

## 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  $1$  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).