

## 9. BASIC STRUCTURAL FEATURES

answer the reverse question: which space groups contain special positions of given point group  $\mathcal{G}$ ? Some general points may be noted.

(i) Special positions of symmetry  $\mathcal{G}$  will be found in the symmorphic, but not other, space groups of the geometric class  $\mathcal{G}$ . Thus, for example, there are special positions of symmetry  $mmm$  in  $Pnmm$ ,  $Cmmm$ ,  $Fmmm$ ,  $Immm$ , but not in any other space group in the geometric class  $mmm$ .

(ii) A ‘family tree’ of point groups is given in Fig. 10.3.2 of Volume A of *International Tables for Crystallography* (Hahn, 1995). Special positions of symmetry  $\mathcal{G}$  may be sought in space groups of the geometric classes linked to  $\mathcal{G}$  by a line (possibly zigzag) having a generally upwards direction. Thus, to take the same example, special positions of symmetry  $mmm$  are found in certain space groups of  $4/mmm$  ( $P4/mmm$ ,  $P4/mbm$ ,  $4_2/mmc$ ,  $P4_2/mcm$ ,  $P4_2/mmm$ ,  $I4/mmm$ ,  $I4/mcm$ ), in  $6/mmm$  ( $P6/mmm$ ), in  $m\bar{3}$  ( $Pm\bar{3}$ ,  $Im\bar{3}$ ), and in  $m\bar{3}m$  ( $Pm\bar{3}m$ ,  $Fm\bar{3}m$ ).

(iii) Obviously, the higher up the tree the symmetry  $\mathcal{G}$  is, the fewer will be the space groups in which it can occur – special positions of symmetry  $m\bar{3}m$  can occur only in the three symmorphic space groups of the corresponding geometric class. The lower symmetries ( $2$ ,  $m$ ,  $\bar{1}$ ,  $\bar{3}$ ), with nothing below them but  $1$ , can be traced upwards along many branches, and so can occur in many space groups, but not all are equally favoured. Special positions of symmetry  $2$  can be sought in all higher geometric classes except  $\bar{6}$ ,  $3m$ , and  $\bar{3}$ , but those of symmetry  $3$  could occur only in the classes of the trigonal, hexagonal, and cubic systems. An approximate count\* (Table 9.7.2.1) shows that special positions of symmetry  $2$  occur in 167 space groups, of  $m$  in 99, of  $\bar{1}$  in 38, and of  $3$  in 57. The only other special positions with space-group frequencies of this order are  $2/m$  (39),  $222$  (30), and  $mm$  (57).

## 9.7.3. Empirical space-group frequencies

Empirical space-group frequencies are based on two major collections of structural data for organic substances, in Cambridge and Moscow, respectively.

The Cambridge Structural Database (Allen *et al.*, 1991) contains assignments of space groups for a variety of different types of organic compounds. The file can be computer searched in many ways; it is easy, for example, to trace all structures having a particular space group, or those having a particular space group and a particular number of formula units per unit cell. For the present purpose, a selection has to be made, omitting space groups not substantiated by a full structure determination or dubious because of disorder in the crystal. The packing considerations discussed in previous paragraphs would not apply to crystals in which the intermolecular binding was ionic rather than van der Waals or the like, so that space groups of ionic structures (for example salts of organic acids) are also rejected. Unfortunately, as it is implemented at present (early 1995), it is not possible to search for structures with molecules occupying general positions or specified special positions, so that, in particular, the frequency data of Wilson (1993*d*) are inflated by the inability to distinguish between single occupation of a general position and multiple occupation of special positions.

The file compiled by V. A. Belsky at the L. Ya. Karpov Institute of Physical Chemistry in Moscow is the source of the data used by Belsky, Zorkaya & Zorky (1995). This file differs in objective from the Cambridge file; the latter includes all

\* Such counts are tedious and subject to error, but the table should be correct within a few units.

reasonably established organic structures short of proteins and high polymers, whereas the former concentrates on structures containing only a single type of molecule (‘homomolecular structures’). It thus contains appreciably fewer entries than the Cambridge file, even if structures of the types mentioned in the previous paragraph are excluded from the latter. The Moscow file is, of course, the primary source for the data of Belsky, Zorkaya & Zorky (1995), in which the occupation of general and special positions is explicitly presented.

## 9.7.4. Use of molecular symmetry

It has long been recognized that in many crystal structures molecules with inherent symmetry occupy Wyckoff special positions, so that molecular and crystallographic symmetry elements coincide, but until recently systematic data have been lacking. Now the occurrence of molecules of particular symmetry in structures of various space-group types can be traced in the data of Belsky, Zorkaya & Zorky (1995), and will be discussed briefly.

## 9.7.4.1. Positions with symmetry 1

The empirical results for ‘homomolecular structures’ with one molecule in the general position are given in Table 9.7.1.2. The classification by arithmetic crystal class and degree of symmorphism follows Wilson (1993*d*); the numerical data are taken from Belsky, Zorkaya & Zorky (1995). Space groups symmorphic in the technical sense (Wilson, 1993*d*) are prefixed by an asterisk (\*), and in each arithmetic crystal class the space group most nearly antimorphic is followed by an obelus (†). The number of known structures having precisely one molecule in the general Wyckoff position is given as a superscript in brackets. It will be noticed immediately that structures with space groups ‘fully symmorphic’ or ‘tending to symmorphism’ are extremely rare. Most have no examples; three ( $P4_2$ ,  $P4/n$  and  $P\bar{3}$ ) are credited with a single example each. The frequency of space groups increases rapidly with increasing antimorphism. In the monoclinic system, the ‘fully symmorphic’ space group  $P2/m$  has no examples with one molecule in the general position, the ‘equally balanced’  $P2/c$  has 11 examples, the ‘tending to antimorphism’  $C2/c$  has 587, and the ‘fully antimorphic’  $P2_1/c$  has 5951. Other systems have fewer examples, but the trend is the same; the really popular space groups are the ‘fully antimorphic’ plus  $P1$  and  $P\bar{1}$ .

All space groups, of course, possess general positions of symmetry 1, and the data in Table 9.7.2.1 show that 116 of them exhibit structures of some kind, and that 57 exhibit structures in which one or more general positions are used. 13 space groups ( $P1$ ,  $P2_1$ ,  $Pc$ ,  $Cc$ ,  $P2_12_12_1$ ,  $Pca2_1$ ,  $Pna2_1$ ,  $P4_{1,3}$ ,  $P3_{1,2}$ ,  $P6_{1,5}$ ) have no positions with symmetry higher than 1. These space groups contain no syntropic symmetry elements, and all are relatively popular.

9.7.4.2. Positions with symmetry  $\bar{1}$ 

Many space groups are centrosymmetric (all those in the geometric classes  $\bar{1}$ ,  $2/m$ ,  $mmm$ ,  $4/m$ ,  $4/mmm$ ,  $\bar{3}$ ,  $\bar{3}m$ ,  $6/m$ ,  $6/mmm$ ,  $m\bar{3}$ ,  $m\bar{3}m$ ), but comparatively few of them possess special positions of symmetry  $\bar{1}$ , as the centres of symmetry are often encumbered by other symmetry elements. All centres of symmetry in  $P\bar{1}$ ,  $P2_1/c$  and  $Pbca$  are free, as are some of those in  $P\bar{3}$  and  $R\bar{3}$ . When the encumbrance is an antitropic symmetry element, the special position can still be occupied by a molecule of symmetry  $\bar{1}$  only, but when the encumbrance is syntropic or atropic the position cannot accommodate such a molecule. Table

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

 Table 9.7.2.1. Statistics of the use of Wyckoff positions of specified symmetry  $\mathcal{G}$  in the homomolecular organic crystals, based on the data by Belsky, Zorkaya & Zorky (1995)

$\mathcal{G}$	Space groups with positions of symmetry $\mathcal{G}$	Space groups actually occurring	Space groups using such positions
1	230	116	57
2	167	79	38
$m$	99	42	25
$\bar{1}$	38	28	10
3	57	18	7
4	24	6	4
$\bar{4}$	29	17	8
222	50	15	5
$mm2$	57	18	5
$2/m$	39	21	6
6	5	1	0
$\bar{6}$	8	3	1
32	22	5	1
$3m$	22	8	5
$\bar{3}$	14	7	4
422	9	1	1
$\bar{4}2m, \bar{4}m2$	19	5	3
$4mm$	8	3	0
$4/m$	7	2	1
$mmm$	16	3	2
23	12	3	2
622	2	1	0
$\bar{6}2m, \bar{6}m2$	5	1	1
$6mm$	2	0	0
$6/m$	2	1	0
$\bar{3}m$	8	3	1
$4/mmm$	4	2	1
432	5	0	0
$\bar{4}3m$	6	2	1
$m\bar{3}$	5	2	1
$6/mmm$	1	0	0
$m\bar{3}m$	3	2	2

9.7.2.1 indicates that there are 38 space groups with special positions of symmetry  $\bar{1}$ , that 28 of them have examples of structures of some kind, and that ten have structures in which the centre of symmetry is actually used by a molecule.

The three space groups with no special positions except those of symmetry  $\bar{1}$  are very popular, whether or not the centre of symmetry is actually used by the molecule. The single criterion 'no special positions except possibly free centres of symmetry' thus selects the space groups favoured by structures in which inherent molecular symmetry is not used.

## 9.7.4.3. Other symmetries

Table 9.7.2.1 gives statistics for the number of space groups possessing Wyckoff positions of symmetry  $\mathcal{G}$ , where  $\mathcal{G}$  is one of the 32 point groups, the number exhibiting structures of some kind, and the number in which the special position of symmetry  $\mathcal{G}$  is actually used. It has to be remembered that this table represents the state of knowledge in 1994, that there may be small errors in the counts in the second column, and that new structures will gradually increase the numbers in the third and

fourth columns. Nevertheless, some trends are clear. The arrangement of the point groups is in ascending order of their 'order' (Hahn, 1995, p. 781), and all numbers show a general decrease with increasing order. When molecular symmetry is used, the favourite is the diad axis 2, closely followed by the mirror plane  $m$ , with the centre of symmetry  $\bar{1}$ , the triad axis 3 and the tetrad inversion axis  $\bar{4}$  trailing. It must also be remembered that these data are for numbers of *space groups*, not numbers of *structures*.

## 9.7.4.4. Positions with the full symmetry of the geometric class

The symmorphic space groups are in a one-to-one correspondence with the arithmetic crystal classes, and each has at least one Wyckoff position with the full symmetry of the geometric crystal class. It would thus be possible for each symmorphic space group to accommodate molecules with the full symmetry of the point group corresponding to the geometric crystal class. With the obvious exceptions of  $P1$  and  $P\bar{1}$ , there seem to be no symmorphic space groups with primitive cells and one molecule only in the cell that do so, but the data of Belsky, Zorkaya &

## 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  $l$  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).