1118 *Chem. Mater.* **1994,** *6,* 1118–1127

# Towards a Grammar of Crystal Packing[1]

Carolyn Pratt Brock

*Department of Chemistry, University of Kentucky, Lexington, Kentucky 40506-0055*

Jack D. Dunitz

*Organic Chemistry Laboratory, Swiss Federal Institute of Technology,*
*CH-8092 Zurich, Switzerland*

*Received February 15, 1994. Revised Manuscript Received April 22, 1994*[⊗]

Compilations and analyses of space-group frequencies are reviewed. A new analysis of the distribution of space groups for molecular crystals brings out the role of molecular symmetry in influencing crystal packing. Some space groups occur predominantly, and some almost exclusively, when molecules of appropriate symmetry occupy special positions. The relative importance of various molecular symmetry elements in governing crystal packing is discussed.

## Introduction

A grammar of language: the rules by which words are put together in sentences. A grammar of crystal packing: the rules by which molecules are put together in periodic ordered arrays.[2] The metaphor seems apt, but there are obviously great differences. Most obvious, sentences are built from different words, whereas most crystals are built from identical (or at least isometric) molecules. Yet there are analogies between the study of inscriptions in unknown ancient languages and the patterns found in crystals. The grammar of crystal packing needs to be deciphered from the observed regularities concerning the combined operations of the various symmetry elements that make up the 230 three-dimensional space groups (or 219 if the 11 pairs of enantiomorphic groups[3] are counted only once each). The study of space-group frequencies is a first step in this direction.

One practical application of these rules could be to develop and test strategies for increasing the probability of occurrence of noncentrosymmetric space groups (necessary for nonlinear optical materials)[4] and chiral space groups (necessary if spontaneous resolution is to occur). More generally, any rules found could be useful as guides to the synthesis of designed molecular materials (i.e., for "crystal engineering").

Since the long-range periodicity of crystals results from the net effect of specific short-range interactions among neighboring molecules, packing rules and regularities should be of interest not only to crystallographers but to all scientists concerned with aspects of what has come to be known as supramolecular chemistry. Furthermore, such principles could be useful to scientists determining structures. The structures of mono- and bilayers are difficult to work out because scattering data are limited; any rules of molecular aggregation would be helpful. The same is true of any material that forms crystals large enough for a powder diffraction study but not large enough for a single-crystal structure determination.

We start with an enumeration of the principles of crystal packing and continue with a review of previous studies of space-group frequencies. A more extensive study of the role of molecular (or ionic) symmetry on space-group probabilities is then described, and the results are discussed in light of the principles of crystal packing.

## Principles of Crystal Packing

**Maximize density; minimize free volume.** This is the primary packing rule for molecular crystals; it is Kitaigorodskii's Principle of Close Packing.[5] Void space in crystals is always unfavorable. Melting is nearly always accompanied by a decrease in density. For polymorphic systems this rule should point to the crystal form that is thermodynamically stable at $T = 0$. At higher temperatures the entropy term $TS$ becomes increasingly important; as the temperature is raised, crystals stable at 0 K either melt or transform to another crystal form in which there is greater molecular motion. In either case a positive $T\Delta S$ term offsets a positive $\Delta H$ term.

Pauling and Delbrueck,[6] in an earlier, but less well-known, statement of this principle emphasized complementarity of molecular surfaces (cf., the "bumps and hollows" of Kitaigorodskii[5]). They asserted that molecular complementarity and self-complementarity, rather than molecular identity, is the important principle governing stability of crystals.[7]

---

[⊗] Abstract published in *Advance ACS Abstracts,* August 15, 1994.

(1) Dedicated to the memory of Margaret C. Etter.

(2) Meaning 4 of *grammar* as given by Webster's Ninth New Collegiate Dictionary (Merriam-Webster Inc.: Springfield, MA, 1985: "the principles or rules of an art, science, or technique").

(3) For example, $P3_1$ and $P3_2$; $P4_12_12$ and $P4_32_12$.

(4) For example, Whitesell, J. K.; Davis, R. E.; Wong, M.-S.; Chang, N.-L. *J. Am. Chem. Soc.* **1994,** *116,* 523–527.

(5) Kitaigorodskii, A. I. *Organic Chemical Crystallography;* Consultant's Bureau: New York, 1961. (English translation of the Russian original published by Press of the Academy of Sciences of the USSR, Moscow, 1955).

(6) Pauling, L.; Delbrueck, M. *Science* **1940,** *92,* 77–79. "...in order to achieve the maximum stability, the two molecules must have complementary surfaces, like die and coin, and also a complementary distribution of active groups.... The case might occur in which the two complementary structures happened to be identical; however, in this case also the stability of the complex of two molecules would be due to their complementariness rather than their identity."

(7) A still earlier statement of this principle is to be found in: Lucretius, *De Rerum Natura.* "Things whose textures have such a mutual correspondence, that cavities fit solids, the cavities of the first the solids of the second, the cavities of the second the solids of the first, form the closest union." English translation by: Munro, H. A. J. In *The Stoic and Epicurean Philosophers;* Oates, W. J., Ed.; Random House: New York, 1940.

0897-4756/94/2806-1118$04.50/0

**Satisfy H-bond donors and acceptors and any other special types of intermolecular interactions.** Etter[8] expanded on Donohue's rule that in the solid state all available acidic H atoms participate in hydrogen bonds;[9] Etter pointed out that H-bond acceptors are utilized to the extent that H-bond donors are available[10] and that in a crystal containing multiple donors and acceptors a bond is preferentially formed between the best (i.e., most acidic) donor and the best (i.e., most basic) acceptor.[11] Because the energy of a hydrogen bond is much greater than the energy of a van der Waals interaction, H bonds in molecular crystals are structure determining.

Other kinds of important donor–acceptor interactions[12] include those involving halogen atoms as electron acceptors (electrophiles) and oxygen, nitrogen, sulfur, and selenium atoms as electron donors (nucleophiles).[13] Some of these interactions have energies comparable to those of hydrogen bonds and also have comparably strong directional dependencies.

**Minimize electrostatic energy.** Electrostatic interactions are dominant in ionic crystals, but are also very important in molecular crystals. The magnitudes of any localized charges in the latter are small, but the electrostatic energy is often still large relative to the energies of the van der Waals interactions. To a first approximation electrostatic interactions can be optimized by avoiding like–like interactions in favor of like–unlike interactions, variously depicted as bumps against hollows, donors against acceptors, positive ends of bond dipoles against negative ends. Favorable interactions among overall molecular dipoles are usually much less important than local interactions among bond dipoles. In any case, the dimensions of most molecules are much larger than the shortest distances between molecules so that inferences based on the dipole–dipole approximation are invalid.[14]

## Review of Previous Work on Space-Group Frequencies

The first papers on space-group frequencies were published by Nowacki in 1942[15] and 1943,[16] when the space groups were known for only about 1000 organic crystals and for only about 3000 crystals of all types. Revised lists were published in 1951[17] and 1954.[18] The dominance of the distribution for organic crystals by a few space groups was clear from the beginning; groups $P2_1/c$, $P2_1$, and $P2_12_12_1$[19] accounted for 44% of the organic structures. The only symmetry elements present in these dominant space groups were inversion centers, 2-fold screw axes, and glide planes. Nowacki considered the zigzag chains of molecules produced by the screw and glide operations to be favorable electrostatically.

Kitaigorodskii's first papers on the systematics of crystal packing appeared in Russian in the mid-1940s[20] but did not become widely known until his 1955 book appeared in English translation in 1961.[5] His seminal "Theory of Close Packing of Molecules" was based on the premise that a "bump" in one molecule should adjoin a "hollow" in its neighbor so that the contact between molecules would be maximized and the volume per molecule minimized. His predictions about space-group frequencies, both for symmetric and asymmetric molecules,[21] have proved remarkably accurate.[25] Kitaigorodskii pointed out that location of a molecule on an inversion center imposed no restrictions on the orientation of that molecule with respect to its neighbors, and he predicted that molecules able to conform to inversion symmetry would usually be located on $\bar{1}$ sites. Kitaigorodskii expected, however, that location of molecules on 2-fold rotation axes or mirror planes would be less likely, and that location of molecules on sites of still higher symmetry would be rare.

Nowacki and Kitaigorodskii were clearly aware of each other's work but referenced each other only selectively.[26]

In a comment on the statistics of Nowacki's 1954[18] distribution of space groups, Mackay[27] observed that if $M_t$ was the number of different space groups that occurred at least $t$ times, then $1/M_t$ varied almost linearly with $t$. Mackay used this observation to estimate the number of unobserved space groups and hence the total number of possible space groups. He obtained a value of 220, to be compared with the correct number of $219 = 230 - 11$. Mackay concluded that as the number of unobserved space groups agreed with the statistical estimate, "their absence is due to chance and not to the intrinsic impossibility for physical reasons of the occurrence of certain space groups". This conclusion would appear to be marred, however, by his own observation that when the analysis was applied to organic and inorganic crystals separately, the resulting estimates of the number of space groups were too low, 185 and 209, respectively. The absence of certain space groups

---

(8) Etter, M. C. *J. Phys. Chem.* **1991**, *95*, 4601–4618.

(9) Donohue J. *J. Phys. Chem.* **1952**, *56*, 502–510.

(10) Etter, M. C. *J. Am. Chem. Soc.* **1982**, *104*, 1095–1096.

(11) Etter, M. C. *Acc. Chem. Res.* **1990**, *23*, 120–126.

(12) See: Desiraju, G. R. *Crystal Engineering: The Design of Organic Solids*; Elsevier: Amsterdam, 1989, and references therein.

(13) We reluctantly adopt established practice in using the word "donor" both for proton donors and electron donors, and similarly for "acceptors".

(14) For a refutation of the idea that molecules with a permanent dipole moment are especially likely to crystallize in centrosymmetric space groups see: Whitesell, J. K.; Davis, R. E.; Saunders, L. L.; Wilson, R. J.; Feagins, J. P. *J. Am. Chem. Soc.* **1991**, *113*, 3267–3270.

(15) Nowacki, W. *Helv. Chim. Acta* **1942**, *25*, 863–878.

(16) Nowacki, W. *Helv. Chim. Acta* **1943**, *26*, 459–462.

(17) Nowacki, W. *Helv. Chim. Acta* **1951**, *34*, 1957–1962.

(18) Nowacki, W. In *Crystal Data*; Donnay,J. D. H., Nowacki, W., Eds.; Geological Society of America: Boulder, CO, 1954; pp 85–104.

(19) Space group $P\bar{1}$ is underrepresented in early tabulations because of the difficulty of doing the necessary calculations in a triclinic coordinate system without the aid of an electronic computer.

(20) See the text on p 66 of ref 5 and the list of references at the bottom of that page.

(21) Statements that Kitaigorodskii predicted the most frequent space groups would be those that permit close packing of triaxial ellipsoids appear in a number of papers in the literature (e.g., refs 22 and 41). We cannot find the phrase "close packing of triaxial ellipsoids" in any English-language version of Kitaigorodskii's most cited writings. Professor A. J. C. Wilson has pointed out to us that in an early paper[23] Kitaigorodskii described the molecular shapes of benzene, naphthalene, anthracene, and the like as triaxial ellipsoids and predicted that the space groups most often observed would be just those most favorable for packing of such ellipsoids.[24] On the other hand, most of the drawings in Kitaigorodskii's 1955 book *Organic Chemical Crystallography*[5] show highly asymmetric molecules with partially concave surfaces. In any event, no ellipsoid-shaped molecule can have a "hollow" into which a "bump" could fit, and it is the fitting together of molecular bumps and hollows that Kitaigorodskii emphasized in later writings.

(22) Giacovazzo, C.; Monaco, H. L.; Viterbo, D.; Scordari, F.; Gilli, G.; Zanotti, G.; Catti, M. *Fundamentals of Crystallography*; Oxford University Press: Oxford, 1992.

(23) Kitaigorodskii, A. I. *J. Phys.* (*USSR*) **1945**, *9*, 351–352.

(24) The identification of the plane and space groups most favorable for the packing of ellipses and ellipsoids was pursued by Nowacki: Nowacki, W. *Schweiz. Min. Petr. Mitt.* **1948**, *28*, 502–508 [*Chem. Abstr.* 42:7595]; Matsumoto, T.; Nowacki, W. *Z. Kristallogr.* **1966**, *123*, 401–421. See also: Tanemura, M.; Matsumoto, T. *Z. Kristallogr.* **1992**, *198*, 89–99, and references therein.

(25) Wilson, A. J. C. *Acta Crystallogr.* **1993**, *A49*, 210–212. See also: Wilson, A. J. C. *Acta Chim. Hung.* **1993**, *130*, 183–196.

(26) For a discussion of their interaction, see: Buergi, H.-B.; Dunitz, J. D. *Helv. Chim. Acta* **1993**, *76*, 1115–1166.

(27) Mackay, A. L. *Acta Crystallogr.* **1967**, *22*, 329–330.

for these two classes of compounds might not, therefore, be entirely due to chance but might be an indication that organic and inorganic structures tend to obey different packing rules. This was probably the case for Mackay's sample (based on Nowacki's[18] data), where the inorganic group of structures consisted largely of elements, alloys, and simple inorganic salts with spherically symmetric or tetrahedral structural units. In Nowacki's[18] list, 917 of the 2603 inorganic structures (35%) occur in the highly symmetric cubic space groups.

Bel'skii Tables of space-group frequencies based on larger numbers of structures were published by Belsky and Zorkii in 1970[28] and 1977.[29]

The advent of large structural databases facilitated the investigation of space-group distributions. Statistics for both organic and inorganic structures in the NBS Crystal Identification File[30] were published by Mighell and Rodgers;[31] a revised version of the organic distribution that included additional checks for possible errors in space-group assignments appeared a few years later.[32] The distribution of organic structures is dominated by a few space groups: $P2_1/c$ accounts for 36% of the structures and five groups ($P2_1/c$; $P\bar{1}$, 14%; $P2_12_12_1$, 12%; $P2_1$, 7%; $C2/c$, 7%) account for 75%. The distribution for inorganic structures was found to be much flatter, with no group accounting for more than 9% of the total.

A discussion of how to count the entries in chiral space groups[33] was initiated by Donohue.[34] He proposed that the statistical weight of all such groups should be doubled since the databases were unlikely to list the enantiomeric structures (e.g., of L- and of D-alanine) separately in the unlikely event that both had been determined experimentally. Brock and Dunitz[35] argued that only nonisometric structures should be included in a list of space-group frequencies; no extra weight should be given to structures in chiral groups like $P2_12_12_1$ and the entries in each member of the 11 pairs of enantiomorphic space groups[3] should be combined and their frequencies added together. Srinivasan[36,37] maintains, however, that the assumption of equal probability of occurrence of enantiomeric structures is unwarranted.

In an elegant series of papers, Wilson[38–41] developed mathematical models relating the number and kind of symmetry elements in the space group to the frequency of the space group for organic crystal structures archived in the Cambridge Structural Database (hereafter, the CSD).[42] As expected, the presence of pure rotation axes and, especially, mirror planes is unfavorable. Wilson found the 13 groups with no special positions[43] and the three groups with special positions of $\bar{1}$ symmetry only[44] to be, taken as a set, especially favorable. Beginning with the 1991 paper,[40] he recognized the importance of the number of molecules in the asymmetric unit ($Z' = Z/n$, where $Z$ is the number of molecules in the unit cell and $n$ is the order of the group). In the tabulation of Mighell et al.[32] $Pnma$ ranks seventh, but essentially all the structures have $Z' = {}^1/_2$; that is, the molecules are located on mirror planes. Since the subgroups of $Pnma$ that lack the mirror plane (i.e., $P2_1/c$, $P2_12_12_1$, and $Pna2_1$) all occur frequently, it is not surprising that molecules that conform to mirror symmetry in the solid state often crystallize in $Pnma$. Wilson's 1993 paper[41] discusses the role of the variable $Z'$ in detail. The importance of the variable $Z'$ was also recognized by Scaringe[45] and by Chernikova et al.,[46] who have published a most interesting compilation of space-group frequencies in terms of "structural classes". Members of a given structural class share not only the same space group but also the same distribution of molecules over systems of equivalent positions (orbits). Thus, in an obvious notation, different structural classes of the space group $P2_1/c$ are characterized by $Z = 2(\bar{1})$, $Z = 4(1)$, $Z = 4(\bar{1},\bar{1})$, $Z = 6(1,\bar{1})$, $Z = 8(1,1)$, etc. From this compilation[46] it is clear that certain space groups occur almost exclusively with molecules in special positions.

Baur and Kassner[47] produced a table of space-group frequencies for inorganic crystals from the entries in the Inorganic Crystal Structure Database[48] and the Metals Datafile.[49] While Mighell and Rodgers[31] had found $Fm\bar{3}m$ to be the most common group for these kinds of structures, the newer study found $Pnma$ (8%) and $P2_1/c$ (8%) to be at the top of the list. Groups $Fm\bar{3}m$, $P\bar{1}$, $C2/c$, and $P6_3/mmc$ followed with 4% each; groups $C2/m$, $I4/mmm$, and $Fd\bar{3}m$ (3% each) and $R\bar{3}m$ (2%) rounded out the top 10. The overlap between the lists of most frequent space groups for organic and inorganic structures is notable.

Baur and Kassner[47] also reviewed the literature on errors in space-group assignments. $Cc$ seems to be the group most often assigned incorrectly; other problem groups are $P1$, $Pc$, and $C2$, all noncentrosymmetric. The 3-fold axes of rhombohedral groups are often missed, as are the $d$ glide planes of $Fdd2$.

Padmaja et al.[50] reported space-group frequencies for the structures in the 1987 version of the CSD; the distribution agreed well with that found by Mighell et al.[32] Padmaja et al. also considered the role of $Z'$; they found $Z' > 1$ structures to be concentrated in the lower symmetry crystal systems, with an especially high incidence in group $P1$. Finally, they compared the space-group frequencies for the macromolecules in the Protein Data Bank (hereafter, the PDB)[51] with those for the

---

[28] Bel'skii, V. K.; Zorkii, P. M. *Kristallografiya* **1970**, *15*, 704–709.

[29] Bel'skii, V. K.; Zorkii, P. M. *Acta Crystallogr.* **1977**, *A33*, 1004–1006.

[30] Distributed by the International Centre for Diffraction Data, 12 Campus Blvd., Newtown Square, PA 19073-3273.

[31] Mighell, A.; Rodgers, J. R. *Acta Crystallogr.* **1980**, *A36*, 321–326.

[32] Mighell, A. D.; Himes, V. L.; Rodgers, J. D. *Acta Crystallogr.* **1983**, *A39*, 737–740.

[33] *I.e.*, space groups having proper symmetry operations (the identity element, rotation axes, and screw axes) only.

[34] Donohue, J. *Acta Crystallogr.* **1985**, *A41*, 203–204.

[35] Brock, C. P.; Dunitz, J. D. *Acta Crystallogr.* **1991**, *A47*, 854.

[36] Srinivasan, R. *Acta Crystallogr.* **1991**, *A47*, 452.

[37] Srinivasan, R. *Acta Crystallogr.* **1992**, *A48*, 917–918.

[38] Wilson, A. J. C. *Acta Crystallogr.* **1988**, *A44*, 715–724.

[39] Wilson, A. J. C. *Acta Crystallogr.* **1990**, *A46*, 742–754.

[40] Wilson, A. J. C. *Z. Kristallogr.* **1991**, *197*, 85–88.

[41] Wilson, A. J. C. *Acta Crystallogr.* **1993**, *A49*, 795–806.

[42] Allen, F. H.; Kennard, O.; Taylor, R. *Acc. Chem. Res.* **1983**, *16*, 146–153.

[43] $P1$, $P2_1$, $Pc$, $Cc$, $P2_12_12_1$, $Pca2_1$, $Pna2_1$, $P3_1/P3_2$, $P4_1/P4_3$, and $P6_1/P6_3$.

[44] Groups $P\bar{1}$, $P2_1/c$, and $Pbca$.

[45] Scaringe, R. P. In: *Electron Crystallography of Organic Molecules*; Fryer,J. R., Dorset, D. L., Eds.; Kluwer: Dordrecht, 1991; Vol. 328, pp 85–113.

[46] Chernikova, N. Yu.; Bel'skii, V. K.; Zorkii, P. M. *J. Struct. Chem.* **1991**, *31*, 661–666 (*Chem. Absstr.* 114:92077d).

[47] Baur, W. H.; Kassner, D. *Acta Crystallogr.* **1992**, *B48*, 356–369.

[48] Bergerhoff, G.; Hundt, R.; Sievers, R.; Brown, I. D. *J. Chem. Inf., Comput. Sci.* **1983**, *23*, 66–69.

[49] Calvert, L. D. *Acta Crystallogr.* **1981**, *A37*, C343–C344.

[50] Padmaja, N.; Ramakumar, S.; Viswamitra, M. A. *Acta Crystallogr.* **1990**, *A46*, 725–730.

[51] Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535–542.

molecules in the CSD that crystallize in chiral space groups. Tetragonal and trigonal groups are much more highly populated in the PDB (13% and 16%) than in the subset of chiral space groups in the CSD (3% and 2%). Groups $C2$ and $C222_1$ are also more frequent in the PDB (13% and 5%) than in the CSD (4% and 1%).

A particularly bothersome problem with the databases of molecular structures is the difficulty of determining whether crystals in a chiral space groups were (1) grown from enantiomerically pure material (as is the case for most natural products), were (2) the product of spontaneous resolution, or were (3) grown from achiral material or from material that racemizes rapidly. Without this information it is essentially impossible to use the databases to predict the probability of spontaneous resolution.[52] With patience and a good library it is usually possible to determine which of the possibilities is correct, but the process is so time-consuming that it can be undertaken only for small subsets of structures.

## New Compilation of Space-Group Frequencies

A commonly stated aim of the several available compilations of space-group frequencies is to provide a basis for a better understanding of molecular packing in crystals, i.e., of the general principles of molecular aggregation. It has been found that neighboring molecules in crystals tend to be related by screw axes, glide planes, inversion centers, and translations rather than by pure rotation axes or mirror planes. There are obvious reasons why these "pure" symmetry elements lead to unfavorable packing arrangements, but nevertheless, a number of common space groups (e.g., $C2/c$ and $Pnma$) do include them. One flaw in most of the previous compilations is the failure to take full account of molecular symmetry and its expression in crystals. The location of symmetric molecules at special positions of the appropriate site symmetry leads to fractional values of $Z'$. We therefore began a more extensive study of the importance of the variable $Z'$. After completing the compilation we discovered that Wilson[41] had produced a very similar set of tables. Following discussions with him we decided to proceed independently. His acceptance criteria for structures differ somewhat from ours, and his goal was a mathematical model.

## Procedure

Searches of the July 1991 (MDAT 910508) version of the CSD[42] included the following acceptance criteria: RFAC 0.010–0.070 (conventional $R$ factor reported and less than 7.0%); COOR $\geq 2$ (coordinates present for at least two atoms); SIGF $\leq 2$ (average standard deviation for a C–C bond $\leq 0.010$ Å); ZVAL $\geq 0$ (number of formula units in the cell reported). These criteria were intended to eliminate structures for which the space group may be in doubt. The SIGF criterion is particularly strong and has undoubtedly led to the systematic exclusion of structures containing atoms with large numbers of electrons. Structures of organometallic compounds are therefore probably underrepresented in our sample. No test was made of the DISO flag because it is sometimes set as a consequence of disorder of a single hydrogen atom. In any event, no structure affected significantly by disorder would have passed the RFAC and SIGF tests, which are substantially stricter than those in any previous compilation.

Although most of the structures in our sample consist of neutral molecules, we see no reason to exclude structures containing ionic moieties. The packing of organic and organometallic ions is governed far more by the intricacies of interaction between local polarizable charge distributions than by the classic patterns of simple inorganic salts. Indeed, only very few of the structures in the CSD include ions with high charge/volume ratios, and to eliminate these would have meant excluding hydrochloride salts of alkaloids and other nitrogen bases, alkali-metal complexes with neutral ionophores, and many other compounds in which the crystal packing is dominated by the same kind of intermolecular interactions as are operative in crystals of electrically neutral molecules.

Duplicate REFCODEs were located and checked by hand. While such duplicate entries usually refer to the same structure, sometimes they do not. One statistically insignificant consequence of this step is that polymorphs with the same space group and the same REFCODE but different cell constants may have been counted incorrectly.

Calculations of $Z'$ values for triclinic and monoclinic structures allowed for possible variation in the number of lattice points per cell resulting from the choice of an unconventional centered cell (e.g., $C\bar{1}$ or $F\bar{1}$ rather than $P\bar{1}$). The relationship between $Z$ and $Z'$ values for rhombohedral structures depended on the axial system (rhombohedral or hexagonal) used as well as on the symmetry.

Of the 31 770 entries (i.e., 35% of the file) that satisfy our criteria, 99% were published after 1969 and 82% after 1979. The results of the survey are presented in Table 1[53] for all space groups occurring at least five times. A table of results for all space groups has been deposited in the supplementary material (see paragraph at end of paper).

A preliminary account of some of this work has been published.[54]

## Results

**Inversion centers are favorable.** Structures in group $P2_1/c$ account for 38% of the 31770 entries—an overwhelming preference! Such structures almost always have three unoccupied *pairs* of inversion centers and usually have four. Structures in group $P\bar{1}$ account for 20% of the entries. Structures in $P\bar{1}$ almost always have seven unoccupied inversion centers and usually have eight. There is clearly no penalty associated with an unoccupied inversion center.

Centrosymmetric[55] molecules, when present, usually occupy one set of inversion centers ($Z' = {}^1/_2$), occasionally two sets ($Z' = 2 \times {}^1/_2 = 1$), or sometimes one set of inversion centers and one set of general positions ($Z' = {}^3/_2$), but there are always at least several sets of unoccupied

---

[52] Brock, C. P.; Schweizer, W. B.; Dunitz, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 9811–9820.

[53] Space group numbers according to Hahn, T. *International Tables for Crystallography*; The International Union of Crystallography: Dordrecht, 1983; Vol. A: Space-Group Symmetry.

[54] Brock, C. P.; Dunitz, J. D. *Mol. Cryst. Liq. Cryst.* **1994**, *242*, 61–69.

[55] I.e., molecules that are intrinsically centrosymmetric and that show only minor deviations from inversion symmetry in the crystal.

**Table 1. Frequencies of Space Groups as a Function of *Z′* for Groups Having at Least Five Entries[a] (*Z′* Values for Darkly Shaded Areas Are Impossible; *Z′* Values for Lightly Shaded Areas Can Occur Only if Two or More Sets of Positions Are Occupied)**

Triclinic System:

| | | | Z'= | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| SPG# | | Ord | 1/2 | 1 | 3/2 | 2 | 3 | 4 | >4 | |
| 1 | P1 | 1 | | 120 | | 93 | 2 | 4 | 1 | 220 |
| 2 | P-1 | 2 | 1295 | 4460 | 19 | 566 | 19 | 12 | 1 | 6372 |
| | | | 1295 | 4580 | 19 | 659 | 21 | 16 | 2 | 6592 |

Monoclinic System:

| | | | Z'= | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SPG# | | Ord | 1/4 | 1/2 | 3/4 | 1 | 3/2 | 2 | 3 | 4 | >4 | |
| 4 | P2(1) | 2 | | | | 1556 | | 256 | 9 | 7 | | 1828 |
| 5 | C2 | 4 | | 50 | | 163 | | 22 | | | | 235 |
| 7 | Pc | 2 | | | | 61 | | 17 | 1 | 1 | 1 | 81 |
| 8 | Cm | 4 | | 12 | | | | | | | | 12 |
| 9 | Cc | 4 | | | | 192 | | 19 | 2 | 3 | | 216 |
| 10 | P2/m | 4 | 6 | | | | | | | | | 6 |
| 11 | P2(1)/m | 4 | | 180 | | 10 | 1 | | | | | 191 |
| 12 | C2/m | 8 | 102 | 38 | | 1 | | | | | | 141 |
| 13 | P2/c | 4 | | 74 | | 49 | 2 | 8 | | | | 133 |
| 14 | P2(1)/c | 4 | | 1897 | | 9702 | 15 | 560 | 21 | 6 | | 12201 |
| 15 | C2/c | 8 | | 1270 | | 1023 | 16 | 32 | 7 | 2 | 1 | 2351 |
| | Other | | | | | 2 | | 1 | | | | 3 |
| | | | 108 | 3521 | | 12759 | 34 | 915 | 40 | 19 | 2 | 17398 |

Orthorhombic System:

| | | | Z'= | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SPG# | | Ord | 1/8 | 1/4 | 1/2 | 3/4 | 1 | 3/2 | 2 | 3 | 4 | >4 | |
| 18 | P2(1)2(1)2 | 4 | | | 41 | | 67 | | 5 | | | | 113 |
| 19 | P2(1)2(1)2(1) | 4 | | | | | 3215 | | 139 | 7 | 5 | | 3366 |
| 20 | C222(1) | 8 | | | 12 | | 15 | | | | | | 27 |
| 29 | Pca2(1) | 4 | | | | | 148 | | 40 | 1 | 1 | | 190 |
| 31 | Pmn2(1) | 4 | | | 15 | | 1 | | | | | | 16 |
| 32 | Pba2 | 4 | | | 1 | | 3 | | 1 | | | | 5 |
| 33 | Pna2(1) | 4 | | | | | 337 | | 23 | 4 | 1 | | 365 |
| 34 | Pnn2 | 4 | | | 2 | | 4 | | 1 | | | | 7 |
| 36 | Cmc2(1) | 8 | | | 37 | | 2 | | | | | | 39 |
| 41 | Aba2 | 8 | | | 18 | | 5 | 1 | | | | | 24 |
| 43 | Fdd2 | 16 | | | 56 | | 48 | | | | | | 104 |
| 45 | Iba2 | 8 | | | 6 | | 12 | | 2 | | 1 | | 21 |
| 52 | Pnna | 8 | | | 16 | | 7 | | | | | | 23 |
| 54 | Pcca | 8 | | | 10 | | 2 | 1 | | | | | 13 |
| 55 | Pbam | 8 | | 2 | 1 | | 3 | | | | | | 6 |
| 56 | Pccn | 8 | | | 49 | | 41 | 1 | 1 | | | | 92 |
| 57 | Pbcm | 8 | | | 35 | | | | | | | | 35 |
| 58 | Pnnm | 8 | | 20 | 1 | 1 | | | | | | | 22 |
| 59 | Pmmn | 8 | | 11 | 1 | | | | | | | | 12 |
| 60 | Pbcn | 8 | | | 192 | | 101 | 2 | 1 | | | | 296 |
| 61 | Pbca | 8 | | | 187 | | 1143 | | 37 | | | 1 | 1368 |
| 62 | Pnma | 8 | | | 440 | | 12 | 1 | | | | 1 | 454 |
| 63 | Cmcm | 16 | | 25 | | | | | | | | | 25 |
| 64 | Cmca | 16 | | 27 | 17 | | | | | | | | 44 |
| 68 | Ccca | 16 | | 8 | 2 | | | 1 | | | | | 11 |
| 70 | Fddd | 32 | | 8 | 5 | | | | | | | | 13 |
| 72 | Ibam | 16 | | 9 | 1 | | | | | | | | 10 |
| 73 | Ibca | 16 | | | 9 | | | | | | | | 9 |
| | Other | | 1 | 16 | 7 | | 12 | | | | | | 36 |
| | | | 1 | 126 | 1161 | 1 | 5178 | 7 | 250 | 12 | 8 | 2 | 6746 |

Tetragonal System:

| | | | Z'= | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SPG# | Space Grp | Ord | 1/8 | 1/4 | 1/2 | 3/4 | 1 | 3/2 | 2 | 3 | 4 | >4 | |
| 76, 78 | P4(1), P4(3) | 4 | | | | | 57 | | 3 | | | | 60 |
| 82 | I-4 | 8 | | 27 | 4 | | 8 | | | | | | 39 |
| 85 | P4/n | 8 | | 22 | 1 | | | | | | | | 23 |
| 86 | P4(2)/n | 8 | | 15 | 20 | | 19 | | | | | | 54 |
| 87 | I4/m | 16 | 9 | 1 | 3 | | | | | | | | 13 |
| 88 | I4(1)/a | 16 | | 27 | 23 | | 40 | | | | | | 90 |
| 92, 96 | P4(1)2(1)2, P4(3)2(1)2 | 8 | | | 70 | | 49 | | 1 | | 1 | | 121 |
| 110 | I4(1)cd | 16 | | | 6 | | 4 | | | | | | 10 |
| 114 | P-42(1)c | 8 | | 10 | 4 | | 12 | | | | | | 26 |
| 121 | I-42m | 16 | 4 | 1 | | | | | | | | | 5 |
| 122 | I-42d | 16 | | 5 | 6 | | 1 | | | | | | 12 |
| 142 | I4(1)/acd | 32 | | 8 | 5 | | 1 | | | | | | 14 |
| | Other | | 14 | 23 | 11 | | 7 | | 3 | | | | 58 |
| | | | 27 | 139 | 153 | | 198 | | 7 | | 1 | | 525 |

Trigonal System:

| | | | Z'= | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SPG# | Space Grp | Ord | 1/12 | 1/6 | 1/3 | 1/2 | 2/3 | 1 | 2 | 4/3 | 3/2 | >2 | |
| 143 | P3 | 3 | | | | | | 4 | 1 | | | | 5 |
| 144, 145 | P3(1), P3(2) | 3 | | | | | | 37 | | | | | 37 |
| 146 | R3 | 3 | | | 13 | | 4 | 6 | 5 | | | | 28 |
| 147 | P-3 | 6 | | 11 | 13 | 3 | | 1 | | | | | 28 |
| 148 | R-3 | 6 | | 57 | 38 | 8 | | 38 | | | | 1 | 142 |
| 152, 154 | P3(1)21, P3(2)21 | 6 | | | | 26 | | 13 | 1 | | | | 40 |
| 155 | R32 | 6 | | 9 | | 2 | | | | | | | 11 |
| 160 | R3m | 6 | | 7 | | | | | | | | | 7 |
| 161 | R3c | 6 | | | 23 | | 2 | 3 | 1 | | | | 29 |
| 163 | P-31c | 12 | | 7 | | | | | | | | | 7 |
| 165 | P-3c1 | 12 | | 2 | 5 | 4 | | | | | 1 | | 12 |
| 167 | R-3c | 12 | | 23 | 3 | 1 | | | | | | | 27 |
| | Other | | 5 | 1 | 2 | 2 | | | | | | | 10 |
| | | | 5 | 117 | 97 | 46 | 6 | 102 | 8 | | 1 | 1 | 383 |

Hexagonal System:

| | | | Z'= | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SPG# | | Ord | 1/12 | 1/6 | 1/4 | 1/3 | 1/2 | 2/3 | 1 | 2 | 4/3 | 3/2 | >2 | |
| 169, 170 | P6(1), P6(5) | 6 | | | | | | | 15 | | | | | 15 |
| 173 | P6(3) | 6 | | | | 7 | | | 2 | | | | | 9 |
| 176 | P6(3)/m | 12 | | 30 | | | 4 | | | | | | | 34 |
| 178, 179 | P6(1)22, P6(5)22 | 12 | | | | | 4 | | 2 | | | | | 6 |
| | Other | | 3 | 7 | 1 | | 2 | | | | | | | 13 |
| | | | 3 | 37 | 1 | 7 | 10 | | 19 | | | | | 77 |

Cubic System:

| | | | Z'= | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SPG# | | Ord | 1/24 | 1/12 | 1/8 | 1/6 | 1/4 | 1/3 | 1/2 | 1 | >1 | |
| 198 | P2(1)3 | 12 | | | | | | 12 | | | | 12 |
| 205 | Pa-3 | 24 | | | | 8 | | 5 | | 1 | | 14 |
| | Other | | 4 | 5 | 3 | 4 | 1 | 6 | | | | 23 |
| | | | 4 | 5 | 3 | 12 | 1 | 23 | | 1 | | 49 |

[a] $Z'$ is $Z/n$, where $Z$ is the number of molecules (or, formula units) in the unit cell and $n$ is the order of the group. A few fractional values of $Z'$ in the high-symmetry groups have been omitted, but there is no structure in the list that corresponds to those $Z'$ values.

inversion centers in the unit cell.[56] In general, a set of molecules in these space groups can always be regarded as a set of centrosymmetric dimers (see, e.g., the discussion of Wilson[25]). In many cases the classification makes chemical sense; that is, there are typical kinds of attractive interactions (e.g., hydrogen bonds) between pairs of molecules related by one particular set of inversion centers. But in other cases this classification seems more formal than chemically significant.

Another example of the strong tendency of molecules to pack across inversion centers is shown by the tendency of many "slightly" chiral molecules to pack across pseudo-inversion centers.[57] Pseudo-inversion centers may also occur in centrosymmetric space groups, e.g., in $P\bar{1}$ with $Z' = 2$.[59] The pseudosymmetry of molecular aggregates (chains, layers, etc.) in crystals has been studied by Zorkii and co-workers[60] for symmetrical molecules that do not display their full point-group symmetry in crystals.

**Mirror planes are always occupied.** Unoccupied mirror planes are especially unfavorable because they require like–like interactions between adjacent molecules. The frequencies for space groups having special positions of symmetry *m* and no other special positions are shown in the top half of Table 2. There is an overwhelming preponderance of structures with $Z' = {}^1/_2$, and checks of the structures having $Z' = 1$ revealed that *all* contain two independent sets of mirror-symmetric molecules. Space

---

[56] There are even a few structures, sometimes termed hypercentric, in which centrosymmetric molecules occupy only general positions (i.e., the asymmetric unit of the centrosymmetric structure is itself centrosymmetric). The hypercentric intensity-statistics distribution was described by: Lipson, H.; Woolfson, M. M. *Acta Crystallogr.* **1952**, *5*, 680–682.

**Table 2. Space Groups Having Special Positions with Mirror and Inversion Symmetry Only[a]**

| | | $Z'$ | | | | |
|---|---|---|---|---|---|---|
| no. | group | $^1/_2$ | 1 | $^3/_2$ | 2 | subgroups (mirrors removed) |
| 36 | $Cmc2_1$ | 37 | 2 | | | $P2_1$; $Cc$; $Pna2_1$; $Pca2_1$ |
| 31 | $Pmn2_1$ | 15 | 1 | | | $P2_1$; $Pc$ |
| 8 | $Cm$ | 12 | | | | $P1$; $Pc$ |
| 26 | $Pmc2_1$ | | 5 | | | $P2_1$; $Pc$ |
| 62 | $Pnma$ | 440 | 12 | 1 | | $P2_12_12_1$; $P2_1/c$; $Pna2_1$ |
| 11 | $P2_1/m$ | 180 | 10 | 1 | | $P2_1$; $P\bar{1}$ |
| 57 | $Pbcm$ | 35 | | | | $P2_12_12$; $P2_1/c$; $P2/c$; $Pca2_1$ |
| | other (5) | 3 | 1 | | 1 | (various) |

[a] Groups 62, 11, and 57 also have $\bar{1}$ positions; group 57 has an additional set of positions of symmetry 2. All groups except 26 have one set of mirror positions; group 26 has 2 sets.

group $Pmc2_1$, which has two independent sets of mirror planes rather than one, is represented *exclusively* by structures with $Z' = 2 \times {}^1/_2 = 1$. The frequencies for space groups having special positions of symmetry both $m$ and $\bar{1}$ are shown in the bottom half of Table 2. Checks reveal that most of the $Z' = 1$ structures contain two independent sets of mirror-symmetric molecules. It is not possible to determine without individual scrutiny whether the molecules in the $Z' = {}^1/_2$ structures in groups $P2_1/m$, $Pbcm$, and $Pnma$ have symmetry $m$ or $\bar{1}$ (or, in $Pbcm$, 2), but given the preponderance of $Z' = {}^1/_2$ structures in the other four groups it seems likely that the molecular symmetry is usually $m$.

**Table 3. Space Groups Having Special Positions of Threefold Rotation Symmetry Only[a]**

| | | $Z'$ | | | | |
|---|---|---|---|---|---|---|
| no. | group | $^1/_3$ | $^2/_3$ | 1 | 2 | subgroups (3-fold axis removed) |
| 161 | $R3c$ | 23 | 2 | 3 | 1 | $Cc$ |
| 146 | $R3$ | 13 | 4 | 6 | 5 | $P1$ |
| 198 | $P2_13$ | 12 | | | | $P2_12_12_1$ |
| 173 | $P6_3$ | 7 | | 2 | | $P2_1$ |
| 143 | $P3$ | | | 4 | 1 | $P1$ |
| | others (3) | 1 | | | | (various) |

[a] Groups 161, 146, and 198 have one set of positions of site symmetry 3; groups 173 and 159 have 2 sets, and group 143 has 3 sets.

Molecules with apparent crystallographic mirror symmetry often show some atomic displacement ellipsoids that are elongated perpendicular to the mirror plane. These displacement parameters may indicate that the apparent mirror symmetry is only statistical in nature (a space or time average over slightly different molecular orientations) or even that it arises from twinning of a lower symmetry structure. Unfortunately, since displacement parameters are still not included in the CSD, it is not possible to detect such behavior without recourse to the original literature.

If all the special positions of a given site symmetry (e.g., $m$) are occupied, the packing is effectively that of a lower symmetry group. The subgroups[61] formed by removing the mirror planes from the groups listed in Table 2 are shown in the last column of that table. At least one of the subgroups of each of the space groups occurs frequently for molecules of arbitrary shape. Specific subgroups, e.g., $P2_1$, occur on several lines because there are a number of ways a mirror plane can be added. Starting with $P2_1$, addition of a mirror plane perpendicular to the screw axes gives $P2_1/m$, while addition of a mirror parallel to the screw axes gives $Pmc2_1$, $Pmn2_1$, or $Cmc2_1$ depending on whether the mirror goes through adjacent axes, halfway between adjacent axes, or diagonally across the $P2_1$ cell. The four possibilities are not equally favorable.

**Groups with 3-, 4-, and 6-fold axes do not usually occur unless the axes are located within molecules of the appropriate symmetry.** The frequencies for space groups having special positions of symmetry 3 ($Z' = {}^1/_3$) and no other special positions with $Z' < {}^1/_3$ are shown in Table 3. There is a strong preference for the 3-fold axis to be occupied. There are only a few structures in space groups in which the highest site symmetry is 4 and in which no $\bar{4}$ axis is present:[62] two in $P4$, two in $I4$, and four in $P4nc$; five of these eight structures have $Z' = {}^1/_4$. There are *no* structures in groups in which the highest site symmetry is 6. It is difficult for the space around a 3-fold rotation axis to be filled efficiently because molecules are seldom wedge-shaped; the space-filling problem is increasingly acute for 4- and 6-fold axes. Exceptions to this generalization are structures in which hydrogen-bonded rings are formed[63] and inclusion complexes in which channels are filled by small solvent molecules. Protein crystals, which are highly hydrated, could be considered as part of this latter group.

**Twofold rotation axes are sometimes occupied and sometimes not.** The frequencies for space groups having only special positions of symmetry 2 are shown in Table 4. Centrosymmetric space groups are not shown because of the problem of knowing whether it is the 2 or the $\bar{1}$ sites that are occupied.[64] Nevertheless, it appears that molecules are located on the 2-fold rotation axes in roughly

---

(57) Although chirality is from one point of view a "yes or no" property, depending on the absence or presence of a symmetry element of the second kind (a rotatory-inversion axis, $\bar{1}$, m, etc.), from another viewpoint one can speak of greater or lesser degrees of chirality of a given geometric figure (or molecule), even if this may not be exactly quantifiable. (For a review of recent work on this problem see: Buda, A. B.; auf der Heyde, T.; Mislow, K *Angew. Chem., Int. Ed. Eng.* **1992**, *31*, 989–1007). For a different approach see: Zabrodsky, H.; Peleg, S.; Avnir, D. *J. Am. Chem. Soc.* **1992**, *114*, 7843–7851; **1993**, *115*, 8278–8289.) According to one possible criterion, a slightly chiral figure will be largely, although not completely, superimposable upon its mirror image. Thus, it makes sense to speak, at least in a qualitative way, of slightly chiral molecules, some examples being 3,3-diethyl-5-methylpiperidin-2,4-one, where the enantiomers are almost completely miscible in the solid state (Oberhaensli, W. E. *Helv. Chim. Acta* **1982**, 65, 924–933), chiral metals formed by *S,S,S,S*-bis-(dimethylethylenedithio) tetrathiafulvalene, a chiral tetramethyl derivative of TMET (Karrer, A.; Wallis, J. D.; Dunitz, J. D.; Hilti, B.; Mayer, C. W.; Buerkle, M.; Pfeiffer, J. *Helv. Chim. Acta* **1987**, *70*, 942-953), (*S*-1-cyanoethyl)bis(dimethylglyoximato)(pyridine)cobalt(III) (Ohashi, Y.; Yanagi, K.; Kurihara, T.; Sasada, Y.; Ohgo, Y. *J. Am. Chem. Soc.* **1982**, *104*, 6353–6359). Such molecules have a tendency to crystallize in pseudocentrosymmetric structures; i.e., inversion of the structure through a point (periodically repeated) leads to near coincidence of most of the atoms with their inverted images. Intensity-statistics tests for such crystals often yield distributions that are indistinguishable from those of truly centrosymmetric crystals. Indeed, for the TMET derivatives mentioned above, if one did not know from other evidence that the crystals *must* be chiral, one would assign them to the centrosymmetric space group $P\bar{1}$ whereas they are all in $P1$ with $Z' > 1$. It is interesting that the common chiral low-symmetry space groups show remarkably high proportions of crystals with $Z' = 2$ or 4 ($P1$, 45%; $P2_1$, 15%, $C2$, 9%). We believe, although we have not made the necessary checks, that many of these crystals will prove to have pseudocentrosymmetric structures.[58]

(58) R. E. Davis and K. A. Wheeler (personal communication) found that 27% of the $Z' > 1$ structures in seven common space groups have approximate $\bar{1}$, 2, $2_1$, or translation operations relating the independent molecules. They found 55 structures in $P2_1$ that are pseudo-$P2_1/c$.

(59) Desiraju, G.; Calabrese, J. C.; Harlow, R. L. *Acta Crystallogr.* **1991**, *B47*, 77–86.

(60) Zorkii, P. M.; Potekhin, K. A.; Dashevskaya, E. A. *Acta Chim. Hung.* **1993**, *130*, 221–233.

(61) Classes I and IIa from Vol. A of the *International Tables for Crystallography*.[53]

(62) $\bar{4}$ axes are often occupied (see groups 82, 86, and 88).

(63) See, e.g.: Brock, C. P.; Duncan, L. D. *Chem. Mater.*, this issue.

(64) From the compilation of structural classes[46] based on a dataset of 15 282 crystal structures, it appears that in space groups for which both $\bar{1}$ and 2 sites are available, the latter are occupied preferentially: e.g., by 294 to 79 in $C2/c$, by 16 to 6 in $Pccn$, and by 86 to 6 in $Pbcn$.

**Table 4. Space Groups Having Special Positions of 2-fold Rotation Symmetry Only**[a]

| no. | group | $Z'$ $^1/_2$ | $Z'$ 1 | $Z'$ $^3/_2$ | $Z'$ ≥2 | subgroups (2-fold axis removed) |
|---|---|---|---|---|---|---|
| 5 | $C2$ | 50 | 163 | | 22 | $P1$; $P2_1$ |
| 92<br>96 | $P4_12_12$<br>$P4_32_12$ | 70 | 49 | | 2 | $P4_1$; $P2_12_12_1$ |
| 18 | $P2_12_12$ | 41 | 67 | | 5 | $P2_1$ |
| 43 | $Fdd2$ | 56 | 48 | | | $Cc$ |
| 152<br>154 | $P3_121$<br>$P3_221$ | 26 | 13 | | 1 | $P3_1$ |
| 20 | $C222_1$ | 12 | 15 | | | $P2_1$; $P2_12_12_1$ |
| 41 | $Aba2$ | 18 | 5 | 1 | | $Cc$; $Pc$; $Pca2_1$; $Pna2_1$ |
| 45 | $Iba2$ | 6 | 12 | | 3 | $Cc$; $Pca2_1$ |
| 110 | $I4_1cd$ | 6 | 4 | | | $P4_1$ |
| 178<br>179 | $P6_122$<br>$P6_522$ | 4 | 2 | | | $P6_1$ |
| 34 | $Pnn2$ | 2 | 4 | | 1 | $Pc$ |
| 32 | $Pba2$ | 1 | 3 | | 1 | $Pc$ |
| | (others) | 9 | 9 | | | (various) |

[a] Groups 92/96, 43, 41, and 110 have one set of positions of site symmetry 2; all the others have two sets.

**Table 5. Space Groups Having No Special Positions**

| group no. | space group | deg of freedom | $Z' = 1$ | $Z' > 1$ | fraction $Z' > 1$ |
|---|---|---|---|---|---|
| 1 | $P1$ | 9 | 120 | 100 | 0.45 |
| 4 | $P2_1$ | 9 | 1556 | 272 | 0.15 |
| 7 | $Pc$ | 8 | 61 | 20 | 0.25 |
| 9 | $Cc$ | 8 | 192 | 24 | 0.11 |
| 19 | $P2_12_12_1$ | 9 | 3215 | 151 | 0.04 |
| 29 | $Pca2_1$ | 8 | 148 | 42 | 0.22 |
| 33 | $Pna2_1$ | 8 | 337 | 28 | 0.08 |
| 76<br>78 | $P4_1$<br>$P4_3$ | 7 | 57 | 3 | 0.05 |
| 144<br>145 | $P3_1$<br>$P3_2$ | 7 | 37 | 0 | 0 |
| 169<br>170 | $P6_1$<br>$P6_5$ | 7 | 15 | 0 | 0 |

**Table 6. Space Groups Having Special Positions of Symmetry $\bar{1}$ Only**

| group no. | space group | deg. of freedom | $Z' < 1$ | $Z' = 1$ | $Z' > 1$ | fraction $Z' > 1$ |
|---|---|---|---|---|---|---|
| 2 | $P\bar{1}$ | 12 | 1295 | 4460 | 617 | 0.10 |
| 14 | $P2_1/c$ | 10 | 1897 | 9702 | 602 | 0.05 |
| 61 | $Pbca$ | 9 | 187 | 1143 | 38 | 0.03 |

half the structures. Twofold rotation axes seem to be neither particularly favorable nor particularly unfavorable as far as crystal packing is concerned. They fulfill the same function in producing dimers from pairs of homochiral molecules as inversion centers do for enantiomeric pairs. In particular, 2-fold axes are effective in bringing homochiral bean-shaped molecules, i.e., molecules with partially concave surfaces, into close contact. Protein molecules are often roughly bean shaped and often occur as dimers across 2-fold axes. Space groups containing 2-fold axes are more prevalent in the PDB than in the CSD since the former contains only structures consisting of homochiral molecules.[50]

**Symmetry operators that include translation: $2_1$ screw axes are more favorable than glide planes, which are comparable to pure translations.** It is generally believed that $2_1$ screw axes are more favorable than glide planes, which are in turn more favorable than pure translations; this belief has been bolstered recently by semiempirical energy calculations.[65] A comparison of the frequencies for groups without any special positions (Table 5) provides several indications that $2_1$ axes are more favorable than pure translations (*t*). Group $P2_1$ ($2_1,t,t$) is ca. 10 times more frequent than group $P1$ ($t,t,t$), and group $P2_12_12_1$ ($2_1,2_1,2_1$) is ca. twice as frequent as group $P2_1$. All three groups have the same number of degrees of freedom, i.e., the same number of choices to specify the size and shape of the unit cell and the orientation and location of the molecule. 2-fold screw axes seem more favorable than higher order axes, but each of the groups $P4_1$, $P3_1$, and $P6_1$ has two fewer degrees of freedom than does group $P2_1$, and it is difficult to allow for that difference.

We find that glide planes (*g*) are comparable to translations in their effect on crystal packing. Here we face the difficulty that translation is an operation of the first kind (preserving sense of chirality) while glide reflection is one of the second kind (reversing sense of chirality) that can only be present in racemic crystals. Ignoring this complication, comparison of $P1$ (chiral; $t,t,t$) with $Pc$ and $Cc$ (racemic; $g,t,t$) suggests the glide operation is more favorable (similar number of entries with one fewer degree of freedom), but comparison of $P2_1$ ($2_1,t,t$) with $Pca2_1$ ($g,g,2_1$) and $Pna2_1$ ($g,g,2_1$) suggests the reverse. The prevalence of the group $P2_1/c$ over $P\bar{1}$ might suggest that glide operations are favorable, but the prevalence of $P2_1/c$ over $Pbca$ would indicate the opposite. The high frequency of $P2_1/c$ could also result from the combination of the $2_1$ axes with the inversion centers, with the glide planes being regarded as dependent rather than independent operations. The large number of degrees of freedom (see Tables 5 and 6) in $P2_1/c$ must be another important factor in its popularity.

Table 5 also shows that the fraction of structures in $P1$ with $Z' > 2$ is exceptionally high.[66] The fractions for $Pc$ and $Pca2_1$[68] are also notable. The presence of two molecules in the asymmetric unit may compensate for the less favorable translations in $P1$ and $Pc$. The tendency towards pseudo-centrosymmetric structures in $P1$ has already been mentioned. The requirement of two perpendicular translations in $Pc$ is clearly unfavorable; the same requirement exists in $Cc$, but the centering operation lessens its impact because interactions to next-nearest neighbors are not usually structure-determining in molecular crystals. We do not mean to imply, however, that translations are generally unfavorable; if they were, there would be many more structures with $Z' > 1$. Translation, i.e., periodicity, is a governing principle of the solid state, but it usually occurs in conjunction with some other operation(s) relating adjacent molecules.

***Z* is relatively constant across the crystal systems.** The average value of *Z*, the total number of molecules (or formula units) in the unit cell, is remarkably constant over the seven crystal systems even though the order of the

---

(65) Filippini, G.; Gavezzotti, A. *Acta Crystallogr.* **1992**, *B48*, 230–234.

(66) Some authors have argued that crystals with $Z' > 1$ are "are more ordered" than those with $Z' = 1$ and hence are disfavored entropically; others have regarded an "entropy of mixing" in crystals with $Z' > 1$ as a contribution towards their greater stability. As Pauling and Tolman pointed out in 1925[67] and as we reiterated for the specific case of spontaneous resolution,[52] the entropy of all *ordered* crystals is zero at 0 K—as it must be according to the third law of thermodynamics. As far as we are aware, there is no evidence that the vibrational entropy of a crystal with $Z' > 1$ is necessarily greater or smaller than that of a crystal with $Z' \leq 1$.

(67) Pauling, L.; Tolman, R. C. *J. Am. Chem. Soc.* **1925**, *47*, 2148–2156.

(68) In group $Pca2_1$ the screw axis is located on a glide plane (it is "encumbered" according to Wilson's[41] terminology), while in group $Pna2_1$ it is not. Clearly the latter arrangement is more favorable.

**Table 7. Average Number of Molecules in the Unit Cell**

| system | structures | order | $\langle Z \rangle$ |
|---|---|---|---|
| triclinic | 6592 | 1–2 | 2.0 |
| monoclinic | 17398 | 2–8 | 4.0 |
| orthorhombic | 6746 | 4–16 | 5.2 |
| tetragonal | 525 | 4–32 | 6.0 |
| trigonal | 383 | 3–12 | 2.8 |
| hexagonal | 77 | 6–24 | 3.6 |
| cubic | 49 | 12–192 | 6.7 |
| all | 31770 | | 3.85 |

groups rises dramatically as the lattice symmetry is increased (see Table 7). Space groups with many symmetry operations do not occur unless the molecular symmetry is high and $Z'$ is small, in which case the packing is effectively that of a lower symmetry group. Structures in low-symmetry groups are more likely than most to have more than one molecule in the asymmetric unit.

## Discussion

**Unoccupied Special Positions and Forbidden Space.** Every special position that is not occupied creates a kind of forbidden zone around itself. Around each $\bar{1}$ this is a roughly spherical region of diameter $R$ (roughly the van der Waals diameter, approximately 3–3.5 Å), around each rotation axis (of whatever order) it is an infinite cylinder of diameter $R$, and around each mirror plane it is an infinite sheet of thickness $R$. Thus, unoccupied mirror planes create more empty space than do unoccupied rotation axes, which in turn create more empty space than do unoccupied centers. Hence, it is on the whole most advantageous from the point of view of molecular packing to occupy the mirror planes (if compatible with the molecular symmetry). In the same way, it is more important to occupy rotation axes than centers of inversion.

The contrast between the populations of the chiral groups $P3_121$ ($P3_221$) and $P3_112$ ($P3_212$) in our listing (Table 1) is striking. The former pair has 40 entries, the latter none, although the numbers and types of the symmetry elements are the same in the two pairs. The difference is in the amount of space occupied by the forbidden zones associated with the 2-fold axes. In the former pair of groups the total length of the 2-fold axes within the trigonal cell having $a = b = 1$ is 6, whereas it is $6\sqrt{3}$ in the latter pair of groups. There is almost 75% more forbidden space in the unpopulated space group. Exactly analogous arguments apply to the two space groups $P31m$ and $P3m1$, the amount of forbidden space associated with the mirror planes being greater by a factor of $\sqrt{3}$ in the latter. Here, however, the frequency statistics are not compelling since both space groups are unpopular. Neither appears in our listing, while the score is 2–0 according to less restrictive admission criteria.[50]

In the higher symmetry space groups in general the rotation axes are no longer parallel and become so numerous that a molecule in a general position will have a hard time avoiding all of them. Thus, space groups such as $R32$ occur almost always with molecules of the appropriate symmetry situated on the rotation axes. This may also be a factor in the relative popularity of space groups such as $P3_121$ and $P4_12_12$ in crystalline proteins[50] and oligonucleotides,[69] which often crystallize as dimers.

**Table 8. Occurrence of Noncentrosymmetric and Chiral Space Groups**

| system | % of all structures | % of structures in this system that are noncentrosymmetric | % of structures in this system that are chiral |
|---|---|---|---|
| triclinic | 21 | 3 | 3 |
| monoclinic | 55 | 14 | 12 |
| orthorhombic | 21 | 64 | 52 |
| tetragonal | 2 | 55 | 36 |
| trigonal | 1 | 42 | 32 |
| hexagonal | <1 | 52 | 44 |
| cubic | <1 | 49 | 31 |
| all | (100) | 23 | 19 |

**Rules for crystal packing are the same for highly symmetric molecules and ions as for molecules of arbitrary shape.** Most organic molecules crystallize in low-symmetry space groups and most elemental metals and many inorganic salts crystallize in high-symmetry groups, but this difference is a consequence of the shapes of the packing units and of the presence or absence of strong electrostatic interactions rather than of any fundamental difference in the packing rules followed by organic and inorganic materials. The vast majority of the structures archived in the CSD are of low-symmetry molecules, so the preponderance of low-symmetry space groups is to be expected. On the other hand, it would have been a surprise if buckminsterfullerene, $C_{60}$, had *not* crystallized in a high-symmetry group.[70] The same could be said of adamantane ($P\bar{4}2_1c$),[71] hexamethylenetetramine ($I\bar{4}3m$),[72] cubane ($R\bar{3}$),[73] and dodecahedrane ($Fm\bar{3}$).[74] Even secododecahedrene crystallizes in a disordered cubic structure indistinguishable from that of its parent compound.[75] The failure of benzene, and of many symmetric benzene derivatives, to retain 6-fold rotational symmetry in the crystal may be related to its planarity.[76]

**Inversion centers are especially favorable for crystal packing because they diminish like–like interactions and are uniquely compatible with translation.** The $\bar{1}$ operation is unique in that it changes the direction, but not the orientation, of intermolecular vectors. If the van der Waals surface of a molecule were an ellipsoid, it would unchanged by the inversion operation. The prevalence of small unit cells ($Z \le 4$; $Z' \le 1$) shows the importance of periodicity, which may be modified by screw or glide operations, but is the dominant principle of crystal packing. In very few structures are there more than four orientations of the molecular inertia tensor, which is unchanged by inversion, and in most structures there is only one ($P\bar{1}$, $Z \le 1$) or two ($P2_1$, $Z'=1$ or $P2_1/c$, $Z' \le 1$).

---

(69) Egli, M., personal communication

(70) David, W. I. F.; Ibberson, R. M.; Dennis, T. J. S.; Hare, J. P.; Prassides, K. *Europhys. Lett.* **1992**, *18*, 219–225. Buergi, H.-B.; Blanc, E.; Schwarzenbach, D.; Liu, S.; Lu, Y.-J.; Kappes, M. M.; Ibers, J. A. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 640–642. Buergi, H.-B.; Restori, R.; Schwarzenbach, D. *Acta Crystallogr.* **1993**, *B49*, 832–838 and references therein.

(71) Nordman, C. E.; Schmitkons, D. L. *Acta Crystallogr.* **1965**, *18*, 764–767.

(72) (a) Dickinson, R. G.; Raymond, A. L. *J. Am. Chem. Soc.* **1923**, *45*, 22–29. (b) Becka, L. N.; Cruickshank, D. W. J. *Proc. R. Soc. London, Ser. A* **1963**, *273*, 455–465.

(73) Fleischer, E. B. *J. Am. Chem. Soc.* **1964**, *86*, 3889–3890.

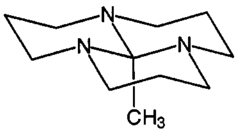
(74) Gallucci, J. C.; Doecke, C. W.; Paquette, L. A. *J. Am. Chem. Soc.* **1986**, *108*, 1343–1344.

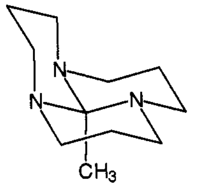
(75) Ermer, O. *Angew. Chem., Int. Ed. Engl.* **1983**, *22*, 251–252.

(76) The structure of benzene ($Pbca$; $Z' = {}^1/_2$; molecular $\bar{1}$ symmetry) is a slightly deformed version of a possible higher symmetry structure in $Pa\bar{3}$ ($Z' = {}^1/_6$) in which the molecular centers are cubic close packed and located on $\bar{3}$ (i.e., $S_6$) axes. Hexaaminobenzene and hexacyanobenzene do crystallize in the cubic structure. (See: Mak, T. C. W.; Zhou, G.-D. *Crystallography in Modern Chemistry*; Wiley-Interscience: New York, 1992; pp 158–162. Druck, U.; Kutoglu, A. *Acta Crystallogr.* **1983**, *C39*, 638–640).

**The fraction of structures in noncentrosymmetric and chiral space groups is considerably greater in the higher symmetry crystal systems.** If the goal is a noncentrosymmetric structure, the primary task is to avoid crystallization in $P\bar{1}$ and $P2_1/c$. While the absolute number of structures in the higher symmetry systems is low, the fraction of structures in noncentrosymmetric and chiral groups rises dramatically with lattice symmetry (see Table 8). This magnification would be seen to be even greater if structures of crystals grown from enantiomerically pure material could be removed from the list. Molecules with imposed 3-fold rotational symmetry account for 133 of the structures in our list; of these, 68 (51%) are in noncentrosymmetric groups and 38 (29%) are in chiral groups. The corresponding percentages for the CSD as a whole are 23% and 19%. If crystals grown from enantiomerically pure material are removed from the sample,[77] the percentages drop to 48% and 24%, but the percentages for the CSD as a whole would drop much more. One might guess that roughly half of the chiral structures in the CSD were grown from enantiomerically pure material.[78]

**High-symmetry groups are more likely in the case of salts and solvates.** By the end of 1992 (MDAT 920701) there were 16 structures in group $P2_13$ archived in the CSD and meeting the criteria given above; 14 of the 16 are salts, and one of the remaining two is a hydrate. The problem of avoiding like–like interactions in the presence of multiple symmetry elements is diminished if more than one kind of molecule or ion is present. Small counterions and solvent molecules not only separate the larger molecular species but also fill space that would otherwise have to be empty. Consider the tricyclic orthoamide





which crystallizes as the trihydrate in the cubic group $Pa\bar{3}$ in the *all-trans* configuration with a 3-fold rotation axis ($Z' = 1/3$), but in $P2_1/c$ with $Z' = 2$ when water is excluded.[79] Furthermore, in the monoclinic structure one of the two independent molecules adopts the less favorable *cis,cis,-trans* configuration, presumably so that space can be filled more effectively.

## Comments and Unanswered Questions

**Structures that are very similar may occur in different space groups.** As far as packing is concerned, a structure that is only pseudo-$P2_1/c$ corresponds essentially to one that is truly $P2_1/c$. Tetracene and pentacene are both triclinic, $P\bar{1}$, $Z' = 1 = 2 \times 1/2$, but they have nearly the same packing as naphthalene and anthracene, $P2_1/c$, $Z' = 1/2$. The observation that the $P\bar{1}$ structures are pseudo-$P2_1/c$ allows recognition of the common thread that unites the packing of these four aromatic hydrocarbons. Consider also chiral structures with $Z' = 2$ that have pseudo-inversion symmetry (see the examples cited in ref 57).

**The list of structures includes a larger proportion of centrosymmetric molecules than expected.** In the space groups $P\bar{1}$, $P2_1/c$, and $Pbca$ there are more than 3400 crystal structures with $Z' = 1/2$ or $3/2$ that must contain centrosymmetric molecules, and there must be another few hundred or so such structures scattered through the remaining space groups. Thus, approximately 12% of the total sample of 31 770 structures contains molecules located on inversion centers. It is inconceivable, however, that anything like one in eight of the several millions of known organic and organometallic molecules is centrosymmetric; almost all must be devoid of any kind of symmetry whatsoever.[80] Our sample is clearly affected by bias although the source of this bias is not so obvious. Crystals containing centrosymmetric molecules are likely to occur in centrosymmetric space groups (indeed practically certain, according to Kitaigorodskii[81]), thus allowing easier solution of the phase problem and generally leading to more accurate molecular parameters than for noncentrosymmetric crystals. The convenience of centrosymmetric space groups was probably an important factor before powerful, automatic direct-methods programs were available but has been much less important recently. The vast majority of structures in our list were published in the last 15 years (99% after 1969 and 82% after 1979).[82] In any event, although the origin of the bias toward centrosymmetric molecules is problematic, its existence seems undeniable.

**Exactly what molecular symmetries are retained in the crystal?** Consider the $Z' = 1/2$ entries for $C2/c$ in Table 1. These could correspond to molecular symmetry 2 or $\bar{1}$, but it is possible only to determine which by only examination of the atomic coordinates in the individual structures. Or consider the $Z' = 1$ entries for $C2/c$, $P\bar{1}$, etc. What fraction of these structures actually contain two independent molecules, each located at a special position? This question has been answered to some extent by Zorkii,[46] and a computer program designed to analyze structures in the CSD in terms of their molecular symmetry has recently been completed.[83]

**Is it true that chiral molecules with potential symmetry 2 are more likely to crystallize in chiral space groups (i.e., to undergo spontaneous resolution) than molecules without a potential 2-fold axis?** This notion has appeared from time to time in the chemical literature.[84] On the other hand, the centrosymmetric group $C2/c$, which includes a position with site symmetry 2, is essentially the centered version of $P2_1/c$ and occurs more

---

(77) A library exercise only possible for small samples.

(78) The frequency of spontaneous resolution has been estimated to be in the range 5–10% (Jacques, J.; Collet, A.; Wilen, S. H. *Enantiomers, Racemates, and Resolutions*; Wiley-Interscience: New York, 1981; p 81), i.e., substantially lower than the frequency of chiral groups in the CSD (19%).

(79) (a) Seiler, P.; Weisman, G. R.; Glendening, E. D.; Weinhold, F.; Johnson, V. B.; Dunitz, J. D. *Angew. Chem., Int. Ed. Engl.* **1987**, *26*, 1175–1177. (b) Seiler, P.; Dunitz, J. D. *Helv. Chim. Acta* **1989**, *72*, 1125–1134.

(80) However, small molecules are more likely to be symmetric than large ones, and also more likely to be subjected to X-ray analysis. Moreover, symmetric molecules gain undue prominence in books and courses because they are useful illustrations for such topics as molecular orbital calculations, spectroscopy, and gas-phase electron diffraction where the power of the method diminishes rapidly with the number of independent atoms. And one should not underrate the aesthetic appeal of symmetric molecules for both chemists and crystallographers.

(81) Kitaigorodskii, A. I. *Sov. Phys. Crystallogr.* **1958**, *3*, 393–394.

(82) Attitudes implanted 20 years ago in young chemists may still influence present-day research supervisors: students may still be directed to submit compounds containing centrosymmetric molecules for X-ray analysis in preference to compounds containing noncentrosymmetric molecules. And there is always the aesthetic aspect.[80]

(83) Howard, J. A. K., Cole, J. C., Allen, F. H., personal communication.

frequently than does the combination of all the chiral groups with special positions of symmetry 2 (see Table 4).

**What is the incidence of spontaneous resolution and cocrystal formation?** Unfortunately neither question can be answered on the basis of the information archived in the various structural databases. It is generally believed that crystallization from a racemic solution is likely to produce a racemic compound D:L but that crystallization from a 1:1 solution of two molecules A and B is unlikely to produce a solid-state compound A:B unless there is a strong, specific interaction (e.g., an acid–base or hydrogen-bonding interaction) between A and B. From the point of view of the phase rule both D + L and A + B are two-component systems (assuming that the enantiomers do not interconvert rapidly). So what is the difference between them? The enantiomers have the same melting points so that the racemate can be isolated if it has a higher melting point than the 1:1 eutectic mixture. The melting point of a 1:1 mixture of A and B will, however, normally be higher than the melting point of the eutectic, and a compound A:B with a melting point slightly above the eutectic temperature may not be discovered. But symmetry may also play a role. The enantiomerically pure compound must crystallize in a chiral space group with its limited repertoire of symmetry operations, while the racemic compound has many more choices available and can crystallize in one of the popular centrosymmetric space groups. This factor favoring formation of D:L compounds is rarely present in the more general case of two molecules A and B.

**How important are kinetic factors?** This is an extremely important question about which little is known. We note simply that because of the importance of surface effects the order of stability for crystal nuclei of different structures is not necessarily the same as for bulk crystals, and that the rate of nucleation can be *much* lower than the rate of crystal growth.[85] This means that crystallization of polymorphs will sometimes be under kinetic rather than thermodynamic control.

**Acknowledgment.** We thank our institutions for providing access to the Cambridge Structural Database, and we acknowledge very useful discussions with Professor A. J. C. Wilson. We are also grateful to Professor J. A. K. Howard and Mr. J. C. Cole (University of Durham, England) and Dr. F. H. Allen of the Cambridge Crystallographic Data Centre for communicating their complete compilation of molecular point group symmetries in the various space groups, which confirms some of the conclusions of this paper and answers many of the questions that we have left unanswewred. C.P.B. thanks the Kentucky EPSCoR program (NSF Grant EHR-91-08764) and the Center for Computational Sciences for financial support and Laura L. Duncan for help with several of the searches.

**Supplementary Material Available:** Frequencies of space groups as a function of $Z'$ (8 pages). Ordering information is given on any current masthead page.

---

(84) Collet, A.; Brienne, M.-J.; Jacques, J. *Bull. Soc. Chim. Fr.* **1972**, 127-142. Collet, A. In *Problems and Wonders of Chiral Molecules*; Simonyi, M., Ed.; Akademiai Kiado: Budapest, 1990; pp 91–109.

(85) McBride, J. M.; Carter, R. L. *Angew.Chem.* **1991**, *103*, 298–301; *Angew. Chem., Int. Ed. Engl.* **1991**, *30*, 293–296.