



Identifying and characterizing translationally modulated molecular crystal structures

Carolyn Pratt Brock^{a*} and Robin Taylor^b

^aDepartment of Chemistry, University of Kentucky, 505 Rose Street, Lexington, KY 40506-0055, United States, and
^bCambridge Crystallographic Data Centre, 12 Union Road, Cambridge, CB2 1EZ, United Kingdom. *Correspondence
e-mail: cpbrock@uky.edu

Received 12 May 2020

Accepted 10 June 2020

Edited by A. J. Blake, University of Nottingham, United Kingdom

Keywords: pseudosymmetry; approximate translations; modulations; Z' .

Supporting information: this article has supporting information at journals.iucr.org/b

Most structural (*i.e.* displacive) modulations make molecules independent that had been related by translation in a phase having a smaller or centered unit cell. In the modulated structure the independent molecules are differentiated by small translations, rotations, and/or conformational changes but an approximate translational relationship is normally retained. A program has been written to identify such pseudotranslations because they can be difficult to find by eye and because they combine with each other and with lattice translations in ways that can be confusing. To characterize the pseudotranslations the program calculates their fractional translational, orientational, and conformational components as well as several quality indicators. While many pseudotranslations are obvious, others are borderline; setting tolerances for identifying a pseudotranslation proved difficult. Defaults were chosen to reproduce experience-based judgment but they can be varied in the program input. The program was run for organic and for metallo-organic structures with $R \leq 0.075$ in the 2019 release of the Cambridge Structural Database. The frequency of pseudotranslations increases with Z' and is approximately 50% for $Z' > 4$. Some structures were found in which an identified pseudotranslation cannot correspond to a modulation. These include structures in which some but not all of the molecules are related by pseudotranslations and structures in which pseudotranslations in different parts of the unit cell have different directions.

1. Introduction

In a modulation¹ small translations, rotations, and/or conformational changes make molecules² independent that would otherwise be related by symmetry, which is usually crystallographic translation (Fig. 1). It has been observed (Brock, 2016) that *ca.* half of the 'organic' structures archived in the Cambridge Structure Database (CSD) (Groom *et al.*, 2016) and having more than four independent formula units in the unit cell ($Z' > 4$) are modulated.³

We sought to use the $Z' > 4$ structures that had already been investigated to develop an algorithm to find translational modulations automatically. Doing so is desirable because it can be easy to miss seeing them, even if using a sophisticated graphics display program like the CCDC's program *Mercury* (Macrae *et al.*, 2020). The modulation directions are not



¹ More precisely, in a displacive modulation (rather than occupational or ADP modulations; Schönleber, 2011).

² Unless the context clearly indicates differently, the term 'molecule' is intended to include any molecular ions, counterions, and/or solvent molecules incorporated in a structure.

³ All of those structures were refined assuming commensurate modulations (otherwise they would not be in the CSD) although it is certainly possible that some would be better described as incommensurate.

necessarily parallel to a cell axis [e.g. Fig. 1(a)] and the distances between related molecules may be long.

We also wanted to know how frequent translational modulations are for $Z' \leq 4$ structures, which are too numerous to consider individually, and how the modulations influence space-group distributions. An additional goal was to quantify the major geometrical components of the modulations, *i.e.* transverse and longitudinal displacements, molecular rotations, and conformational differences.

Note that not all displacively modulated structures include an approximate translation. The simplest examples of those that do not are the $P1$, $Z = Z' = 2$ structures that seem to be derived from $P2_1$ or $P\bar{1}$ structures having $Z = 2$ and $Z' = 1$ (see the supporting information). If molecules related by 2_1 or $\bar{1}$ symmetry become independent they are seldom also related by a pseudotranslation.

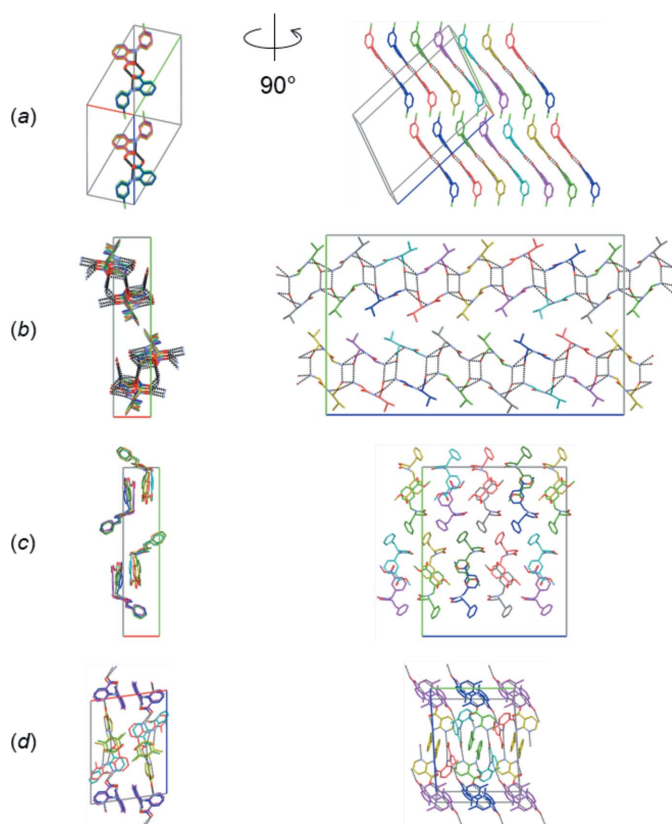


Figure 1

Some examples of modulated structures. In this and all following drawings the **a**, **b** and **c** axes are colored red, green and blue, respectively, and the C atoms are colored by symmetry, *i.e.* a different color used in each symmetry independent molecule. In this and some of the following drawings the non-C atoms are colored by element type. (a) The $[111]/6$ pseudotranslation in KUTMAP02 ($P\bar{1}$, $Z' = 6$; López-Mejías & Matzger, 2015). (b) The $[001]/7$ pseudotranslation in WEVWOK ($P2_12_12_1$, $Z' = 7$; Görbitz *et al.*, 2007). (c) The $[001]/4$ pseudotranslation in NANBEL ($P2_1$, $Z' = 8$; Fresno *et al.*, 2011); the two independent pseudotranslational columns in NANBEL are related by approximate screw axes with the basic cell having approximate $P2_12_12_1$ symmetry ($\beta = 90.4^\circ$). (d) The $[010]/2$ pseudotranslation in CUNKUS01 ($P\bar{1}$, $Z' = 6$; Long & Li, 2009), in which the three pairs of related molecules have very different orientations.

2. Modulations, pseudotranslations and pseudotranslational columns

Most modulations include an approximate translation. If the modulation increases the volume of the unit cell then molecules that were related by translation become independent; in the larger cell at least some of the molecules are related by approximate translation, *i.e.* by pseudotranslation.

A *pseudotranslational column* is a column of molecules of the same residue that are related by approximate translation. [The term ‘residue’ is used to indicate a chemically unique component of a structure. For example, an asymmetric unit containing six phenylethylammonium and six dihydrogen phosphate ions has twelve molecules (as defined above) and two residues.] The column will lie along a lattice direction $[uvw]$ and hence will have a repeat unit determined by the length of that lattice vector.

If all the symmetry independent molecules of the residue lie in pseudotranslational columns, this is said to be a *pseudotranslation* and the residue is *pseudotranslated*. If there are s independent molecules of the residue in the asymmetric unit and s has a factor f , the pseudotranslation may consist of a single pseudotranslational column containing s independent molecules in its repeat unit but alternatively may consist of s/f columns containing f molecules each, *i.e.* a pseudotranslation may consist of ≥ 1 pseudotranslational columns.

If all the residues in a crystal structure are pseudotranslated in the same direction, the structure is modulated. If a structure is modulated it should have a corresponding basic structure (smaller and/or higher-symmetry unit cell) that can at least be imagined to exist. Since the modulation usually increases the volume of the unit cell, molecules that were related by crystallographic translation become independent; in the larger cell the molecules are instead related by pseudotranslation. (*Crystallographic* symmetry is used to refer to the symmetry of the time- and space-averaged structure. The term is used in preference to terms like exact or perfect symmetry because no symmetry operation in a crystal is exact in the absence of averaging, especially if there is disorder.)

Modulations that lower the symmetry of the unit cell with or without increasing its size may lead to screw, glide, or inversion operations becoming approximate. If the molecular orientations are sufficiently similar a remnant screw or glide operation is a pseudotranslation.

The basic cell of a modulated structure is the unit cell that would be expected if all the approximate symmetry relationships were to become crystallographic. A transformation from a modulated cell to its basic cell is sometimes observed if the crystal is heated and is occasionally observed with cooling (see the supporting information). If no basic cell can be identified then the structure is probably not modulated.

It is sometimes reasonable to regard a structure as modulated even if some of the residues are not pseudotranslated. If there is pseudotranslation of the larger residue(s) but not of small residues like H_2O and NH_4^+ , then there is probably a modulation because those small residues could almost certainly reorient and translate during a transition.

Pseudotranslations can also exist in structures that are not modulated. Some pseudotranslations occur in only certain regions of the unit cell. There are some salts and co-crystals built from two ions or molecules of similar size in which a pseudotranslation relates only one of the two residues. There are some layered structures in which a pseudotranslation has different directions in different layers (*e.g.* [100] and [010] in a $P4_1$ structure). In most cases, however, a pseudotranslation does correspond to a modulation; some exceptions are listed in the supporting information.

3. Methodology

3.1. Outline of algorithm

A full description of the program for finding modulations is included in the supporting information. Here we give only a brief summary. The algorithm looks for modulations along each of a set of directions $[uvw]$, where $|u|$, $|v|$, $|w|$ are less than a user-specified value m . An array of $(m+1) \times (m+1) \times (m+1)$ unit cells is generated and the centroids of the molecules in the array calculated.

If there is more than one residue in the structure, the program then considers each in turn as follows. In the rest of this section, ‘molecules’ means molecules of the residue currently being processed, and ‘centroids’ means centroids of those molecules.

Modulations are found by searching for columns of centroids that lie approximately along the $[uvw]$ direction currently being examined. To be acceptable, a column must contain $n+1$ centroids, where $n \geq 2$. The first (1) and last ($n+1$) centroids must be separated by the length, d , of the $[uvw]$ lattice vector, *i.e.* related by crystallographic translational symmetry. The other centroids must lie in intermediate positions.

The objective is to assign every molecule in the unit cell (not just those in the asymmetric unit) to columns. If this is achieved and at least one of the columns contains ≥ 2 symmetry-independent molecules, the residue might be pseudotranslated (only might be – extra checks are required, see next paragraph). If only some of the molecules are in columns, the residue might be partially pseudotranslated (*i.e.* only in certain regions of the unit cell) provided that the columns that exist contain at least one example of each of the symmetry-independent molecules and at least one of the columns contains ≥ 2 symmetry-independent molecules. Otherwise, the residue is not pseudotranslated.

If pseudotranslation is a possibility, its quality is assessed. Most importantly, it is determined whether the molecules in a column are acceptably aligned. This is tested by taking each of the $n(n-1)/2$ pairs in turn (not just the pairs involving adjacent molecules, and excluding pairs involving molecule $n+1$) and moving one of the molecules towards the other by translating it along $[uvw]$ by the distance expected for ideal translational symmetry (translation c in Fig. S1).⁴ After this translation, the fit of the two molecules is measured by the root-mean-square deviation (rmsd) of their matching non-

hydrogen atoms. This rmsd must be less than a user-specified value. Further, av_rmsd , the average rmsd over all the columns must also be less than a user-defined value. If so, the set of columns is identified as a pseudotranslation provided various other checks (see the supporting information) are passed.

If a residue is pseudotranslated in the current search direction, but only in some regions (*e.g.* layers) of the unit cell, the pseudotranslation is marked as *incomplete* and the structure will not be classified as modulated, although the incomplete pseudotranslation will be output. The problem of layered structures is discussed more fully in the supporting information. (A sample output file, which is a spreadsheet, is available with the supporting information. The file includes the results for all the refcodes mentioned in the text and supporting information.)

3.2. Data sets

The program was run on organic structures from the CSD (2019 release) with R -factor ≤ 0.075 . Runs were also performed on polymeric (*catena*) and non-polymeric metallo-organic entries using the same criteria.

3.3. Working definition of translational modulation

A precise definition of modulation that could be implemented within the program was required for the results analysis. While the best evidence for a displacive modulation is the presence of a smaller and/or higher symmetry basic cell, parts of that criterion cannot yet be coded. The definition given below for the more restrictive term *translational modulation* is necessarily somewhat arbitrary.

Firstly, all residues having > 1 non-hydrogen atoms and with > 1 symmetry-independent molecules must be pseudotranslated in the same direction. Secondly, the pseudotranslation must describe the entire unit cell, *i.e.* a pseudotranslation that is confined to a layer (or set of layers) for any of the residues is unacceptable. The program rmsd criteria (given in the supporting information) determines what constitutes a pseudotranslation.

Some problems with this definition are discussed in the supporting information.

4. Parameters used to describe pseudotranslations

4.1. Order

Pseudotranslation vectors are written here as $[uvw]/n$, where all four values are integers. The vector $[uvw]$ describes the direction and period of the pseudotranslation; n gives the number of molecules related [*e.g.* (111)/6 in Fig. 1(a)]. A pseudotranslational column of $n+1$ molecules found by the program defines the repeat unit but only the first n are regarded as being in the pseudotranslation as molecules 1 and

⁴ In a set of $n+1$ molecules spanning a total distance d , the centroids of any two of them would ideally be separated by $(p+1)d/n$, where p is the number of intervening molecules. For a pair of adjacent molecules this reduces to d/n since $p=0$, but non-adjacent pairs ($p>0$) are also considered.

$n + 1$ are related by crystallographic symmetry. Alternatively, these two molecules could both be included but at half weight. An alternative form of $[uvw]/n$ is $[u_1u_2u_3]/n$, which will be useful later.

The order n of the pseudotranslation is normally either Z' [Figs. 1(a) and 1(b)] or a factor of Z' [Figs. 1(c) and 1(d)]. The few examples of $Z' > 4$ structures in which different residues, or two sets of the same residue, are related by different pseudotranslations are described in the supporting information, as are the few self-inclusion compounds in which n is not a factor of Z' .

When running the program, the upper limit on $|u_i|$ was set at 4 because values ≥ 5 sometimes cause calculation times to rise steeply and very rarely find pseudotranslations of interest. Only one important pseudotranslation with a higher u_i was found: $([015]/10$ in SEMPEH (Scott *et al.*, 2012; Fig. S2).⁵

4.2. Number of independent molecules

The n molecules in a pseudotranslational column are usually all symmetry independent but occasionally two or more are symmetry related (Figs. S3 and S4). The quantity n_{ind} is the number of independent molecules in a column; max_{ind} and min_{ind} are, respectively, the maximum and minimum values of n_{ind} in the columns of a pseudotranslation. Max_{ind} and min_{ind} can differ only if a pseudotranslation contains >1 column [e.g. an $n = 4$ pseudotranslation in a $Z' = 6$ structure (Figs. S4) with one of the two columns including a symmetry element so that $max_{ind} = 4$, $min_{ind} = 2$].

4.3. Deviations from ideal translational symmetry

We described earlier how the parameter av_{rmsd} is used to measure the overall deviation of a pseudotranslation from ideal translational symmetry.⁶ The av_{rmsd} value can be broken down into components:

conformational: the molecules may adopt different shapes;

orientational: a rigid-body rotation may be needed to position one molecule optimally on another;

transverse: the centroids of the molecules in a column may not lie exactly in a straight line but may instead have perpendicular ('transverse') displacements from their median $[uvw]$ vector (the median vector lies along the $[uvw]$ direction and is positioned so as to minimize the transverse deviations);

longitudinal: the projections of these centroids onto their median vector may not be evenly spaced.

We show in the supporting information that the relative contributions of these components to av_{rmsd} can be calculated as

$fract(conf)$, $fract(orient)$, $fract(trans)$ and $fract(long)$, where:

$$fract(conf) + fract(orient) + fract(trans) + fract(long) = 1.$$

⁵ The manual search that found the SEMPEH $n = 5$ pseudotranslation was not exhaustive.

⁶ We also use max_{rmsd} , the maximum rmsd calculated for any molecule pair, when assessing the quality of the pseudotranslation (Section 3.1).

5. Other features of pseudotranslations

5.1. Combinations with lattice vectors; the base pseudotranslation

As is illustrated in Fig. 2, pseudotranslations are almost invariably detected in several directions because a pseudotranslation vector $[u_1u_2u_3]/n$ can be combined with any lattice vector $[l_1l_2l_3]$ (e.g. $[100]$ or $[011]$). If there is one pseudotranslation then there must be an infinite number of derived pseudotranslations (as is illustrated in Fig. S5). Almost all of the pseudotranslations, however, are too long to be obvious visually; the longer the pseudotranslation, the more likely it will be obscured by interleaving molecules. Different combinations of a pseudotranslation vector with lattice vectors can have different ratios of the transverse and longitudinal components.

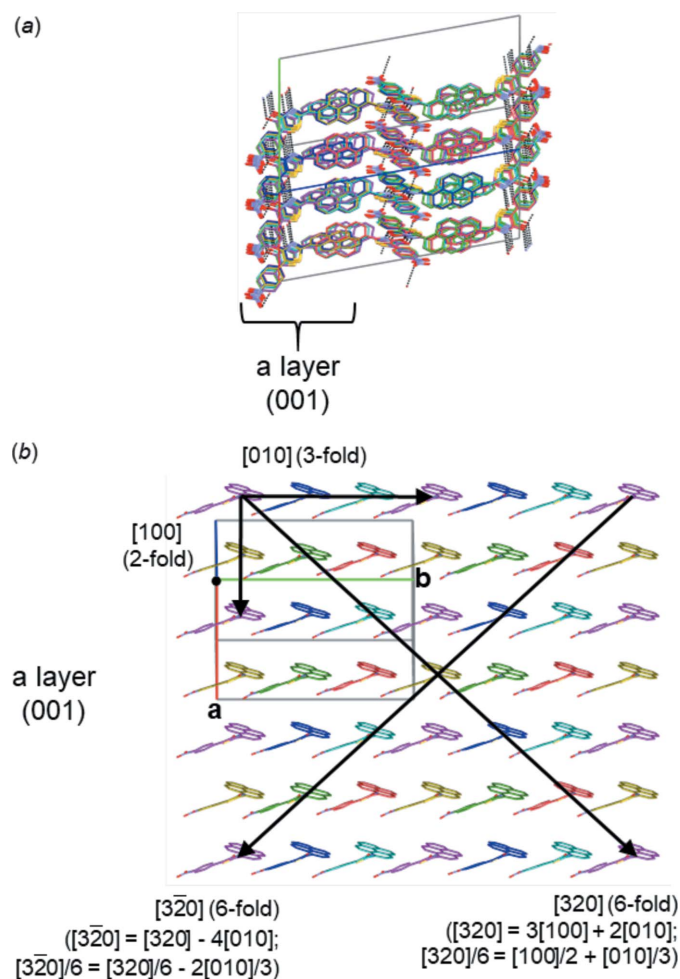


Figure 2

Two views of the structure MUKWIA ($P2_1/c$, $Z' = 6$; Wrona-Piotrowicz *et al.*, 2015). The disorder, which involves only small atomic displacements in three of the molecules, is not shown. (a) A view along $[320]$. (b) A view of a layer (001) showing that the $[320]/6$ and $[3\bar{2}0]/6$ pseudotranslations are related by addition of translations along \mathbf{b} but are not related by any other symmetry operation. The orientations of the molecules relative to the two directions are not the same. Note also that the pseudotranslation $[320]/6$ can be viewed as the combination of the pseudotranslations $[100]/2$ and $[010]/3$, and that $[001]$ is the only direction in which the related molecules are in direct contact.

When combined with a lattice vector the modulation vector $[u_1u_2u_3]/n$ can be multiplied by any integer k so that the set of possible pseudotranslations becomes:

$$k[u_1u_2u_3]/n + [l_1l_2l_3]$$

Visually obvious pseudotranslations usually have $k = \pm 1$, but values of $n/2$ and $(n-1)/2$ can also be important (e.g. $k = 2$ if $n = 4$ or 5).

If n has an integer factor $p \neq 1$ (e.g. 2 and 3 for $n = 6$) then there is also a set of pseudotranslations that relate n/p molecules:

$$k[u_1u_2u_3]/(n/p) + [l_1l_2l_3]$$

(e.g. a pseudotranslation composed of two columns of three molecules each if $n = 6$ and $p = 2$). There is a p -fold pseudotranslation coincident with each n -fold pseudotranslation just as a sixfold rotation axis must have a coincident threefold rotation axis.

In a commensurate structure the shortest pseudotranslation that relates the most molecules (n_{\max}) is almost always unique and is termed the *base pseudotranslation*. That pseudotranslation usually seems to be the most worthy of attention [e.g. (111)/6 for KUTMAP02; López-Mejías & Matzger, 2015; Fig. 1(a)]. There are some structures, however, in which that pseudotranslation seems to be the result of combining two more meaningful lower- n modulations (e.g. MUKWIA; Wrona-Piotrowicz *et al.*, 2015; Fig. 2).

It is possible, but unusual, for there to be two independent pseudotranslations; in such cases the n values for the two pseudotranslations are prime and are usually the same (e.g. 2 and 2 in the case of $Z' = 4$). Some examples are listed in the supporting information. The sum of the values n_i for independent pseudotranslations that are modulations should not exceed Z' (see Section 7.5 of the supporting information). If two pseudotranslations are listed with relatively prime n_i , then a pseudotranslation of higher order ($n = n_1n_2$) may be missing. In the case of SEMPEH ($Z' = 10$) the [015]/10 pseudotranslation undetected by the program (see Section 4.1) can be inferred from the listing of [010]/5 and [011]/2; [015]/10 = [011]/2 - 2*[010]/5.

When $Z' \geq 4$ the shortest pseudotranslation for a factor of n (e.g. 2 or 3 for $n = 6$) is occasionally collinear with $[uvw]/n$. In those cases the lower- n pseudotranslation is not listed.

5.2. Symmetry relationships

Symmetry operations relating pseudotranslations are not necessarily the same as those relating lattice vectors. All pseudotranslations can be inverted, even in noncentrosymmetric space groups, because inversion does not affect the orientation of the vector, which is the only thing that matters. Secondly, an operation that is allowed for a lattice vector does not necessarily relate two pseudotranslations. In MUKWIA ($P2_1/c$, Fig. 2) the pseudotranslations $[320]/6$ and $[\bar{3}20]/6$ are not the same because there is neither a mirror perpendicular to **b** nor a twofold along **b**. The two modulation vectors in MUKWIA are not, however, independent because they are related by addition of a lattice vector: $[320]/6 - [100] = [\bar{3}20]/6 - [600]/6 = [\bar{3}20]/6$, which is the same (by inversion) as $[320]/6$.

The distance between identical molecules is the same in the two directions as is the ordering of the independent molecules along the pseudotranslation, but the orientations of the molecules relative to the two vectors are different.

5.3. Cell centering

If the cell is centered then the centering vector is a lattice vector and can be combined with the pseudotranslation vector. A complete period of a pseudotranslation of order n in a side- or body-centered cell may include two sets of molecules that are equivalent by translation; such a structure could be transformed to a primitive cell in which there would be a pseudotranslation $[u'v'w']/(n/2)$. The value n_{ind} (the number of independent molecules in the pseudotranslational columns; Section 4.2) will therefore be different from n if the pseudotranslation includes a centering vector. For C - or I -centering $n/n_{\text{ind}} = 2$; for R centering $n/n_{\text{ind}} = 3$. In such cases *av_rmsd*

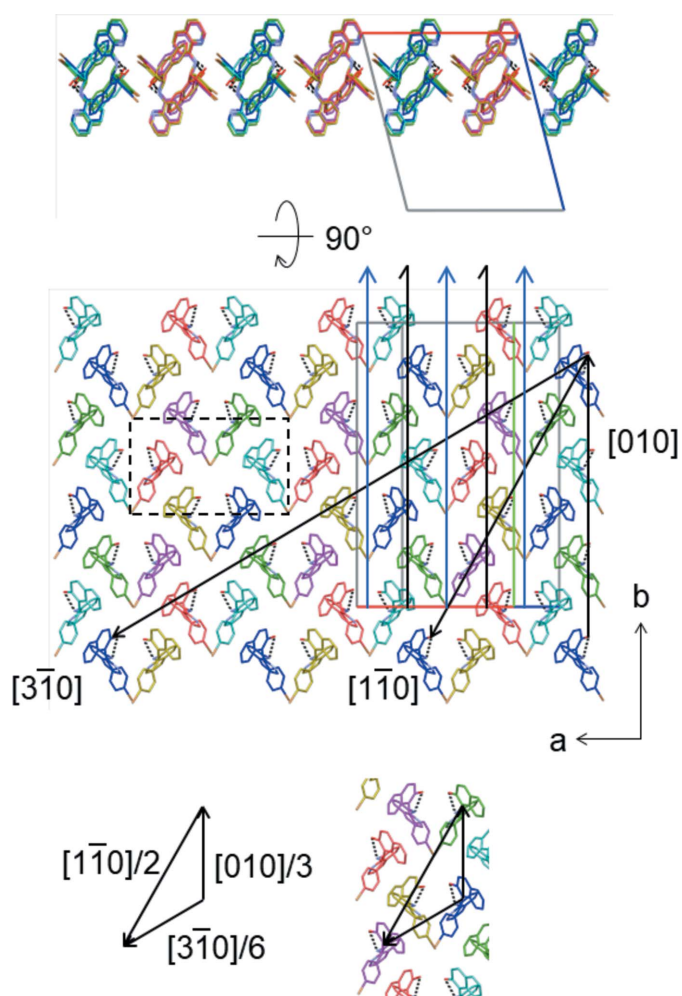


Figure 3
A layer (001) of the structure NAHCOQ (Mukhopadhyay *et al.*, 2010). The three shortest pseudotranslations ($[3\bar{1}0]/6$, $[010]/3$ and $[1\bar{1}0]/2$) are shown; only two of them are independent. The $P2_1/c$, $Z' = 6$ structure seems to be derived from a $C2/c$, $Z' = 1$ basic cell (shown as dashed) by a threefold lengthening of the axis **b**. That modulation is accompanied by the loss of the exact twofold axes (shown in blue) and of the centering, which becomes the pseudotranslation $[110]/2$ (or, $[1\bar{1}0]/2 = [100] - [110]/2$).

should arguably be based only on pairs of molecules that are not related by crystallographic translations.

5.4. Approximate cell centering

If the centering is retained only approximately then the centering vector becomes a modulation vector. For nine of the $Z' = 6$ structures (e.g. NAHCOQ; Fig. 3; other examples are listed in the supporting information) it seems that the axis \mathbf{b} of a $C2/c$ basic cell has been lengthened by a factor of three. In these structures the approximate centering vector is $[1\frac{1}{3}0]/2$ so that there is a pseudotranslation $[310]/6$; $[110]/2 = [330]/6 = [310]/6 + [010]/3$ is then usually listed as well. Also listed in the supporting information are five structures in which there is a similar modulation $[010]/n$, $n \neq 3$ of a basic cell that is centered.

5.5. Symmetry elements within the pseudotranslation

There are some structures in which the pseudotranslation includes a symmetry operation other than centering (Fig. 4). If the aligned molecules in a pseudotranslational column have approximate inversion symmetry then a crystallographic inversion center can be included in the vector $[uvw]$. Similarly, $[uvw]$ may include a glide plane, a screw axis, a mirror, or a perpendicular twofold axis if the molecules have appropriate approximate symmetry and orientation. If a twofold special position is occupied then $n_{ind} = (n + 1)/2$; if not, $n_{ind} = n/2$. Some examples are listed in the supporting information.

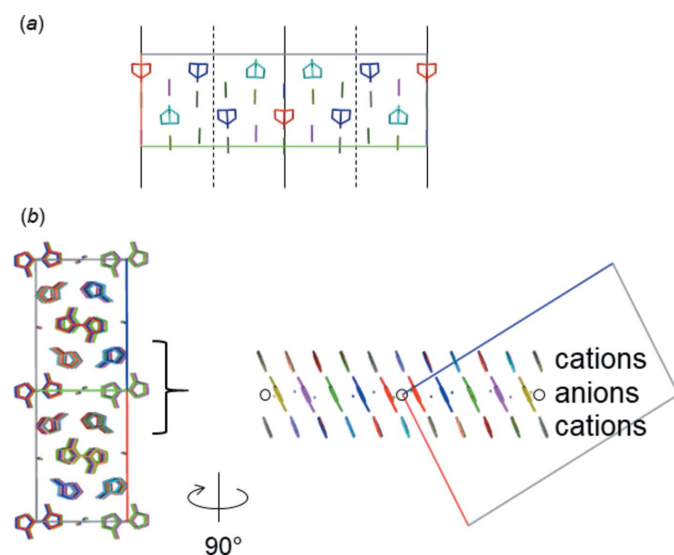


Figure 4
(a) View along \mathbf{c} showing the $[010]/5$ modulation in HEKZOO ($\text{C}_6\text{H}_{12}\text{N}_2 \cdot 2\text{I}_2$; Cm , $Z' = 2.5$; Peuronen *et al.*, 2012). One of the three independent DABCO cages and two of the six I_2 molecules lie on mirror planes. (b) Two views showing the modulation along $[101]$ in OMOBUP (Shang *et al.*, 2016), a hydrated 2:1 salt in $P2_1/c$ with $Z' = 5$. The pseudotranslation relates the ten independent cations. The larger anions also have a $[101]/10$ pseudotranslation but it includes inversion centers and so relates two sets of the five anions. The equivalence of inversion and pseudotranslation is possible for the anions because they are nearly centrosymmetric.

6. Results

This section principally describes the results obtained for organic structures. The metallo-organic structures were used only in analyses of modulation frequencies and space-group distributions. Moreover it was found that polymeric metallo-organic structures very rarely have $Z' \geq 2$ (only 308 of 68 213 structures, $< 0.5\%$) so they were sidelined in this work. (They seem worthy of future study as a very high proportion of these $Z' \geq 2$ structures contain pseudotranslations.) Further references to metallo-organic structures therefore refer only to the non-polymeric subset.

6.1. Most too-good modulations are easy to spot

Some structures appear to be modulated because they have been refined in a too-large unit cell; their pseudotranslations are better described as true translations. If both max_rmsd and av_rmsd are below 0.10 \AA for all residues the modulation is very likely too good. If the largest value is in the range $0.10\text{--}0.18 \text{ \AA}$ the modulation is suspect; a final decision about its validity cannot be made without consulting the original paper,⁷ knowing the fraction of superlattice reflections having $I > 3\sigma(I)$, and examining the atomic displacement ellipsoids. For the $Z' > 4$ structures examined we found that all av_rmsd values for reliable modulations are at least 0.15 \AA for at least one residue and are usually much larger. [Marsh (1999) suggested that in a superstructure at least some of the corresponding atoms should be shifted by 0.5 \AA or more. We believe that criterion is too strict.] Additional discussion of av_rmsd values is available with the supporting information.

A structure that seems to have a too-high Z' value may still have a real pseudotranslation. Consider KUTCUY (Beckett *et al.*, 2010). *PLATON* (Spek, 2009) recommends a change from $P1$, $Z' = 5$ to $I\bar{4}$, $Z' = 1.25$ but in the tetragonal cell there would still be an $n = 5$ pseudotranslation consisting of four symmetry-related molecules on general positions and one molecule on a 4 site ($n_{ind} = 2$).

If a structure has a too-good modulation then that $[uvw]/n$ is a lattice vector that can be combined with any real pseudotranslation. XECPON (Gunawardana *et al.*, 2017) has a too-good modulation $[110]/2$ so that the cell is actually C -centered. The genuine pseudotranslation $[101]/2$ for residue 1 could then also be written as $[011]/2$ because $-[101]/2 + [110]/2 + [001] = [011]/2$.

6.2. Frequencies of occurrence of modulations

Table 1 shows that the frequency of modulations increases with Z' ; about 50% of the structures with $Z' > 4$ are modulated. The small number of modulated $Z' < 2$ structures necessarily have molecules located on special positions (e.g. ZZZVTY12, $P\bar{3}$, $Z' = 1 = 3 \cdot \frac{1}{3}$; $[1\bar{1}\bar{1}]/3$, with all three molecules lying on a threefold axis; Dunand & Gerdil, 1982). The second line of data in the table gives figures for the structures that meet this condition.

⁷ In the case of a very subtle modulation it is the authors' responsibility to justify the choice of the larger unit cell.

Table 1
Frequency of modulated structures as a function of Z' .

Z'	Organic			Metallo-organic		
	Number of structures	Number modulated	% modulated	Number of structures	Number modulated	% modulated
<2	288 363	908	0.3	312 399	1025	0.3
<2†	7656	908	11.9	14 062	1025	7.3
2	35 043	4996	14.3	22 766	3084	13.5
3	2037	523	25.7	1130	308	27.3
4	2220	782	35.2	821	266	32.4
>4	441	213	48.3	143	72	50.3
Total	328 104	7422	2.3	337 259	4755	1.4

† Structures with $Z' < 2$ and at least one molecule in a special position.

Comparison of the right and left halves of the table suggests that metallo-organic structures are less likely to be modulated than organic structures, though not when $Z' = 3$ or > 4 . The difference may however be deceptive. The satellite reflections that indicate the presence of a modulation are likely to be relatively weaker in metallo-organic structures because the more strongly scattering metal atoms are near the molecular centroid and so are less likely to contribute substantially to the satellite reflections. If the satellite peaks were not noticed (or were ignored) then no modulation (or pseudotranslation) would have been reported.

Table 2 shows that base pseudotranslations with $n > 6$ are rare for organic structures; values of $n < 4$ are most common, even when Z' is larger. Unsurprisingly, $n > 4$ is much more common when $Z' > 4$; $n = 2$ is overwhelmingly preferred when $Z' \leq 4$.

Pseudotranslations are more common in structures determined at low temperature,⁸ but about 20% of the $Z' > 4$ structures with modulations were determined at or above 290 K.

In $Z' > 4$ structures, if a pseudotranslation is identified for the largest residue, it is very often found for all the other residues present. Some layered structures (see below) are an exception to this trend. Some other exceptions are listed in the supporting information. The tendency becomes much less pronounced if $Z' \leq 4$.

6.3. Effect on space-group frequencies

The space-group frequencies for all structures considered and for modulated structures found are shown in Fig. 5, where they are grouped by type (organic and metallo-organic). The graphs show that modulation favors lower-symmetry groups. For organic structures there is a decrease in the frequencies of the most common groups No. 14 ($P2_1/c$, etc.) and higher, while there is an increase in the frequencies of groups Nos. 1, 2, 3, 5 and 7. The frequency of No. 9 (Cc) is unchanged. For metallo-organic structures the frequency decrease begins with No. 9, skips No. 14 (which increases some), and continues with No. 15

⁸ Median(modulated) = 173 K, median(unmodulated) = 200 K, statistical significance < 0.001 (Mann-Whitney test).

Table 2
Distributions of the orders of base pseudotranslations in modulated organic structures with $Z' \leq 4$ and $Z' > 4$ (with percentages in brackets).

Order	2	3	4	5	6	>6
$Z' \leq 4$	7834(80.4)	711(7.3)	917(9.4)	20(0.2)	222(2.3)	41(0.4)
$Z' > 4$	57(22.5)	67(26.5)	23(9.1)	47(18.6)	40(15.8)	19(7.5)

($C2/c$). Additional information is available with the supporting information.

In both sets of structures, the ordering of the first pair of the most common space groups [*i.e.* No. 14 ($P2_1/c$, etc.) and No. 2 ($P\bar{1}$)] is reversed; for organic structures the ordering of the second pair [No. 19 ($P2_12_12_1$) and No. 4 ($P2_1$)] is also reversed.

The space group of a modulated structure is usually a subgroup of the known or assumed basic cell. Maximal subgroups of $P2_1/c$ are $P2_1$, $P\bar{1}$ and Pc (No. 7), all of which are more frequent for the modulated structures. $P2_1$ is also a maximal subgroup of No. 5 ($C2$) and No. 19, $P\bar{1}$ is also a maximal subgroup of $C2/c$ (as is $C2$), and Pc is also a maximal subgroup of No. 9, No. 29 ($Pca2_1$), and No. 33 ($Pna2_1$). $P2_1/c$ is a maximal subgroup of No. 15, No. 60 ($Pbcn$) and No. 61 ($Pbca$), which may explain its increased frequency in the modulated metallo-organic structures. $P1$ (No. 1) is a maximal subgroup of Nos. 2, 4, 5 and 9. Other subgroup relationships probably explain the decrease with modulation of the frequencies of the 'other' space groups (*i.e.* those not in the top ten).

6.4. Special types of patterns of deviations from translational symmetry

In most pseudotranslational columns the variation of the displacements over the period could be described by a Fourier series having a small number of terms. Discontinuous patterns (Fig. 6), however, are also important, as they are in incom-

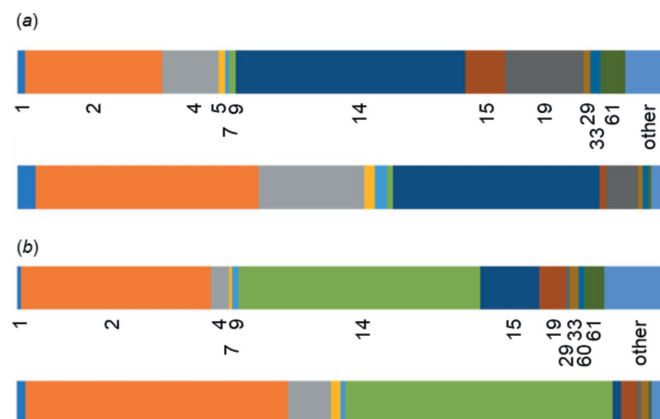


Figure 5
Space-group percentages for all structures considered and for the modulated structures identified for (a) organic structures, and (b) metallo-organic structures. The ten most common space groups in each set are identified by number; within each set the color code for the space groups is the same. For the metallo-organics group No. 60 ($Pbcn$) replaces No. 5 ($C2$) so that the colors for groups Nos. 7–33 differ between the two sets.

mensurate modulations (see Petříček *et al.*, 1995). Most common are:

Sawtooth patterns. These patterns are characterized by a sequence of approximately linear changes in the transverse or orientational deviations followed by a larger change in the opposite direction.

Crenel patterns. In a crenel modulation a block of molecules quite well related by pseudotranslation is followed by a block of molecules with a different transverse deviation or different conformation or orientation.

Ordered faults. An ordered fault is a special kind of crenel modulation in which one of the blocks contains only one molecule.

6.5. Pseudotranslations often have more than one important component

Fig. 7 shows histograms of the distributions of $fract(conf)$, $fract(orient)$, $fract(trans)$ and $fract(long)$ based on pseudotranslations of residues having at least six non-hydrogen atoms. The longitudinal component is almost always small, probably because modulated structures remain close packed along the pseudotranslation direction. The distributions are highly skewed, all of them peaking at or near zero.

In *ca.* 30% of the modulations one component dominates, *i.e.* it accounts for at least 80% of the av_rmsd . The dominant

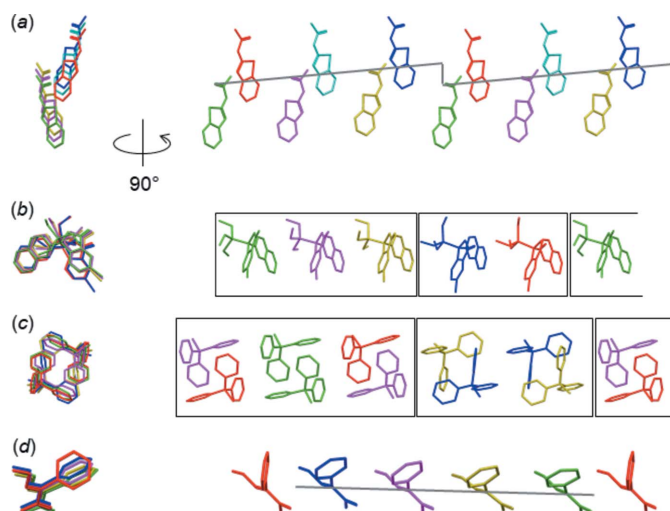


Figure 6

(a) The sawtooth modulation $[01\bar{1}]/3$ in OYETET (Balijapalli *et al.*, 2017; $P1$, $Z' = 6$); neither the included water molecules nor the hydrogen bonds are shown. (b) The crenel modulation $[201\bar{1}]/5$ in UDICIU (Shintani *et al.*, 2007; $P2_1$, $Z' = 5$). The molecules in the two groups have significantly different orientations and somewhat different conformations. (c) The crenel modulation $[12\bar{1}]/5$ in ZZZVTY04 (Kahr & Carter, 1992; $P1$, $Z' = 5$). The Ph_3CCl molecules in the two groups are conformational isomers. There is also a strong longitudinal component with the spacings in the group of three molecules being substantially shorter than the spacing in the group of two. (d) The $[01\bar{1}]/5$ modulation in XAPKOQ (Wu *et al.*, 2010; $P1$, $Z' = 5$) showing the ordered fault. Molecule 3 (in red) has a quite different conformation and is *ca.* 0.8 Å closer to its adjacent molecules than the other four molecules are to each other. Molecule 3 is also at the discontinuity of a subtle sawtooth modulation.

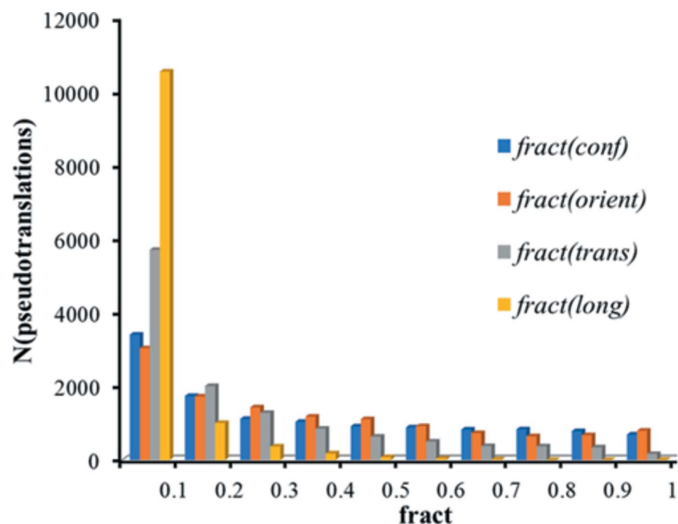


Figure 7

Distributions of conformational, orientational, transverse and longitudinal components.

term is more often conformational or orientational than transverse (see Fig. 7). In fewer than 1% of the structures does the longitudinal component account for more than half of the av_rmsd . In the structures having $fract(long) > 0.70$ there is a clear reason such as a Peierls-like distortion, a difference in the way the molecules form hydrogen bonds, or an important conformational or orientational variation.

A number of the pseudotranslations for which only the conformational component is important have a phenyl ring that has different orientations in the n molecules (see the supporting information). This modulation seems to allow the spacing of the Ph rings to better match the spacing needed for the rest of the molecule. Rotations of methoxy, isopropyl, and carboxyl groups are also common.

7. Problem structures and interesting outliers

Additional information about members of the classes of structures listed below is available with the supporting information.

7.1. Positioning of H atoms

No pseudotranslation is identified if it would relate residues such as an acid and its anion or an amine and its protonated cation. The H atoms, however, may not have been located correctly or may be disordered. Even if H atom positions were not reported there is still a possible problem because the matching algorithm used to decide if molecules are chemically identical takes all atoms into account, including H atoms, even if they were not located crystallographically.⁹ Consider co-crystals of phenazine and chloranilic acid: QOPVAU ($F1$, $Z' =$

⁹ If a structure contains, *e.g.* one acetic acid molecule and one acetate ion but the hydrogens were not located, the database building software guesses arbitrarily where each H atom should go and adds it, though without 3-D coordinates so it is not shown by any display program. See the discussion of KOHRAA in the supporting information.

2 at 100 K; Noohinejad *et al.*, 2014) and MAMPUM–MAMPUM07 ($P2_1/n$, $Z' = \frac{1}{2}$ at room temperature and above; $P2_1$, $Z' = 1$ at lower temperatures; see Noohinejad *et al.*, 2015). MAMPUM and QOPVAU are related by a single-crystal-to-single-crystal transition but have different refcodes because in the $P2_1$ structure there is no proton transfer while in the $F1$ structure proton transfer has been reported in one of the two independent phenazine/chloranilic acid pairs. If there were no proton transfer, Z' for the $F1$ structure would be 4 rather than 2, the refcode would be in the MAMPUM family, and a twofold pseudotranslation would be identified.

7.2. Molecules that are related by symmetry

If all members of each column in a pseudotranslation are also related by crystallographic symmetry then no pseudotranslation is identified. If members of only some of the columns are so related then a pseudotranslation is identified. An example is TARNEG ($P\bar{1}$, $Z' = 6$; Fig. S3; Gómez *et al.*, 2005), in which the $[001]/2$ pseudotranslation relates molecule numbers 1&1' and 5&5' but also 2&4 and 3&6.

The conventions can lead to a pseudotranslation being identified for only some of the residues when the whole structure is almost certainly modulated. In FUGMIF01 ($P\bar{1}$, $Z' = 2$ at 413 K; piperazin-1-ium perchlorate; Wojtaś *et al.*, 2015) the $[1\bar{1}1]/2$ pseudotranslation is found for the anions but not for the cations, which are related by inversion along $[1\bar{1}1]$ with the two independent cations in different stacks (note that the paper did not report any distinction between protonated and neutral N atoms.)

7.3. Pseudoglides and pseudoscrews may be identified as pseudotranslations

If a modulation lowers the symmetry of the unit cell a glide operation of the basic cell may be identified as a pseudotranslation if the molecular centroid is close to the glide plane and the molecule does not deviate much from mirror symmetry. The displacements in such pseudotranslations are often primarily orientational. Examples are NAKNEV and NAKNOF (1,4-diazoniabicyclo[2.2.2]octane perchlorate monohydrate and the corresponding BF_4^- salt), which have been studied carefully (Szafranski, 2017) both above and below a $Pca2_1 \rightarrow P1$ transition during which Z ($= 4$) is constant and the crystal remains single. The a glide of the orthorhombic cell becomes a $[101]/2$ pseudotranslation in the conventional triclinic cell, with the displacement of the DABCO^{2+} ions being more than 90% orientational. The remnants of the shorter c glide and the screw are not identified as pseudotranslations because the av_rmsds are too large.

7.4. Layered structures

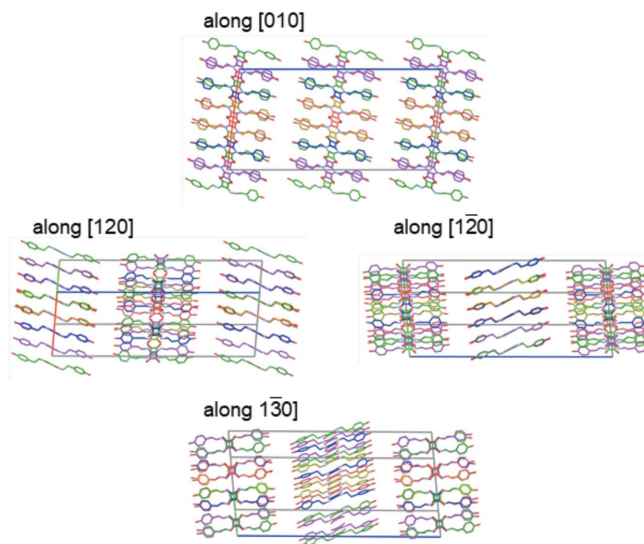
In some structures there are pseudotranslations $[uvw]/n$ that have different directions ($[u'v'w'] \neq$ either $[uvw]$ or $[\bar{u}\bar{v}\bar{w}]$) in adjacent layers. Even if the layers are related by space-group symmetry the pseudotranslation for one may not work well, or even at all, for the other (see Fig. 8). The layers are, however, sometimes independent, with different packing

arrangements (Fig. 9). In a few cases the pseudotranslations are confined to one of two different kinds of layers. Most of the layered structures are marked in the output as 'incomplete' to indicate that they are not considered modulated even though all the molecules of all the residues may lie in pseudotranslational columns.

7.5. Extra pseudotranslations

The program occasionally identifies more independent pseudotranslations than there are possible modulations. The number of formula units related by modulations cannot be larger than Z' unless the molecules have imposed symmetry. For eight isostructural HEYJOK0n structures ($P2_12_12_1$, $Z' = 5$; quinoxaline; Sathishkumar *et al.*, 2010) the program lists $[001]/5$ [$av_rmsd = 0.6 \text{ \AA}$] and also the very slightly longer $[101]/5$ [$av_rmsd = 1.2 \text{ \AA}$]. The two sets of $\text{C}_8\text{H}_6\text{N}_2$ molecules are the same except for translations along axis \mathbf{a} , which is much shorter than \mathbf{c} ($a/c = 0.11$). Both pseudotranslations, which are not related by a lattice vector, just happen to fit within the

(a) Views of the structure



(b) Layer at $x = 0$

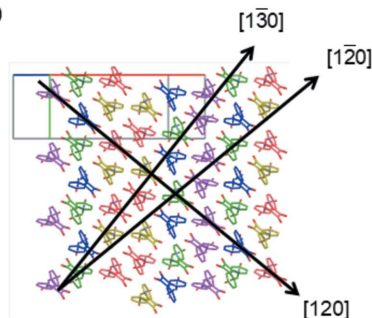


Figure 8

The layered structure of WOPVOO ($P2_1/c$, $Z' = 5$; Prohens *et al.*, 2014); the solvent molecules have been deleted for clarity. (a) Views of the structure showing that the pseudotranslation direction is specific to the layer, $[120]/5$ for the layer at $z = 0$ and $[\bar{1}20]/5$ for the layer at $z = \frac{1}{2}$. (b) Projection of the layer at $z = 0$ showing that for it $[120]/5$ and the longer $[\bar{1}30]/5 = [120]/5 - [010]$ are pseudotranslations but that $[\bar{1}20]/5$ is not. In the program output both $[120]/5$ and $[\bar{1}20]/5$ are marked as 'incomplete'.

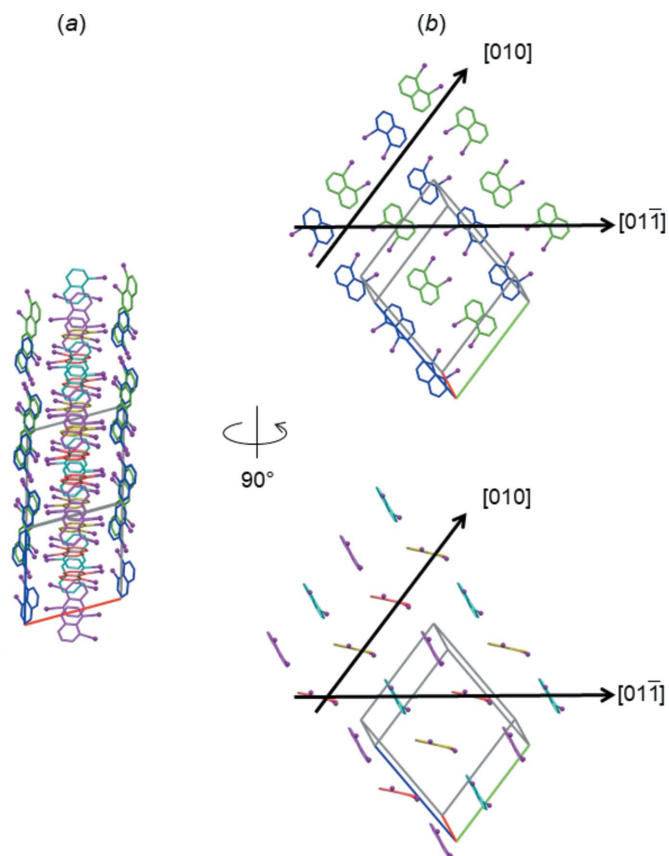


Figure 9

The layered structure of PIPQUA ($P\bar{1}$, $Z' = 6$; 1,8-diiodonaphthalene; Bock *et al.*, 1998). (a) View of the structure along $[01\bar{1}]$. (b) Views of the layers at $x = 0$ and the half layer at $x = 0.49$. (The inversion-related half layers at $x = 0.49$ and 0.51 interleave.) The $[01\bar{1}]/2$ pseudotranslation is excellent for the layer at $x = 0$ but for the layer at $x = 0.5$ the $[010]/2$ pseudotranslation is much better. The program lists $[01\bar{1}]/2$ only; its *qual_range* = 0.67 Å. No $[010]/2$ pseudotranslation is listed because the very different molecular orientations in the layer at $x = 0$ result in a value of *av_rmsd* larger than the default tolerance.

tolerances set. This complication is rare. When it does occur the molecules related by the two pseudotranslations are the same and one pseudotranslation has a lower *av_rmsd* than the other.

Very occasionally two listed pseudotranslations have the same numerical descriptors (e.g. $[10\bar{1}]/3$ and $[01\bar{1}]/3$ in trigonal SEUREA03/04; Luo & Dauter, 2017). In such cases only one of the pseudotranslations is independent.

7.6. Independent pseudotranslations that are correct

A special molecular orientation occasionally leads to identification of a pseudotranslation that is not related to the base pseudotranslation and so is marked as *none found*.¹⁰ An example is the $[001]/2$ pseudotranslation in HMHOCN05 ($P2_1/c$, $Z' = 4$; Reany *et al.*, 2009). The base pseudotranslation is $[211]/4$ ($n_ind = 3$), which generates $[011]/2$ ($= 2*[211]/4 + [\bar{1}00]$). An extra pseudotranslation, $[001]/2$, is listed. The orientation of the mean molecular plane (normal parallel to

¹⁰ Meaning no relationship to the base found.

\mathbf{a}^*) and the very similar x coordinates of the centroids of four independent molecules are responsible for the identification of the extra pseudotranslation.

7.7. High-symmetry space groups

Space groups having three-, four- and sixfold axes present special problems because of the number of directions that are equivalent by symmetry. If the axis \mathbf{c} is unique and $[uvw]/n$ has a non-zero value of u or v , then pseudotranslations in these groups are sometimes marked as *none found* or *incomplete* when they are not. Problems with identifying linear combinations of identified pseudotranslations may also occur. Pseudotranslations that are not parallel to the highest-order rotation axis in these group should be considered carefully.

8. Discussion

8.1. The program identifies pseudotranslations that might otherwise be missed

When looking for pseudotranslations with a program like *Mercury* it is often necessary to specify coordinate ranges $0 - |u_i| + \frac{1}{2}$ (e.g. 0–1.5, 0–2.5, 0–3.5 for $[12\bar{3}]/n$). If the coordinate range is smaller than at least one of the related molecules is often missing. Too-small ranges often render the pseudotranslation invisible.

Finding the $[234]/12$ modulation in LUXYOU (Brandel *et al.*, 2015) by eye would require looking at a large number of unit cells because the modulation is long and the direction is not at all simple.¹¹ The $[232]/6$ pseudotranslation had been found by inspection (Brock, 2016) but the $[234]/12$ modulation (Fig. S6) had not. That study also missed the $[\bar{1}2\bar{1}]/4$ pseudotranslation in LABQOX (Blaschette *et al.*, 2016; Fig. S7). When there are several pseudotranslational columns in a pseudotranslation, as they are in LABQOX, the columns may interpenetrate in a way that can make the pseudotranslation hard to identify. It seems unlikely that the $[\bar{1}\bar{4}]/12$ pseudotranslation in VANFUO ($R\bar{3}$, $Z' = 288/18 = 16$; Frampton *et al.*, 2017; Fig. 10)¹² would have been spotted even though the relationship of the $Z' = 16$ form to two lower- Z' forms was described in the publication.

8.2. Criteria for identifying pseudotranslations are not easy to set; judgment is required

We found this project much more challenging than anticipated. It seems unlikely that a completely reliable black-box algorithm will ever be possible. Experience-based judgment often seems essential, especially for picking out the important details of the pattern. Furthermore, the concept of modulation is inherently fuzzy; if the pseudotranslations in a model of a modulated structure were made progressively worse, there would not be a precise point at which the structure ceased to

¹¹ The transformation matrix relating the cells at RT ($C2/c$, $Z' = 1$) and 100 K ($P2_1/c$, $Z' = 12$) is given in the original paper; that matrix can be used to show that the pseudotranslation in the $Z' = 12$ form arises from the C -centering of the disordered RT cell.

¹² The refcode for VANFUO has very recently been changed to VANFOI03.

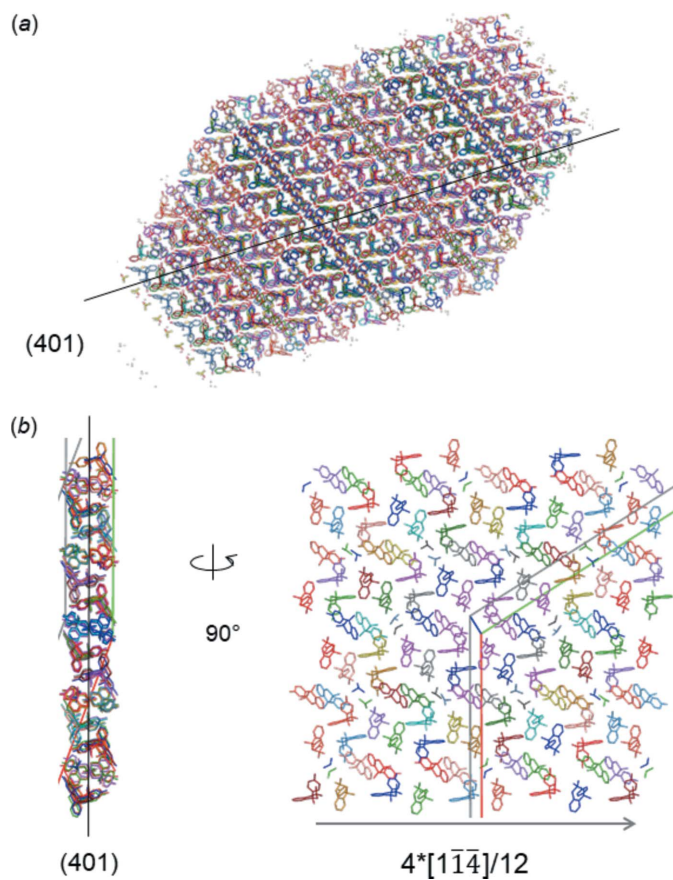


Figure 10
 The $[1\bar{1}4]/12$ pseudotranslation ($n_{ind} = 4$) in VANFUO ($R\bar{3}$, $Z = 288$, $Z' = 16$; Frampton *et al.*, 2017). (a) View along $[1\bar{1}4]$. (b) Two views of a layer (401) showing $4/12$ ths = $1/3$ rd of the pseudotranslation along $[1\bar{1}4]$. Any specific horizontal line of the second drawing in (b) shows only $1/3$ rd of the pseudotranslation but the other $2/3$ rds, which are related by the rhombohedral centering, can be seen along other horizontal lines.

be modulated. There are mathematical definitions of true symmetry but not of pseudosymmetry.

8.3. A few pseudotranslations are not modulations

A pseudotranslation that affects only the more numerous component of a 2:1 salt or co-crystal is not a modulation. An example is XEGKIG (Ambika, 2017), L-proline bis(2,4-dichlorobenzoic acid) hemihydrate. A $[100]/2$ pseudotranslation relates pairs of the four independent acid molecules but not the two proline zwitterions, which are related by pseudo-inversion. Another example is XECPON (Gunawardana *et al.*, 2017), a 2:1 co-crystal in which the second component is 1,2,3,4-tetrafluoro-5,6-diiodobenzene. While the $[110]/2$ pseudotranslation is too perfect (the cell volume should be halved), there is a $[101]/2$ pseudotranslation found for the first component but not for the second; the substituted benzenes are related instead by a pseudoglide along $[011]$ with mirror perpendicular to **a**. In order to obey the $[101]/2$ pseudotranslations half of the substituted benzenes would have to be rotated by *ca.* 120° around their plane normal. While such motions are known to occur in crystals the alternating orien-

tations seem to allow $N \cdots I$ interactions, which are known to be favorable.

An n -fold pseudotranslation in a structure having a single solvent molecule in the asymmetric unit probably cannot be a modulation because the position of the one solvent molecule would have almost certainly been established at the time of crystal nucleation.

8.4. Identifying the basic cell

For all the modulated structures it is possible to identify the basic cell and describe its symmetry. That task is sometimes easy (*i.e.* pseudotranslation along a cell axis that is the same in both structures) but can be difficult. Consider LUXYOU (Brandel *et al.*, 2015) in which a sixfold modulation along the centering vector of the $C2/c$, $Z' = 1$ cell found at room temperature (LUXYOU01/02) leads to the $P2_1/c$, $Z' = 12$ cell (LUXYOU) found at 100 K.

The symmetry of the basic cell is quite often higher than that of the modulated cell. If it is, the lost symmetry becomes approximate. If the modulated cell is triclinic but the basic cell would be monoclinic or higher the choice of a conventional triclinic unit cell sometimes obscures the approximate symmetry, especially if the basic cell is centered (Fig. S8).

For 27 of the modulated structures having $Z' > 4$ or space group $P1$ a very closely related structure with a smaller asymmetric unit was found in the CSD. In almost all cases (including LUXYOU) the smaller and larger unit cells were reported to be related by a reversible phase transition during which the crystal remained single.

In 11 of the 27 pairs of phases there is no change in the space-group symmetry but the modulated unit cell is larger; in seven pairs there is no important change in the unit-cell volume but the modulation lowers the symmetry. In nine pairs the transition changes both the volume and the symmetry. In almost all cases, the higher- Z' structure was determined at a lower temperature than the lower- Z' structure but there are three well documented exceptions to that generalization [JAWQIH01/JAWQIH03 (Hao *et al.*, 2005; Siegler *et al.*, 2008), VIVSAV04/VIVSAV01 (Ślepokura, 2016), and XOJYEC02/XOJYEC01 (McKellar *et al.*, 2014)]. More information about relationships between specific supercells and their basic cells is given in the supporting information.

8.5. The interactions responsible for modulations are not easy to determine

Because modulations extend over large distances and many molecules they are probably established during the early stages of crystal growth rather than at the time of nucleation. Modulations may also appear during the crystal contraction that accompanies cooling or an increase in the pressure. Understanding how modulations are established, however, is likely to be difficult because the molecules in a pseudotranslational column are often not in close contact. When they are not, they are in contact through molecules in adjacent columns, with the adjacent columns usually related by crys-

tallographic (Fig. S9) or approximate symmetry (Fig. S10), especially screw axes or glides.

8.6. There is no simple way to identify the most important modulation(s)

It is often impossible to identify the most important pseudotranslations without investigating the intermolecular interactions in detail, probably by calculating energies. Consider UZILIA (*Cc*, $Z' = 8$; Lobkovsky & Porco, 2016; Fig. S11) in which the shortest pseudotranslations are $[010]/2$ and $[102]/4$ ($length/n = 9.60$ and 13.79 Å; $[010]/2 = [001] - 2*[102]/4 + [110]/2$). Molecules related by $[010]/2$ are in close contact but those related by $[102]/4$ are not; in the two independent $[102]/4$ columns there is no contact significantly shorter than the sum of the van der Waals radii. Molecules related by $[102]/4$ ($= [102]/4 - [001]$) are, however, in good contact. The $length/n$ value for $[102]/4$ (20.61 Å) is long because the units related are RCOOH dimers lined up along their longest dimension; the effective $length/n$ value is more like $20.6/2 = 10.3$ Å.

In XIFMOQ (*Cc*, $Z' = 6$; Villemin *et al.*, 2013) the shortest pseudotranslation is $[10\bar{1}]/3$ (11.61 Å) but the related $[102]/3 = [10\bar{1}]/3 + [001]$ (21.82 Å) is much more informative. There are obvious approximate twofold axes along $[102]$, which makes an angle of 89.8° with $[100]$ and (because the space group is monoclinic) 90° with $[010]$. There also seems to be centering in a layer (010) if the axes are $[100]$ and $[102]$. With increased tolerances *PLATON* suggested the space group *Fdd2* with twice the cell volume but only three independent (although somewhat disordered) molecules. In *Fdd2* the obvious pseudotranslation would be $[001]/3$ (*i.e.* $[102]/3$ in the *Cc* cell), although $[30\bar{1}]/6$ (*i.e.* $[10\bar{1}]/3$ in the *Cc* cell) is shorter. Other examples are described in the supporting information.

9. Summary

Approximate symmetry is pervasive in structures with $Z' \geq 4$ (Brock, 2016; Baggio, 2019, 2020; this work) but until now there has been no program that identifies approximate symmetry both automatically and reliably. (*PLATON* spots overlooked symmetry well but that is a less complicated task. *PLATON* notes some other approximate relationships but does not do so reliably.) The program described here finds approximate (*i.e.* pseudo) translations and quantifies their four components (transverse, longitudinal, conformational, and orientational). The program could be expanded to find other approximate symmetry.

Acknowledgements

The project would have been impossible without the CCDC's program *Mercury*. Especially useful was the quite new option of viewing a structure along the direction of a general lattice vector $[uvw]$. We are also grateful to the authors who recognized and described modulations and phase relationships, and who devised informative atom-numbering schemes. RT thanks the CCDC for their provision of an Emeritus Research Fellowship.

References

- Ambika, R. (2017). CCDC 1564129: *CSD Communication*, <https://doi.org/10.5517/ccdc.csd.cc1phlsk>.
- Baggio, R. (2019). *Acta Cryst.* **C75**, 837–850.
- Baggio, R. (2020). *Acta Cryst.* **C76**, 258–268.
- Balijapalli, U., Udayadasan, S., Panyam Muralidharan, V., Sukumarapillai, D. K., Shanmugam, E., Gopal, A. P., Rathore, R. S. & Kulathu Iyer, S. (2017). *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **173**, 572–577.
- Beckett, M. A., Horton, P. N., Hursthouse, M. B. & Pszolla, C. (2010). *Acta Cryst.* **E66**, o833.
- Blaschette, A., Jones, P. G., Lozano, V. & Freytag, M. (2016). CCDC 993324: *CSD Communication*, doi: 10.5517/cc12bmqr.
- Bock, H., Sievert, M. & Havlas, Z. (1998). *Chem. Eur. J.* **4**, 677–685.
- Brandel, C., Cartigny, Y., Couvrat, N., Eusébio, M. E. S., Canotilho, J., Petit, S. & Coquerel, G. (2015). *Chem. Mater.* **27**, 6360–6373.
- Brock, C. P. (2016). *Acta Cryst.* **B72**, 807–821.
- Dunand, A. & Gerdil, R. (1982). *Acta Cryst.* **B38**, 570–575.
- Frampton, C. S., Ketuly, K. K., Ali, H. B. M., Azizan, A. H. S., Gall, J. H. & MacNicol, D. D. (2017). *CrystEngComm*, **19**, 2653–2659.
- Fresno, N., Pérez-Fernández, R., Goya, P., Jimeno, M. L., Alkorta, I., Elguero, J., Menéndez-Taboada, L. & García-Granda, S. (2011). *Tetrahedron*, **67**, 9104–9111.
- Gómez, C., Maciá, B. & Yus, M. (2005). *Tetrahedron*, **61**, 9325–9330.
- Görbitz, C. H., Agustsdóttir, S. & Bleken, F. (2007). *Acta Cryst.* **C63**, o58–o60.
- Groom, C. R., Bruno, I. J., Lightfoot, M. P. & Ward, S. C. (2016). *Acta Cryst.* **B72**, 171–179.
- Gunawardana, C. A., Desper, J., Sinha, A. S., Đaković, M. & Aakeröy, C. B. (2017). *Faraday Discuss.* **203**, 371–388.
- Hao, X., Parkin, S. & Brock, C. P. (2005). *Acta Cryst.* **B61**, 675–688.
- Kahr, B. & Carter, R. L. (1992). *Mol. Cryst. Liq. Cryst. Sci. Technol. Sect. A*, **219**, 79–100.
- Lobkovsky, E. & Porco, J. (2016). CCDC 757114: *CSD Communication*, doi: 10.5517/ccdc.csd.cctdv12.
- Long, S. & Li, T. (2009). *Cryst. Growth Des.* **9**, 4993–4997.
- López-Mejías, V. & Matzger, A. J. (2015). *Cryst. Growth Des.* **15**, 3955–3962.
- Luo, Z. & Dauter, Z. (2017). *Plos One*, **12**, e0171740.
- Macrae, C. F., Sovago, I., Cottrell, S. J., Galek, P. T. A., McCabe, P., Pidcock, E., Platings, M., Shields, G. P., Stevens, J. S., Towler, M. & Wood, P. A. (2020). *J. Appl. Cryst.* **53**, 226–235.
- Marsh, R. E. (1999). *Acta Cryst.* **B55**, 931–936.
- McKellar, S. C., Kennedy, A. R., McCloy, N. C., McBride, E. & Florence, A. J. (2014). *Cryst. Growth Des.* **14**, 2422–2430.
- Mukhopadhyay, C., Rana, S. & Butcher, R. J. (2010). *Arkivoc*, 291–304.
- Noohinejad, L., Mondal, S., Ali, S. I., Dey, S., van Smaalen, S. & Schönleber, A. (2015). *Acta Cryst.* **B71**, 228–234.
- Noohinejad, L., Mondal, S., Wölfel, A., Ali, S. I., Schönleber, A. & van Smaalen, S. (2014). *J. Chem. Crystallogr.* **44**, 387–393.
- Petríček, V., van der Lee, A. & Evain, M. (1995). *Acta Cryst.* **A51**, 529–535.
- Peuronen, A., Valkonen, A., Kortelainen, M., Rissanen, K. & Lahtinen, M. (2012). *Cryst. Growth Des.* **12**, 4157–4169.
- Prohens, R., Font-Bardia, M. & Portell, A. (2014). CCDC 1017985: *CSD Communication*, doi: 10.5517/cc13597t.
- Reany, O., Kapon, M., Botoshansky, M. & Keinan, E. (2009). *Cryst. Growth Des.* **9**, 3661–3670.
- Sathishkumar, R., Mahapatra, S., Thakur, T. S. & Desiraju, G. R. (2010). *Curr. Sci.* **99**, 1807–1811.
- Schönleber, A. (2011). *Z. Kristallogr.* **226**, 499–517.

- Scott, J. S., Birch, A. M., Brocklehurst, K. J., Broo, A., Brown, H. S., Butlin, R. J., Clarke, D. S., Davidsson, Ö., Ertan, A., Goldberg, K., Groombridge, S. D., Hudson, J. A., Laber, D., Leach, A. G., Macfaul, P. A., McKerrecher, D., Pickup, A., Schofield, P., Svensson, P. H., Sörme, P. & Teague, J. (2012). *J. Med. Chem.* **55**, 5361–5379.
- Shang, Y., Jin, B., Peng, R., Guo, Z., Liu, Q., Zhao, J. & Zhang, Q. (2016). *RSC Adv.* **6**, 48590–48598.
- Shintani, R., Takatsu, K. & Hayashi, T. (2007). *Angew. Chem. Int. Ed.* **46**, 3735–3737.
- Siegler, M. A., Hao, X., Parkin, S. & Brock, C. P. (2008). *Acta Cryst.* **B64**, 738–749.
- Ślepokura, K. A. (2016). *Acta Cryst.* **B72**, 117–125.
- Spek, A. L. (2009). *Acta Cryst.* **D65**, 148–155.
- Szafranski, M. (2017). *CrystEngComm*, **19**, 179–182.
- Villemin, E., Robeyns, K., Robiette, R., Herent, M.-F. & Marchand-Brynaert, J. (2013). *Tetrahedron*, **69**, 1138–1147.
- Wojtaś, M., Gağor, A., Czupiński, O., Piecha-Bisiorek, A., Isakov, D., Medycki, W. & Jakubas, R. (2015). *CrystEngComm*, **17**, 3171–3180.
- Wrona-Piotrowicz, A., Plažuk, D., Zakrzewski, J., Métivier, R., Nakatani, K. & Makal, A. (2015). *Dyes Pigments*, **121**, 290–298.
- Wu, Y.-J., Huang, S., Ng, A., Gao, Q., Roy Kimura, S. & Langley, D. R. (2010). *Tetrahedron Lett.* **51**, 2144–2147.