



STRUCTURAL SCIENCE
CRYSTAL ENGINEERING
MATERIALS

Volume 78 (2022)

Supporting information for article:

Pervasive approximate periodic symmetry in organic *P1* structures

Carolyn Pratt Brock

Pervasive Approximate Periodic Symmetry in Organic $P1$ Structures

BY CAROLYN PRATT BROCK

Department of Chemistry, University of Kentucky, Lexington, KY 40506-0055, USA

EMAIL address: CPBROCK@UKY.EDU

Abstract

The goal of this project was to identify the prevalence of approximate symmetry in organic $P1$ structures. In the November 2019 version of the CSD there are 2592 organic, $P1$, $R \leq 0.050$ structures; complete, unique entries are available for 1407 $Z=Z' > 1$ and 1049 $Z=Z'=1$ 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.

Files

1. Additional commentary (pdf file; 73 pp.)
2. Lists of the 169 illustrated structures as sorted alphabetically and by type of approximate symmetry (pdf file; 5 pp.)
3. Illustrations of 169 selected structures (pdf file; 179 pp.)
4. Spreadsheet of the 1456 $P1$, $Z>1$ and $Z_{eff}>1$ structures (spreadsheet file named $Z>1$)
5. Spreadsheet of the 236 $P1$, $Z=1$ structures of potentially symmetric molecules (spreadsheet file named $Z=1$)
6. Spreadsheet showing cell transformations and calculations of layer offsets (spreadsheet file named TrnsfrmsEtc)

NB/ So much handwork was necessary for this project that errors in what follows and in the spreadsheets are inevitable. The material has been largely checked but it is certain that mistakes were missed. Homonymic misspelling of refcodes was a recurring problem.

1. Introduction

The most recent set of *P1* space-group corrections (Henling & Marsh, 2014) was based on the February 2013 version of the CSD (Groom *et al.*, 2016). The number of *P1* structures in the CSD is now much larger; during the years 2013 – 2019 (this study is based on the 2019 version of the CSD) more than 2000 *P1*, $Z > 1$ structures were added to that database, with more than 1500 of them being organic.

2. Methodology

2.1 Database searches

Details of the counts

The *R*-factor criterion was imposed both to limit the number of structures to be considered and to reduce the frequency of structures complicated by disorder.

The total number of structures in the 2019 version of the CSD (Groom *et al.*, 2016) that meet the basic criteria [organic (C but not H required), $R \leq 0.050$, coordinates, no errors, not polymeric, not a powder structure] is 219,293. Of these 2592 (1.2%) are in space group *P1*.

Duplicate structures include those determined at a different temperature or pressure, or containing the other enantiomer.

Twenty-two structures were discarded because they contain enantiomers of structures retained in the list. In almost all cases the members of such a pair have different refcodes (the exception is CEGLCA/CEGLCA01) although the refcodes are usually, but not always, similar. In the following list the underlined refcode is the one in the list.

$Z > 1$ structures (18):

GEVBAN/GETZUD; GIGDAB01/QODJEA; HEDQEP/HEDQOZ; LAKJIR/LAKJOX;
MECRAP/VOXMAY (MECRAP was determined from a powder sample); PARZOA/UCIYAI;
PATWUE/PATXAL; RESRAK/RESMEO; SEMBAP/SEMBET; SUHCOO/SUHCUU;

TAVGAB/TAVGEF; TOQCEK/TOQCIO; UBAMES/UBAMIW; UXOQAA/UXOQEE;
VIWVEF/VIWVIJ; VOQGUF/VOQHAM; XOCTOZ/XOGZAV; XOGBAX/XOGBIH
Z=1 structures (4):

CEGLCA/CEGLCA01; CUWBIG/CUWBOM; HEVZEP/HEVZIT; QESPUZ/QESQEK
but of these only CUWBIG (an interesting *p*211 mimic) is symmetric enough to have been included in the list of Z=1 structures investigated.

The structures classified as too inorganic are

FABPUW Complicated aggregate of two Ph₂P(=S)C(Li)₂S(=O)(Ph)NMe units with two additional Li atoms, four THF molecules, and two Me₃SiO units. The aggregate has approximate twofold symmetry. THF and EtOEt solvent molecules are present at low occupancy. If FABPUW had been included it probably would have been classified as having no approximate symmetry although a view of a layer (111) suggests there may actually be some approximate symmetry.

VUSGAT Structure of the iodide salt of a Ph₃Te⁺ cation; Te...I contacts connect the four independent sets of ions, which form an aggregate with approximate inversion symmetry. If VUSGAT had been included it would have been classified as having distorted *P* $\bar{1}$ symmetry.

XOMKOB Z'=2 structure of [Ca(cytidine)₂(H₂O)₄]Cl₂·3H₂O, where cytidine is C₉H₁₃N₃O₅. If XOMKOB had been included it would have been described as having layers (010) (axes [101], [10 $\bar{1}$]; angle 92.0°) with good approximate symmetry *c*211. The layers are separated by water molecules that form H bonds to both adjacent layers.

The one structure eliminated for other reasons is OKELOG, which is a channel inclusion complex formulated as 1:0.17 with Z=6. The one included guest molecule is quite large (diethyl succinate, C₇H₆O₄). A 6:1 formulation with Z=1 would also have been possible. The guest molecule has approximate symmetry *2/m* but that symmetry is not consistent with the symmetry of the host.

Inconsistencies in Z values, etc.

The CSD entry contains the chemical formulation given by the authors. A hemisolvate is therefore sometimes formulated as *A*·½*B* and sometimes as 2*A*·*B*; the Z value for the latter formulation, which is less common, is half that for the former. FOQRIO (*P*1, Z=1) and OKELEW (*Cc*, Z'=6) provide a good example of this type of inconsistency. The two have

nearly the same formula unit: six 3,5-dihydroxybenzoic acid molecules, one included large molecule, and seven water molecules.

A 1:1 compound of diastereomers is almost always formulated as $Z=2$, perhaps because differences in configurations at stereocenters are not considered when a CSD entry is created.¹ COYSOY, however, has $Z=1$ because the two independent molecules differ in the location of the double bond.

Three crystals containing two different tautomers² had been formulated as $Z=1$ compounds (see list below) even though the two tautomers cannot be separated. In five other structures the tautomers were formulated with $Z>1$.

There are also occasional inconsistencies in the assignment of refcodes. QOPVAU is an $A:B:A^+:B^-$ compound; it is a modulated form of a $P2_1$ co-crystal $A:B$ (MAMPUM0 n , $n = 2,3,5,6$) that is stable at higher T, and of a $P2_1/n$ structure with $Z'=1/2$ studied at yet higher T (MAMPUM0 m , $m = 0,1,4$). The refcode changes from MAMPUM to QOPGAU as a result of proton transfer in half the molecules. On the other hand QAWNAD01 ($P1$, $Z=2$ at 153 K) and QAWNAD ($P\bar{1}$, $Z=2$, $Z'=1$ at RT) have the same refcode even though the former is formulated as a salt and the latter as a co-crystal.

Quasiracemates³ seem always to have been formulated as 1:1 compounds with $Z=1$.

Disorder

Disorder was generally ignored. It is clearly unimportant if it affects only small counterions, common solvent molecules, groups like CMe₃ and CF₃, or the Me group of an Et or OMe substituent. Even compositional disorder may be unimportant, as in JAQLAP, which is a mixed crystal of two natural products that differ in the substitution of C(=O)OH by CH₂OH. (Incorporation of the two different components of JAQLAP is possible because the carboxy =O does not participate in the H-bonding pattern.) Compositional disorder was noted because it indicates a solid solution but was not considered when categorizing the approximate symmetry.

¹ Diastereomers are stereoisomers that are not related by inversion; in all cases found in this study other than COYSOY the differences are at the stereocenters; the stereocenter configurations and any axial chirality in the diastereomers are neither all the same nor all different. In the CSD it seems that molecules are considered the same if they have the same 2-D chemical line drawing.

² Tautomers are isomers that cannot be separated.

³ A quasiracemate is a structure containing two molecules that would be enantiomers but for a small difference (*e.g.*, a halo substituent that is Cl in one molecule and Br in the other, or an S atom substituted for a CH₂ group).

Several mixed crystals of molecules and their photoproduct were found but in all cases there was also a *P1* structure of the starting material so that the mixed crystals could be classified as duplicates.

Z=1 entries that were included in the Z>1 list (i.e., that have $Z_{eff} > 1$)

The *P1*, $Z = 1$ search found 47 structures of compounds containing two very similar molecules or large ions and two structures that contain more than two. These 49 structures are:

(1) 22 compounds of very similar molecules, i.e.,

15 quasiracemates

BERNER, DEZLAY, DEZMAZ, IREREE, IRERII, JAQJET, KACXET, LUNRER, NIVYOG, PATXUF, POVZAC, SONCED, URODEL, YOGBON, YOGCOO

[but POVZAC (amino acids with $R = \text{CH}_2\text{Ph}$ and *i*-Bu) and SONCED (amino acids with; $R = \text{CH}_2\text{Ph}$ and *i*-Pr) are at the boundary of what can be considered a quasiracemate; perhaps they should simply be described as 1:1 compounds.]

1 compound of diastereomers (COYSOY)

3 crystals formulated as compounds of tautomers (HOSMUY, OFESAT, TETQAK)

1 ordered crystal (PIQRIT) and 2 solid solutions (JAQLAP, PATXOZ) of two very similar molecules that are not quasiracemates.

(2) 11 $n:m$ solvates with $Z=1$ that could have been formulated as $1:(m/n)$ solvates with $Z=n$.

BEPGOU (2:1), EABHIY (2:3), JOKLOM (2:1:1 solvate hydrate), PEKNEY (2:5), ULUFOY (2:1:2 solvate hydrate), UMABAN (2:1), WUYXIX (2:3 hydrate of a 2:1 salt), XIFCAR (4:3), XYFUAD10 (3:2), YIVJET (2:5 hydrate of a 1:2 salt), ZODHIL (2:1)

(In these 11 structures the solvents that occur an odd number of times and therefore might lie on sites of approximate symmetry or be disordered around such sites are H₂O, MeOH, MeCN, isopropanol, CHCl₃, C₆H₆, C₄H₈N₂, and 4-bromophenol. The C₄H₈N₂ and 4-bromophenol molecules do lie on sites of approximate symmetry.)

(Six $n:m$, n even but m odd, $Z=1$ inclusion complexes were not considered because the included molecule is large and asymmetric.)

(3) 8 1:1 salts of small counterions ($\text{XY}_6^{n\pm}$, XO_4^{n-} , X^- , I_3^- , F_3CCO_2^- , SeX_2CN^-) that also contain a neutral molecule that is the same (except for an H⁺ or e⁻) as the larger ion. Of those 8 the first seven listed below contain both charged and neutral forms of a substituted tetrathiafulvalene.

DIWNUS10, EWOCEA, GIGDAB01, LAKJIR, PEXTER, SUHCOO, YEMCUO01, ROLYIA (a 1:1:1:1 salt hydrate),

(4) 4 $n:m$ salts in which there are two chemically identical but crystallographically independent molecular ions and a small counterion:

OCANUD (2:1:5 SO_4^{2-} hydrate), RALHAO (1:2 $\text{Ca}(\text{H}_2\text{O})_6^{2+}$ salt), WINPER (2:1:1 salt with I_3^- and I^-), YIDKIF (2:1 SO_4^{2-} salt)

(5) 4 other structures that are better understood as having $Z_{\text{eff}} > 1$:

BOHDUA (2:1:1 hydrate plus triflic acid; the large molecules are zwitterions and the triflic acid proton was found to have been transferred to the water molecule so that the cell contains two zwitterions, one triflate anion, and one H_3O^+ cation)

QQQETG01 (two formula units of sodium thyroxine pentahydrate formulated as a single molecule because the two Na^+ ions have different coordination shells)

VUPGUJ (1:1:3:1 salt solvate with large 2+ and 1+ ions that differ by one H atom, three PF_6^- ions, and one solvent that is a mixture of MeCN and water)

ZUQMIJ (the $Z=1$ formulation is inconsistent with the formula given, which is for a 1:5/2 hydrate. There are two crystallographically independent $\text{C}_{15}\text{H}_{11}\text{I}_4\text{NO}_4$ molecules and five independent water molecules in the unit cell.)

Molecular symmetry in $Z=1$ structures

The identification of potentially symmetric molecules requires a rough estimate of whether the symmetric conformation is energy accessible. Salts and co-crystals having two large components were sometimes a problem, such as when both units can have approximate mirror symmetry; the mirrors might be aligned (*e.g.*, AXUVUM, DOXHII) or might have orientations different enough that there is no overall approximate symmetry (*e.g.*, CUMHAT, HUHDUK01). Another possibility is that the two molecules are segregated in layers that have different approximate symmetry (HEZZIV01, SIHLIF).

2.2 Analyses of the structures

Spreadsheet

The spreadsheets include the refcode, year of publication, T , Z , number of residues having more than 7 non-H atoms, the total number of stereocenters (or sites of axial chirality) in each of the residues and the number that differ between the independent molecules that are diastereomers. Also included are the *rmsds* for overlay of the independent residues as found in the crystal, as are the *rmsds* for overlay with inversion, and with flexibility if they are smaller than the previous *rmsds* for that entry. *PLATON* alerts for additional symmetry at or above the 90% level are

listed. The entry notes any special group (see below) to which the structure belongs, gives an overall categorization of the packing, includes a short description of the structure, and reports the view direction that seemed most informative.

The special groups are translationally modulated structures (*M*), kryptoracemates (*K*; crystals of racemic material that are best described by a Sohncke space group), quasiracemates (*CQ*; compounds of different but very similar molecules in which all corresponding stereocenters have opposite configurations), compounds of diastereomers (*CD*), compounds of tautomers (*CT*), and other compounds (*C*). Structures having no stereocenter or axial chirality that would persist in solution are marked as *Achrl*.

Display options

An uncluttered view is important so H atoms were suppressed. Bond types were displayed but the circles in aromatic rings were not because the planarity of aromatic systems is so easy to recognize. The *Capped sticks* display option of *Mercury* (Macrae *et al.*, 2008) seemed best except for the stereocenters, which were usually displayed as small balls. Coloring the C atoms by residue number made it easier to see spatial relationships but it proved important to color the heteroatoms by atom type.

Packing diagrams

It was useful to start by looking at a single unit cell. If the projection seemed to show an inversion relationship then the cell was viewed from two other directions, either along the other two crystallographic directions or by rotating the cell by 90° along the horizontal and vertical directions. If there is approximate inversion symmetry it is obvious in all three views; if there is approximate rotational symmetry the direction of the axis becomes obvious.

If no approximate inversion was found, or if $Z > 2$, then other approximate symmetry seemed likely. It proved best to display a packing diagram of 3x3x3 unit cells; a 2x2x2 view seemed too small to reveal relationships reliably. The display was viewed along all three axes, and perhaps along directions [110], [1 $\bar{1}$ 0], [101], *etc.* and [111], [11 $\bar{1}$], *etc.* If molecules are related by an approximate translation this step should find it.

Analyzing layers

If approximate inversion and/or translation symmetry do not relate all *Z* molecules than it is often easiest to find other approximate symmetry by looking at one or more 3x3x1 layers, or at slices (*hkl*), chosen either

to include an approximate symmetry element already spotted,

to include a cell angle close to a special value (90, 60 or 120°), or

because the layer seems to be identifiable (*e.g.*, because it is separated from the adjacent layers by a gap that would be filled with H atoms or because it includes all the conventional H bonds).

Looking for approximate symmetry in layers is always worthwhile even if they interpenetrate somewhat. If the layer has only one non-zero index (*e.g.*, [001]) then slices with coordinates along the out-of-plane axis in both the ranges 0 – 1 and $\frac{1}{2}$ – $1\frac{1}{2}$ must be considered because they are equally likely to show the approximate symmetry, especially if there is no H bonding. (H-bonds usually lie within layers but occasionally it is the adjacent half-layers that are related by approximate symmetry.) If an approximate 2 or 2₁ relationship in a layer is obvious but the molecular spacing along that axis is uneven then changing the layer boundaries by $\frac{1}{2}$ is very likely to make the spacing even.

Examples of structures in which there are interpenetrating layers too symmetric to ignore:

DISXIP Layer (001) (axes **b**, **a**; angle 89.0°) has very good approximate symmetry *p*2₁11; related molecules are #1&3 as well as #2&4. The terminal (CH₂)₃(*i*-Pr) groups in adjacent layers interpenetrate but it is the 2-D approximate symmetry that is convincing.

KECZOK Layer (001) (axes **b**, [210]; angle 86.8°; 0.5 < *z* < 1.5) has symmetry *c*211 that is more approximate than most but is still convincing. There is interpenetration of layers but slip along **b** is unhindered and that is the direction of the layer offset.

If the layer has general indices (*hkl*) [*e.g.*, (101) or (111)] then a plane must be defined and a slice calculated and projected along its layer normal. The depth and displacement of the slice should be chosen with the possible layer boundaries in mind. A slice area of *ca.* 40 Å in each direction is usually appropriate but may have to be increased if *Z=Z'* is high or the molecule is large.

The layer should first be viewed along its normal [*e.g.*, along **c*** if the layer is (001) or using the *view perpendicular to plane* option for a more general plane (*hkl*)]. If there is approximate symmetry it will almost always be immediately obvious, but the axes of the approximate layer cell are quite often not those of the *P*1 cell. If approximate symmetry is seen, edge-on views of the layer are often very informative.

Hints in the cell constants

In some of the structures, but by no means in all, the cell constants provide clues about the approximate symmetry. If two of the cell angles (*e.g.*, β and γ) are both near 90° then projection down their common axis (*e.g.*, **a**) may reveal an approximately monoclinic cell. If only one of

the cell angles (*e.g.*, γ) is near 90, 60 or 120° there may be a layer [*e.g.*, (001)] having approximate symmetry. If the lengths of two axes are very similar and those axes make very similar angles with the third axis (*e.g.*, *a* and *b* very similar and also α and β) the sum and difference of those axes [*e.g.*, $(\mathbf{a}+\mathbf{b})/2$ and $(\mathbf{a}-\mathbf{b})/2$] may be the axes of the approximate 2-D or 3-D cell.

UNEVAK01 is an instructive counterexample. It has $\beta = 91.0$ and $\gamma = 90.3^\circ$ but the only approximate symmetry is a 2_1 axis along **b** (*i.e.*, not along **a**). A layer (001), $0.5 < z < 1.5$ has excellent approximate symmetry $p2_111$ but there can be no approximate 3-D symmetry because $\alpha = 99.1^\circ$.

Stereocenters

The option in *Mercury* to *Select Stereocenters* is very useful but not infallible; the marked atoms were always checked against the chemical line drawing. It seems that sulfoxide centers ($R_1R_2S=O$), which do not racemize under normal conditions, are never marked, nor are chiral allenes $R_1R_2C=C=CR_3R_4$. The P atom in an $RP(OH)(=O)(O^-)$ group is sometimes marked as a stereocenter even though proton rearrangement leading to “racemization” would be very rapid. In a few entries none of the obvious C-atom stereocenters is marked; in a few others (*e.g.*, ADGSMF) some but not all are marked. In most *meso* molecules the C atoms with four different substituents are not marked as stereocenters but in a few entries (*e.g.*, MAGLIQ) they are.

The CSD includes no information about axial chirality such as in hindered biphenyls and binaphthyls; in almost all such cases the original literature provided information about the separability of the atropisomers.

Using PLATON

A *PLATON* recommendation for higher symmetry (default tolerances) was usually found to be reliable if the level was above 95% but there are exceptions, most of which are in the following categories:

- (1) Structures that are known to be related by a phase transition to a higher symmetry phase.
- (2) Structures of homochiral molecules that are barely chiral [*e.g.*, have one stereocenter $RCH(Me)R'$]. (The *PLATON* output mentions this possibility.)
- (3) Structures in which there is disorder around an approximate symmetry element but in which the occupancy factors are far from 0.5 (*e.g.*, DIWGAS01, DIWGEW, DIWGIA).

If there was any question about the *PLATON* recommendation a *.res* file was generated and displayed in *Mercury* to see if the averaging resulted in significant disorder. The original paper was consulted to see if the authors had carefully justified the choice of space group *P1*.

PLATON is optimized for finding overlooked symmetry, especially overlooked inversion, but is a less reliable tool for finding approximate symmetry.⁴ The recommendations are, however, sometimes very helpful. If the molecule has at least one obvious stereocenter a recommendation using standard tolerances for a $P\bar{1}$ description suggests that the structure is a $P\bar{1}$ mimic. When the tolerances were increased (*ADDSYM 2 1 1 1*), however, *PLATON* sometimes identified approximate inversion symmetry when a description including a twofold rotation seems better (*e.g.*, FONXIQ, TEGXEI, XUWLOS, XUPLEB).

The *PLATON FIT* instruction was also very useful. If two homochiral molecules are related by approximate inversion but have similar conformations they are related more precisely by an approximate twofold axis.⁵ That axis, however, often points in a direction $[uvw]$ that is not simple (*i.e.*, that includes integers u, v, w larger than 1), and there is often an associated translation that is not a simple fraction of $[uvw]$. The *PLATON* instruction *FIT n m*, where n and m are molecule numbers, finds the direction of the best rotation axis relating n and m , the magnitude of the rotation, and gives an *rmsd* indicating the quality of the fit. The *FIT* instruction sometimes led to the identification of approximate rotational symmetry and other times confirmed the finding that there is no approximate periodic symmetry (*e.g.*, *rmsd* > 1.5 Å for the fit, $[uvw]$ not simple, rotation angle not near a special value, and/or a translation that is not a simple fraction of a simple vector $[uvw]$). Approximate symmetry that had been missed was sometimes found after running *PLATON FIT*. (*e.g.*, in LUSMAN, NAXRIQ).

It seems that *PLATON* recommendations above the 95% level should always be investigated, even if the tolerances have been increased. Recommendations above *ca.* the 85% level with standard tolerances sometimes provide useful information. If no approximate symmetry has been found it can be useful to run the *FIT* instruction.

Additional comments about *PLATON* output are given in Appendix 1 at the end of the supplementary material.

⁴ While the results given in this study are based on a version of *PLATON* that is more than several years out of date, Ton Spek confirmed that the *ADDSYM* routine has not been changed recently.

⁵ If two molecules have the same conformation then they are related by six variables, which are usually thought of as three translations and three rotations. It is also possible, however, to describe the relationship with three translations and a rotation n around an axis specified by two orientation angles. The *FIT* instruction in *PLATON* gives that axis in crystal coordinates and the necessary rotation around it. A rotation of close to 180° around an axis near $[uvw]$ suggests an approximate 2 or 2₁ axis along $[uvw]$.

Molecular overlay algorithm

The *rmsd* for overlay of molecules as found in the crystal is always present in the spreadsheet; *rmsds* for overlay with inversion and with flexibility are only present if they are smaller. If the *rmsd* with inversion is significantly (>20%?) smaller than the *rmsd* without inversion then there is a very good chance that the molecules are related by an approximate inversion or glide.

It is important to check whether the overlay with inversion causes some or all of the stereocenters to be overlaid with their inverted forms. If inversion allows good overlay of all stereocenters then the structure is a kryptoracemate (a compound of separable enantiomers that is described by a Sohncke space group). If some but not all of the stereocenters are superimposable then the structure is of a compound of diastereomers.

The *rmsd* sometimes rises modestly when flexibility is allowed, which suggests an undiagnosed problem with the algorithm. The information on the CCDC website about the overlay function (see <https://www.ccdc.cam.ac.uk/support-and-resources/ccdcresources/mercury.pdf>) says

Rings are not flexed during the overlay. Hydrogen atoms are ignored. The RMSDs reported for multiple overlays of the same pair of molecules may vary. The algorithm is not exhaustive and between overlays, may find different, closely related solutions.

An extreme example of the limitations of the overlay algorithm is SANNUS, an enantiomerically pure $P\bar{1}$ mimic containing two molecules having an *RCHMeR'* stereocenter in a C_4N ring that contains an N atom that can invert. The *rmsds* are 2.84 and 1.37 Å for overlays without and with flexibility but only 0.51 Å for overlay with inversion. Enantiomerically pure WASVOB is also instructive: the two molecules have one C-atom stereocenter but the configurations at two of the four N atoms differ. The overlay *rmsds* are 1.88 for the molecules as observed, 1.51 Å if flexibility is allowed, but only 0.37 Å if one of the molecules is inverted.

The overlay algorithm occasionally allows a conformational change that cannot occur at room temperature (*e.g.*, racemization of a hindered 1,1'-binaphthyl). QASNIJ is an example of a resolvable 1,1'-binaphthyl for which the flexibility option allows racemization. The bianthryl VOJBZ is another example.

Additional comments about the overlay option in *Mercury* are given in Appendix 2 at the end of the supplementary material.

Confirming the absence of approximate symmetry

Approximate symmetry is very unlikely if there are very different molecular conformations that are not approximately related by inversion or a mirror. Approximate symmetry is also unlikely if

the *FIT* instruction in *PLATON* finds the molecules to be related by a rotation that is not near 0 or 180° and/or is not along a simple crystallographic direction.

2.3 Criteria for identifying approximate symmetry

Small solvent molecules and counterions

The positions and orientations of solvent molecules and counterions having no more than seven non-H atoms were not usually considered. Such residues very seldom influence the overall approximate symmetry and are often disordered.

Deviations of angles from special values

If the approximate symmetry corresponds to a space group other than $P\bar{1}$ and is not purely translational then at least two angles in the approximate 3-D cell must be close to a special value (usually 90° but very occasionally 60 or 120°). A special angle is also required for most of the layer groups, although not for those in which the only symmetry is a twofold rotation axis normal to the layer and/or a mirror or glide mirror within the layer (*i.e.*, groups $p112$, $p11m$, $p11a$, $p112/m$, and $p112/a$).

The deviations from the special angle are usually $\leq 3^\circ$. After looking at many hundred $P1$ structures the initial decision was made to limit the possible deviation to 5° but later that choice seemed too conservative; a few structures have such obvious approximate symmetry that it was noted in the spreadsheet (although usually marked as marginal or borderline) even though the deviation was somewhat larger. Drawings of all structures having deviations $>6^\circ$ are provided.

	Have approximate 3-D symmetry other than $P\bar{1}$ (2, or very occasionally 3, special angles)	Have approximate 2-D symmetry and one special angle (<i>e.g.</i> , not $p112$ or $p11a$)
Number of structures	181	302
Average deviation from special angle (°)	2.0	1.5
Median deviation from special angle (°)	1.8	1.0
Maximum deviation from special angle (°)	7.5	9.7
(All) deviations $\leq 3^\circ$	59%	85%
(All) deviations $\leq 5^\circ$	85%	95%
(All) deviations $\leq 7^\circ$	98%	99%
(All) deviations $\leq 10^\circ$	100%	100%
(At least) one deviation $> 5.0^\circ$	27	16
(At least) one deviation $> 6.0^\circ$	9	10
(At least) one deviation $> 7.0^\circ$	3	3

The nine 3-D structures with a deviation $> 6.0^\circ$ are

IHOXUY (7.9, 7.4, 7.3°)	The cell angles (82.6, 82.7, 82.1° at 173 K) are all far from 90° but the approximate symmetry $P2_122_1$, $Z'=\frac{1}{2}$ (standard setting $P2_12_12$) is too compelling to ignore. No layer has better approximate symmetry than does the structure as a whole.
----------------------------	---

YULFIV (7.6, 6.2°)	Transformation (012 010 100) gives a <i>C</i> -centered cell with angles 82.4, 89.3, and 83.8°. Angles α and γ deviate by quite a lot from 90° but the approximate <i>C</i> 2 symmetry (2 along b) is visually obvious (see a view along a*) and seems to be the best description.
MOMVOA (7.3°)	A view along a suggests a 3 or 3 ₁ axis although the spacing along a is a little uneven (centroids separated by 0.37, 0.30, 0.33 in <i>x</i>), Angles are 112.5, 91.5, and 91.8°; <i>b</i> , <i>c</i> are 13.3, 13.6 Å. MOMVOA looks like a very distorted, marginal <i>P</i> 3 ₁ structure except for the included solvent.
EPIYUY (6.6°)	Approximate axes 2 and 2 ₁ are approximately parallel to b . Transformation (012 010 100) gives an approximate <i>C</i> 2 cell with angles 87.3, 91.9, and 96.6°. A view along a* _{<i>P</i>1} = c* _{<i>C</i>2} is convincing in spite of the large deviation of γ from 90°.
GEGYAV (6.6°)	Layer (001) (axes [110], [110]; angle 92.1°) has very good approximate symmetry <i>c</i> 211. Transformation (112 110 110) gives an approximate <i>F</i> 2 cell with angles 87.9, 95.4, and 96.6° so the structure has borderline 3-D approximate symmetry. See a view along [110].
ULUHUE (6.4°)	Transformation (021 001 100) gives an approximate <i>C</i> 2 cell with angles 90.2, 92.1 and 83.6°. The angle γ makes the description borderline, but there seems to be no better description given the obvious approximate symmetry and the H-bonding pattern. Views along c and a* are convincing.
ADMPT10 (6.3°)	Transformation (110 110 001) gives a borderline <i>C</i> 2 cell with angles 86.4, 103.6, and 83.7°. The deviation of γ from 90° is large but the approximate axes are quite obvious so classify as marginal 3-D.
AFONUY (6.3°)	Two independent layers (001) [axes b , a ; angle 90.0°; layers of #1&2 (0.5< <i>z</i> <1.0) and of #3&4 (1.0< <i>z</i> <1.5)] have quite good approximate symmetry <i>p</i> 2 ₁ 11. The layers are related by approximate 2 ₁ s along a and c so the structure is probably distorted from <i>P</i> 2 ₁ 2 ₁ 2 ₁ (α , β are 95.0, 96.3°).
VOJXEA (6.2°)	There are approximate 2 ₁ axes along a . If the space group is <i>P</i> 2 ₁ the two monoclinic angles differ from 90° by 2.9 and 6.2°. If there are layers (001) the layer angle is 96.2°. If layers (010) the angle is 92.9° but the spacing is uneven. VOJXEA seems to be best described as a quite distorted, marginal <i>P</i> 2 ₁ structure. See a view along b* .

The ten 2-D structures with a deviation $>6.0^\circ$ are

PEQGAU (9.7°)	Very obvious H-bonded layer (001) (axes [110], [1 $\bar{1}$ 0]; angle 99.7°; $0.5 < z < 1.5$) has very distorted, but still very obvious, approximate symmetry $c211$. The two $Bz_2NH_2^+$ ions would each lie on a twofold axis. The presence of (only) one water molecule is part of the distortion. An interesting if marginal example.
ZODHIL (7.8°)	Layers (001) (axes [1 $\bar{1}$ 0], [110]; angle 97.8°) are separated by the included chloroform molecules and have quite obvious approximate symmetry $c211$. Given the layer angle the approximate symmetry is clearly borderline but it still seems to be the best description.
ZUWDIG (7.5°)	H-bonded layer (100) (axes b , [012]; angle 82.5°) has quite good approximate symmetry $c211$ even though the angle is far from 90°. Borderline.
FOYZUO (6.8°)	H-bonded layer (001) (axes a , [$\bar{1}$ 20]; angle 83.2°) has approximate symmetry $c211$ but is borderline because of the deviation of the angle from 90°.
XIKMAG (6.8°)	Very good 2_1 along b but any 2-D approximate symmetry is marginal. Layer (10 $\bar{1}$) (axes b , [212]) has a good 2_1 and a less good 2 but the layer angle (83.2°) is rather far from 90°. Still, the 2-D $c211$ description seems better than a 1-D 2_1 description.
HIBTES (6.4°)	Obvious layer (01 $\bar{1}$) has good 2_1 axes along [011] and recognizable symmetry $p2_111$ but the layer angle is 83.6° so the 2-D approximate symmetry is borderline. (The approximate $c211$ symmetry in layer ($\bar{1}$ 1 $\bar{1}$) is considerably less good.)
IDUYAK (6.3°)	Layer (010) (axes [101], [10 $\bar{1}$]; angle 83.7°) has very obvious approximate symmetry $c211$ although the angle is quite far from 90° and the 2s are rotated away from [101] by 16°. The $c211$ description is marginal but seems to be the best choice.

SEYNOB (6.3°)	Rods along $[11\bar{1}]$ having an approximate 2_1 axis are arranged in an approximately centered rectangular array. Layer (101) (axes $[11\bar{1}]$, $[1\bar{1}\bar{1}]$) has quite good approximate symmetry $c211$ but the angle is 83.7° . Borderline $c211$ structure.
EABHIY (6.2°)	H-bonded layer (010) has an approximate 2_1 along c . A $c211$ layer cell (axes c , $[201]$) is easily recognizable but the angle is 96.2° . Borderline.
PIVZAW (6.2°)	Layer (100) (axes c , b ; angle 83.8°) has approximate symmetry $p2_111$. A $p2_111$ layer (010) is also possible but the angle would be even smaller (73.7°).

In a number of structures that have approximate symmetry other than inversion the approximate layer symmetry is very good but it is not clear whether that symmetry should be described as 2-D or 3-D. If three layers are viewed along their layer normal then the 3-D description is the obvious choice if the offsets of translation-related molecules are small. It is not straightforward, however, to set numerical criteria for a 3-D vs. a 2-D description because the magnitude of the offset depends on the length of the third axis as well as on the angles of the third axis with the special layer axis (or axes in the case of higher symmetry). In any case that angle is $>5.0^\circ$ in 98% of the 302 structures described as layered, and is $>7.0^\circ$ in 90%. In 29 of the 302 that angle is in the range $5.0 - 7.0^\circ$ so that the distinction between approximate 3-D and 2-D symmetry was a judgement call.

Phase relationships

The $P1$ structures that are clearly related by a phase transition to structures with higher symmetry are listed below. In most cases the two structures were published in the same paper; in the others the two structures are very similar but the subtle distortions in the lower-temperature phase are easy to see. In almost all cases *PLATON* found the phase relationship, and almost always at the 100% level.

***PLATON* recommendations that agree with the higher-symmetry cell described in the literature (even though the $P1$ description of the lower-T form is almost certainly correct)**
100% $P1$

CAZLAR02 ($P1$, $Z=2$ at 20 K) CAZLAR01 ($P\bar{1}$, $Z=2$, $Z'=1$ at 100 K)

CILHIO38 ($P1$, $Z=2$ at 20 K) CILHIO21 ($P\bar{1}$, $Z=2$, $Z'=1$ at RT)

Other CILHIO structures in the CSD with $R \leq 0.050$ of that $P\bar{1}$ phase (there is also a different $P\bar{1}$ phase) are CILHIO12 (at 100 K) and CILHIO19 (at RT)

FUGMIF ($P1$, $Z=4$ at 200 K) FUGMIF01 ($P\bar{1}$, $Z=4$, $Z'=2$ at 413 K)

RIDFOA01 ($P1$, $Z=2$ at 50 K) RIDFOA ($P\bar{1}$, $Z=2$, $Z'=1$ at RT)
[The temperature (RT) given in the CSD for RIDFOA01 is incorrect.]

100% C2

POPGUW01 ($P1$, $Z=2$ at 123 K) POPGUW ($C2$, $Z=4$, $Z'=1$ at RT)

100% P2_{1/n}

QOPVAU ($F1$, $Z=8$, $Z'=2$ at 100 K) MAMPUM06 ($P2_1$, $Z=2$, $Z'=1$ at 100 K)
MAMPUM01 ($P2_{1/n}$, $Z=2$, $Z'=1/2$ at RT)

For comparisons with MAMPUM0 n , Z' and Z for QOPVAU should be 4 and 16 rather than 2 and 8 because there is proton transfer in half the formula units of QOPVAU but not in those in MAMPUM0 n family.

100% P3₂

IBEPUC06 ($P1$, $Z=3$ at 200 K) IBEPUC0 n ($P3_2$, $Z=3$, $Z'=1$ at 100 K to RT)

(It seems likely that the $P1$ phase is actually a solid solution that includes substitution by F for H atoms at two sites)

100% P1 but with a smaller cell

JAKKUB01 ($P1$, $Z=2$ at 85 K) JAKKUB ($P1$, $Z=1$ at 150 K)

other

EJUQIK ($P1$, $Z=4$ at 123 K) EJUQIK01 ($P1$, $Z=2$ at RT)
PLATON finds the approximate translation $[110]/2$ but also recommends $P\bar{1}$, but only at the 88% level,

***PLATON* recommendations that are not the same as the phase relationship described in the literature**

100% Pc

NAKNEV04 ($P1$, $Z=4$ at 200 K) NAKNEV02 ($Pca2_1$, $Z=4$, $Z'=1$ at 233 K)

For NAKNEV04 *PLATON* recommends $Pca2_1$ but only at the 81% level

For NAKNEV03 at 150 K *PLATON* recommends Pc at the 100% level

The paper discusses the Pc possibility but rejects it

NAKNOF04 ($P1$, $Z=4$ at 220 K) NAKNOF02 ($Pca2_1$, $Z=4$, $Z'=1$ at 250 K)

For NAKNOF04 *PLATON* recommends $Pca2_1$ but only at the 86% level

For NAKNOF03 at 150 K *PLATON* recommends Pc at the 100% level

The paper discusses the Pc possibility but rejects it

none

ZEBVEJ ($P1$, $Z=4$ at RT)

ZEBVEJ01 ($P2_1$, $Z = 4$, $Z'=2$ at 90 K)

Perhaps *PLATON* issues no alert because the angle changes are so large (92.3, 90.2, 90.1° to 90, 90.8, 90° for corresponding angles)

2.4 Categories

In the case of a structure with more than one type of approximate symmetry the following hierarchy was applied:

1. *Approximate inversion* If an approximate inversion center was present the category was *P1 mimic* if the crystal contained enantiopure material or *distorted P1* if the crystal contained achiral or racemic material. If an approximate glide, screw, rotation, or translation was also present the term *plus* was appended. The label *distorted P1* implies that transition to a centrosymmetric phase seems possible, although it might require disorder of a solvent or counterion.
2. *Approximate glide or mirror* If no approximate inversion was present but an approximate glide or mirror was present then the category was *glide*, *glide mimic*, *mirror*, or *mirror mimic*. If an approximate screw, rotation, or translation was also present the word *plus* was appended.
[Approximate mirrors were found in $Z=1$ structures only. Approximate C_m and $cm11$ symmetry was categorized as *glide plus (mimic or not)* because C_m and $cm11$ include both mirrors and glides.]
3. *Approximate rotation* If no approximate inversion or glide symmetry is present but approximate rotation was found the categories were 3 , 3_1 , $2&2_1$ (parallel axes 2 and 2_1), 2_1 , $P2/p211$ (for a twofold axis in a 3-D structure or lying in a 2-D layer), and $p112$ (for a twofold axis perpendicular to a 2-D layer). In the case of a rotation or screw axes in a second direction the term *plus* was appended. That term was also added if there was an approximate translation other than the centering required by the $2&2_1$ combination.
4. *Translation* This term was applied if there was an approximate translation but no other approximate symmetry.

The term *alt layers* indicates the presence of parallel layers that differ in composition, approximate symmetry, packing and/or orientation. In a few cases the layers are connected by hydrogen bonds.

The label *hybrid packing* was used for structures in which different parts of a single molecule pack differently (*e.g.*, the parts of the molecules containing the stereocenters are related by a 2_1 screw and the achiral parts by inversion). A structure with *hybrid packing* can be thought of as having *alt layers* that are connected by covalent bonds.

If no approximate n -D symmetry was found the label *none found* was assigned.

In all cases the dimensionality of the approximate periodic symmetry was given in the spreadsheet column immediately following the category column.

Alternating layers

In 51 structures, 36 of which have $Z > 2$, there are several independent layers so that no layer contains all the independent molecules. In some of those structures very similar layers have different orientations that are related by local approximate symmetry (*e.g.*, CICTIT, GIPLID, HOCYEG, ZZZVXQ06). In others the layers differ markedly (*e.g.*, CIFGEE, WIYSAZ). In yet a third group one layer has good approximate symmetry while the other does not (*e.g.*, FUXYUS, GUVROF).

Hybrid packing

In 75 structures at least one part of the molecule forms layers that have approximate periodic symmetry while the remaining part of the molecules forms layers that either have different periodic symmetry or no approximate periodic symmetry.

Complications

Identification of approximate inversion symmetry is often easy but is sometimes a matter of judgement. If the material is enantiomerically pure then at least some parts of the molecules cannot be related by approximate inversion. Which groups should match sufficiently for approximate inversion to be declared?

Those that have the most effect on the van der Waals surface of the molecule?

Those with the most, and perhaps heaviest, heteroatoms?

Those that can form H bonds?

If the molecule is mostly planar then approximate inversion does not differ much from a twofold rotation around the normal of its molecular plane followed by a translation.

2.5 Dimensionality of the approximate symmetry

Conventions for layer-group symbols

In the tables of layer groups in Vol. E of *International Tables* (Kopský & Litvin, 2010) the symbol always starts with an italicized lowercase letter, the axis *c* is always perpendicular to the layer, and symmetry associated with all three directions is always specified. The symbol *p11a* then describes a primitive cell with an *a* glide for which the mirror lies in the layer and *p112/a*

describes a cell that also has a twofold axis perpendicular to the layer. If there is one rotation axis that lies within the layer it is described as parallel to **a** (e.g., $p2_111$, $p211$, or $c211$). If there is one glide (or mirror) with its normal in the layer the normal is described as parallel to **a** (e.g., $pb11$ or $pm11$). If an axis and a glide normal or a mirror normal are parallel then the symbol is, e.g., $p2_1/b11$. When there are symmetry elements parallel to both layer axes the group is described with a symbol like $p2_12_12$ or $pba2$, with glides taking precedence over axes (e.g., $pb2_1a$ rather than $p2_1ab$). If there is symmetry associated with two axes then there is necessarily symmetry associated with the third.

Symbols for rod groups

Vol. E of *International Tables* also contains the symbols for the 75 rod groups having axes of order 1, 2, 3, 4, and 6, but these symbols were not used because they did not seem informative for the 19 structures classified as having 1-D approximate symmetry.

Checking whether approximate 2-D symmetry might actually be 3-D

If the layer includes all the independent molecules and its only approximate symmetry is inversion or translation then the overall approximate symmetry is necessarily 3-D.

If adjacent layers are related by a rotation other than by *ca.* 180° or 120° , then the classification *alt layers* is likely.

If the approximate symmetry is only 2-D then adjacent layers must either be different or must be offset along a direction perpendicular to the layer normal in a way that precludes extension of the approximate periodic symmetry into the third dimension. The first step in the analysis is finding the best third axis. The simplest method seems to be to view a projection of two layers, with the molecules in the two displayed differently (e.g., one as thick capped sticks and the other as wireframe). The best third axis is then the vector between the translation-related molecules in the two layers that are closest in the projection. If an approximate rotation axis is vertical in the projection then a vertical offset between layers is significant but a horizontal offset is not because it simply indicates a monoclinic angle β significantly different from 90° . The same is true if the normal to the mirror (or glide mirror) is vertical.

If the approximate symmetry is 3-D a projection of three layers along the layer normal [e.g., along \mathbf{c}^* for a layer (001)] should show that the layers are offset only a little along the special axis. Making the special axis vertical in the projection is usually helpful.

If there is approximate symmetry associated with the third axis (e.g., groups $p112$, $p11a$, $p2_12_12$, $pb2_1a$, and $pba2$) then the third axis must be approximately perpendicular to both layer axes and therefore approximately parallel to the layer normal.

The projection of the third axis (*i.e.*, \mathbf{c}_{layer}) on any special layer axis (*e.g.*, $c\cos\beta/a$ for a 2_1 axis along \mathbf{a}_{layer} or a glide with mirror perpendicular to \mathbf{a}_{layer}) gives the offset of adjacent layers in that direction. A value near 0.0 indicates straightforward 3-D symmetry; a value near 0.5 suggests a centered cell. Typically there is an offset between layers in the fractional range 0.15 – 0.35; most values are in the range 0.2 – 0.3. Values far from 0.25 were investigated carefully. Values <0.15 or >0.35 may indicate approximate 3-D rather than approximate 2-D symmetry; values <0.07 or >0.43 essentially always do. In the case of an offset near 0.5 it proved useful to view a projection of three layers to see how well the first and third layers are aligned.

The values of that projection can, however, be misleading because its value depends on the ratio of the lengths of the special axis and the third axis; if those axes have very different lengths then the value of the projection may not be a reliable guide. For that reason the angle of the third axis with any special axis was also tabulated and considered. The angle of the third axis with a special direction in the layer should deviate from 90° by no more than *ca.* $5 - 7^\circ$ unless there is centering in the third direction. The values of both the projection and the angle were considered when deciding on the dimensionality, as was a view along the layer normal of two (or three) adjacent layers.

SOXMAW is a good example of an offset/angle discrepancy. Layer $(01\bar{1})$ has very good approximate symmetry $c222$ and a layer angle of 95.5° . The layer axes (\mathbf{a} and $[011]$ of the $P1$ cell) have very different lengths (10.75, 31.08 Å) so that the projection of \mathbf{b}_{layer} on \mathbf{a}_{layer} is much larger than that of \mathbf{a}_{layer} on \mathbf{b}_{layer} (0.28 vs. 0.03).

It proved impossible to set angle or offset ranges that reliably distinguish 2-D from 3-D approximate symmetry.

3. Results

1456 structures total, of which 1407 have $Z > 1$. All but two of the 49 $Z = 1$, $Z_{eff} > 1$ structures have $Z_{eff} = 2$.

Z	number	% of 1456 (i.e., of all)	% of 1337 (i.e., of all reliable)
1	49	3%	4%
2	1185	81%	81%
3	38	3%	3%
4	149	10%	10%
5	3	0%	0%
6	17	1%	1%
7	0	0%	0%
8	11	1%	1%
>8	4	0%	0%
>4	35	2%	2%

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

119 marked *PSTG*

78 contain achiral material (39.4% of all 198 structures of achiral material)

18 contain racemic material (40.0% of all 45 structures of racemic material)

23 contain enantiopure material (1.9% of all 1213 structures of enantiopure material)

48 marked in the November 2019 CSD as reinterpreted

14 structures for which other information in the CSD supports the *PSTG* designation

57 for which nothing in the CSD suggests a problem

The most common overlooked space groups are $P\bar{1}$ and $P2_1$ (70 and 16 structures, respectively); the next most common are $C2$ (9) and Cc (8).

Likely space groups of the 78 achiral structures and 18 kryptoracemates marked *PSTG*:

$P\bar{1}$	70	(73%)	
Cc/Ia	8	(8%)	
$P2_1$	4	(4%)	
other	14	(15%)	$C2, Pc, P2_1/m, P2_1/n, P2_12_12, P2_12_12_1 (2x), Cmc2_1, Pna2_1 (2x), Fdd2 (2x), Pnma, I\bar{4}$

8 of the 96 are Sohncke groups.

Of the 18 *PSTG* kryptoracemates 14 (78%) should almost certainly have been described in $P\bar{1}$.

Likely space groups of the 23 chiral structures marked *PSTG*:

$P2_1$	12	(52%)	
$C2/I2$	8	(35%)	
Other	3	(13%)	$P2_12_12_1, I4, R3$

Achiral *PSTG* structures marked in the CSD as re-interpreted (33):

APONOJ, APOPAX, APOPOL, APORIH, BAZHAM, CEBZAM, CEPKOY, FEPGOY, FURGAA02, GIPTAA, GOFMOD, HEZJIG, HIDJUA, HIGHAH, HOFVED, IBIXEY, ISECOZ, JUXLIX, KETCAN, NIXQAM, NOQBUQ, PABYIA, PUBTUC, QANCEP, QETTEO, SIDXUY, TEBLUH, TUFTUK, VAVHEH, WEQHEF, XEJNUV, YEBQUR, ZOWFAS

Kryptoracemate *PSTG* structures marked in the CSD as re-interpreted (6):

EKOKOE, HITNED, TOQJAK, VENPOU, WEPWUJ, YAHYIP

Chiral *PSTG* structures marked in the CSD as re-interpreted (9):

CUNBET, GULPUY, JEHWAX, JORMIM, LEZLEI, MBDARP, NICBOQ, POTBAC, VEMSAH

Achiral *PSTG* structures **not** marked in the November 2019 CSD as re-interpreted

(45).ABORER, .AJORAU, BIHKAG, BOCKEK06, CIDHAB, CILHAI, CUNPEI, DCHPHA01, EFEQUD, EYISAI01, FOKDOZ, FRANAC03, FUCTIG03, GASXIK, GIBXOH, GIMTON, JETJUQ, KAYSOT, MAGLIQ, MOZZEG, NEVNOT, POQPIX02, QOJLUY, RACVIB, ROJLUZ, ROKYOH, ROXRUT, SARZOA02, SEFDOW01, SEYVUQ, TEGBEP, TOCKON, UNAXIS, UQIMOY, VEPCOL, XABXUV, XIFCAR, XIZMAW, XOHKOU, XUBPUF, XUQXIQ, YAPBAT, ZEGGOJ, ZEGHEA, ZUCJAK

Kryptoracemate *PSTG* structures **not** marked in the November 2019 CSD as re-interpreted (12).
DAGXUH, EGOXEG, GAGFAX, IHULUU, JEGHUB, JISREI, LUQLUF, LUYMAV,
NITSUH, RUCXAQ, WOKKIT, YUFVAX

Chiral *PSTG* structures **not** marked in the November 2019 CSD as re-interpreted (14).
ACATOQ, CAXBAF, IXENUU, KERVEI, LITQEL, LOPLEJ, MAKJIT, MUTYOP01,
RALHAO, TOJQIU, WUHPOE, XEPNOY, YIQSOI, ZEHFEZ

Achiral structures for which other evidence in the CSD supports the *PSTG* classification (10):
ABORER, CIDHAB, DCHPHA01, EYISA01, FUCTIG03, POQPIX02, SARZOA02,
SEFDOW01, XUBPUF, XUQXIQ

Kryptoracemic structures for which other evidence in the CSD supports the *PSTG* classification
(3):
GAGFAX, JISREI, YUFVAX

Chiral structures for which other evidence in the CSD supports the *PSTG* classification (1):
MUTYOP01

UNIBAW seems to have been described correctly in *P1* but the length of **b** should almost certainly be halved.

3.3 Approximate periodic symmetry is pervasive

Table 1. Types of approximate symmetry

Structures in the inversion category may also include approximate glides, rotations, and/or translation. Structures in the glide category may include rotations or translation but not approximate inversion. Structures in the rotation category may include no other type of approximate symmetry other than translation.

	Number achiral or racemic	Number enantio-pure	Total number	% achiral or racemic (of 147)	% enantiopure (of 1190)	% of all 1337 reliable structures
$P\bar{1}$ plus (distorted or mimicked)	5	19	24	3%	2%	2%
$P\bar{1}$ (distorted or mimicked)	26	353	379	18%	30%	28%
$P\bar{1}?$ (distorted or mimicked)	0	3	3	0%	0%	0%
Glide plus (distorted or mimicked)	8	4	12	5%	0%	1%
Glide (distorted or mimicked)	40	32	72	27%	3%	5%
$2&2_1$ plus	1	8	9	1%	1%	1%
$2&2_1$	6	177	183	4%	15%	14%
2_1 plus	3	17	20	2%	1%	1%
2_1	10	210	220	7%	18%	16%
$P2/p211$ plus	0	1	1	0%	0%	0%
$P2/p211$	0	8	9	0%	1%	1%
$p112$ plus	0	1	1	0%	0%	0%
$p112$	0	23	23	0%	2%	2%
3 plus, 3 , 3_1	0	8	8	0%	1%	1%
Translation	9	49	58	6%	4%	4%
Alternating layers	15	36	51	10%	3%	4%
Hybrid packing	6	72	75	4%	6%	6%
None found	18	169	188	12%	14%	14%
Total reliable structures	147	1190	1456	100%	100%	100%

3.4 The approximate symmetry is often subperiodic

Numbers of structures having approximate n -D periodic symmetry.

	Achiral or Racemic Material			Enantiopure Material		
	3-D	2-D	1-D	3-D	2-D	1-D
Glide or glide mimic	9	35	4	24	11	1
Ratios (# n -D)/(#3-D)		3.9	0.4		0.5	0.0
Screw	6	14	0	140	264	13
Ratios (# n -D)/(#3-D)		2.3	0.0		1.9	0.1
Pure rotation	0	0	0	2	33	1
Ratios (# n -D)/(#3-D)		(NA)	(NA)		16.5	0.5
All	15	49	4	166	308	15
Ratios (# n -D)/(#3-D)		3.3	0.3		1.9	0.1

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

Examples of structures that have approximate 3-D rotational or glide symmetry but at most one angle within 10° of 90°

Approximate $P2_1$ cell but no angle in the $P1$ cell within 10° of 90°

ATEMAO, $Z=4$ 13.58, 13.64, 15.05 Å

$P1 \rightarrow P2_1$, $Z'=1$ **65.3, 65.5, 64.5°** ($\frac{1}{2} \frac{1}{2} 0 \mid \frac{1}{2} \frac{1}{2} 0 \mid 0 0 1$), $\det=\frac{1}{2} \rightarrow 89.8, 119.5, 90.3^\circ$

Approximate *Cc* or *Ia* cell but at most one angle in the *P1* cell within 10° of 90°

DEZLAY, <i>Z</i> =2	5.02, 8.60, 10.21 Å		
<i>P1</i> → <i>Ia</i> , <i>Z'</i> =1	69.0, 77.7, 89.5°	(0 $\bar{1}$ 0 $\bar{1}$ 0 0 1 1 $\bar{2}$), det=2 →	92.3, 94.1, 89.5°
IYIKUW, <i>Z</i> =2	7.82, 9.14, 17.30 Å		
<i>P1</i> → <i>Cc</i> , <i>Z'</i> =1	92.2, 101.4, 113.9°	(0 1 $\bar{1}$ $\bar{1}$ 0 0 1 1 1), det=2 →	92.2, 127.8, 89.2°
KOBQOH, <i>Z</i> =2	5.80, 10.64, 12.96 Å		
<i>P1</i> → <i>Ia</i> , <i>Z'</i> =1	109.4, 102.4, 90.2°	(0 $\bar{1}$ 0 1 0 0 1 1 2), det=2 →	89.6, 94.8, 89.8°
PEMWOU, <i>Z</i> =2	8.03, 10.30, 10.32 Å		
<i>P1</i> → <i>Cc</i> , <i>Z'</i> =1	62.5, 70.8, 81.1°	(0 0 1 0 2 $\bar{1}$ $\bar{1}$ 0 0), det=2 →	89.4, 109.2, 92.5°
QOSLOA, <i>Z</i> =2	7.24, 10.13, 15.94 Å		
<i>P1</i> → <i>Ia</i> , <i>Z'</i> =1	103.2, 96.6, 109.8°	(1 0 0 $\bar{1}$ $\bar{1}$ $\bar{2}$ 0 1 0), det=2 →	90.8, 109.8, 90.2°
ZAMMIM, <i>Z</i> =2	8.88, 11.57, 13.42 Å		
<i>P1</i> → <i>Ia</i> , <i>Z'</i> =1	113.7, 105.5, 97.6°	(1 0 0 $\bar{1}$ $\bar{1}$ $\bar{2}$ 0 1 0), det=2 →	89.8, 97.6, 90.3°

The *Ia* cell, which might have been difficult to find, was recommended by *PLATON*.

Approximate *C2*, *I2*, or *F2* cell but no angle in the *P1* cell within 10° of 90°

FACXIR, <i>Z</i> =4	12.54, 12.60, 13.31		
<i>P1</i> → <i>C2</i> , <i>Z'</i> =½	79.1, 62.0, 65.9°	(½ ¾ ¾ ½ ¼ ¼ 0 ½ ½), det=½ →	91.4, 93.7, 89.9°

A 2:1 sulfate salt so that in *P1* $Z_{eff}=8$ for the cation and in *C2* $Z'_{eff}=1$ for the cation.

GEHJAF, <i>Z</i> =2	5.96, 10.41, 11.74 Å		
<i>P1</i> → <i>I2</i> , <i>Z'</i> =1	108.2, 104.6, 104.1°	(1 1 1 1 0 0 0 1 $\bar{1}$), det=2 →	88.7, 97.3, 87.7°
LITPAI, <i>Z</i> =2	7.43, 13.70, 14.77 Å		
<i>P1</i> → <i>I2</i> , <i>Z'</i> =1	76.1, 77.6, 75.2°	(0 1 $\bar{1}$ $\bar{1}$ 0 0 $\bar{1}$ 1 1), det=2 →	88.0, 95.0, 91.0°
SUQCUE, <i>Z</i> =4	11.37, 12.36, 15.95 Å		
<i>P1</i> → <i>C2</i> , <i>Z'</i> =1	104.2, 100.4, 103.6°	(½ 2 0 ½ 0 0 -½ $\bar{1}$ $\bar{1}$), det=1 →	89.7, 116.0, 90.3°

SUQCUE is a good example of a distorted structure that is also modulated.

In the following examples a nonstandard *F2* cell was chosen because it was more obvious than the standard settings of space group #5 (*C2*, *A2*, or *I2*). The transformation from *F2* to *I2* is (½0½|010|½0½). The transformation to *C2* or *A2* leaves either the **a** or **c** axis unchanged and changes the other to [101]/2 or [10 $\bar{1}$]/2. The cell choice may affect the deviations of α and γ

from 90°. In BEVVEE the $F2$ cell has angles 91.3, 121.7, and 86.8° while the $C2$ cell has angles 87.4, 138.5, and 93.2°.

BEVVEE, $Z=2$ 8.85, 10.20, 11.93 Å

$P1 \rightarrow F2, Z'=1$ **79.9, 87.4, 65.6°** ($1\bar{2}2 | 100 | \bar{1}20$), $\det=4 \rightarrow$ 91.3, 121.7, 86.8°
 Obvious layers (011) (axes \mathbf{b} , $[1\bar{2}2]$; angle 86.8°) have very good approximate symmetry $c211$,
 A view of three layers along the (011) normal reveals the F centering. A C -centered cell is also
 possible; the transformation is $(\bar{1}22|\bar{1}00|001)$, $\det=2$ but then $\beta=138.5^\circ$.

CUGMAT, $Z=2$ 7.21, 7.28, 12.18 Å

$P1 \rightarrow F2, Z'=1$ **104.6, 94.9, 112.7°** ($110 | 1\bar{1}0 | \bar{1}\bar{1}2$), $\det=4 \rightarrow$ 95.9, 91.5, 90.6°
 (The 2 and 2_1 axes are along $[1\bar{1}0]$. The $F2$ description is convincing even given the angle α that
 deviates significantly from 90°.)

KAGNEN, $Z=2$ 7.64, 22.62, 17.78 Å

$P1 \rightarrow F2, Z'=1$ **77.0, 78.4, 74.8°** ($1\bar{2}0 | 100 | \bar{1}02$), $\det=4 \rightarrow$ 90.8, 100.6, 86.1°

YULMOI, $Z=2$ 7.72, 11.90, 14.30 Å

$P1 \rightarrow F2, Z'=1$ **96.9, 103.5, 108.3°** ($120 | \bar{1}00 | 102$), $\det=4 \rightarrow$ 92.2, 102.1, 90.6°

Angles near 90° do not guarantee approximate symmetry

GUHFOE The cell angles are 87.9, 74.9, 86.7° but no approximate symmetry was found.

PANFOA Layer (01 $\bar{1}$) (axes \mathbf{a} , $[122]$, angle 90.5°) has excellent approximate symmetry $c211$. The 3-D cell resulting from the transformation $((122|100|01\bar{1}))$ has angles 91.0, 93.1, 90.5° and approximate 2 and 2_1 axes along the new \mathbf{b} (*i.e.*, \mathbf{a}_{P1}) but no approximate 3-D symmetry (see towards the end of Section 3.5, where the similar ZIYNOW is also discussed).

RIYVIE Angles 93.0, 91.5, and 108.6° but the only symmetry is in the layer (011) (axes \mathbf{a} , $[01\bar{1}]$), which has approximate symmetry $p11a$.

SEZRIY Angles 104.5, 90.2, 92.6° but the approximate 2_1 axes are along \mathbf{b} in layers (001). The angle β of 90.2° has no significance.

XOGWOH Angles 90.7, 109.5, 91.4° but described as a $p2_111$ structure with layers (001) (axes \mathbf{a} , \mathbf{b} ; angle 91.4°). Could also be described as a $P\bar{1}$ mimic.

YOYFAT Angles 90.0, 101.8, and 90.2° but there is no approximate symmetry along or perpendicular to **b**. Layers (001) do have approximate 2_1 axes along **a**.

Layers that are centered and that have an offset along the special axis of ca. $\frac{1}{4}$

The number of structures described as having layers of approximate symmetry $c211$ is 114. If the layer offset in a centered cell along the approximate twofold axis is $0.25+q$ the next-nearest layer (*i.e.*, $\Delta z_{c211}=2$ if **c** relates adjacent layers) will be offset by $0.50+2q$. If q is small (maybe, $-0.02 - +0.02$) then the centering allows molecules in the original and next-nearest layers to be approximately superimposed by a rotation around the special (*e.g.*, twofold) axis (*i.e.*, by a change in the angle β). That rotation angle may be small or quite large, Taking the vector between the two superimposed molecules as the third axis, a cell with the third axis and the two layer axes (switched so that the twofolds are along **b**) has approximate $C2$ symmetry. The superimposed molecules are related in the layer cell by $[uv2]$ (**a** unique; adjacent layers related by **c**; u usually 0, v usually ± 1). The superimposed molecules are related in the new, C -centered cell by $[001]$ (**b** unique; α, γ near 90°).

The length of **c** in the 3-D C -centered cell will be doubled (because it goes to the next-nearest layer) so that Z' will be twice as large as in the layer cell (larger cell but no additional approximate symmetry). There will then be two independent layers because no matter what rotation is needed to superimpose translationally related molecules in the first and third layers the related molecule in the second layer will lie halfway between them.

There is a molecule in the third layer at $[\frac{1}{2}\frac{1}{2}1]_{C2}$ so that there is a molecule in the second layer is at $[\frac{1}{4}\frac{1}{4}\frac{1}{2}]_{C2}$ or $[\frac{1}{4}\frac{1}{4}\frac{1}{2}]_{C2}$. The molecule in the second layer is then be related to those in the first and third layers by translations that do not occur in any monoclinic group. The layer description is then better even though a cell can be found with two angles near 90° . Two examples are described below, but there are many others; 35% of the 114 structures described as having $c211$ layers have layer offsets in the range 0.23 – 0.27 and 46% have offsets in the range 0.22 – 0.28.

PANFOA, $Z=2$ 8.49, 10.24, 16.65 Å Layers (01 $\bar{1}$) are offset along [100] by 0.24.

$P1 \rightarrow C2$, $Z=8$, $Z'=2$ 114.3, 100.9, 103.1° (122|100|01 $\bar{1}$) gives angles 91.0, 93.1, 90.5°

A rotation of 3° around an approximate twofold axis of PANFOA lines up next-nearest layers. The transformation shown gives a C -centered cell with α and γ near 90° , $\beta=93^\circ$, and obvious approximate 2 and 2_1 axes along $\mathbf{b}_{C2}=[100]_{P1}$. That $C2$ cell, however, has $Z=8$ and $Z'=2$ but no additional approximate crystallographic symmetry so that there would be two

crystallographically independent, but very similar, layers $(001)_{C2} = (01\bar{1})_{P1}$. Adjacent layers would be offset by $\frac{1}{4}$ along the two layer axes ($\mathbf{a}_{C2}=[122]_{P1}$ and $\mathbf{b}_{C2}=[100]_{P1}$).

ZIYNAW, $Z=2$ 9.67, 9.49, 6.19 Å Layers (111) are offset along $[0\bar{1}1]$ by 0.24.
 $P1 \rightarrow C2$, $Z=8$, $Z'=2$ 75.0, 74.2, 61.2° $(2\bar{1}\bar{1}|0\bar{1}1|\bar{1}0\bar{1})$ gives angles 87.8, 118.9, 91.4°

A rotation of 29° around an approximate twofold axis of ZIYNAW lines up next-nearest layers. The transformation shown gives a *C*-centered cell with α and γ near 90°, $\beta=119^\circ$, and obvious approximate 2 and 2₁ axes along $\mathbf{b}_{C2}=[0\bar{1}1]_{P1}$. That *C2* cell, however, has $Z=8$ and $Z'=2$ but no additional approximate crystallographic symmetry so that there would be two crystallographically independent, but very similar, layers $(001)_{C2} = (111)_{P1}$. Adjacent layers would be offset by $\frac{1}{4}$ along the two layer axes ($\mathbf{a}_{C2}=[2\bar{1}\bar{1}]_{P1}$ and $\mathbf{b}_{C2}=[0\bar{1}1]_{P1}$).

Drawings of PANFOA and ZIYNAW illustrating this situation are available.

The only centered layer groups having glide planes but no rotation axes also have mirrors. No *P1*, $Z>1$ structure having an approximate mirror was found.

3.6 R. E. Marsh was right

If the *PSTG* structures are included then the percentage of chiral *P1*, $Z=2$ structures that are approximately centrosymmetric goes from 33% to 32% for $Z=2$ structures and 31% to 30% for all $Z>1$ structures.

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

While 35% of the structures of enantiopure material mimic either inversion (32%) or glide (3%) symmetry, 38% (453) have approximate screw and/or rotation axes instead: 18% have a single set of parallel screw axes, 15% have parallel 2 and 2₁ axes only, and 3% have a single set of parallel rotation axes. In the rest of the 453 the approximate screw or rotation is accompanied by an approximate translation (17 structures) or by additional rotational symmetry in the perpendicular directions (13 structures).

Enantiopure material; approximate rotation or screw axis accompanied by an approximate translation (17) (approximate rotational symmetry is 3-D except as noted)

AJOTID, ATEMAO, FACXIR, HAMFOR, IYAVEK, KILGOE (2-D), LAQFOZ, MAYKOP (2-

D), SUQCUE, TIVCEH (2-D), TORHAK (2-D), VIQCIK, WEYZOQ, WIPTAT, XEVCEJ, XODGAA, YONVOM (2-D)

Enantiopure material; approximate higher symmetry (13)

AFONQY ($P2_12_12_1$), AZALIY ($p321$), AZALOE ($p321$), BIHVUJ ($p2_12_12$), BIVLIC ($p222$), DAYREA06 ($P2_12_12_1$), DOVXIW ($p2_122$), IHOXUY ($P2_12_12$), QACCII ($P622$), QEFRAX ($pb2_1m$), RAKBUC ($p2_12_12$ and also an approximate translation), SOXMAW ($c222$), UJADOZ ($p2_122$)

The 20 reliable structures of achiral or racemic material that have rotational symmetry only are in the categories:

$2\&2_1$ (7) (2-D $c211$ except as noted)

BADSAB, DAKBIC, DMTCUN10, ITAVOQ, MMANCN, PEQGAU,

ZEBVEJ (3-D; approximate $P2_12_12$)

and

2_1 (13) (2-D except as noted)

ACEHEX (3-D), CODRES (3-D), FABDIY, FINMIA, GUZVIG, IKAHOS, JULTUF01 (3-D;

also a translation), LAFHEH, MEFJIS, NORMET, QOPVAU (3-D, and also a translation),

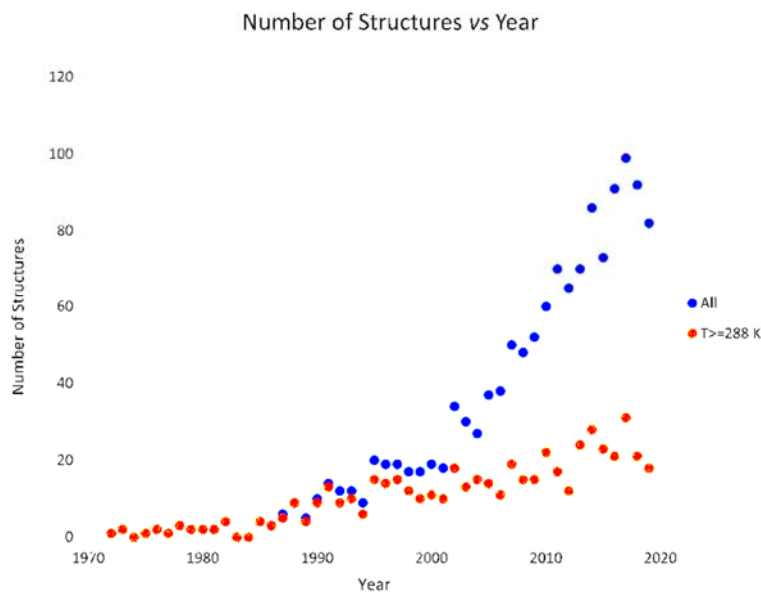
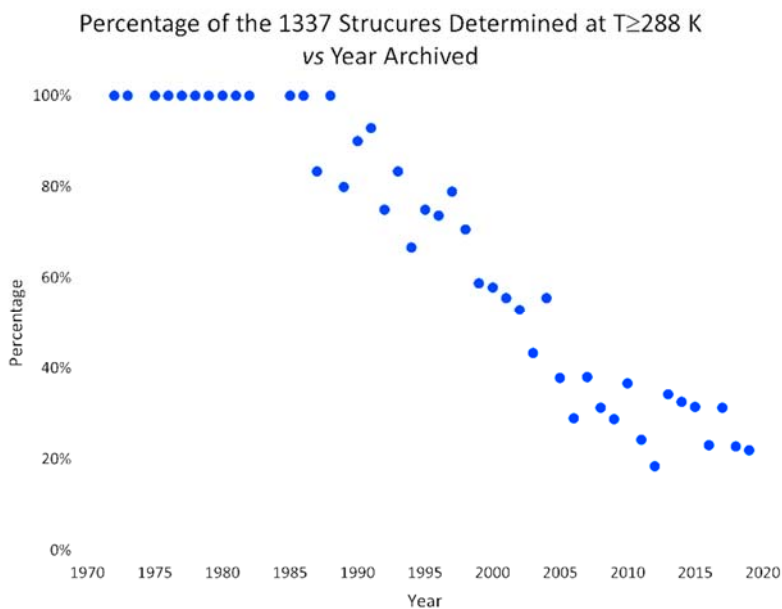
WUKSOL (3-D, and also a translation), ZODWOE

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

A temperature given as “RT” in a *ConQuest* or *Mercury* hit indicates a temperature in the range 283 – 303 K. The actual temperature, however, is almost always shown if the CSD entry is exported to a spreadsheet. The three structures for which no specific temperature was given are HEDQEP, JEGHEL, and JEGHUB; their temperatures were entered as 293 K.

In the list of 1337 reliable, $P1$, $Z>1$ structures there are only three with $273 < T < 290$ K: QUBPIN at 283 K, JIVNEI at 288 K, and GIFDIL at 289 K. QUBPIN was determined at a temperature not far above the beginning of its transition to a very similar structure in which the approximate $pb2$ layers are offset differently; for the other two structures no information about the temperature is given in the papers. Consequently the cutoff used for compiling statistics was 288 K (15° C).

No structure investigated in this study was determined at low temperature before 1987. After 1987 the percentage of low-temperature structures increased approximately linearly, but even very recently more than 20% of the deposited structures were determined at a temperature above 288 K. Of all 1337 structures 39% were determined at ≥ 288 K.



More information about the structures determined at ≥ 288 K is given in the following table. Structures in the categories *alt layers*, *hybrid packing*, and *none found* were excluded from that

table because it seems very unlikely that there is a very similar structure with a smaller or higher symmetry unit cell. The structures of enantiopure material that mimic inversion or glide symmetry were excluded for the same reason.

Approximate symmetry	Achiral or Racemic Material			Enantiopure Material		
	All	Number with $T \geq 288$ K	% with $T \geq 288$ K	All	Number with $T \geq 288$ K	% with $T \geq 288$ K
inversion	31	6	19%			
glide	48	23	48%			
screw	20	10	50%	417	161	39%
rotation	(none)			36	14	39%
translation	9	4	44%	49	11	22%
All	108	42	39%	502	186	37%

***P1* structures known to be related to a higher symmetry phase by a reversible transition during which the crystal remains single**

For eight substances it has been shown that the *P1* structure is related by a reversible transition to a phase at higher temperature in which approximate symmetry is crystallographic. For one more the two very closely related phases were published independently. The space groups of the nine higher-temperature phases are split nearly evenly among $P\bar{1}$ (3), groups with glides but no inversion (2), *P1* with $Z_{HT}=Z_{LT}/2$ (2), and Sohncke groups other than *P1* (2).

CAZLAR02 (co-crystal of TTF and tetrabromo-*p*-benzoquinone) The *P1*, $Z=2$ structure studied at 20 K transforms at *ca.* 53 K to a $P\bar{1}$, $Z'=1$ structure. Seems to have been very careful work.

EJUQIK (4-ethylanilinium hydrogen (2*R*,3*R*)-tartrate) The *P1* structure has $Z=4$ at 123 K and $Z=2$ at RT. A phase transition at *ca.* 186 K was documented. The layer (001) of achiral cations has approximate inversion symmetry but the layer of the hydrogen tartrates has

translational symmetry only, The transformation matrix ($\bar{1}10|\bar{1}\bar{1}0|001$) relates the $Z=4$ phase to the $Z=2$ phase.

FUGMIF (piperazin-1-ium perchlorate) The structure is $P1$ at 200 K; in the very similar $P\bar{1}$ structure at 413 K the ClO_4^- ions are disordered. Superstructure reflections were observed at 200 K.

JAKKUB01 (1:1 2-propanol solvate of $\text{C}_{18}\text{H}_{25}\text{O}_8\text{PS}_3$) The $P1$, $Z=2$ phase studied at 85 K transforms to a $P1$, $Z=1$ phase at *ca.* 90 K; the transformation matrix is ($\frac{1}{2}\frac{1}{2}\frac{1}{2}|010|100$). The modulation is quite subtle but the work seems to have been careful.

NAKNEV [$\text{DABCOH}_2^{2+}\cdot 2(\text{ClO}_4^-)\cdot \text{H}_2\text{O}$] The $P1$, $Z=4$ structure transforms in the range 230–250 K to a $Pca2_1$, $Z'=1$ structure. Seems to have been very careful work.

NAKNOF [$\text{DABCOH}_2^{2+}\cdot 2(\text{BF}_4^-)\cdot \text{H}_2\text{O}$] The $P1$, $Z=4$ structure transforms in the range 230–250 K to a $Pca2_1$, $Z'=1$ structure. Seems to have been very careful work.

QOPVAU (co-crystal of phenazine and choranic acid) The $F1$, $Z=8$ (or, $P1$, $Z=2$) structure studied at 100 K transforms at *ca.* 137 K to a $P2_1$, $Z'=1$ phase (*e.g.*, MAMPUM06, MAMPUM07). In the transition to QOPVAU **b** of MAMPUM is doubled and the 2_1 axis becomes approximate. At RT the structure is $P2_1/n$, $Z'=1/2$. Seems to have been very careful work.

RIDFOA01 (co-crystal of 5,5'-dimethyl-2,2'-bipyridine and iodanic acid with transfer of one proton). The $P1$, $Z=2$ structure at 50 K transforms at *ca.* 270 K to a $P\bar{1}$, $Z'=1$ structure (RIDFOA). (The temperature is given as 50 K in the paper but as RT in the CSD. The cell volume for RIDFOA is 2.5% greater than the volume for RIDFOA01.)

POPGUW01 (histidine acetate dihydrate) The $P1$, $Z=2$ structure at 123 K is very similar to the $C2$, $Z'=1$ structure POPGUW at RT that was published by different authors. The transformation from the $P1$ to the $C2$ cell is ($011|0\bar{1}1|100$), $\text{det}=2$, the transformation gives angles 90.4, 94.1, and 89.8°.

Note:

The possibility that the structures of NAKNEV03/04 (150 and 200 K) and NAKNOF03/04 (150 and 220 K) should have been described in Pc with $Z'=2$ rather than in $P1$ with $Z=4$ must be considered. *PLATON* recommends Pc for the 150 K structures and $Pca2_1$ for the higher-T structures, but the paper reporting the four structures says:

The solution and refinements of these structures in the higher symmetric monoclinic space groups Pc and $P2_1$ resulted in models with much higher reliability R -factors, and with the

cations disordered even at 150 K, where both crystals should be well-ordered, according to the calorimetric data (see Fig. 1 and S1†).

The study of these crystals seems to have been both very careful and very thorough but it is not possible to see any deviations from Pc symmetry even for the structures determined at 150 K. The averaged model output by *PLATON* shows no disorder. On the other hand the atomic ellipsoids in the $P1$ structures are all very reasonable and the CIF shows that no restraints were applied. For the structures at 200 (NAKNEV04) and 220 K (NAKNOF04) *PLATON* suggests $Pca2_1$ at the 81% (NAKNEV04) and 86% (NAKNOF04) levels although thermodynamic data demonstrate that a phase transition took place. The averaged $Pca2_1$ structure of NAKNEV04 is ordered except for one H atom on the water molecule and looks normal except for one H atom on the $DABCO^{2+}$ ion. In the averaged $Pca2_1$ structure of NAKNOF04 one F atom of the two BF_4^- ions is disordered over two sites 0.7 Å apart and one H atom on a C atom of the $DABCO^{2+}$ ion is disordered. In the end it was decided to accept the space-group assignment for NAKNEV03/04 and NAKNOF03/04 because the work seems to have been so careful.

There is a similar problem with BOCKEK06 (140 K), which is the FHF^- salt of $DABCO^+$, which is $Cmc2_1$, $Z=1$ (m imposed) above *ca.* 150 K. The structure was reported in $P1$ but no deviations from Pc symmetry could be spotted. Since neither the paper nor its supplementary material addressed the $P1/Pc$ ambiguity BOCKEK06 was marked as *PSTG*, but with regret because the work seemed to have been done very carefully and the displacement ellipsoids are unremarkable.

3.10 Some structures are composed of different layers that alternate

The *alt layers* description was applied to 15 crystals of achiral or racemic material and 36 crystals of enantiopure material. In the lists that follow the refcode is starred if the different layers are composed of different ions or, in one case, of the different molecules of a co-crystal.

All layers are approximately the same but their orientations differ (14):

ACEZOX, CEHYEW, CICTIT, COVBEV, ECEHAY, EGOZAE, EGUWUB, GIPLID, HASYUX, HINKUN, HOCYEG, UMAHAT, WAZSAT, ZZZVXQ06

The two layer types have the same composition but different approximate periodic symmetries (4):

EVOREM, IBIJAF01, QACXUN, TOJRAM

The two layer types have the same composition but one has approximate periodic symmetry while the other does not (13, 10 of which have $Z=3$):

EYUJIT, FINTIF, FONVUC, FUXYUS, GUVROF, IDOQEA, IFAFUS, PONTAN, RICHUG, SEZPUJ, VEHLAY, XAGMOH, XURWEO

The two layers have different compositions but the same approximate periodic symmetry (1):
XAFPEB*

The two layers have different compositions and different approximate periodic symmetries (8):
GADPUX, HUZDOV*, TOLRUK*, WAWLEM*, WIYSAZ, YEMCUO01*, ZIKDEC*, ZINPIY* (the only co-crystal)

The two layers have different compositions; one has approximate periodic symmetry but the other does not (9):

ACUBAC*, CIFGEE, HIPKAS, IMUBAV*, JUFVAJ*, OCANUD*, OFAXUP*, QQQETG01*, WUYXIX*

Neither layer type has approximate periodic symmetry but there is obvious approximate local symmetry (2):

ADUNAP, SURYEK

Layers connected by H bonds

Special attention was given to structures in which alternating layers are linked by hydrogen bonds.

Z=3 structures in which there are chains of H bonds that connect the two layer types

ADUNAP The three molecules form an H-bonded chain along **c**. Molecules #2 and 3 are very well related by a 2 along **c*** plus a translation that suggests a local 2₁ axis but the next molecule in the chain is #1, which has an unrelated orientation and is in the next layer.

EYUJIT Three molecules plus a water molecule form a helical H-bonded chain along **c**. Molecules #1 and 3 are in a layer (010) that has approximate symmetry *c*211 but the orientation of #2 is unrelated.

FINTIF Layers (001) of molecules #1 and 3 alternate with layers of #2. An approximate 2₁ axis along **b** relates #1 and 3 but γ is 82.4° so that the approximate 2-D symmetry, while easily recognized, is borderline. Each molecule participates in two H-bonded chains that run along **a**.

HIPKAS This 3:2 hydrate of D-glucitol has a 3-D pattern of H bonds. Molecules #2 and 3 and the water molecules are in a layer (001) (axes **b**, **a**; angle 95.9°; $-0.4 < z < 0.3$)

that has good approximate symmetry $p211$. Molecule #1 is in a layer by itself that has translational symmetry only. The orientation of molecule #1 is related to that of #2 by an approximate twofold rotation around $[011]$.

- RICHUG Layers (001) (axes **b**, $[210]$; angle 84.0° ; $0.0 < z < 0.8$) of molecules #1 and 2 alternate with layers of #3. The layer of #1 and #2 has approximate symmetry $c211$ except for the orientation of one OH group. The layer of molecule #3 ($0.9 < z < 1.0$) has translational symmetry only. There are five H bonds within layers and one between layers.
- VEHLAY Layers (010) (axes **c**, **a**; angle 89.2° ; $0.7 < y < 1.2$) of molecules #1 and 3 alternate with layers of #2 ($0.3 < y < 0.7$). The layer of #1 and 3 has very good approximate symmetry $p2_111$. The layer of #2 has translational symmetry only. The two layers are not so well separated but most H bonds are within the layers, and the few bonds between layers are all mediated by water molecules.

Z=2 and, Z≥4 structures in which there are chains of H bonds that connect the two layer types

- ACUBAC There are many strong H bonds in the layer (001) of sulfate ions and water molecules that has translational symmetry only. The large cations form one H-bond each to the anion/H₂O layer and are related to each other by very good approximate inversion symmetry.
- EGOZAE Complicated $Z=6$ structure. There are two H-bonded linear trimers (#1,4,6 and #3,2,5). Layers (001) of #1-4 and of #5&6 have approximate symmetry $p2_111$ (axes **b**, **a**; angle 93.1°). Layer of #1-4 also has an approximate translation along the layer normal. The two layers are related by an approximate twofold axis along **a**.
- EVOREM A sterol. Layers (001) of #1&4 (axes **a**, $[\bar{1}20]$; angle 92.9° ; $0.8 < z < 1.2$) that have good approximate symmetry $c211$ alternate with layers of #2&3, which mimic inversion symmetry. Layers are linked by helical chains of H bonds that include the MeOH solvent molecules.
- GADPUX Strongly H-bonded layer (001) ($0.0 < z < 0.8$) containing the achiral anions and the two chiral NHPPhMeNH₃⁺ cations mimics approximate symmetry $p\bar{1}$. The neutral NHPPhMeNH₂ molecule (#5) is a little out of the layer although H-bonded to it. The layer of #5 has translational symmetry only.
- IBIJAF01 Molecules of 6-phenyl-1,3,5-triazine-2,4-diamine are in two layers (001). The layer of molecules #1 and 2 has approximate symmetry $p\bar{1}$. The layer of molecules #3 and 4 has either an approximate 2_1 along **a** or an **a**/2 glide with

mirror (010)] but that layer has approximate 1-D symmetry only ($\gamma=78.3^\circ$).
Layers are connected by longish NH...N bonds.

- IMUBAV 2:1:1 co-crystal of S-brucine and 1,1'-binaphthalene-2,2',7-triol that includes MeOH. Layer (001) (axes [110], [$\bar{1}\bar{1}0$]; angle 82.1° ; $0.3 < z < 1.0$) of brucine has obvious approximate symmetry $c211$ even though the angle is far from 90° ; the binaphthalene layer has translational symmetry only. The two layers are connected by two OH...N bonds and one OH...O= bond.
- OCANUD 2:1:5 hydrated sulphate salt of a large cation. Layers (001) of the achiral cations ($0.5 < z < 1.0$) alternate with a layer containing the anion and the water molecules. The cation layer has very good approximate inversion symmetry; the anion/water layer has many strong H bonds but no approximate symmetry. The layers are connected by H bonds.
- QQQETG01 Sodium thyroxine pentahydrate. The two thyroxine anions mimic an inversion relationship. The layer of Na^+ ions and water molecules has no symmetry other than translation. The two layer types are connected by H bonds and are not well separated but a layer description seems preferable to a description of hybrid packing. (The two sodium thyroxine pentahydrate units are formulated in the CSD as one because of the Na-O bonds).
- TOLRUK Layers (001) (angle 89.0°) are connected by H bonds but the bonds in the anion/solvent layer are stronger than the bonds between layers. Cation layers have very good approximate 2_1 axes along **b**; the anion/MeOH layers have approximate inversion symmetry.
- WAWLEM The 1:1:3 hydrated salt brucinium dihydrogen citrate. The layer (001) (axes [$\bar{1}\bar{1}0$], [110]; angle 87.2° ; $0.5 < z < 1.0$) of the brucinium ions has excellent approximate symmetry $c211$. The layer of citrates and water molecules has approximate 2_1 axes along **a** but $\gamma=81.6^\circ$ so the approximate 2_1 symmetry is best described as 1-D.
- WUYXIX 4:2:3 hydrate of melaminium selenate. Alternating layers (001). The cation layer ($0.5 < z < 1.0$) has an approximate translation **a**/2 and approximate inversion centers. The layer of anions and water molecules has no approximate symmetry. There are H bonds within and between the layers.
- XAFPEB (Although it is formulated as a 2:1:1 tartrate salt that also includes a tartaric acid molecule, the structure could be a 1:1 hydrogen tartrate salt. Scant crystallographic information is available.)
Layers (001) of cations and of $\text{C}_4\text{H}_n\text{O}_6$ units alternate. The cation layer (axes **a**,

$[1\bar{2}0]$; angle 88.4° ; $0.7 < z < 1.4$) has excellent approximate symmetry $c211$. The tartrate/tartaric acid layer ($0.4 < z < 0.7$) has less good approximate symmetry $c211$ but the axes are **b** and $[2\bar{1}0]$ (angle 84.1°). Layers are connected by H bonds but there are more and stronger H bonds within the tartrate layer.

ZINPIY Co-crystal of achiral 5-fluorocytosine and proline. Components are segregated in layers (001); there are many H bonds within the layers but only one between them. The 5-fluorocytosine layer (axes $[1\bar{1}0]$, $[110]$; angle 95.1° ; $0.0 < z < 0.5$) has good approximate symmetry $c211$; the proline layer has less good 1-D twofold axes along **b**.

3.11 Some structures are best described as having hybrid packing

Types of structures with hybrid packing

In most structures classified as having hybrid packing the molecule has two regions that pack differently, with the different regions segregated into layers that alternate. Many of those molecules have an achiral part that packs with approximate inversion and a part with stereocenters that packs with approximate 2_1 axes or in a centered array with both 2 and 2_1 axes. Examples are:

AXEKAP, CAKQUC, EGOTOK, IXOJAG, PAXNIL02, POXDUB, TOQWAZ, UMAPUT, UQISIY, VOMFOS, WAXLEM, YACTEC01, YITMAR, YUNTAD

In a few structures the chiral part of the molecule lies between two achiral parts, *e.g.*:

BINHEO There is one stereocenter in the central 4-oxo-2,3,4,5-tetrahydro-1,5-benzothiazepine unit of the $C_{30}H_{25}N_3O_4S$ molecule. The (2-nitrophenyl)methyl groups at one end of the two molecules are related by a **b**/2 glide with mirror (001). The N-benzylbenzamide groups at the other ends are related by a $[110]/2$ glide with the same mirror. The central fused rings are in two different layers in which they related by translation only.

SIYLOE The 1H-isoindole-1,3(2H)-dione fragments are related by good approximate inversion centers as are the 4-fluorophenyl ring of one molecule and the phenyl ring of the other, but the three sets of approximate inversion centers do not make a periodic array. Local twofold axes along $[201]$ relate pairs of molecules.

Salts and co-crystals in which the different components pack in different layers having different symmetry can often be described either as having alternating layers or as having hybrid packing.

The choice was made based on the extent of H-bonding between the two layers and the differences between the two layers. Examples of the difficulties associated with distinguishing between the two categories are:

FINTIF Classified as *alt layers* because $Z=3$ and there are layers (001). In one layer two of the molecules are well related by a 2_1 axis along **b**; the other layer has translational symmetry only. On the other hand the layers are well connected by strong H bonds.

HURCIH Classified as *hybrid packing* because of the large number of H bonds between the cation and anion layers. HURCUT is very similar; it has the same anion and a (disordered) cation that differs from the cation in HURCIH by the substitution of CH₂OH for CH₃.

IBIJAF01 Classified as *alt layers* because the four independent molecules are in two layers that are clearly different. The layers are, however, connected by NH...N bonds.

QACXUN Classified as *alt layers* because there are two different layers, only one of which has hybrid packing.

QUSQOL Classified as *hybrid packing* because two large molecules and one 4,4'-bipyridine form H-bonded units.

A few structures that could have been described as having *hybrid packing* were classified as something other than *alt layers*:

CAKHIF C₁₅H₂₀N₂O₆·H₂O. Classified as 3-D 2_1 because of an obvious 2_1 relationship along **b** and because α and γ are 91.8 and 92.1°. The two Ph rings (6 of 24 non-H atoms) are parallel and so related by approximate inversion but overall CAKHIF looks like a $P2_1$ structure.

GIGNOB C₁₇H₁₉NO₅. Classified as 2-D 2_1 because the only inversion-related groups are the -C(=O)Me units and that relationship is quite imperfect. The configuration of the N atoms to which those groups are attached does, however, differ between the two molecules.

And some structures were classified as having *hybrid packing* when they might have been described in some other way other than as *alt layers*:

KUKZEX A hydrated salt. Layers (001) of the two ion types alternate but the structure was classified as *hybrid packing* because there are H-bonded chains along **a** containing all three components. The cations are well related by inversion but the anions are related by a $[210]/2$ glide with mirror ($\bar{1}21$), *i.e.*, a mirror that is neither within nor perpendicular to the layer.

PAXNIL02 C₈H₁₀N₂O₆ [cyclo-(L-aspartyl-L-aspartyl)]. All atoms other than those in one of the two COOH groups lie in a layer having excellent approximate symmetry $c211$. The two COOH groups that do not fit the $c211$ description are related by inversion. The choice of *hybrid packing* over a $C2$ or $c211$ description was based on the obvious difference in the approximate

symmetries of the two regions.

No claim is made that the classification of borderline structures is completely consistent. JIGTEY and PAXNIL02 have similar features but JIGTEY was classified as $2\&2_1$ while PAXNIL02 was labeled as having *hybrid packing*.

Salts and co-crystals (i.e., structures having more than one component with >7 non-H atoms) having *hybrid packing* (approximate 2_1 and $\bar{1}$ unless otherwise specified):

Achiral components (1):

KUKZEX (glide, $\bar{1}$)

At least one chiral component (11):

CELWIB04 ($\bar{1}$, none), DIZSOX (2_1 , $\bar{1}$), EJUQIK01 (translation, $\bar{1}$), QUSQOL ($2\&2_1$, $\bar{1}$), TEFMEW01 (none, $p112/b$). UBAMES (2 , 2_1), UYOTEI, UYOTOS (none, $\bar{1}$), VEVMUH, VEVMUH01 (2_1 , glide), YIZGAP (none, $\bar{1}$)

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

Of the 1115 $Z=1$, $R\leq 0.050$ structures 61 were discarded as incomplete (1), a duplicate (54), too inorganic (4), or for other reasons (2). The number examined was then 1054 of which 49 had already been investigated as $Z_{eff}>1$ structures so that the number of new $Z=1$ structures was only 1005.

Those 1005 structures were then scanned to find the ones composed of molecules that could have symmetry in a crystal. This admittedly subjective assessment found 236 candidates. They were labeled with the symmetry that seemed most likely. Molecules that could have either symmetry $\bar{1}$ or 2 were labeled with inversion because it is so much more likely to be retained in the crystal. Similarly, molecules that could have either symmetry 2 or m were labeled as most likely having twofold symmetry. The possible molecular symmetry in the 236 identified structures is roughly evenly split between the point groups $\bar{1}$, 2 , and m . (The labels did not influence the investigation of approximate crystal symmetry.)

Of those 236 structures 25 were classified as *PSTG*, with 24 of the 25 being achiral. Of the 24 achiral structures 21 should probably have been refined in $P\bar{1}$; the others are DANGOR

(probably $C2/m$), KECYBU17 (probably $I4/m$), and SOLTAR (probably Cm). The one chiral $PSTG$ structure is MIGREC (probably $C2$).

[Of the 49 $Z_{eff}>1$ structures, 2 (4%) were classified as $PSTG$.]

Of the 236 structures 129 are achiral while 107 have resolvable enantiomers. The fractions of $PSTG$ structures for the two groups are then $24/129=0.19$ and $1/107=0.01$; the overall fraction of $PSTG$ structures is $25/236=0.11$. The $PSTG$ fraction for all $P1$, $Z=1$, $R\leq 0.050$ structures, however, is much lower because a $Z=1$ structure cannot be $PSTG$ unless $Z_{eff}>1$ or the molecule can lie on a symmetry element. The $PSTG$ fraction for all $P1$, $Z=1$, $R\leq 0.050$ structures is then $0.025(1)$ whether or not the $Z=1$, $Z_{eff}>1$ structures are included [$25/1005$ or $(25+2)/(1005+49)$].

For the 236 $Z=1$ structures the ratio of achiral to chiral molecules is $0.99(=105/106)$ if the $PSTG$ structures are excluded. That ratio is much higher than the ratio of $0.12(=147/1190)$ found for the Z and $Z_{eff}>1$ structures.

The possible molecular symmetries are distributed among the point groups as follows:

$\bar{1}$	80 = 32 that could have inversion symmetry plus 48 possible mimics
$2mm$	3 (BOHGUC, IFEKEL, PMNTBZ01)
2	90 = 85 plus 5 possible mimics (CUWBIG, FAZQAY, FIRHUM, PAPLAU, SITFEH)
m	61 = 59 plus 2 possible mimics (BAPGIK, VEVNIW)
3	2 (FESCOW01, MOKSUC)

Molecules that could lie on exact or approximate inversion centers:

The 32 compounds that could have exact inversion symmetry were categorized as follows:

22	$PSTG$ (20 $P\bar{1}$, 1 each $C2/m$ and $I4/m$)
5	distorted $P\bar{1}$ (all determined at ≤ 193 K)
3	$P\bar{1}$ mimic NAWYIV, PAPDER, VEYXAY [Symmetry is broken by the included solvent molecule(s)]
1	no approximate periodic symmetry found (TUVDOE)
1	conformation observed is asymmetric [CANXIA; $(O_2N)(HOCH_2)_2CC(CH_2OH)_2(NO_2)$]

The 48 compounds that could mimic inversion symmetry are divided as follows:

1	$PSTG$ (KARKAP, probably disordered in $P\bar{1}$)
42	$P\bar{1}$ mimics
2	alternating layers (GUQKEK, GUQKOU)

- 1 $p112$ (MOLMUV, which can also be seen as a $P\bar{1}$ mimic)
- 2 hybrid packing (IMOMAA, REBNAP)

Molecules that could have symmetry $2mm$

BOHGUC and IFEKEL are essentially isostructural; the large molecule, the NMe_4^+ ions and included MeCNs are the same but the anions (Br^- and NO_3^-) differ. Both structures have approximate $p211$ layers. The third structure, PMNTBZ01, is of pentamethylnitrobenzene, in which there is a close-packed layer ($11\bar{1}$) with very good approximate symmetry $p11m$.

Molecules that could mimic symmetry 2

PAPLAU and SITFEH are $C2$ mimics. CUWBIG and FAZQAY have approximate $c211$ layers. FIRHUM has no approximate symmetry.

- | | |
|--------|--|
| PAPLAU | Has a good twofold axis but for the difference between fused C_4O and C_5O rings |
| SITFEH | A twofold would exchange N atoms and CH groups in the central pyrazine ring |
| CUWBIG | The molecule has a good twofold axis but for the difference between naphthyl and 1H-benzimidazole ring systems |
| FAZQAY | There is a good approximate twofold but for the kink in the $-\text{O}-\text{P}(\text{NMe}_2)-\text{O}-$ group |
| FIRHUM | A 2:1 malate salt; the anion could have an approximate twofold that would relate H and OH |

Molecules that could mimic symmetry m

BAPGIK and VEVNIW are both Im mimics.

- | | |
|--------|--|
| BAPGIK | Has good mirror symmetry but for the difference between an O atom and a CH_2 group |
| VEVNIW | Has good mirror symmetry but for the difference between NH_2^+ and CH_2 groups |

Molecules that could have symmetry 3

- | | |
|----------|--|
| FESCOW01 | The authors report that the $P1$, $Z=Z'=1$ structure at RT transforms without loss of crystallinity to a $P3$, $Z=3$, $Z'=1=3(\frac{1}{3})$ structure at 30 K |
|----------|--|

MOKSUC The cation has very good threefold symmetry but the axis does not point in a simple direction $[uvw]$

Structures of achiral molecules that can only mimic inversion symmetry (14 of 48 $\bar{1}$ mimics; all structures are ordered and were classified as $P\bar{1}$ mimics unless otherwise specified)

Structures in which the 1 conformation corresponds to a low-energy transition state (2)

FAWMAU $(HO)R_2Si-O-SiR_2(OH)$, $R=1$ -naphthyl. The observed Si-O-Si angle is 168° and the ellipsoids look normal. The Si-O-Si group is very flexible but the linear conformation that would allow inversion symmetry has a higher energy than does a bent conformation.

OGUQAK $(HO)_2RSi-O-SiR(OH)_2$. The observed Si-O-Si angle is 132° and the ellipsoids look normal. The Si-O-Si group is very flexible but the linear conformation that would allow inversion symmetry has a higher energy than does a bent conformation.

Structures in which only one conformational enantiomer is present (3)

GOFTIF No stereocenter; would have inversion symmetry but for the one NO_2 substituent on the central phenyl ring.

PATVIR No stereocenter; racemization would take place through a conformation with mirror symmetry. The paper and its supplementary material indicate that the enantiomers are conformers.

XEBWEJ No stereocenter. A surprisingly good $P\bar{1}$ mimic except for the disulfide group at the center of the long, flexible molecule.

Structures in which the noncentrosymmetric part of the molecule is in a region that has better approximate mirror symmetry than inversion symmetry (9)

GOCLIV The 9-methylanthracene molecule in this co-crystal is planar but not centrosymmetric. The other molecule is centrosymmetric.

GUGHOH The Me and lone pair that would be superimposed by inversion are both in the mirror plane.

PAXTIS Both residues are planar. The bipyridine has approximate symmetry $2/m$. In the other molecule the approximate $2/m$ symmetry is broken by the difference between the *para* COOH and B(OH)₂ substituents, which form similar sets of H bonds.

TOHLOU The 4-phenylpyridinium cation has $2/m$ symmetry but for the CH/NH₂⁺ difference but the atomic ellipsoids are unremarkable. The possible $2/m$ symmetry of the

benzene-1,2,4,5-tetracarboxylate anion is broken by the rotation of one COOH group.

VIHTUD Hydrated ammonium salt of a planar anion. The approximate inversion relates pairs N/CH, COOH/COO⁻ and NH₄⁺/H₂O.

WIKRUE Planar molecule that would have inversion symmetry but for the difference of a CH₂ group and an O atom near the center of the molecule.

XAVHAF 6,13-bis[(*i*-Pr)₃CC]pentacene in which two of the pentacene CH groups are replaced by NH groups.

XAVHEJ 6,13-bis[(*i*-Pr)₃CC]pentacene in which two of the pentacene CH groups are replaced by N atoms.

KARKAP Acridine 1,2,4,5-benzenetetracarbonitrile co-crystal of two planar molecules. A re-refinement in $P\bar{1}$ by R. E. Marsh indicated 50:50 N/CH disorder in the acridine molecule. Classified as *PSTG*.

Structures in which the inversion symmetry is broken by an included solvent (5; all structures were reported as ordered; in this study all were classified as $P\bar{1}$ mimics unless otherwise specified)

NAWYIV The inversion symmetry is broken by the included solvent molecules CHCl₃ (occupancy 0.25) and H₂O (occupancy 0.5). The solvents separate molecular layers (001). The ellipsoids are not suspicious.

PAPDER The included, ordered cyclopent-2-en-1-one is the acceptor of a good H bond from the larger, centrosymmetric molecule.

REVNIQ The inversion symmetry is broken by the one included toluene. The solvent was reported to be ordered but its atoms were refined isotropically. A number of ellipsoids are very eccentric. REVNIQ was deposited as a *CSD Communication* so no crystallographic details available. It was concluded that the choice of space group *P1* was insufficiently justified and so REVNIQ was classified as *PSTG*.

Structures in which approximate inversion relates a small ion and a water molecule

VEYXAY The centroid of the EDTA monocation is separated from the centroid for the Br⁻ ion and the water molecule by 0.001, 0.022, -0.012, which differs from 0, 0, 0 by 0.005, 0.15, 0.13 Å. No displacement ellipsoids are available but if there were significant Br⁻ disorder the *R* factor would almost certainly have been greater than 0.050.

VIHTUD The centroid for the two 5-carboxypyridine-2-carboxylate ions is separated from the centroid for the NH₄⁺ ion and the water molecule by 0.540, -0.054, 0.004, which differs from

$\frac{1}{2}, 0, 0$ by 0.15, 0.33, 0.04 Å. The displacement ellipsoids for the C₅N ring atoms show no sign of N/C disorder. The ellipsoid for the water O atom suggests possible minor disorder, either in its position or in a switch with the NH₄⁺ ion.

4. Discussion

4.2 Comparisons with previous work

Papers by Ricardo Baggio

Baggio (2019) reported analyses of 18 $Z'=4$ structures, of which 3 were reported in space group $P1$. Two (CUNBET, LOPLEJ) were found by both Baggio and Brock to have too-good pseudosymmetry ($P2_1$, $Z'=2$ and $P2_12_12_1$, $Z'=1$, respectively). MAYKOP was found by both authors to have an approximate translation $[011]/2$ (or, equivalently $[01\bar{1}]/2$) and an approximate 2_1 axis along **a**. Brock found MAYKOP to have layers (011) with axes $[100]$ and $[01\bar{1}]/2$ and approximate symmetry $p2_111$.

Baggio (2020) reported having studied *ca.* 300 $Z'=4$ structures of which 41 are described in either the text or the supplementary material. He found that *ca.* 8% of the structures investigated have good 3-D pseudosymmetry. The only $P1$ structure documented is RIZGIR. Baggio and Brock both found RIZGIR to be a slightly distorted $P2_1$, $Z'=2$ structure.

Paper by Toms Rekis

Rekis (2020) considered organic, $2 \leq Z \leq 6$ Sohncke-group structures archived in the CSD as of January 2019 and containing a C-atom stereocenter. As in this study structures of polymers and structures determined from powders were excluded.

The most important differences between the two studies are that Rekis considered structures in all Sohncke groups rather than just those described in $P1$ and that he allowed R factors up to 0.075 rather than to 0.050. The latter difference increased the number of structures analyzed by a factor of *ca.* 1.8.

Another important difference between the two studies is that Rekis excluded salts, solvates, and co-crystals (including compounds of diastereomers), as well molecules containing elements other than C, H, O, N, S, F, Cl, Br, and I, and molecules having a molar mass outside the range 70 – 500 g/mol. He did not consider molecules having an S-atom stereocenter (*e.g.*, an $R_1R_2S=O$ group) or molecules having axial chirality.

Only 168 structures appear on both the list of 1456 considered in this study and the Rekis list of the 1200 best centrosymmetric mimics. Five more $P1$, $Z>1$ structures categorized in this study

as $P\bar{1}$ mimics appear on Rekiš's supplementary list of 4264 structures

The 168 structures identified by Rekiš as $P\bar{1}$ mimics were described in this study as follows:

120 of the 168 were categorized as $P\bar{1}$ mimics

48 of the 168 were categorized as something other than a $P\bar{1}$ mimic.

11 were categorized as having *hybrid packing* (numbers in parentheses give the position of the structure on the Rekiš list); in all 11 there is a region that has approximate inversion symmetry

YACTEC01 (#384), GEHZOK (#448), WAXLEM (#498), OBIPUN (#625),
NUVMEY (#656), PAXNIL02 (#928), XISYAA (#766), AXEKAP (#922),
KEZRUC (#1021), SUTNAX (#1032), EHAROV (#1188)

1 was categorized as a 3-D glide mimic
PEXLEK (#806)

10 were categorized as $2\&2_1$ (4 as 3-D and 6 as 2-D)

3-D: WIPSOG (#830), AJEXAQ (#879), CUGMAT (#931), ULUHUE (#1061)
2-D: CANBEB (#337), DEBMIJ (#621), LEPBEO (#679), ZUWDIG (#919),
HIRHIC (#1039), FONXIQ (#1154)

7 were characterized as 2_1 (1 as 3-D, 6 as 2-D)

3-D: CUHZAHA (#471)
2-D: XOGWOH (#198), EHUGIZ (#344), XUSXIS (#556), VUGWUP (#772),
XEHPUY (#857), SEZRIY (#1167)

8 were characterized as $p112$

(distinguishing between a $p112$ and a $P\bar{1}$ description can be a judgement call,
especially if the molecule is quite planar)

CECYIU (#408), JIKHOA (#431), AJOVIG (#590), CONGIX (#663),
QAKVUT (#695), RUKHAG01 (#708), JITTUB (#963), XILHIL (#1079)

2 were categorized as modulated by an approximate translation
LASRAY (#1086), RUDFUT (#1162)

7 were categorized as *none found*

MALTOI (#601), FOPPAD (#670), TABLEN (#677), XAKJIF (#682),
OHIVEG (#1014), WAPHUR (#1127), BOXQUC (#1199)

2 were categorized as $PSTG$, with both identified as $P2_1$

ZEHFEZ (#564), LITQEL (#932)

Structures identified in this study as $P\bar{1}$ mimics that were not identified by Rekiš:

18 structures categorized in this study as a $P\bar{1}$ mimic plus or $P\bar{1}$ mimic

$P\bar{1}$ mimics plus: ABONEM, AXIBIT, HEXZIW, SEVKUB

$P\bar{1}$ mimics: DAXSIH, FEKJOW, MAYLIK, MOVDOR, NIVNOY, PAZFUR,
PIJSOR, QAMQIG, SANBAN, TUQWIM, UPENEJ, YAVJIR, YUFYED, ZEBSEH

(AXIBIT, SEVKUB, FEKJOW, MAYLIK, and NIVNOY are less good *P1 mimics* but the other 13 seem to be very good mimics)

5 structures categorized in this study as a *P1 mimic* or *P1 mimic?* are on the supplemental Reki list of 4264

P1 mimics: EDEHAX (#760), WUJCIN (#1075), TITLUE (#1442), HUSLUE (#1509)

P1 mimic?: JUSZOO (#109)

Of the 48 structures on the Reki list of 1200 that were not described as $P\bar{1}$ mimics in this study only 12 are in the top half of his list. The average and median ranks are #778 and #769.

Note that the distinction between hybrid packing and a $P1$ mimic is sometimes a matter of opinion.

YACTEC01 (#384 in Reki list but classified here as having hybrid packing)

In layer (01 $\bar{1}$) (axes **a**, [011]) the pyrimidine rings form dimers that mimic inversion symmetry.

The remaining parts of the molecules form H-bonded chains along **a** with very good approximate 2₁ axes.

JKHOA (#431 in Reki list but classified here as having hybrid packing)

The three stereocenters are all in a C₄O ribose ring. The cytosine rings are well related by inversion but the ribose rings in a layer (012) are related by an approximate twofold perpendicular to the layer. There are H bonds that link the layers but most of the H bonds lie within the layers.

WAXLEM (#498 in Reki list but classified here as having hybrid packing)

The molecule and the packing are very similar to YACTEC01. The Me group in YACTEC01 is missing in WAXLEM. The layer in WAXLEM is (011) (axes **a**, [01 $\bar{1}$]) rather than (01 $\bar{1}$).

Comments on the remaining structures that are in the top half of the Reki list that were not categorized in this study as $P1$ mimics

XOGWOH (#198 in the Reki list)

The description here is as a layered, $p2_111$ structure [layer (001), axes **a**, **b**; angle 91.4°; $-0.1 < z < 0.9$] but the choice between the two descriptions may be a matter of taste (see summary below). The view that most clearly shows the deviations from $P\bar{1}$ is along [01 $\bar{1}$].

CANBEB (#337 in the Reki list)

A description of layers (001) [axes **a**, [120]; angle 89.1°] having approximate symmetry $c211$ seems better. Possible inversion symmetry is broken by the orientations of three of the four carbonyl groups and by the positions of the O atoms in the C₃NO rings.

EHUGIZ (#344 in the Reki list)

In this study the description was as an approximate $p2_111$ [layer (001); axes **a**, **b**; angle 91.7°] but description as a $P\bar{1}$ mimic is about equally good because the molecules are essentially flat.

CECYIU (#408 on the Reki list but described here as having $p112$ layers)

The planar sections of the molecule are related by good approximate inversion but the relationship of the C₄O rings is better described by a twofold axis. Layer (1 $\bar{2}$ 1) has approximate symmetry $p112$ (see a view along [10 $\bar{1}$]).

GEHZOK (#448 in the Reki list but classified here as having hybrid packing.

The terminal 4-MeO-Ph groups of the two molecules are related by good approximate inversion. There is an approximate screw axis near [210] in layer (001) that relates the rest of the atoms although it is 1-D only.

CUHZAH (#471 in the Reki list)

The structure is described here as distorted $P2_1$ with **a** unique and angles β and γ 93.5 and 91.2°. The choice of $P2_1$ over $P\bar{1}$ was based on a single O atom but the molecule is small (C₉H₇BrO₂). See views along **b** and **c**.

XUSXIS (#556 in the Reki list)

Layer (011) (axes [01 $\bar{1}$], **a**; angle 91.8°) has very good approximate symmetry $p2_111$. An inversion relationship between the two molecules seems less convincing, especially because of the relationships of the O atoms.

ZEHFEZ (#564 in the Reki list)

The structure is almost certainly described better by the space group $P2_1$ with **b** unique; α and γ are both 89.9°.

AJOVIG (#590 in the Reki list)

Layer (102) has approximate symmetry $p112$ although the twofold (approximately along [4 $\bar{2}$ 3]) is tipped away from the layer normal by about 12°. Still, there is obvious approximate symmetry. The other possibility is description as a $P\bar{1}$ mimic but that choice seems less good.

Summary of comparisons with the Reki lists

The agreement of this study with the Reki list is very good. Most of the discrepancies involve molecules that are thinner in one direction than in the other two, especially if there is an approximate mirror plane normal to that direction (see Section 3.3 of the paper; examples are XOGWOH, #198 on the Reki list, and EHUGIZ, #344). See also ZEHFEZ (#564), which should almost certainly have been described in $P2_1$ rather than in $P1$.

Comments about discrepancies between the Reki list and *PLATON* recommendations are given towards the end of Appendix 1.

4.4 Distorted forms of higher-symmetry structures dominate

Percentages for all distorted higher-symmetry structures

(i.e., all non-PSTG structures less those categorized as alt layers, hybrid packing, none found, or some kind of mimic)

	Achiral or a kryptoracemate	Chiral
Z or $Z_{eff} > 1$	73% (108 of 147)	42% (502 of 1190)
$Z=1$	50% (52 of 105)	34% (36 of 106)
All Z	63% (160 of 252)	42% (534 of 1296)

Percentages for structures categorized as mimics

	Achiral or a kryptoracemate	Chiral
Z or $Z_{eff} > 1$	(none)	35% (411 of 1190) (375 $P\bar{1}$ mimics; 36 glide mimics)
$Z=1$	23% (24 of 105) (22 $P\bar{1}$ mimics; 2 glide mimics)	34% (36 of 106) (32 $P\bar{1}$ mimics; 2 glide mimics; 2 $2\&2_1$ mimics)
All Z	10% (24 of 252)	34% (447 of 1296)

4.5 Possible relationship between 2-D and 3-D symmetries

Statistics for all structures in 2019 CSD that meet the criteria (other than space group and Z) of this study

(i.e., organic, coordinates present, $R \leq 0.050$, no errors, not polymeric, not PXRD)

These statistics are essentially the same as those reported by Rekiş (2020) for the January 2019 version of the CSD and those for the 2016 version of the CSD that are available on the CCDC website: (<https://www.ccdc.cam.ac.uk/support-and-resources/ccdcresources/a343010d2dfe48f1a577211a2e3e055d.pdf>)

Sohncke Groups	Spc Grp #	Z'>1	Z'=1	Z'<1	All Z'	%
P2(1)2(1)2(1)	19	1940	27927	2	29869	49.8%
P2(1)	4	4592	15916	4	20512	34.2%
P1	1	1477	1115	0	2592	4.3%
C2	5	374	1744	351	2469	4.1%
P2(1)2(1)2	18	125	645	281	1051	1.8%
P4(n)2(1)2, n=1-3	92,94,96	26	464	340	830	1.4%
P4(n), n=1-3	76,77,78	57	468	2	527	0.9%
P3(1)/P3(2)	144,5	82	430	0	512	0.9%
P6(n), n=1-5	169-73	28	368	56	452	0.8%
P3(1)21/P3(2)21	152,4	14	161	137	312	0.5%
C222(1)	20	15	177	105	297	0.5%
R3	146	14	104	112	230	0.4%
cubic	>=195	1	4	69	74	0.1%
P6(n)22, n=1-5	178-82	1	12	45	58	0.1%
I4(1)	80	10	16	19	45	0.1%
C222/F222/I222	21-23	3	10	8	21	0.0%
P2	3	10	8	0	18	0.0%
P222(1)	17	0	1	3	4	0.0%

P222	16	0	0	0	0	0.0%
Sum of above		8769	49570	1534	59873	99.8%
All Sohncke		8780	49621	1574	59975	
Other		11	51	40	102	0.2%

non-Sohncke Groups	Spc Grp #	Z'>1	Z'=1	Z'<1	All Z'	%
#14 [P2(1)/c, etc]	14	5027	57978	7381	70386	47.1%
P1-bar	2	5714	28487	4665	38866	26.0%
C2/c, I2/a, etc	15	490	6991	4619	12100	8.1%
Pbca	61	295	6880	583	7758	5.2%
Pna2(1)	33	527	3337	3	3867	2.6%
Cc, Ia, etc	9	363	2058	1	2422	1.6%
Pca2(1)	29	558	1598	2	2158	1.4%
Pnma	62	23	51	1512	1586	1.1%
Pbcn	60	43	599	699	1341	0.9%
Pc, Pa, Pn	7	312	792	1	1105	0.7%
Fdd2	43	28	485	285	798	0.5%
P2(1)/m	11	5	42	684	731	0.5%
P2/c	13	77	311	289	677	0.5%
R3-bar	148	12	287	353	652	0.4%
Pccn	56	17	313	175	505	0.3%
I4(1)/a	88	16	256	150	422	0.3%
C2/m, I2/m, etc	12	1	12	388	401	0.3%
Aba2	41	16	111	94	221	0.1%
Sum of above		13524	110588	21884	145996	97.7%
All Non- Sohncke		13619	111266	24579	149464	
Other		95	678	2695	3468	2.3%

Approximate and mimicked layer groups found for all structures described as having 2-D or 2-D/3-D packing except for those in the categories *Alt Layers* and *Hybrid Packing*.

	Achiral & kryptoracemic	Chiral
All layer groups	49	312
All Sohnke layer groups	14	297
<i>p</i> 2 ₁ 11 (#9)	8	151
<i>c</i> 211 (#10)	6	108
<i>p</i> 11 <i>a</i> (#5)	28	5
<i>p</i> 112 (#3)	0	24
<i>p</i> 211 (#8)	0	7
<i>p</i> 2 ₁ / <i>b</i> 11 (#17)	0	4
<i>p</i> <i>b</i> 2 ₁ <i>a</i> (#33)	3	4
<i>p</i> <i>b</i> 11	3	2
other*	1	11

*The other group for structures of achiral and kryptoracemic material is *pba*2 (#25).

The other approximate layer groups for structures of enantiopure substances are *p*2₁/*b*11 mimic (#17)(4x), *p*2₁22 (#20)(2x), *p*2₁2₁2 (#21)(2x), *c*222 (#22), and *p*321 (#68)(2x).

The five structures identified as having 2-D/3-D packing; all are composed of enantiopure material

ABONEM	<i>p</i> 2 ₁ / <i>b</i> 11 mimic (<i>z</i> '=1)/ <i>P</i> $\bar{1}$ mimic (<i>Z</i> '=2)
AXIBIT	<i>p</i> 2 ₁ / <i>b</i> 11 mimic (<i>z</i> '=1)/ <i>P</i> $\bar{1}$ mimic (<i>Z</i> '=2)
CIMCIN	<i>p</i> 2 ₁ / <i>b</i> 11 mimic (<i>z</i> '=1)/ <i>P</i> $\bar{1}$ mimic (<i>Z</i> '=2) (2:1 salt)
ROJREQ	<i>p</i> 2 ₁ / <i>b</i> 11 mimic (<i>z</i> '=1)/ <i>P</i> $\bar{1}$ mimic (<i>Z</i> '=2)
DOVXIW	<i>p</i> 2 ₁ 22 [<i>z</i> '=4(½)=2]/ <i>P</i> 2 ₁ (<i>Z</i> '=2)

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

Fábián & Brock (2010) estimated that only 6% of achiral and racemic materials crystallize in Sohncke groups. Rekiş's (2020) estimate was 9.5%.

*P*1, *Z*>1 structures of enantiopure material examined in this study”

1190	reliable structures
371	<i>P</i> $\bar{1}$ mimics
72	described as having hybrid packing’=

4.7 Molecular features that favor inversion mimicry

Quasiracemates found in this study (15)

*P*1 mimics (13)

BERNER	Et vs. <i>i</i> -Pr
IREREE	Cl vs. Br
IRERII	Cl vs. Br
JAQJET	CH ₂ CHMe ₂ vs. CH ₂ CH ₂ SMe
KACXET	Me vs. Et
LUNRER	NH vs. O
NIVYOG	H vs. OH
PATXUF	C ₅ H ₈ vs. C ₆ H ₁₀ in a spiro compound; the other ring in both molecules is C ₃ O ₂
POVZAC	CHMeEt vs. CH ₂ Ph
SONCED	<i>i</i> -Pr vs. CH ₂ Ph
URODEL	CHMeEt vs. CH ₂ (<i>i</i> -Pr)
YOGBON	Me vs. Br (twice)
YOGCOO	Cl vs. Br (twice)

glide mimics (2)

DEZLAY	H vs. OH
DEZMAZ	H vs. OH

Compounds of diastereomers found in this study (34)

Most of the molecules are heterochiral in the central part of the molecule with the homochiral stereocenters on the periphery. Exceptions are FARNUI, OJUSIW, TEVGEG (both on the periphery), XADVED (very small molecule), IYIKUW (both on the periphery), and FOMDIW.

P1 mimics (29)

ABONEM	Homochiral at three C-atom stereocenters in a menthoxy [<i>i.e.</i> , (1R,2S,5R)-5-methyl-2-(propan-2-yl)cyclohexyl]oxy] ring (H/Me, H/ <i>i</i> -Pr; H/OR), Heterochiral at central P atom
BEBHUN	Homochiral at one C-atom stereocenter (H/ <i>i</i> -Pr) Heterochiral at central P atom
HEROZ	Homochiral at three of four C-atom stereocenters in a [2.2.1] cage Heterochiral at the point of attachment of the B(C ₆ F ₅) ₂ group to the cage
FARNUI	Homochiral at two C-atom stereocenters in a quite flat system of fused rings Heterochiral at a ring C atom with substituents <i>n</i> -Bu and OH
FONVOW	Homochiral at one C-atom stereocenter with substituents H and Et Heterochiral at a central C atom
GALDEE	Homochiral at two C-atom stereocenters in a [2.2.1] cage Heterochiral at the six C atoms of a saturated C ₆ ring
HATZUX	Four stereocenters in a system of fused rings. Three are heterochiral The one homochiral stereocenter has substituents H and Me
IDIMUD	Homochiral at two C-atom stereocenters in a [2.2.1] cage Heterochiral at a central C atom
IDISUJ	Homochiral at a central CHRR'R" stereocenters The hindered biphenyl groups have opposite twists; the configuration at the three-coordinate N atoms differs as well
JODNIC	Two C-atom stereocenters separated by a CH ₂ group. The one with the larger substituents is heterochiral; the one with substituents C(=O)OMe and CH ₂ Ph is homochiral
JOSHIL	Homochiral at a C-atom stereocenter with substituents H and C(Me)=CH ₂ Heterochiral at the two C atoms where a C ₃ and a C ₅ O ring are fused
KELNIA01	Homochiral at three C-atom stereocenters in a [3.1.1] cage Heterochiral sulfoxide group with substituents CH ₂ (cage) and CH ₂ CH ₂ OH
KICDOR	Homochiral at two C-atom stereocenters in a [2.2.1] cage Heterochiral at the point of attachment (an sp ³ C atom) to the cage
LIPZEP	Homochiral at one C-atom stereocenters with substituents H and Me Heterochiral at the two C atoms where C ₄ O and C ₄ O ₂ rings are fused
LOYMER	Homochiral at two C-atom stereocenters in a [2.2.1] cage Heterochiral at the point of attachment (an sp ³ C atom) to the cage
MIJBEN	Homochiral at three C-atom stereocenters in a menthoxy ring (H/Me, H/ <i>i</i> -Pr; H/OR), Heterochiral at central C atom with substituents CN and <i>i</i> -Pr

MIQCAZ	Homochiral at three C-atom stereocenters in a menthoxy ring (H/Me, H/ <i>i</i> -Pr; H/OR), Heterochiral at two more central C-atom stereocenters
OJUSIW	Homochiral at a central C-atom stereocenter with three large substituents two of which are OS(=O) ₂ Me and CH ₂ OS(=O) ₂ Me Heterochiral at a C-atom stereocenter with substituents H and Me
QASNIJ	Homochiral at a hindered binaphthyl Heterochiral at the central P atom
QIKSIM	Homochiral at the two C atoms where a C ₅ and a C ₄ N ring are fused Heterochiral at the central Si-atom stereocenter
REBLIU	Homochiral at the two C atoms where a C ₃ O ₂ ring is fused to a C ₆ P ₂ ring Heterochiral at the two P atoms
RUSZAG	Homochiral at three C-atom stereocenters in a [2.2.2] cage Heterochiral at the C atom that bonds to that cage
TEFFEP	Homochiral at two C-atom stereocenters, both with substituents H, Me, Ph Heterochiral at the very hindered biphenyl in the middle of the molecule
TEVGEG	Homochiral at one C atom with substituents Me and CH ₂ C(H)=Me Heterochiral at the adjacent C-atom that has substituents H and OH
VIJXES	Homochiral at one C-atom stereocenter that has substituents H and Me Heterochiral at the central Si atom
VOJBAZ	Homochiral at four C-atom stereocenters in a C ₆ ring; three have substituents H and OH while the fourth has substituents Me and OH Heterochiral at the very hindered bianthryl
WAFBOV	Homochiral at four C-atom stereocenters in a C ₄ N ring Heterochiral at one C-atom stereocenter with substituents H and OH
XADVED	S,S-2- and R,S-amino-3-methylpentanoic acid Homochiral at the C-atom stereocenter with substituents Me and Et Heterochiral at the C-atom stereocenter with substituents NH ₃ ⁺ and COO ⁻
XUTPOS	Homochiral at one C-atom stereocenter with substituents H and CH ₂ C(H)=C(Me) ₂ Heterochiral at the C-atom stereocenter at the fusion of C ₅ N and C ₄ N ₂ rings
<i>glide mimics (2)</i>	
IYIKUW	Homochiral at one C-atom stereocenter with substituents H and Me Heterochiral at one C-atom stereocenter with substituents H and <i>i</i> -Pr
MIYGAC	Homochiral at one C-atom stereocenter with substituents H, Me, and Et Heterochiral at a central C-atom stereocenter

2&2(1) (1)

FOMDIW Homochiral at the two C-atom stereocenters at the fusion of two C₆ rings
Heterochiral at a C-atom stereocenter with substituents H and OH
In this case it is the homochiral part of the molecule that determines the packing

hybrid packing (1)

WAGBUB *n*-Octyl α -D-glucopyranoside *n*-octyl β -D-glucopyranoside
Homochiral at the four C atoms of the C₅O ring that have OH substituents
(three also have H substituents; the fourth has a CH₂OH substituent)
Heterochiral at the C atom of the C₅O ring that has substituents H and O(CH₂)₇Me

none found (1)

COYSOY Molecules are the same except for the position of one double bond

Compounds found in this study that have homochiral *sec*-Bu groups (4)

CISDOA, PIJSOR, and TICLIC (no other stereocenter)

XADVED (one other, heterochiral, stereocenter)

All four are *P* \bar{I} mimics.

Sulfoxides in which the S atom is the only stereocenter (5)

Non-Sohncke mimics

CIYPAC, KISCUM, VUKHEQ, XOJFUX

Other

FOPYER (no approximate symmetry found)

Compounds in which an enantiomer switch involves atoms not bonded to the same or to adjacent C atoms (6)

ARAXUP H/Me on C atoms separated by CH₂ in a saturated C₃OP ring

COTBIY H/OH on C atoms separated by CHR in a saturated C₅N ring

DUMYEQ H/OH on adjacent C atoms

KASPUP H/Me on C atoms separated by CHR in a saturated C₅N ring

PEXLEK H/Me on C atoms quite far apart but related by near twofold molecular symmetry

QAMQIG H/F on Ph and 4-F-Ph groups bonded to the same C atom

Two stereocenters that are adjacent; no other stereocenters (18)

In all structures listed one of the substituents is an H atom and there is a staggered conformation in which the H atoms are *trans*.

ENUPEJ	Each C atom has two large substituents.
EWOCEA	H/Me on both; the cation and uncharged molecule have approximate twofold axes
GIGDAB01	H/Me on both (and at both ends of the cation and molecule, which have approximate twofold axes)
IVEMAY	H/NMe ₂ on both; the molecule has an approximate twofold axis
KUMLAG	H/OH on both; the molecule has an approximate twofold axis
LIVQIT	the two stereocenters are part of a C ₃ ring
MOCSEF	NH ₂ /Ph and NH ₃ ⁺ /Ph; the cation has an approximate twofold axis
QODKIF	H/Me on both (and at both ends of the cation)
QODKOL	H/Me on both (and at both ends of the cation)
QODKUR	H/Me on both (and at both ends of the cation) QODKIF, QODKOL, and QODKUR are closely related.
RARKEC	the adjacent C atoms are part of a C ₅ N ring
SELQAC	H/OMe on both; the molecule has near mirror symmetry
SIKBAQ	CHRR' on both; the molecule could have an approximate twofold axis
SUYLEF	H/Me on both; the H atoms are <i>trans</i>
YEQHAD	H/OH on both; the C-CH(OH)-CH(OH)C- unit has an approximate twofold axis
YIPRIA	the H atoms are <i>trans</i>
YOFDEF	linked by OCO, where the central spiro C atom is also part of a C ₆ H ₁₀ ring; the H atoms are <i>trans</i>
ZAQLOS	linked by OC(Me) ₂ O; the H atoms are <i>trans</i> group

And also

GIVXUF	C ₅ O rings with three H/OH C atoms, one H/CH ₂ OH C atom, and one H/R C atom; H atoms on adjacent C atoms are always <i>trans</i> so that all non-H substituents are equatorial.
POSPUL	The substituents on the adjacent stereocenters are large and different but the <i>rmsd</i> for overlay with inversion is 0.39 Å.

Compounds having one stereocenter in which switch of C, N, and/or O atoms in a ring would convert enantiomers (14)

AGIYOG	CH ₂ /O switch in a saturated C ₄ NO ring
ARUCEV	CH ₂ /NH switch in a saturated C ₅ N ring (piperidine-2,4-dione)
DEWZAJ	CH ₂ /NH switch in a 1,4-diazaisotwistane cation
DUDMUL	CH/CH ₂ switch in a C ₆ ring that has one double bond
EVUROE	N/O switch in a C ₃ NO ring (4,5-dihydro-1,3-oxazole)
IBENEK	CH ₂ /C=O switch in a C ₆ , otherwise saturated ring
JEQGOE	CH ₂ /S switch in two positions of a saturated C ₅ S ₂ ring
MEDNAM	-CH=-/CH ₂ - switch in a [3.2.1] cage
NAQMEA	CH=CH/CH ₂ CH ₂ switch in a C ₆ , otherwise saturated ring
NARZUE	CH ₂ /NH switch in a saturated C ₅ N ring
PACTUL	-O-/CH= switch in an otherwise saturated C ₅ O ring
PACVEX	-O-/CH= switch in an otherwise saturated C ₅ O ring
PIMFUM	-O-/CH= switch in an otherwise saturated C ₅ O ring
SOXSAC	-CH ₂ -/CH= switch in an otherwise saturated C ₇ ring

Percentage of non-Sohncke mimics vs. number of stereocenters

(106 of the 1190 reliable structures of enantiopure material are salts or co-crystals in which both residues have >6 non-H atoms. For the purposes of this table the number of stereocenters was taken as the larger of the values for the two residues.)

Number of stereocenter	Number of reliable (<i>i.e.</i> , non-PSTG) structures	Number of structures that are mimics	Percentage of structures that are mimics
0*	2	2	100%
1	389	220	57%
2	285	88	31%
3	150	41	27%
4	141	33	23%
5	88	17	19%
6	43	4	9%
7	23	2	8%
8	24	2	8%
>8	45	2	4%

*DORHOK is one of the two structures listed as having no stereocenters; it is a 1:1 solvate of S-2-butanol (5 non-H atoms). The other structure, IRABUA, is a 2:1 co-crystal that contains one independent molecule of L-malic acid, C₄H₆O₅, which is disordered 0.54/0:46 by an approximate twofold rotation around an approximate inversion center).

Some of the non-Sohncke mimics with n stereocenters, $n > 5$, are

$n=6$ EKINIX, IDENIO, IDPYRS, UDUCAA

$n=7$ RAKXEH, XAMCAQ

$n=8$ COZKEI, GALDEE

$n=9$ BEPXEK, TUPCUD

Compounds containing an enantiomerically pure cage substituent (e.g., norbornyl) (25)

CABBIT, CAMCBX10, DUMYEQ, EHEROZ, GALDEE, GIJGUE, HEDQEP, IDENIO, IDIMUD, JAGXEU, KELNIA01, KICDOR, KIXVOF, LOYMER, MIBDEJ, NILQOP, QUHWAS, RUSZAG, TUTZEN, UQEVIX, VAVRUG, VIWGOY, WUKYIL, XOLCAC, ZORPOL

Compounds having a ring system with multiple stereocenters that is flat enough to mimic non-Sohncke symmetry (11)

BEPXEK	four fused rings; nine stereocenters
COZKEI	four fused rings; eight stereocenters
FARNUI	three fused rings; three stereocenters
KIHSAY	five fused rings; five stereocenters
MIZQAO01	two sets of four fused rings; 18 stereocenters
MODWAE	eight stereocenters
RIWZIG	an isomaltoside derivative; ten stereocenters in the two C ₅ O rings; 3 pairs are heterochiral, 2 pairs are homochiral)
TUHWAW	two [2.2.1] cages; six stereocenters
TUPCUD	four fused rings; nine stereocenters
ULUFIS	four fused rings; five stereocenters
XAMCAQ	four fused rings; seven stereocenters

Examples of compounds in which the homochiral part is near the center and the attached, achiral groups are quite flexible (7)

COJZOS, EGEVUH, FEJVUN (best example), HIYDEZ, LEMSOO, TEMJON, XONSOL

4.9 Implications for the formation of solid solutions

The 411 $P\bar{1}$ and glide mimics were examined to see if any disorder is shown, if the atomic ellipsoids strongly suggest disorder near the stereocenters, and if any of the C-atom stereocenters are suspiciously flat. (While 85 % of the CSD entries examined included atomic ellipsoids, 58 of 375 $P\bar{1}$ mimics and 4 of 36 glide mimics did not.) Good evidence of a solid solution was found in ten of the 411 structures; six of the ten were described by the original authors.

10 non-Sohncke mimics; only one is a glide mimic. The six described by the original authors are starred

CEXPOK** (glide mimic) From the CIF: "... 2 molecules in the unit cell: molecule A is a superposition of 79(1)% of (S)-isomer and 21(1)% of (R)-isomer, molecule B is (S)-isomer only."

GIDWEW Probable diastereomeric disorder; there is *ca.* 6:4 and 5:5 disorder in the two molecules at one of the four stereocenters. No information about that disorder is given in the paper or the supplementary material but diastereomeric contamination seems very likely.

GOMFOG One of the two molecules seems to be disordered about 3:1 with its enantiomer and/or a diastereomer. There is no mention of the disorder in the paper.

IJIHOZ02** Enantiomeric disorder (65:35?) in one of the four independent molecules was reported and some of the ellipsoids are very suspect or even non-positive definite. The paper discusses the formation of the solid solutions in some detail.

JISREI** There is 59:41 H/OH disorder in one of the two molecules at the one stereocenter so the material is a mixed crystal. The racemate (JISQUX, $P\bar{1}$, $Z'=1$) is isostructural and has 74:26 disorder at the same site in one molecule (and presumably has 26:74 disorder in the symmetry-related molecule).

MAYLIK There is good evidence (disordered atom; eccentric atomic ellipsoids) of 65:35 diastereomeric contamination at one of the five stereocenters of molecule #2.

MOVDOR There is Me/Br disorder at the one stereocenter, and therefore enantiomeric contamination, in three of the four molecules; the occupancy factors in the three are 85:15, 83:17, and 94:6. There is no mention of the disorder in the paper, which says “The absolute configuration of the sulfones (*R*)-**10** and (*R*)- **11**, respectively, obtained by *m*-CPBA oxidation of the sulfoxides **5A** and **9A**, were unambiguously established by crystallization and X-ray analysis (Scheme 6).”

REMBUI** 3,3'-Dibromo-5,5'-di-iodo-4,4'-bipyridine. All four independent molecules have an approximate twofold axis perpendicular to the central C-C bond. They would have approximate mirror symmetry (torsion angle between rings near 90°) but for the difference between the (disordered) halogens. The main disorder is around the long (N...N) axis but the authors say there is also enantiomeric contamination, which seems possible given the extensive disorder.

From the supplementary material:

The crystal structure of **3** (2nd eluted peak) was solved in *P*1 space group. The Flack parameter introduced as an inversion twin population refined to 0.269(6), the absolute configuration being *P*. The R_1 values were 0.0496 ($I > 2\sigma(I)$) and 0.0538 (all data). The $wR(F_2)$ values were 0.1481 ($I > 2\sigma(I)$) and 0.1522 (all data). The goodness of fit on F_2 was 1.002. In this model three of the four crystallographically independent molecules, display C-Halogen bonds lengths in between C-Br (1.89Å) and C-I (2.09Å) bond lengths; moreover, corresponding bromine atoms display relative small atomic displacement parameters whereas iodine atoms display relative large ADP values. This indicates that three of these four molecules are disordered. The final model is then built by considering that these three sites are occupied by **3-P** molecules having different orientations; the validity of this model is assessed by the significant decrease of the $wR(F_2)$ value (the final $wR(F_2)$ values were 0.0685 ($I > 2\sigma(I)$) and 0.0707 (all data)). The final Flack parameter (i.e. inversion twin population) is not zero (0.124(5).) and indicates a small enantiomeric *M* contamination, as an enantiomorphous crystal phase. However, since the crystal is disordered, the standard uncertainty on this Flack parameter is certainly larger than the value obtained.

The occupancy factors for molecules #1–4 are 1:0, 0.605:0.395(2), 0.737:0.263(2), and 0.736:0.264(2). The orientations of the pairs #1&2 and #3&4 are not related.

TIMOLM01** The enantiopure and racemic compounds are isostructural and are reported to form a continuous range of solid solutions.

YOGBON** There is 61:39 Me/Br disorder in one of the two molecules. In the paper the material is described as a solid solution in which some of the *S,S*-dimethyl molecules in its quasiracemate with the *R,R*-dibromo molecule are replaced by *S,S*-dibromo molecules.

The structures of enantiomerically pure material that are not mimics and that are marked as disordered were also examined. Four solid solutions were found.

DAYREA06** (*2₁ plus*) A 83:17 mixed crystal of L-asparagine and L-aspartic acid; in the former a C(=O)NH₂ group replaces a C(=O)OH group in the latter. DAYREA06 is approximately isostructural with pure L-asparagine [ASPARM05 $P2_12_1$]₂₍₁₎, $Z'=1$ at 100 K].

JAQLAP** (*none found*) Mixed crystal (65:35) of two steroids that differ in one substituent [C(=O)OH vs. CH₂OH] in one of the two molecules. The substitution is possible because the =O atom of the carboxy group does not participate in the H-bonded chains that extend along [22 $\bar{1}$].

PATXOZ** (*none found*) (R)-2-butylamine solvate of two homochiral *vic*-diols that differ by one CH₂ group (C₆H₁₁ vs. C₅H₉ ring attached to a spiro C atom). There is 55:45 disorder in one of the two molecular sites; the other is ordered.

WIYSAZ (*alt layers*) 1:5 compound of enantiomers but at the site of the different enantiomer there is 59:41 diastereomeric contamination at the C(Me) atom of the C₅ ring. It is possible, but not certain, that the contaminant is the diastereomer in BIVJAS ($P2_1/n$, $Z'=2$).

A few obvious possibilities

KELNIA01 The ellipsoids in one of the two diastereomers (#1) in KELNIA01 are large and unreasonable physically. ZOCGIJ ($P2_1$, $Z'=1$) is the structure of the other diastereomer by itself; ZOCGIJ shows no indication of any disorder. Both structures were reported in the same 2012 paper but the journal (*Zh. Obshch. Khim.*) is not widely available.

QAWNAD/01 Racemic and enantiomerically pure cloperastine fendizoate. The one stereocenter has Ph and 4-ClPh substituents. In QAWNAD01 ($P1$, $Z=2$ $T=153$ K), a $P\bar{1}$ mimic, the Cl atoms are ordered and all atomic ellipsoids are unremarkable. In QAWNAD ($P\bar{1}$, $Z'=1$ at RT; no atomic ellipsoids available) the Cl atom is described as disordered 4:1 over the two 4-Ph sites. There is no evidence that solid solutions form but they certainly seem possible.

OBZXPF The molecule could have mirror symmetry but for one F substituent that replaces an H atom. The structure is an excellent $P\bar{1}$, $Z'=1$ mimic. There are obvious layers (011) with only one F...H contact within the layer (the other F and corresponding H are at the layer surface). There are no displacement parameters in the CSD but there is an *ORTEP* drawing in the paper and all the ellipsoids are normal. Still, formation of a solid solution seems possible. (OBZXPF is #13 on the Rekiş list of 1200 approximately centrosymmetric, Sohncke-group structures.)

And

ADGSMF, ADGSMH (both $P1$, $Z=1$ at RT) are isostructural diastereomers that differ at one of the six stereocenters (a CMeOH atom in a C₄O ring). Structures were published in 1980; no atomic ellipsoids are available.

HUMWUI and HUMXAP (both $P1$, $Z=1$ at 150 K), diphenyl(1,3,3-trimethylbicyclo[2.2.1]hept-2-yl)phosphine oxide) The molecules in the two structures differ in the way the P(=O)Ph₂ group is attached to the bicyclic cage (*endo*- α -fenchyldiphenylphosphane oxide vs. *exo*- β -fenchyldiphenylphosphane oxide). The two structures have very similar layers (001) but the atomic ellipsoids in both structures are all unremarkable. Still, formation of solid solutions seems possible.

4.10 Implications for crystal nucleation

Reliable kryptoracemates that have approximate $P\bar{1}$ symmetry

- ABAPUQ (at 296 K) An included *n*-PrOH molecule occupies the approximate $\bar{1}$ site but there is no indication of C/O disorder, the C-C-C-O torsion angle is 140°, and the structure was published in the 2011 volume of *Acta Cryst. E*.
- RIGSEF (at 295 K) The approximate center that relates the two cations does not relate the three F atoms of the F₃CC(=O)O⁻ anions (see a view along **a**). The cations also differ in small, but clear, ways [in a view along **b** compare C6, C7 as numbered in the CSD entry with C17, C18; in a view along **a** compare the rotations of the C(=O)Me groups around the C-O bonds].
- IBOmaq (at 193 K) $P\bar{1}$ symmetry would require the dipole of one of the two F₃CSO₃⁻ ions be reversed but there is no indication of significant C/S disorder. The small differences between the cations are most obvious in a view along **c**.
- XEPHUX (at 173 K) The orientations of the Et group of the C(=O)OEt substituent are clearly different.
- GUMCEY (at 120 K) The one included ethyl acetate molecule is ordered. The two larger molecules have slightly different conformations (*rmsd* 0.44 Å for overlay with inversion). When viewed along **b** there is an obvious difference between the two molecules in the relationship of the C(=O)Me group and Br atom.
- KETWAK (at 100 K) The orientation of the *i*-Pr substituent differs between the two molecules as does the orientation of the C=O group (see views along **a** and **c**).

The only two of those structures that produce *PLATON* alerts above the 90% level with standard tolerances are ABUPUQ (94%) and RIGSEF (91%).

Kryptoracemates that have neither approximate $P\bar{1}$ nor approximate glide symmetry

- CEHYEW *alt layers* Layers (001) (axes **a**, **b**; angle 88.7°) of #1&3 and of #2&4 both have very good approximate symmetry $pb11$. The two layers are related by a twofold around **b** plus a translation of *ca.* **b**/4 and by an approximate glide (mirror in **c***, translation *ca.* **b**/4). If the layers were not offset there would be approximate orthorhombic symmetry.
- GIPLID *alt layers* There are three layers (001) that each have approximate symmetry $p\bar{1}$. Layers of #1&2 and of #3&4 are related by an approximate translation that would be compatible with a 3-D $P\bar{1}$ cell but the layer of #5&6 is related to the other two by an approximate twofold screw axis or glide.
- EYUDAF *none found* Views in some directions (*e.g.*, along **a** or **b**) suggest approximate symmetry but the two molecules are not related in any simple way. Maybe layer (100) has an approximate *c* glide with mirror perpendicular to **b** but α is 99.1°.
- NUCLIG *none found* The molecular conformations are very different; the *rmsd* for overlay with inversion is 1.30 Å (but is 0.18 Å with both inversion and flexibility).
- PTHAZO10 *none found* The molecular conformations differ by a *ca.* 40° rotation of a Ph ring. The one H-bond acceptor is O= in one molecule and -N= in the other.

Comparison with the Rekiş list of kryptoracemates

The Rekiş list of kryptoracemates has 21 more structures in $P1$ than does the CPB list but 20 of them have $0.050 < R \leq 0.075$. The 21st is a duplicate (KOVBIG10) of a structure on the CPB list.

8 $PSTG P\bar{1}$ kryptoracemates on the CPB list (DAGXUH, EKOKOE, GAGFAX, IHULUU, JEGHUB, LUYMAV, RUCXAQ, WEPWUJ) are on the Rekiş list of kryptoracemates but are not on either of his two main lists of approximately centrosymmetric structures. *PLATON* recommends $P\bar{1}$ at the 100% level for all 8 of those $PSTG$ structures.

2 reliable, approximately centrosymmetric kryptoracemates (KETWAK, XEPHUX) found by CPB are not on the Rekiş list of kryptoracemates.

1 other discrepancy: kryptoracemate NITSUH, which was reported to be in $P1$ but is probably better described in $Pna2_1$, is not on the Rekis list.

References

Groom, C. R., Bruno, I. J., Lightfoot, M. P. & Ward, S. C. (2016). *Acta Cryst.* **B72**, 171-179.

Henling, L. M. & Marsh, R. E. (2014). *Acta Cryst.* **C70**, 834-836.

Kopský, V. & Litvin, D. B. (2010). Editors, *International Tables for Crystallography*, Vol. E, 2nd ed., *Subperiodic Groups*, Chichester: Wiley.

Macrae, C. F., Bruno, I. J., Chisholm, J. A., Edgington, P. R., McCabe, P., Pidcock, E., Rodriguez-Monge, L., Taylor, R., Streek, J. v. d. & Wood, P. A. (2008). *J. Appl. Cryst.* **41**, 466-470.

Appendix 1: Additional comments about the *PLATON* and Rekis algorithms

Comments about the use of *PLATON*

A *PLATON* recommendation for higher symmetry (default tolerances) was often found to be reliable if the level was above 95% but there are exceptions.

(1) Structures that are known to be related by a phase transition to a higher symmetry phase
E.g., CAZLAR02, CILHIO38, FUGMIF, IBEPUC06, NAKNEV04, NAKNOF04, QOPVAU, RIDFOA01, and possibly EHILOY, POPGUW01, VAHSIJ

(2) Structures of homochiral molecules that are barely chiral
[*e.g.*, have one stereocenter $RCH(Me)R'$]:
at least 40 that would be racemic if Me groups were switched or disordered with H atoms in half the molecules

at least 16=7+4+5 that would be racemic if OH, vinyl, or *i*-Pr groups were switched or disordered with H atoms

at least 10 in which the molecules have adjacent stereocenters so that the overall VDW surface is very similar for the molecule and its enantiomer

e.g., ELEUJUB01, REBLIU, SIKBAQ

achiral molecules that have a preferred conformation that is barely chiral:|

e.g., FAWMAU, OGUQAK (central Si-O-Si),

Note that the *PLATON* output says:

“Proposed Inversion or (Glide) Planes do NOT apply for Chiral Molecules.”

(3) Structures in which there is disorder but the occupancy factors are far from 0.5
e.g., AXUVUM, DIWGEW, DIWGAS01, DIWGIA, PETPEL, POGKEB

In some cases there may be ambiguity:

RIGSEF ($Z=2$) Marsh (1999) suggested $P\bar{1}$ and *PLATON* suggests $P\bar{1}$ at the 91% level but the approximate center that relates the two cations does not relate the F atoms of the two trifluoroacetate anions. Unfortunately there are no atomic ellipsoids available in the CSD and little information about the structure determination is available.

Other problems:

PLATON sometimes seems oversensitive to the possibility of inversion symmetry. *E.g.*:

SEYVUQ The recommendation is for $P2_1/n$, $Z=1/2$ at the 92% level but Pc seems to be a better description. The molecule cannot have inversion symmetry without disorder but one of the two independent molecules was reported to be fully ordered.

VAVHEH The recommendation is for *Pnma* at the 93% level but Marsh & Henling (2014) recommended *Pna2₁*. There are obvious, albeit small, deviations from mirror symmetry.

While *PLATON* was designed to recognize overlooked symmetry rather than to find approximate symmetry it seemed worthwhile increasing the tolerances to see if the program would find approximate symmetry. Increasing the tolerances (*ADDSYM 2 1 1 1*) sometimes led to identification of approximate symmetry that might otherwise have been missed but other times suggested *P $\bar{1}$* when approximate rotational symmetry is more obvious.

When run with increased tolerances the *PLATON* recommendations for space groups other than *P $\bar{1}$* , namely *Pc* (e.g., LAVZAK) and *Pn* (e.g., QOMNEM), *P2₁* (e.g., JOWYEB, ZANFEC), *C2* (e.g., EHUGIX) and *I2* (e.g., KARNOI), and *Cc* (e.g., ARAZUP) and *Ia* (e.g., HISPAB, KOBQOH) usually identified approximate symmetry. Those recommendations were helpful because of the transformation matrices *PLATON* provides.

When run with increased tolerances the *PLATON* recommendations for space group *P $\bar{1}$* were sometimes misleading:

AXEKAP (Z=2)	100% <i>P$\bar{1}$</i> but while parts of the molecules are related by local inversion centers other parts are related by local 2 ₁ axes
FACXIR (Z=4)	96% <i>P$\bar{1}$</i> but approximately <i>C2</i> .
FOLTAC (Z=2)	97% <i>P$\bar{1}$</i> but the approximate symmetry is a glide
GOPQEJ (Z=3)	100% <i>P$\bar{1}$</i> but the approximate symmetry is a pseudotranslation [<i>1$\bar{1}$1</i>]/3
XOPTED (Z=2)	100% <i>P$\bar{1}$</i> but very nearly <i>P2₁</i>
ZIVDEN (Z=1)	98% <i>P$\bar{1}$</i> but the large cation has no approximate inversion symmetry at all

Sometimes even the increased tolerances failed to issue any alert for a structure having approximate symmetry that is obvious visually, e.g.:

EGIXUO (Z=2) Transformation (102|100|120) gives an *F*-centered cell with angles 88.7, 94.9, and 88.9° and obvious axes 2 and 2₁ along the new **b**. The *F2* cell could be transformed to a *C2* or *I2* cell with a volume reduction of 50%.

JETJOK (Z=2) A hindered biphenyl that is reported to be resolvable. The two molecules would be enantiomeric if the 2- and 6- phenyl-ring substituents (Me and I) were switched in one of them. *PLATON* with increased tolerances finds no additional symmetry but if the I atoms are changed to C atoms *PLATON* recommends *P $\bar{1}$* at the 100% level.

JEYFUO01 (Z=2) Transformation [*110|1 $\bar{1}$ 0|00 $\bar{1}$*] gives a *C*-centered cell with angles 91.5, 91.3, 93.0° and obvious 2 and 2₁ axes along the new **b**. The structure has approximate symmetry *C2*.

KISCUM (Z=2) A sulfoxide. The molecules would be enantiomers if the S=O and S-*LP* groups were switched in one of them. There is an approximate *b* glide with mirror perpendicular

to **a**; the angles β and γ are 89.8 and 89.3°. If the O atoms are removed *PLATON* still issues no alert even with increased tolerances even though the approximate *Pa* symmetry is obvious visually.

OBZXP ($Z=2$) Possible mirror symmetry of the molecule is broken by an F atom replacing an H atom in a C₅O ring. *PLATON* with increased tolerances finds no additional symmetry but if the F atoms are changed to H atoms *PLATON* recommends $P\bar{1}$ at the 100% level.

PACVEX ($Z=2$) The molecules would be enantiomers if a CH=C(Me)-O group in a C₅O ring were switched to O-C(Me)=CH in one of them. *PLATON* with increased tolerances finds no additional symmetry but if the O atoms are changed to C atoms *PLATON* recommends $P\bar{1}$ at the 100% level.

QAWNAD01 ($Z=2$) The cations would be enantiomers if the Ph and 4-ClPh substituents at the one stereocenter were switched in one of them. *PLATON* with increased tolerances finds no additional symmetry but if the Cl atoms are removed *PLATON* recommends $P\bar{1}, Z'=1$ at the 100% level).

[The structures of **QAWNAD01** ($P1, Z=2$) and **QAWNAD** ($P\bar{1}, Z'=1$) are very similar. In the 1999 $P\bar{1}$ structure an 80:20 disorder was reported; no disorder was reported in the 2011 $P1$ structure. It seems likely that the difference between the two structures is the enantiomeric purity of the crystal studied. A 50:50 H/Cl disorder in the two phenyl rings at the one stereocenter would give a $P\bar{1}, Z'=1$ structure.]

QOQXAX ($Z=2$) The molecules would be enantiomers if the 4-MePh and 4-ClPh substituents were switched in one of them. *PLATON* with increased tolerances finds no additional symmetry but if the Cl atoms are changed to C atoms *PLATON* recommends *Cc* at the 100% level.

It is important to remember that *PLATON* was designed to find overlooked crystallographic symmetry; it was not designed to find approximate symmetry. Even given that *caveat* it seems that *PLATON* when run with increased tolerances quite often identifies approximate inversion symmetry when a description including a twofold rotation would be better.

Comments about discrepancies between the *PLATON* and Reki algorithms

*Structures in the top 200 of the Reki list of 1200 that give no *PLATON* alert $\geq 90\%$: for possible inversion symmetry. Structures were categorized as $P1$ mimics unless stated otherwise.*

PACVEX is #4 on Reki list but gives no *PLATON* alert even with increased tolerances. The chirality depends on the difference between -O- and -CH= groups at the 3 and 5 positions of a C₅O ring with 4-Me and 2, 5 carbonyl substituents.

OBZXP is #13 on the Reki list but gives no *PLATON* alert unless the F atom is changed in each molecule to an H atom, in which case the alert is at the 100% level. (With that change the molecule could have mirror symmetry but it does not.)

WASVOB is #96 on the Reki list but gives a *PLATON* alert at only the 85% level. The one stereocenter (not marked in the CSD) is at the junction of the C₅N and C₃N₂ rings. WASVOB is an excellent example of a structure in which inversion of N atoms is important.

SAJGAN is #151 on the Reki list but gives no *PLATON* alert even with increased tolerances. The substituents related by the approximate inversion are CF₃ and CH₂NO₂. If the Fs and the H₂NO₂ are removed *PLATON* issues an alert but only at the 85% level.

NILQOP is #180 on the Reki list but gives a *PLATON* alert at only the 85% level. The approximate inversion could be crystallographic if an O atom in the C₃O₂ ring and the adjacent C=NPh were switched in one of the molecules. There is no question that the inversion is only quite approximate but the formation of a solid solution of the two enantiomers seems possible.

PIMROV is #199 on the Reki list but gives no *PLATON* alert even with increased tolerances. The molecule could have twofold symmetry. The *rmsds* for the overlays without and with inversion are 1.78 and 0.62 Å. The two stereocenters are both CH(OH)(*t*-Bu); inversion leaves the OH groups pointing in different directions. The paper indicates that rotation around the C-C bond of this 2,2'-bipyridine derivative is not highly restricted.

Structures that give a PLATON alert ≥95% for approximate inversion symmetry but are not on either Reki list

PAZFUR *PLATON* issues an alert at the 95% level but PAZFUR is not on either Reki list. *Rmsds* for overlap as found and with inversion are 0.84 and 0.31 Å for pair #1&2 and 0.57 and 0.32 Å for pair #3&4. No relationship between the two nearly centrosymmetric pairs was found.

PIJSOR *PLATON* issues an alert at the 96% level but PIJSOR is not on either Reki list. *Rmsds* for overlap as found and with inversion are 1.47 and 0.94 Å for pair #1&4 and 1.48 and 0.97 Å for pair #2&3. The two π...π stacks (#1&4; #2&3) are clearly different but both have good approximate inversion symmetry

QAMQIG *PLATON* issues an alert at the 95% level but QAMQIG is not on either Reki list. *Rmsds* for overlap as found are 0.91 Å for #1&3 and 0.93 Å for #2&4. Molecules are achiral but for one H/F substitution; if the F atoms are changed to H atoms the *rmsds* for the pairs are 0.92 and 0.94 as found and 0.06 and 0.05 Å with inversion. The orientations of the two nearly centrosymmetric H-bonded dimers (#1&3, #2&4) are not related in any simple way.

UPENEJ *PLATON* issues an alert at the 95% level but UPENEJ is not on either Rekis list. *Rmsds* for overlap as found are 0.16 and 0.24 Å for pairs #1&4 and #2&3. There are two stereocenters at the two points of attachment of the cyclopropyl rings but the molecules have approximate mirror symmetry. The pairs of molecules related by approximate inversion have orientations that are not related in any simple way

YAVJIR *PLATON* issues an alert at the 96% level but YAVJIR is not on either Rekis list. The *rmsd* for molecular overlay without inversion is 0.16 Å. The molecules have good approximate mirror symmetry except for the two -CH(*i*-Pr)- stereocenters that are at the far end of five-membered rings.

Appendix 2: Additional comments about the molecular overlay function in *Mercury*

Rmsds for overlay of pairs molecules as found in the crystal, and both with and without inversion and flexibility are often informative.

Allowing flexibility usually lowers the overlay *rmsd* for a chiral molecule more than does allowing inversion unless both the *rmsds* are ≤ 0.30 Å or the molecule has a degree of freedom that the flexibility algorithm cannot adjust.

In a non-Sohncke mimic inverting one of the molecules usually reduces the *rmsd* below 0.30 Å and almost always does so if flexibility is also allowed.

For co-crystals of diastereomers the *rmsd* for overlay with inversion and the *rmsd* for overlay with flexibility should both be larger than *ca.* 0.25 – 0.30 Å.

Note that the flexibility option does not allow for (*e.g.*) inversion of sp^3 N atoms or for conformational changes in *n*-membered rings. It *does* allow internal rotations that are almost certainly too hindered to occur at normal temperatures (see JETJOK and TEFEP). And there is occasionally a problem (almost always minor) with reproducibility.⁶

There are occasional exceptions to the guidelines given above:

GUMCEY, a kryptoracemate that has an *rmsd* of 0.44 Å with inversion (*i.e.*, > 0.30 Å) and 0.39 Å with both inversion and flexibility. The ring conformations are all nearly the same but there are many small conformational differences that the flexibility algorithm cannot handle.

⁶ From the *Molecular Overlay* section of the *Mercury* website accessed on 2March2021:

“Rings are not flexed during the overlay. Hydrogen atoms are ignored. The RMSDs reported for multiple (*i.e.*, repeated) overlays of the same pair of molecules may vary. The algorithm is not exhaustive and between overlays, may find different, closely related solutions.”

Experience indicates that the non-reproducibility problem is rarely significant but that when it is encountered it can be confusing.

LIPZEP, a compound of diastereomers that has an *rmsd* of 0.19 Å with inversion (*i.e.*, < 0.25 Å) and 1.65 Å with flexibility. There are three stereocenters. The one homochiral center has substituents H and Me; that center makes little difference to the overlay. The fit with inversion of 31 of the 32 non-H atoms is excellent. The paper reporting LIPZEP and the related LIPZUE ($R=0.53$) is a very interesting read.

Because of the limitations of the flexibility algorithm the *rmsd* for overlay with inversion is occasionally considerably smaller for homochiral molecules than is the overlay with flexibility. In the following examples the *rmsds* are listed in the order as observed, with inversion, with flexibility allowed.

EVUROE ($P\bar{T}$ mimic) One stereocenter; *rmsds* 2.16, 0.25, 0.66 Å/

PAZFUR ($P\bar{T}$ mimic) One stereocenter; *rmsds* 0.84, 0.31, 0.54 Å for molecules 1&2 and 0.57, 0.32, 0.53 Å for molecules 3&4/

SIKBAQ ($P\bar{T}$ mimic) Two stereocenters; *rmsds* 2.18, 0.18, 0.56 Å for molecules 1&2 and 2.05, 0.26, 0.33 Å for molecules 3&4/

YAZMOB ($P\bar{T}$ mimic) One stereocenter; *rmsds* 3.15, 0.28, 0.45 Å/

Because the flexibility option does not allow for the inversion of sp^3 N atoms the following structures look like compounds of diastereomers (because the *rmsds* for overlay with inversion and with flexibility are both large) but are not:

HIYDEZ ($P\bar{T}$ mimic) One homochiral stereocenter and one heterochiral sp^3 N atom; *rmsds* 2.52, 0.38, 0.61 Å/

VICSOR ($P\bar{T}$ mimic) $C_1 H_{10} Br N_3 O$. One homochiral stereocenter and two sp^3 N atoms one of which is very obviously heterochiral. The *rmsd* of 0.96 Å is not lowered by either inversion or flexibility.

WASVOB ($P\bar{T}$ mimic) One homochiral stereocenter and four sp^3 N atom, two of which are pyramidal and connected to three different groups. Those two N atoms have opposite configurations in the two molecules. *Rmsds* are 1.88, 0.37, 1.51 Å. Differences in the conformation of the C_5N ring that contains the C-atom stereocenter also increase the *rmsds*.

There are other limitations of the overlay algorithm. Overlay of the molecule in VAWNEO suggests that rotations around the single bonds adjacent to a triple bond are not handled well. The conformational change needed to overlay the homochiral pentapeptides in NEZVAQ is not found because it requires a change in the direction of coiling.