub>2</sub> <i>m</i>	...	3
4/ <i>mmm</i>	<i>I</i> 4/ <i>mmm</i>	...	1
3	<i>R</i> 3	...	8
$\bar{3}$	<i>R</i> $\bar{3}$	...	6
32	None	<i>R</i> $\bar{3}$ <i>c</i>	5
3 <i>m</i>	<i>R</i> 3 <i>m</i>	...	10
$\bar{3}$ <i>m</i>	<i>R</i> $\bar{3}$ <i>m</i>	...	2
6	None	None	None
$\bar{6}$	None	<i>P</i> 6 <sub>3</sub> / <i>m</i>	12
6/ <i>m</i>	None	None	None
622	None	None	None
6 <i>mm</i>	None	None	None
$\bar{6}$ <i>m</i> 2	None	<i>P</i> 6 <sub>3</sub> / <i>m</i> <i>m</i> <i>c</i>	1
6/ <i>mmm</i>	None	None	None
23	None	<i>F</i> 4 <sub>3</sub> <i>c</i>	1
<i>m</i> $\bar{3}$	<i>F</i> <i>m</i> $\bar{3}$	...	2
432	None	None	None
$\bar{4}$ 3 <i>m</i>	$\bar{I}$ 4 <sub>3</sub> <i>m</i>	...	4
<i>m</i> $\bar{3}$ <i>m</i>	<i>F</i> <i>m</i> $\bar{3}$ <i>m</i>	...	9
	<i>I</i> <i>m</i> $\bar{3}$ <i>m</i>	...	2

in the packing. In about 90% of crystalline compounds, the molecules crystallize in low-symmetry space groups, so that a given molecule has a 12-point contact with neighbouring molecules. As 12 corresponds to the number of nearest neighbours in cubic and hexagonal closest packing of spheres, the periodic assembly of most molecular structures can be regarded as the closest packing of distorted spheres, where symmetry ensures the interlocking of complex shapes (Gavez-zotti, 1994).

For the relatively infrequent cases where high molecular symmetry is reflected in high crystal symmetry, the packing of molecules can be derived from the appropriate, though not necessarily the densest, packing of spheres. For example, ten-point, eight-point and six-point molecular contacts can be

achieved, respectively, by tetragonal close packing (*I*4/*mmm*), by *I*-centred cubic packing (*I**m* $\bar{3}$ *m*), and by primitive cubic packing (*P**m* $\bar{3}$ *m*). For a review and some derivations of the densest packing of equal spheres, see Chapter 9.1 and Patterson & Kasper (1959), Coutanceau Clarke (1972), and Smith (1973); and for packing of clusters of unequal spheres, see Williams (1987).

With spheres having infinite point symmetry  $\mathcal{K}_{\infty h}$ , every sphere can be located on syntropic symmetry elements at special positions with high symmetry up to the symmetry of the lattice. The lattice translations, pertinent to the fully symmorphic space group, are then able to generate the entire crystal structure. When spheres are deformed, symmetry is removed and the non-lattice translations involved with antimorphic space groups

## 9. BASIC STRUCTURAL FEATURES

(which must be subgroups of the sphere-packing groups) become necessary to ensure space filling with the repetitive patterns of complex molecular shapes. In a similar manner, other objects with infinite elements of symmetry (*e.g.* rods; Lidin, Jacob & Andersson, 1995) can be subjected to a rigorous analysis of close packing.

### 9.7.7.2. *The hydrogen bond and the definition of the packing units*

The variously shaped molecular packing units of organic crystal structures are not necessarily identical with the individual molecule. The molecule (of a shape defined by chemical bonds on the inside and van der Waals forces on the outside) can be subjected to clustering under formation of intermolecular hydrogen bonds. Although far weaker than the chemical bond, hydrogen bonds are strong enough to alter the shape of the packing units of the crystal structure significantly. This may have far-reaching consequences for the adopted packing and symmetry. An extreme example is represented by the clustering of H<sub>2</sub>O molecules, where two hydrogen bonds and two regular O—H bonds create a 43m point symmetry at each O atom, and a highly symmetrical structure emerges with an infinite bond network, similar to that in quartz, SiO<sub>2</sub>. From the point of view of an individual H<sub>2</sub>O molecule, the structure is very open. In contrast, a pseudo-close-packed structure of crystalline water, assuming an effective H<sub>2</sub>O radius of 1.38 Å, would have specific density of 1.8 g cm<sup>-3</sup>.

Analogous principles apply to organic structures with hydrogen bonding. CH<sub>3</sub>OH, for example, forms hydrogen-bonded zig-zag chains in its crystal structure. Obviously, the shape of the hydrogen-bonded cluster of molecules depends on the number and orientations of the hydrogen bonds relative to the size and shape of the molecule, causing three-dimensional, planar and linear 'polymers', or the formation of dimers and trimers. As in the example of water, this introduces additional symmetry elements and decreases the degree of space filling.

There is a general rule that ensures that this phenomenon is widespread. The principle of maximum hydrogen bonding states that all the H atoms in the active (polar) groups of a molecule are employed in hydrogen-bond formation (Evans, 1964). Therefore, as the O···HO and O···HN hydrogen bonds are both energetic and common, they are also of the greatest importance in this respect. Although most pronounced in smaller molecules, the symmetry-altering influence of hydrogen bonding also applies to relatively large molecules with a lower proportion of hydrogen bonding as, for example, in long-chain carboxylic acids that are linked in pairs. In large molecules with many active groups, however, the hydrogen bonds merely become the new delimiters of the shape of the individual molecule. The perils of the symmetry-statistical treatments of the hydrogen-bonded structures are well recognized and, for some purposes, the strategy adopted is to exclude such systems from the statistical pool (Filippini & Gavezzotti, 1992).

### 9.7.8. *A priori predictions of molecular crystal structures*

As physical properties of a molecular compound are a function of the spatial arrangement of molecules, an important goal of the

structural chemist is to predict the space group and crystal structure from the molecular shape. On the basis of the observation that many structures of organic compounds are formed on the principle of periodic close packing of variously shaped molecules, it seems that such prediction would be a more or less straightforward computational task. However, the task of predicting the crystal structure of a specific molecular solid is complicated owing to the occurrence of hydrogen bonding (Subsection 9.7.7.2) and the widespread phenomenon of polymorphism (Gavezzotti, 1994). With only subtle differences in their Gibbs free energies, the occurrence of the structural modifications can be influenced by various non-equilibrium factors during crystallization.

In spite of the above problems, experience has shown that prediction algorithms can often be used to generate several reasonable structures for any given molecule and that in many cases the correct structure is among them. There are two general strategies that have been adapted for structure prediction. In the first one, developed by Kitajgorodskij (1955, 1973), the molecular shape is physically constructed from models of atoms having van der Waals radii, resulting in the calot model. The physical calot model is then used for an analogue calculation of the space filling using a mechanical instrument that relates the molecules in three-dimensional space so that the projection of one molecule fits into the voids of other molecules. When the unit-cell dimensions are known, the entire crystal structure can be derived in this way. In the second approach, the same yet abstract 'fused sphere model' is analysed for its symmetry by what can be called a 'morphic' (as opposed to metric) transformation by the methods of molecular topology (Mezey, 1993). The abstract topological molecular shapes can in principle be treated more rigorously and are computable into probable crystal structures.

Such *a priori* predictions of molecular structures are still in a relatively early stage of development. Several recent studies are indicative of the current progress in the field. For layered structures, good predictions can be obtained using construction techniques, symmetry probabilities, and potential energy functions (Scaringe, 1991). An algorithm for the generation of crystal structures by the optimization of packing potential energy over several possible space groups has been devised by Gavezzotti (1991, 1994). In a third approach, energy minimization without symmetry constraints is used for determining molecular crystal structures (Gibson & Scheraga, 1995).

In spite of recent progress, the conceptual link between the molecular and crystal structures still relies to a large extent on the chemical intuition of scientists. The space-group statistics have played a critical role, as they provide the researcher with a summary of what happens in nature. It is likely that the prediction process can be enhanced by calculating statistics of the space-group frequencies and symmetry for molecules that are the most closely related to the shape or chemistry of the molecule under study. As such statistical subsets are often significantly different from the overall statistics, they may prove more valuable in the *a priori* prediction of the crystal structure for a specific molecule. The space-group frequencies and symmetry statistics remain one of the important strings in this link.