The goal of this project was to identify the prevalence of approximate symmetry in organic P1 structures. In the November 2019 version of the Cambridge Structural Database (CSD), there are 2592 organic, P1, \( R \leq 0.050 \) structures; complete, unique entries are available for 1407 \( Z > 1 \) and 1049 \( Z = 1 \) structures. All the \( Z > 1 \) structures can have approximate symmetry; the \( Z = 1 \) structures were scanned to find those composed of molecules or ions that might lie on a special position and those that have two or more large molecules or ions that are very similar. The number of \( Z = 1 \) structures so identified was 285, of which 49 were grouped with the \( Z > 1 \) structures because \( Z_{\text{effective}} > 1 \). The packing in each of the 1407 + 285 = 1692 structures was investigated. The 144 that should almost certainly have been described in a smaller or higher-symmetry unit cell were removed from the list; 120 of the 144 are composed of achiral or racemic material. (About half of the \( Z = 1 \) and 89% of the \( Z > 1 \) structures are composed of enantiopure material.) Approximate periodic symmetry was found in 86% of the 1337 remaining \( Z > 1 \) structures and in 72% of the 211 remaining \( Z = 1 \) structures. About a third of the enantiomerically pure structures mimic inversion symmetry; 38% have approximate rotational symmetry. For the structures of achiral and racemic material, distorted glide or mirror symmetry is more common than is distorted inversion symmetry. Approximate rotational and glide symmetry was found to be periodic in two dimensions considerably more often than in three. In 4% of the structures, different layer types alternate or layers are related by approximate local rotations, as well as by small translations. In 5% of the structures, different parts of the molecule are segregated into two-dimensional regions that have different approximate symmetries. More than a third of the structures that are a distorted version of a higher-symmetry structure were determined at \( T \geq 288 \) K.

1. Introduction

\( P1, Z > 1 \)\(^1 \) structures are always suspected of having more symmetry than was reported; there is always the question of whether the independent molecules\(^2 \) might actually be related by symmetry that had been overlooked. That distrust is rooted in the trove of more than 250 space-group corrections published for \( P1 \) structures (Marsh, 1999, 2005, 2009; Marsh et al., 2002; Clemente, 2003; Marsh & Clemente, 2007; Henling & Marsh, 2014). Recognition of the problem of missed symmetry led to the addition of the ADDSYM routine (Marsh & Spek, 2001) to the widely used program PLATON (Spek, 2020), but structures with too-perfect ‘approximate’ symmetry are still being published.

\(^1\) \( Z \) is the number of formula units in the unit cell. In \( P1, Z \), the number of crystallographically independent formula units, is the same as \( Z \) because there is no space-group symmetry other than translation.

\(^2\) The term molecule is used here to refer to any covalently bonded group of atoms and so includes polyatomic ions.
This survey of $P1, Z = Z' > 1$ structures was begun in connection with the session honoring the memory of Dick Marsh that took place during the 2018 meeting of the American Crystallographic Association. The aim of the project was to determine the prevalence of true approximate symmetry in $P1, Z > 1$ structures. Previous studies of $Z' > 4$ and of layered structures (Brock, 2016, 2020) had identified approximate symmetry in many of them. No previous study of $P1$ structures has reported a careful look at symmetry other than approximate inversion.

The most important finding of this work is that independent molecules in $P1$ structures are usually related by approximate, but not crystallographic,\(^4\) symmetry. It may be difficult to decide how to best describe that approximate symmetry, and it may be periodic in fewer than three dimensions, but it is very often present. Structures in which independent molecular conformations and orientations are both unrelated are much less common. This work also determined that the percentage of $P1, Z > 1$ structures that would almost certainly be described better in a higher-symmetry space group or a smaller unit cell varies strongly with structure type, but is still 8% overall, as was found previously (Marsh, 2009).

Less detailed observations of approximate symmetry have been made previously. Zorky (1996) described hypersymmetry in $Z > 1$ structures, which is a screw relationship of independent molecules around a crystallographic axis \([uvw]\), where \(u, v\) and \(w\) are small integers; Nespolo (2008) recast Zorky’s idea in terms of local and partial symmetry operations. Marsh (1999) noticed that about a third of $P1, Z = 2$ structures of enantiopure material mimic inversion symmetry. That estimate was confirmed by Somov & Chuprunov (2009), who devised an algorithm designed to find Sohncke-group structures\(^4\) that are approximately centro-symmetric. Rekis (2020) described an alternative method for finding such structures. To keep the length of this report reasonable, many details have been placed in the supporting information, in which the sections are numbered as they are below.

2. Methodology

2.1. Database searches

The November 2019 version of the Cambridge Structural Database (CSD; Groom et al., 2016) was searched using the following criteria: space group $P1$, archived coordinates, no errors, $R \leq 0.050$, structure determined from a single crystal, compound marked as organic and non-polymeric. The resulting number of refcodes was 1477 for $Z > 1$ and 1115 for $Z' = 1$. Group $P1$ is very unusual in having more $Z > 1$ than $Z = 1$ structures.

Seventy $Z > 1$ structures were discarded as being incomplete (13), duplicates (53), too inorganic (3) or otherwise problematic (1). Disordered structures were considered carefully but all were retained. If the same structure was determined at more than one temperature the highest-$T$ structure was kept because cooling sometimes results in a symmetry-lowering transition.

The existence of quasiracemates (compounds of different but very similar molecules in which all corresponding stereocenters have opposite configurations; Fredga, 1960; Whitesell et al., 1994; Zhang & Curran, 2005) prompted a check of the $Z = 1$ structures to find those that could be described as having $Z_{\text{effective}} > 1$. That search also revealed inconsistencies in the $Z$ values of solvates. Forty-nine $Z_{\text{eff}} > 1$ structures were identified. They are:

(i) Twenty 1:1 compounds plus two solid solutions of two very closely related but not identical molecules;  
(ii) Fifteen of the 20 ordered compounds are quasiracemates;  
(iii) Eleven solvates formulated as, for example, 2$A\cdot S$ with $Z = 1$, rather than as $A\cdot 2S$ with $Z = 2$ (the latter formulation being much more common);  
(iv) Eight salts like $A^+\cdot A^-\cdot B^-$ that contain an extra neutral molecule and a small counterion;  
(v) Four 2:1 and 1:2 salts of small symmetric counterions like $SO_4^{2-}$ and $[Ca(H_2O)_6]^{2+}$;  
(vi) Four additional structures that are better understood as having $Z > 1$.

Those 49 $Z_{\text{eff}} > 1$ structures, 47 of which have $Z_{\text{eff}} = 2$, were added to the $Z > 1$ list of 1407 (= 1477 − 70) to give a total count of 1456. Approximate symmetry is possible for all of them. In what follows, the category $Z > 1$ should be understood to include the $Z_{\text{eff}} > 1$ structures unless there is a statement that it does not.

Of the remaining 1115 − 49 = 1066 $Z = 1$ structures, 61 were discarded as duplicates, etc., leaving 1005 to be examined. They were scanned to find those having molecules that could either be symmetric or mimic symmetry. In structures having more than one residue\(^6\) with $\geq 7$ non-H atoms, the possible approximate symmetries were required to be compatible (e.g. $\bar{1}$ and $2/m$ but not 2 and m). A very few structures in which the symmetries were incompatible were retained if the different molecules were segregated into different layers. The number of $Z = 1$ structures identified as having possible approximate molecular symmetry was then 236 (23% of the 1005).

The November 2021 version of the CSD was searched to see if any additional information was available about any of the refcodes identified in the 2019 version (e.g. a space-group correction). Nothing of significance was found.

\(^3\)The adjective crystallographic is used here in preference to exact because the only symmetry that can be observed in a diffraction experiment is that of a time and space average.

\(^4\)Sohncke groups are those groups that contain no symmetry operation that inverts stereocenters. Enantiomerically pure material must crystallize in a Sohncke space group but may mimic a non-Sohncke structure.

\(^5\)If only a single molecule type is present, the number of residues is $Z'$. For $Z' = 1$ 1:1 salt or solvate, the number of residues is two.

\(^6\)A residue is a crystallographically independent set of covalently bonded atoms or an ion or included solvent molecule having one non-H atom (e.g. halide, $NH_4^+$ or water).
2.2. Analyses of the structures

Structures were investigated individually and entered into the spreadsheet for either the $Z > 1$ or the $Z = 1$ structures. Those files are included with the supporting information.

Approximate symmetry was identified visually. A sophisticated visualization program like the CCDC’s Mercury (Macrae et al., 2020) was an essential tool; once the right view was found the approximate symmetry became obvious. It proved very useful to hide the H atoms, to color the non-C atoms by element and the C atoms by symmetry equivalence, and to mark the stereocenters. A packing diagram of $3 \times 3 \times 3$ unit cells was displayed and examined from a wide variety of directions in order to spot approximate symmetry and any obvious layers (Fig. 1). Looking at obvious layers proved very useful because approximate spatial relationships are much easier to recognize when looking at a part of the structure that has only 2D periodicity; once the approximate layer symmetry had been found it was quite easy to check for approximate 3D symmetry (see below).

Hydrogen bonds were displayed during the analysis; they were often helpful in identifying layers. The default criteria in Mercury were used except for N–H⋅⋅⋅X ($X = O$, N or S) hydrogen bonds, which were marked for N⋅⋅⋅X distances as long as the sum of the van der Waals radii plus 0.2 Å.

Pairwise molecular overlays were calculated using Mercury for all residues having at least seven non-H atoms (or nine for very rigid residues). The resulting rmsd (r.m.s. deviations) for overlays with no change, with inversion and with flexibility were added to the spreadsheet. Rmsds less than 0.20 Å normally indicate that the two molecules are essentially identical; values up to ca 0.40 Å are seldom important, although they sometimes indicate a difference in the orientation of a small substituent like the terminal non-H atom of an Et, OMe or CH$_2$OH group. For most structures, the value with flexibility is <0.20 Å unless the molecules are diastereomers, or unless the molecules have a type of conformational flexibility not handled by the algorithm (e.g. inversion at an N atom or conformational flexibility of a five- or six-membered ring). Occasionally, the algorithm limitations result in the rmsd for overlay with inversion being significantly smaller than the overlay with flexibility even though the two molecules are the same enantiomer.

The ADDSYM routine of the program PLATON (Spek, 2020; version of 23 March 2018) was run for each structure with both default and, as necessary, increased tolerances (i.e. ADDSYM 2 1 1 1). Recommendations below the 90% level were not entered in the spreadsheet because they were seldom informative. The FIT n m instruction in PLATON was run for molecule pairs if no approximate symmetry could be found; occasionally that routine led to the discovery of approximate symmetry that had been missed.

2.3. Criteria for identifying approximate symmetry

There are no generally accepted criteria for identifying approximate symmetry. PLATON (Spek, 2020) is an excellent tool for spotting symmetry that is crystallographic, or nearly so, but is less reliable for finding approximate symmetry, especially if it is rotational or is confined to layers. Baggio (2019, 2020) reviewed previous approaches and suggested a promising method that could be made quantitative, but his program is still under development. Brock & Taylor (2020) found that setting cutoffs for the identification of approximate translations in organic structures was problematic. Both Somov & Chuprunov (2009) and Rekis (2020) reported methods for finding approximate inversion but the results were rankings rather than binary decisions. Fig. 1 in that 2009 paper shows a continuum of values for their descriptor. In the end, a judgment has to be made as to whether or not there is approximate symmetry.

Figure 1
Views of the structure of BILJIR (Gelbrich et al., 2013), which has approximate layer symmetry c211. (a) Projection along [110] showing the (001) layers. (b) View of the (001) layer marked in (a) as rotated by 90° around the vertical, as well as views of that layer as projected edge on. The approximate symmetry elements of layer group c211 (axes [110] and [1T0]) are shown. Adjacent layers are offset by approximately [110]/4. In this and the following drawings, the C atoms are colored by residue number and the heteroatoms are colored by atom type. When shown, the axes $a$, $b$ and $c$ of the P1 cell are colored red, green and blue, respectively. Approximate symmetry elements are shown in blue.
While writing an algorithm for finding approximate symmetry may be challenging, in most cases spotting it once the right view is found is not. Usually the approximate symmetry is obvious, and it is very often supported by the presence of one or more angles within several degrees of a special value (i.e. 90, 60 or 120°). Almost all cases were clear cut; fewer than 6% of the descriptions of approximate symmetry were marked as being marginal or borderline. The real percentage may be somewhat higher, but not much.

After a large number of structures had been analyzed, it seemed that approximate symmetry was often obvious for deviations of up to 5° from a special angle. Even that criterion sometimes seemed too conservative. In 43 of the structures, approximate symmetry was identified for angles that differ by more than 5° from a special value, with three of the 43 having a deviation in the range 7–10°. Drawings of the 19 with deviations >6° are available with the supporting information.

It turned out to be quite easy to distinguish approximate symmetry from overlooked crystallographic symmetry. PLATON alerts >90% and rmsds <0.20 Å for molecular overlays were always investigated. If the symmetry is only approximate then obvious systematic discrepancies between pairs of atomic positions that should be related could nearly always be seen in at least one view direction. Very subtle differences are usually associated with phase transitions for which the appearance of extra reflections was documented or other physical evidence (e.g. DSC traces, spectra) was given. In the few questionable cases, the original paper and CIF were consulted; if no convincing justification was given for the refinement in P1 the structure was marked as being PSTG (i.e. PseudoSymmetry Too Good).

2.4. Categories

Each structure was categorized. A structure having approximate inversion was listed as distorted P1 if composed of achiral or racemic material, and as a P1 mimic (Fig. 2) otherwise. Structures having approximate glides (but not inversion) were similarly listed as glide or glide mimic (Fig. 3). If additional approximate symmetry, such as a pseudo-translation or a screw axis, was found, plus was added to the designation.

Structures having approximate rotational symmetry only were categorized as 3i, 3, 2i or 2d2i (Fig. 1), or with a subcategory of 2 (P2/p211 or p112, with the latter reserved for...
Structures having pseudotranslations only were marked as translation. The label alt layers was used for the 51 structures in which there are obvious layers that differ in composition, approximate symmetry, packing and/or orientation (Figs. 4 and 5; see also Brock, 2020). The term hybrid packing was applied to the 78 structures in which different parts of the molecule (cyclo-diaspartic acid) pack. The –CH₂COOH groups form a layer with approximate symmetry p\(^1\); the central C\(_6\)H\(_4\)N\(_2\)O\(_2\) rings (two stereocenters) are in a layer of approximate symmetry c211.

layered structures with the 2 perpendicular to a 2D layer). Structures having pseudotranslations only were marked as translation. The label alt layers was used for the 51 structures in which there are obvious layers that differ in composition, approximate symmetry, packing and/or orientation (Figs. 4 and 5; see also Brock, 2020). The term hybrid packing was applied to the 78 structures in which different parts of the molecule pack with different approximate symmetry (Fig. 6). If no approximate symmetry could be found, the assigned category was none found.

Of the 1456 Z > 1 structures, 119 were classified as PSTG and 187 as none found, so that of the 1337 reliable Z > 1 structures, 1150 (86%) were identified as having approximate periodic symmetry.

2.5. Dimensionality of the approximate symmetry

Approximate inversion symmetry and approximate translations necessarily describe the whole 3D crystal unless it is composed of alternating layers or there is hybrid packing. Approximate rotational and glide symmetry, however, may be confined to a 2D layer or even to a 1D rod, but deciding on the dimensionality was sometimes difficult. In some structures, the approximate periodic symmetry is clearly 3D and in others it is clearly 2D, but in many others the dimensionality is at the 3D/
2D borderline. In only 19 structures, however, was the only approximate symmetry identified as 1D.

Approximate 2D (i.e. subperiodic) symmetry was described with the labels for layer groups found in Vol. E of *International Tables for Crystallography* (Kopsky & Litvin, 2010).

If a layer that has approximate rotational or glide symmetry has been identified, it is then necessary to check whether the approximate symmetry extends into the third dimension (Figs. 1, 7 and 8). Details of the methods for doing so are given in the supporting information. If approximate 2D symmetry has been found, the *Scanning Tables* in Part 6 of Vol. E are very useful for finding the space groups that include a specific layer group.

### 2.6. Uncertainties

The uncertainties in the percentages given are, except as noted, low – probably ca 2%. In each classification there were categorization questions for only a few of the structures analyzed.

The human errors associated with a project that required so much hand work are another source of uncertainty, but it seems very unlikely they affected the conclusions drawn.

### 3. Results

Of the 1456 $Z > 1$, $P1$ structures analyzed, 243 (17%) contain either achiral (198) or racemic (45) material; the other 1213 are enantiopure.7

Most (85%) of the 1456 have $Z = 2$ or $Z_{\text{eff}} = 2$; most of the rest have $Z = 4$ (10%). Structures with $Z = 3$ (3%) and $Z > 4$ (2%) are less frequent.

#### Table 1

Types of approximate symmetry in the $Z > 1$ structures.

<table>
<thead>
<tr>
<th>Number</th>
<th>Number</th>
<th>Total</th>
<th>Total</th>
<th>Enantiopure</th>
</tr>
</thead>
<tbody>
<tr>
<td>achiral</td>
<td>racemic</td>
<td>racemic</td>
<td>percentage</td>
<td>percentage</td>
</tr>
<tr>
<td>or racemic</td>
<td>enantiopure</td>
<td>(of 147)</td>
<td>(of 1190)</td>
<td>of all 1337</td>
</tr>
<tr>
<td>Inversion (distorted or mimicked)</td>
<td>31</td>
<td>375</td>
<td>406</td>
<td>21%</td>
</tr>
<tr>
<td>Glide (distorted or mimicked)</td>
<td>48</td>
<td>36</td>
<td>84</td>
<td>33%</td>
</tr>
<tr>
<td>Rotation</td>
<td>20</td>
<td>453</td>
<td>473</td>
<td>14%</td>
</tr>
<tr>
<td>Translation</td>
<td>9</td>
<td>49</td>
<td>58</td>
<td>6%</td>
</tr>
<tr>
<td>Alternating layers</td>
<td>15</td>
<td>36</td>
<td>51</td>
<td>10%</td>
</tr>
<tr>
<td>Hybrid packing</td>
<td>6</td>
<td>72</td>
<td>78</td>
<td>4%</td>
</tr>
<tr>
<td>None found</td>
<td>18</td>
<td>169</td>
<td>187</td>
<td>12%</td>
</tr>
<tr>
<td>Total reliable structures</td>
<td>147</td>
<td>1190</td>
<td>1337</td>
<td>100%</td>
</tr>
</tbody>
</table>

3.1. Overlooked crystallographic symmetry is common for structures of achiral or racemic material but rare for structures of enantiopure material

Only 119 (8%) of the 1456 structures were identified as PSTG (*Pseudosymmetry too good*), i.e. should almost

![Figure 7](image_url)

Views of two structures that have approximate 3D symmetry. (a) EGOZIK (Forero *et al.*, 2008), viewed along $a^*$. The structure is approximately $P2_1$ but with unique axis $a$ (cell angles 102.9, 92.8 and 91.5°). (b) KADMIL (Flippen-Anderson *et al.*, 1989) viewed along $a^*$, with $0 \leq y \leq 3$. The structure is approximately $P2_1$ but with $c$ unique (cell angles 93.3, 93.7 and 95.9°). (c) KADMIL after rotating the view in (b) by 5.9° around $c$ so that the view is almost along $b$ but not quite because $a = 93.3°$. This rotation removes the apparent horizontal layer offset but leaves the layer offset along the unique axis $c$ visible.
certainly have been refined in a higher-symmetry and/or smaller unit cell. The percentage of PSTG structures is, however, very much higher (40%) for the 243 structures composed of achiral or racemic material than it is (2%) for the 1213 structures of enantiopure material. P1 structures of achiral molecules and of racemic material are equally likely to have too-good pseudosymmetry; the percentages for the two groups differ by less than 1%.

57 of the 119 PSTG structures have publication dates of 2010–2019, while 24 appeared in 2015–2019. The CSD marks 48 of the 119 as having been re-interpreted (Marsh, 1999, 2005, 2009; Marsh et al., 2002; Clemente, 2003; Marsh & Clemente, 2007; Henling & Marsh, 2014; Wojciechowski et al., 2016). Other determinations archived in the CSD support another 14 of the PSTG assignments. About half of the 119, however, seem not to have been questioned previously.

After removing the 119 suspect structures, the list had 1337 entries.

3.2. Most P1 crystals are enantiopure

Of the 1337 reliable structures, only 147 (11%) contain achiral or racemic material.

3.3. Approximate periodic symmetry is pervasive

In 86% of the 1337 reliable P1, Z > 1 structures, there is some kind of approximate symmetry that is periodic.

The qualifier periodic is key. Two molecules that are the same enantiomer and that have the same conformation can always be related exactly by a twofold rotation plus a translation (Mackay, 1984), but that relationship cannot be periodic unless the rotation axis is parallel to a very simple crystallographic direction [uvw] and the translation is [uvw]/2.

A summary of the distribution of the structures among the categories is shown in Table 1. A more detailed table is included in the supporting information.

Note that if the molecule has an approximate mirror plane, as do all molecules that are quite flat, then a structure categorized as having a 2 or 21 axis may also have approximate inversion symmetry. There will be approximate inversion symmetry if, for example, a twofold axis and the normal of the approximate mirror are approximately parallel. In this study, such structures were assigned to a category like 21 or 2j because requirements of those categories are more exacting than those of the category P1 mimic.

3.4. The approximate symmetry is often subperiodic

Approximate inversion and translation necessarily extend through the whole 3D structure, but approximate rotations, screws, mirrors and glides can be confined to 2D layers or 1D columns. A detailed list of the distributions is available with the supporting information.

For Z > 1 structures having approximate rotations, screws and glides, the ratio of structures having 2D to those having 3D approximate periodic symmetry is nearly 2:1 (357 versus 181 structures), but that ratio is imprecise because of the difficulty of deciding on the dimensionality. What is certain is that many more P1, Z > 1 structures having approximate rotations, screws and glides should be described as having 2D approximate symmetry than as having 3D approximate symmetry. In only 19 structures was the approximate symmetry identified as 1D.

3.5. Approximately symmetric structures need not have any cell angle close to 90°

If two angles of the P1 cell are near 90°, approximate monoclinic symmetry is likely although not certain. If one cell angle is near 90° there is often layer symmetry [e.g. a (001) layer in a cell having y near 90° might have approximate screw and/or glide symmetry]. An angle near 90° in the P1 cell, however, is not a requirement for 2D or 3D approximate symmetry. In many structures, the axes of the approximate cell are linear combinations of the P1 cell axes (e.g. [110], [1T0] and c). If the approximate cell is centered, at least one of its axes is necessarily a linear combination of axes of the P1 cell. BILJIR [Fig. 1; layer (001) with approximate symmetry e211] is an example.

3.6. R. E. Marsh was right

In the abstract of Marsh’s (1999) paper titled ‘P1 or P1? Or something else?’ he wrote ‘In approximately one-third of the structures in which chiral molecules crystallize in P1 with Z = 2, the two molecules are related by an approximate center of inversion.’

This study found that percentage to be 33% for the reliable (i.e. non-PSTG) Z = 2 structures of enantiopure material.
The percentage decreases to 28% for the 115 \( Z = 4 \) structures and to 9% for the 23 \( Z > 4 \) structures, but the value for all \( Z \) values remains high (32%) because there are so many more \( Z = 2 \) than \( Z > 2 \) structures and because almost half of the \( Z_{\text{eff}} > 1 \) structures are \( P\bar{T} \) mimics.

The algorithm of Somov & Chuprunov (2009) gave essentially the same result (32%).

3.7. Approximate rotational symmetry is very common for the enantiopure structures and not rare for the others

While 35% of the 1190 reliable enantiopure structures mimic either inversion (32%) or glide (3%) symmetry, even more (38%) have approximate rotation axes only. The primary rotation or screw axis is occasionally accompanied by an approximate translation or by additional approximate rotations in the other two directions.

Of the 147 reliable structures of achiral or racemic material, 20 (14%) have approximate rotational symmetry only.

3.8. For \( P1 \) structures of achiral or racemic material distorted glide symmetry is more common than distorted inversion symmetry

Just over half of the 147 reliable \( P1, Z > 1 \) structures composed of achiral or racemic material are distorted forms of structures that have symmetry of the second kind. Since structures in groups with glide planes but no inversion (e.g. \( Cc \), \( Pna2_1 \)) are less common than groups that do include inversion (e.g. \( P\bar{T}, P2_1/c \)), it might be expected that approximate inversion centers would be more likely than approximate glides in \( P1 \) structures of achiral or racemic material, but that is not the case. Approximate glide symmetry is found more often than approximate inversion symmetry (48 versus 31 structures).

3.9. More than a third of the structures having distorted symmetry were determined at room temperature

It is well known that during cooling crystals sometimes undergo a transition to a phase that has a larger and/or lower symmetry unit cell. Such transitions have been documented for nine of the \( P1, Z > 1 \) structures analyzed in this study. It might then be expected that most of the approximately symmetric \( P1, Z > 1 \) structures were determined below room temperature. This study, however, found that 229 (38%) of the 610 structures that seem to be distorted forms of structures having smaller asymmetric units (i.e. have approximate symmetry other than \textit{alt layers} or \textit{hybrid packing} and are not mimics) were determined at or above 288 K; for the subgroup of layered structures, the value is 41%. Those percentages are nearly the same as for all 1337 structures (39%). While it is possible that some of those crystals were grown from a hot solution it seems very unlikely that they all were.

3.10. Some structures are composed of different layers that alternate

In 51 of the 1337 reliable \( P1, Z > 1 \) structures, there are alternating layers that differ in composition, approximate symmetry, packing and/or orientation (Figs. 4 and 5). Eleven of the 51 were described previously (Brock, 2020). The others were either added to the CSD later, had been missed or had been excluded because of hydrogen bonds between the layers.

Structures with alternating layers of different compositions are particularly interesting because the layers must have been established at the time of crystal nucleation. The examples found are described at some length in the supporting information.

3.11. Some structures are best described as having hybrid packing

In six crystals of achiral or racemic material and 72 crystals of enantiopure material, different parts of the one molecule are related by different types of approximate symmetry that are incompatible. It is quite common, for example, for the part of a chiral molecule containing its stereocenters to be arranged around an approximate \( 2_1 \) axis while the achiral part is arranged around approximate inversion centers (Fig. 6). While the approximate symmetry cannot be periodic in 3D, it can be in 2D and is so in these 78 structures.

The difference between \textit{alt layers} and \textit{hybrid packing} is the strength of the interactions between the layers having different approximate symmetry. If there are only van der Waals contacts, then \textit{alt layers} is the obvious choice; if there

Figure 9
VIHTUD (Gerber & Betz, 2013; ammonium 5-carboxypyridine-2-carboxylate monohydrate) as viewed along \( a, b \) and \( c \). The centroid of the two anions and the centroid of the \( \text{NH}_4^+ \) cation and water molecule are shown. Those centroids are separated by 0.540, −0.054, 0.004, which differs from \( \frac{1}{2}, 0, 0 \) by 0.15, 0.33, 0.04 Å.
are covalent bonds, then hybrid packing is the only possibility. In the case of structures with good hydrogen bonds between layers, distinguishing between the two descriptions is a judgment call. The choice can depend on the relative strengths of hydrogen bonds within versus between layers, on the extent to which the two layers seem to form a single unit, and on the approximate symmetries of the possible layers.

Twelve salts and cocrystals were described as having hybrid packing because while the different ions or molecules are segregated into layers the layers are linked by strong hydrogen bonds.

3.12. Approximate symmetry is common in $Z = 1$ structures of molecules that can lie on special positions

The percentage of PSTG structures for the 236 $Z = 1$ structures identified as having possible approximate symmetry is 11%, i.e. just a little higher than the 8% found for the $Z > 1$ structures. Again, structures of achiral material are much more likely to be PSTG than are those of chiral material (24 versus 1). A difference between the $Z = 1$ and $Z > 1$ groups is that the ratio of achiral to chiral structures is much greater in the former (0.99 versus 0.12 if the PSTG structures are excluded). The 23 achiral molecules that pack with $Z = 1$ and distorted mirror symmetry certainly contribute to that difference.

The $P1$, $Z = 1$ PSTG structures, however, make up only 2% of all the 1115 because overlooked symmetry is impossible unless the molecule can lie on a special position or the crystal has $Z$ or $Z_{\text{eff}} > 1$. It seems safe to say that by looking at all the $P1$, $R \leq 0.050$ structures having $Z > 1$ or possible approximate symmetry, this study found all the PSTG $P1$ structures (8% of the $Z > 1$ and 2% of the $Z = 1$ structures).

Of the 236 $- 25 = 211$ reliable $Z = 1$ structures identified as having possible molecular symmetry, 72% were identified as having approximate periodic symmetry. That percentage is somewhat lower than the 86% found for the $Z > 1$ structures but is still high.

It might seem strange that a structure of achiral molecules can be classified as an inversion or glide mimic but 24 of the $Z = 1$ structures were. In four of those structures, the very good inversion or mirror symmetry of the larger molecule is broken by included ordered solvent; in six others, the structure would be $P\bar{T}$ if the disorder were 50:50 rather than a ratio between 67.33 and 84.16.

The $P\bar{T}$ mimics VEYXAY (Shkol’nikova et al., 1989) and VIHTUD (Gerber & Betz, 2013; Fig. 9) are very interesting 1:1:1 hydrated salts of a small counterion. In these two structures, the small counterion ($\text{Br}^-$ or $\text{NH}_4^+$) is related to the water molecule by the same approximate inversion center on which the larger ion lies. These two structures suggest that hydrate formation might be likely for a 1:1 salt of an approximately centrosymmetric ion and a very simple counterion.

4. Discussion

4.1. Few cases are ambiguous

The entries for 6% of the structures include the word borderline or marginal, with a third of them having borderline 2D/3D approximate symmetry. Even if that percentage were underestimated, it would still be low.

4.2. Comparisons with previous work

Only four structures in this list were also described by Baggio (2019, 2020). The two sets of descriptions agree.

Rekis (2020) used overlap calculations to identify organic, $R \leq 0.075$, $2 \leq Z \leq 6$ Sohnecke-group structures (i.e. not just $P1$ structures) that contain a $C$-atom stereocenter and that have approximate inversion symmetry. The criteria for his study and this one also differed in that he excluded salts, solvates and cocrystals; a few additional differences are described in the supporting information. Only 168 of the 1456 structures considered in this study appear on Rekis’s list of the 1200 most centrosymmetric structures; 120 of the 168 were classified in this study as $P\bar{T}$ mimics. The remaining 48 were described differently, but only 12 of those were in the top half of his list and most of the 12 have an approximate mirror plane (see §3.3 and the supporting information).

This study categorized 18 structures as $P\bar{T}$ mimics that are neither on the Rekis list of 1200 nor on his supplemental list of an additional 4264, even though those 18 seem to meet the requirements of his study. The discrepancies between the two studies attest to the difficulty of setting criteria for approximate symmetry.

4.3. An experienced eye is still the best tool for recognizing approximate symmetry

Methods for finding approximate inversion (Somov & Chuprunov, 2009; Rekis, 2020) and approximate translations (Brock & Taylor, 2020) have been programmed, but there is not yet an effective algorithm for detecting other types of approximate periodic symmetry. A promising method has been proposed (Baggio, 2019, 2020), but needs further development.

The identification of approximate rotational, glide and mirror symmetry is made much more difficult by its often being periodic in only two dimensions. Since identifying molecular crystals that can be deformed by slippage of molecular layers is not yet a solved problem (Bryant et al., 2018; Wang & Sun, 2019), it seems unlikely that software to identify 2D symmetry will be available anytime soon.

For now, the best approach seems to be the combination of a powerful display program like Mercury (Macrae et al., 2020) and a user who is experienced in looking at crystal structures and has a good working knowledge of crystallographic symmetry groups. Since the judgements so made are unavoidably subjective, they should be documented as well as possible, which is why this article is accompanied by so much supporting information.
Note that the algorithm for finding approximate translations in molecular crystals (Brock & Taylor, 2020) was developed by Taylor writing code that reproduced, at least as well as possible, the judgements that Brock had convinced him were correct.

4.4. Distorted forms of higher-symmetry structures dominate

Of the 147 reliable \( Z > 1 \) structures of achiral and racemic materials in the list, 73% (108) seem to be distorted forms of a structure in a higher symmetry or smaller unit cell (i.e. are not in the categories \( \text{alt layers}, \text{hybrid packing or none found} \)). The higher-symmetry group has a glide more often than an inversion center. Twenty have approximate rotational symmetry only; nine have only an approximate translation.

Of the 1190 reliable \( Z > 1 \) structures of enantiopure material, 42% seem to be distorted forms of a structure in a higher-symmetry Sohncke space group. That percentage is substantially lower for the enantiopure materials because so many of their structures (35%) mimic non-Sohncke symmetry.

The most common distortion from higher symmetry is slippage of layers. If the displacement is small the approximate 3D symmetry can be recognized. If the slippage is large the structure is better seen as composed of 2D layers. There are, however, many structures at the 2D/3D borderline. No good way of discriminating between 3D and 2D approximate symmetry could be found.

Layer slippage could result from crystal cooling but might also occur during the early stages of crystal growth. Of the 361 structures that were described as having a single type of offset layer (i.e. as having approximate 2D rotational or glide symmetry), 41% were determined at a temperature of at least 288 K.

4.5. Possible relationship between 2D and 3D symmetries

Identification of 2D approximate symmetry in a layered crystal raises the question of what the 3D symmetry might be in an undistorted crystal. The Scanning Tables in Part 6 of Vol. E of International Tables for Crystallography (Kopsky & Litvin, 2010) give, for each space group, the layer symmetries of possible slices. The 3D groups consistent with a specific layer group can be found by looking through those tables; the resulting list can be trimmed by considering space-group frequencies.

In this study, the layer groups \( p2_{1}11 \) and \( c211 \) together account for 83% of the 312 layered structures of enantiopure materials; the next most frequent group is \( p112 \) (8%). In the 2019 release of the CSD, the most common Sohncke space groups for organic \( R \leq 0.050 \) structures were \( P2_12_12_1 \) (50%), \( P2_1 \) (34%) and \( C2 \) and \( P1 \) (both 4%). If the material is enantiopure, layer group \( p2_{1}11 \) then suggests deformation of a 3D \( P2_1 \), \( P2_12_1 \) or \( C2 \) structure. Group \( c211 \) suggests deformation of a \( C2 \) structure, as does \( p112 \), but the latter might also result from deformation of a \( P2_12_1 \) structure (frequency 2%).

Only 29% of the 49 layered structures of achiral or racemic material are described by the group \( p2_{1}11 \) or \( c211 \), but groups \( pb11 \) and \( p11a \) together account for 63%. The latter two groups suggest a distorted \( Cc \) (or \( Ia \)) structure because \( Pc \) is quite uncommon (<1%) and because a distorted \( C2/c \) structure would probably retain the inversion center. If a \( Cc \) structure is distorted such that \( (010) \) layers are offset, the layer group is \( p11a \); if the distortion leaves a set of layers \( (h0l) \) unchanged, then the layer group is \( pb11 \). Group \( p11a \) is much more common than \( pb11 \) (28 versus 3 structures).

The ratio of \( c211 \) to \( p2_{1}11 \) layers in the enantiopure crystals \((0.72)\) is much greater than the ratio of \( C2 \) to \( P2_12_12_1 \) (0.08) and to \( P2_1 \) (0.12) structures in the CSD as a whole. That observation raises the question of whether a higher-symmetry crystal nucleus might be more prone to distortion than a lower-symmetry nucleus.

The most common non-Sohncke groups in the 2019 release of the CSD for organic \( R \leq 0.050 \) structures of achiral or racemic material were \( P2_{1}c/c \) (47%), \( P\bar{T} \) (26%), \( C2/c \) (8%) and \( Pbca \) (5%). Distortion of a \( P\bar{T} \) cell with symmetry lowering but no volume change results in a \( P1 \) cell with approximate inversion symmetry \((26 \text{ found})\); if the \( P1 \) cell is \( n \) times larger than the \( P\bar{T} \) cell, then there may be a translational modulation as well (5 found).

If a \( P2_{1}c/c \), \( Z = 1 \) or \( \frac{1}{2} \) cell is distorted enough to give a \( P1 \) cell, then it might seem that any layers are likely to have approximate symmetry \( p2_{1}b11 \) \((z = 4 \text{ if } z' = 1)\) because the distortion would seem unlikely to be large and localized enough to obscure all but one of the \( P2_{1}c/c \) relationships. While no \( p2_{1}b11 \) layers were found for \( P1 \) crystals of achiral or racemic material, four mimics were found for structures of enantiopure materials. Similarly, a distorted \( C2/c \) cell can give a \( p2/b11 \) or \( p112/a \) layer, although none was found. A distorted \( Pbca \) cell can give \( pb2/a \) layers, which were found in three structures of achiral or racemic material and in four enantiopure mimics of a racemic structure.

In a very few structures, there is approximate 3D symmetry that is accompanied by higher approximate layer symmetry. An example is DOVXIW (Fischer et al., 2009), which has
approximate $P2_1$ symmetry overall (c unique) and $p2_{1}2_{2}$ symmetry [$2_{1}$ along $c; z = 4(\frac{1}{4}) = 2$] for layers (100).

4.6. If approximate inversion symmetry is possible it is likely to be found

That inversion symmetry is favorable for crystal packing was established long ago. More recent studies of the CSD have shown that less than 10% of racemic and achiral materials crystallize in Sohncke groups (Fábián & Brock, 2010; Rekis, 2020), and that cocrystals of isomers and other closely related molecules are unlikely to form unless they can mimic an inversion relationship (Kelley et al., 2011; Craddock et al., 2021). A third of $P1, Z = 2$ structures of enantiomerically pure material had been estimated to mimic an inversion relationship (Marsh, 1999; Somov & Chuprunov, 2009).

This study is consistent with previous findings and confirms previous estimates. What is new is the observation that achiral parts of chiral molecules (e.g. a flat ring system) adopt an inversion relationship more than occasionally, even when the rest of the molecule cannot; hybrid packing was found in 6% of the enantiopure $P1, Z > 1$ structures. In a few structures (e.g. JEMHEP; Jiang et al., 2006) the non-Sohncke relationship is a glide.

An inversion relationship is favorable because it reverses all three directions of the molecular surface and so allows the fitting together of ‘bumps and hollows’ (Kitaigorodskii, 1961) without generating the excluded volumes that are large for mirrors and significant for rotation axes. Furthermore, an inversion relationship places few restrictions on the relative orientations of the two molecules. A rotation or screw axis requires the two molecules be related by a $2\pi/n$ ($n = 1–4$ or 6) rotation around an axis that is parallel to a very simple translation vector $[uvw]$, and requires that any associated translation be $[uvw]/n$. It is therefore no surprise that molecules often adopt an inversion relationship, even if it can only be approximate.

4.7. Molecular features that favor inversion mimicry

Looking through the list of $P1$ structures classified as mimics of non-Sohncke symmetry shows that there are many ways for pairs of chiral molecules to mimic an inversion or a glide relationship. No single feature accounts for more than 15% of the 411 mimics identified.

Features that favor inversion mimicry are closely related to those that favor the formation of quasiracemates, cocrystals of closely related isomers and solid solutions. Any enantiomerically pure compound that is barely chiral is a good candidate for inversion mimicry. Molecules having a single stereocenter that would be inverted by a switch of two similar substituents (e.g. H/F, H/Me, Me/Et or Me/Cl) are good examples. An H/Me pair accounts for 63 (15%) of the non-Sohncke mimics identified in this study. Ariel and Leiserowitz (Ariel, 1982; van Mil et al., 1982) showed that an enantio-

4.8. Energy differences between conformers found in the same crystal

In most $Z > 1$ structures, all the independent molecules have similar conformations, but the list of $P1$ non-Sohncke mimics provides many examples of structures containing molecules that have very different conformations. In more than 70 structures, the $rmsd$ for overlay of independent molecules is greater than 2.00 Å. Calculating the energy differences of such pairs of molecules (or sets of molecules if $Z > 2$) could give a good estimate of the degree to which the long-range attractions in a molecular crystal can compensate for a higher energy conformation.

4.9. Implications for the formation of solid solutions

Solid solutions of enantiomers are rare (Brandel et al., 2016; Rekis & Běřič, 2018), but are important because they interfere with the industrially important process of purification by fractional crystallization. Features that promote non-Sohncke mimicry also allow incorporation of enantiomeric
and diastereomeric impurities into supposedly enantiopure crystals, and may even result in the formation of ordered compounds of diastereomers. The possibility of diastereomeric contamination seems not to have received much attention.

The 375 $P\bar{T}$ mimics and the 36 glide mimics in the $P1$, $Z > 1$ list were examined for any disorder or eccentric atomic ellipsoids near the stereocenters and for suspiciously flat C-atom stereocenters. Ten structures were identified as solid solutions, six of which [e.g. TIMOLM01 (Bredikhin et al., 2012) and YOGBON (Cross et al., 2014)] had been identified previously. The frequency of solid solutions is probably higher than those numbers imply because a crystal-structure determination is an imperfect tool for identifying an impurity with a mole fraction below 0.20. Sometimes a 5% light-atom impurity can be identified but sometimes a 20% impurity can be missed. Furthermore, the $R < 0.050$ criterion almost certainly reduced the apparent frequency of solid solutions.

Three solid solutions that are not non-Sohncke mimics were identified; two were described in detail previously, but WIYSAZ (*alt layers; Harmata et al., 2006*) was not. Nominally it is a 5:1 ordered array of enantiomers, but the minor enantiomer is disordered 59:41 with a diastereomer.

4.10. Implications for crystal nucleation

The existence of so many structures in which layers have much better approximate symmetry than does the 3D structure suggests nucleation of a higher-symmetry crystal that deforms by slippage of layers during growth and/or cooling.

Consider the 27 reliable structures of kryptoracemates. A kryptoracemate is a 1:1 compound of separable enantiomers that has been described in a Sohncke space group. It might have been expected that most would be distorted $P\bar{T}$ structures that were determined at low temperature, but that is not the case. Only six have approximate $P\bar{T}$ symmetry and two of the six were determined at room temperature. Sixteen, however, have distorted glide symmetry, with seven of them having been determined at $T \geq 288$ K. That observation suggests the possibility that glide planes are more easily lost during cooling and/or crystal growth than are inversion centers, which are understood to be more favorable for crystal packing.

The finding (§3.9) that 38% of the distorted structures were determined at $T \geq 288$ K certainly suggests that the crystal nucleus can have higher symmetry than the macroscopic crystal.

The identification of 51 structures having layers that differ in either packing or orientation is strong evidence that the ordered attachment of molecular layers can be important in the very early stages of crystal growth.

5. Summary

Approximate periodic symmetry was found in 86% of the 1337 reliable organic, $R \leq 0.05$, $P1$ structures in which the asymmetric unit contains two or more formula units and in 72% of the 211 reliable $Z = 1$ structures in which the residues can have or mimic symmetry $\bar{I}$, $m$, 2, or 3. Approximate rotational, glide and/or mirror symmetry is often periodic in only two dimensions and occasionally in only one.

Approximate inversion centers and approximate 2 and/or $2_1$ axes are equally likely. Approximate glides are less common overall, but are more common for the crystals of achiral or racemic material than are approximate inversions.

Nearly 40% of the organic, $P1$, $Z > 1$ structures of achiral material and of kryptoracemates archived in the CSD are almost certainly described better by a smaller or higher-symmetry unit cell. That value is only 2% for structures of enantiopure material. Publication of PSTG structures is still a problem; 20% of those identified appeared during the years 2015–2019.

The identification of structures (most of them inversion or glide mimics) for which the molecular conformations differ substantially could be useful for determining how large a conformational energy penalty crystals can tolerate in order to optimize overall packing energy.

Features that promote non-Sohncke mimicry may also allow incorporation of enantiomeric and diastereomeric impurities into supposedly enantiopure crystals.

The large number of structures determined at $T \geq 288$ K that are approximately symmetric suggests that in the early stages of crystal nucleation and growth the crystallites may have higher symmetry and/or a smaller unit cell than do the macroscopic crystals used for single-crystal structure determinations. The prevalence of quite exact 2D symmetry in structures with identifiable layers suggests deformation of a higher-symmetry structure by slippage of layers.

Acknowledgements

The project would have been impossible without the CCDC program Mercury (Macrae et al., 2020). It is a pleasure to recall discussions of crystal packing with Dick Marsh, Jack Dunitz, Sam Motherwell, Robin Taylor and Laszlo Fabian.

References
