research papers

OD (order–disorder) interpretation and diffuse scattering analysis of an organic polytype with allotwin character: a detailed how-to

David Fröschl,[a] Nicolas Kratena,[a] Berthold Stöger,[b]* Alexandr Virovets[c] and Tobias Wolflehner[d]

[a]Institute of Applied Synthetic Chemistry, TU Wien, Getreidemarkt 9, 1060 Vienna, Austria, [b]X-Ray Centre, TU Wien, Getreidemarkt 9, 1060 Vienna, Austria, [c]Institute of Inorganic and Analytical Chemistry, Goethe Universitaet Frankfurt, Max-von-Laue-Str. 7, 60438 Frankfurt, Germany, and [d]Institute of Chemical Technologies and Analytics, TU Wien, Getreidemarkt 9, 1060 Vienna, Austria. *Correspondence e-mail: bstoeger@mail.tuwien.ac.at

2,3-Dihydroxy-1,3,4-trimethyl-6-oxo-1,4-cyclohexadiene-1-carboxylic acid crystallizes in an order–disorder (OD) structure with a high stacking fault probability. The diffraction pattern features diffuse scattering and broad peaks, which can be attributed to fragments of two polytypes of a maximum degree of order (MDO). Additional weak peaks indicate existence of a non-MDO polytype. A detailed step-by-step instruction is given on how to develop an OD description and interpret the diffuse scattering using a stochastic growth model.

1. Introduction

Polytypes are structures that crystallize with identical (up to minor desymmetrization) layers arranged in different ways (Aksenov *et al.*, 2023). Polytypism is a multifarious phenomenon leading to interesting crystallographic challenges such as twinning (Nespolo & Ferraris, 2004), antiphase domains (Wondratschek & Jeitschko, 1976), allotwins (Nespolo *et al.*, 1999) and diffuse scattering (Welberry, 2010). Often, these appear in the same crystal and may not be separable.

We encountered an example of a such a complex crystallization behavior during a synthetic campaign towards 3,5-dimethylorselinic acid (DMOA)-derived meroterpenoid natural products. During the attempted alkaline hydrolysis of DMOA methyl ester (Powers *et al.*, 2019), a concomitant spontaneous oxidation towards the dearomatized dienone 2,3-dihydroxy-1,3,4-trimethyl-6-oxo-1,4-cyclohexadiene-1-carboxylic acid **1** was observed (Scheme 1).
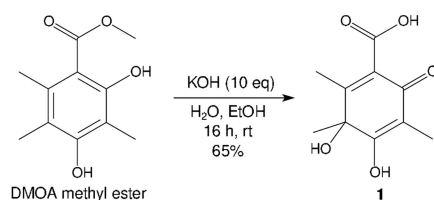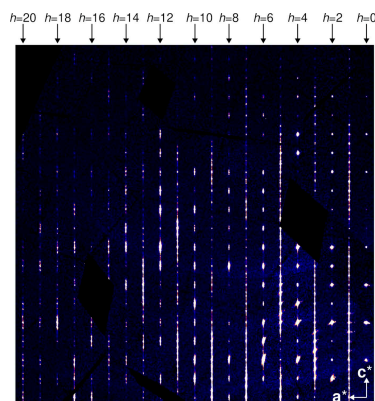
Scheme (1)

Presumably, the oxidative event takes place by nucleophilic attack of a phenolate with molecular oxygen. On routine structural characterization of **1** by single-crystal diffraction, we observed distinct diffuse scattering and broad peaks at positions that indicated fragments of distinct polytypes, *i.e.* an intermediate between a disordered structure and an allotwin. The phenomenon was systematic and observed for different crystals isolated from the same crystallization experiment.

The order–disorder (OD) theory (Dornberger-Schiff & Grell-Niemann, 1961) was created in the 1950s and further developed in the second half of the 20th century to explain the common occurrence of polytypism in all known classes of crystalline matter. It is based on the limited range of interatomic interactions. OD polytypes are an often observed class of polytypes, where layers can connect in different ways, nevertheless forming geometrically equivalent layer pairs. This means that, in theory, an infinite number of ordered polytypes or disordered stacking arrangements may form, which are all locally equivalent and form an OD family.

Herein, we will give a step-by-step introduction on how to develop an OD model based on the structure of **1**. For high-quality diffuse scattering reconstructions, we collected data of a selected crystal at the P24 beamline of the PETRA III synchrotron at DESY. A semi-quantitative interpretation of the diffraction pattern is given using growth models (Welberry, 2010). §3.6 and §3.11 contain detailed derivations, which may be skipped by the reader interested mainly in practical application of OD theory and growth models.

2. Experimental

2.1. Synthesis

The precursor DMOA methyl ester was obtained as per the published procedure by Porco *et al.* (2019).

Procedure: 50 mg of DMOA methyl ester (0.237 mmol, 1 equivalent) were dissolved in ethanol (1 ml) and stirred at room temperature. At the same time, 134 mg of KOH (2.38 mmol, 10 equivalents) were dissolved in deionized water (0.2 ml) and then added to the solution of the starting material. As soon as the base was added, the color of the solution changed from yellow to light red, during the next hour of stirring the solution turned dark red. The solution was stirred at room temperature for 20 h. Work up: the reddish-black solution was made acidic through addition of 3 M HCl (2 ml). The reaction mixture was extracted with ethyl acetate (4×), dried over magnesium sulfate and concentrated *in vacuo*. Purification: the product was purified through column chromatography (DCM/MeOH = 20/1 and 0.5% acetic acid). Yield: 33 mg (0.155 mmol, 65% of theory).

Single crystals were grown by dissolving 10 mg in a 1:1:1 mixture of acetone, ethanol and water and evaporating slowly through a septum with needle over 14 days.

2.2. Single-crystal diffraction

Intensity data of a single crystal of **1** was collected at the high-flux P24 beamline of the PETRA III synchrotron using $\lambda = 0.56$ Å ($E = 22.14$ keV) radiation. Two 360° φ-scans with 0.2° frame width were collected. Data were converted into ESPERANTO format and processed with the *CrysAlis* software (Rigaku Oxford Diffraction, 2022). As discussed below, diffraction data consisted of reflections that could be attributed to multiple polytypes. The major polytype was orthorhombic and reflection intensities of other domains were omitted during structure refinement. The structure was solved

Table 1
Experimental details.

|	
Crystal data	
Chemical formula	$C_{10}H_{12}O_5$
M_r	212.20
Crystal system, space group	Orthorhombic, $Pna2_1$
Temperature (K)	170
a, b, c (Å)	18.2995 (2), 6.5887 (1), 16.7828 (2)
V (Å^3)	2023.50 (4)
Z, Z'	8, 2
Radiation type, λ (Å)	Synchrotron, 0.56000
ρ_{calc} (g cm^{-3})	1.393
μ (mm^{-1})	0.07
Crystal shape, color	Block, colorless
Crystal size (mm)	$0.21 \times 0.14 \times 0.06$
Data collection	
Diffractometer	Huber Eulerian Cradle
Absorption correction	Multi-scan (*CrysAlis PRO*). Empirical absorption correction using spherical harmonics, implemented in SCALE3 ABSPACK scaling algorithm.
T_{min}, T_{max}	0.779, 1.000
No. of measured, independent and observed [$I > 2\sigma(I)$] reflections	66327, 9320, 7252
R_{int}	0.075
$(\sin \theta/\lambda)_{max}$ (Å^{-1})	0.901
Refinement	
$R[F^2 > 2\sigma(F^2)]$, $wR(F^2)$, S	0.082, 0.255, 1.08
No. of reflections	9320
No. of parameters	386
No. of restraints	11
H-atom treatment	H atoms treated by a mixture of independent and constrained refinement.
$\Delta\rho_{max}, \Delta\rho_{min}$ (e Å^{-3})	0.90, −0.32
Absolute structure	Modelled as perfect inversion twin.
Absolute structure parameter	0.5 (fixed)
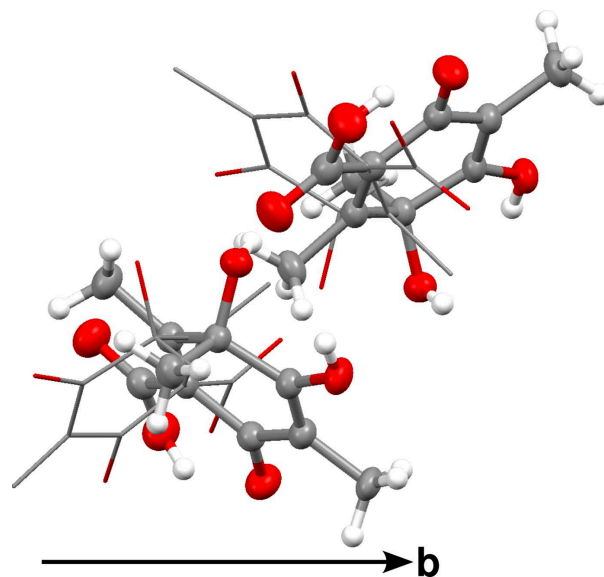
Figure 1
Asymmetric unit of the refined orthorhombic polytype viewed along [100]. Atoms are represented by red (O) and gray (C) ellipsoids drawn at the 50% probability levels; H atoms by white spheres of arbitrary radius. The phantom molecules are shown as wireframe.

with the dual-space approach implemented in *SHELXT* (Sheldrick, 2015*b*) and refined against $|F|^2$ using *SHELXL* (Sheldrick, 2015*a*). The errors owing to ignoring reflections of alternative polytypes produce alternative orientations of the molecules in the difference Fourier maps, sometimes designated as 'phantom molecules'. For each atom one major alternative position could be located in the difference Fourier map and the structure was refined as an overlay of two orientations of each molecule (Fig. 1). The displacement parameters of both orientations were constrained to be equal. Additional electron density peaks are attributed to even more alternative orientations, which were however too weak for inclusion in the refinements. Since the occupancy of the second orientation was only *ca* 11%, some atomic positions were ill-determined and distance restraints were implemented to achieve a reasonable geometry. The hydroxyl and carboxyl H atoms of the major orientation were refined with distance restraints, those of the minor orientation were omitted. More details on data collection and structure refinement are compiled in Table 1. Molecular graphics were prepared using the *Mercury* (Macrae *et al.*, 2020) software.

3. Results and discussion

3.1. Crystal chemistry

The crystal chemistry of **1** will only be sketched briefly based on the major *Pna2*$_1$ polytype. Molecule **1** [Fig. 2(*a*)] possesses three hydrogen bond donors. The carboxylic acid group forms an intramolecular hydrogen bond to the oxo group. Pairs of molecules related by pseudo-inversion symmetry are connected by intermolecular hydroxyl-to-hydroxyl bonds [see Fig. 2(*a*)]. Note that the two molecules are not equivalent according to the *Pna2*$_1$ symmetry (hence $Z' = 2$), which will play a crucial role in the description of the disorder. The remaining hydroxyl group forms an intermolecular bond to a carboxyl group, resulting in infinite sheets extending parallel to the (001) plane [Fig. 2(*b*)].

3.2. Deriving the OD layers

Even though application of OD theory is extremely useful from a descriptive but also a predictive point of view, its is not widely used. We attribute this to a lack of awareness and a perceived impenetrability. Therefore, here we give step-by-step instructions of how to derive the OD model of **1**.

An OD model should always be based on experimental evidence (for example twinning or diffuse scattering). For **1** there were four characteristic signs of stacking disorder.

(i) Pronounced one-dimensional diffuse scattering in **c*** direction.

(ii) 'Phantom molecules' of alternative stacking arrangements.

(iii) Crystallographically independent molecules ($Z' = 2$) with systematic orientation relations.

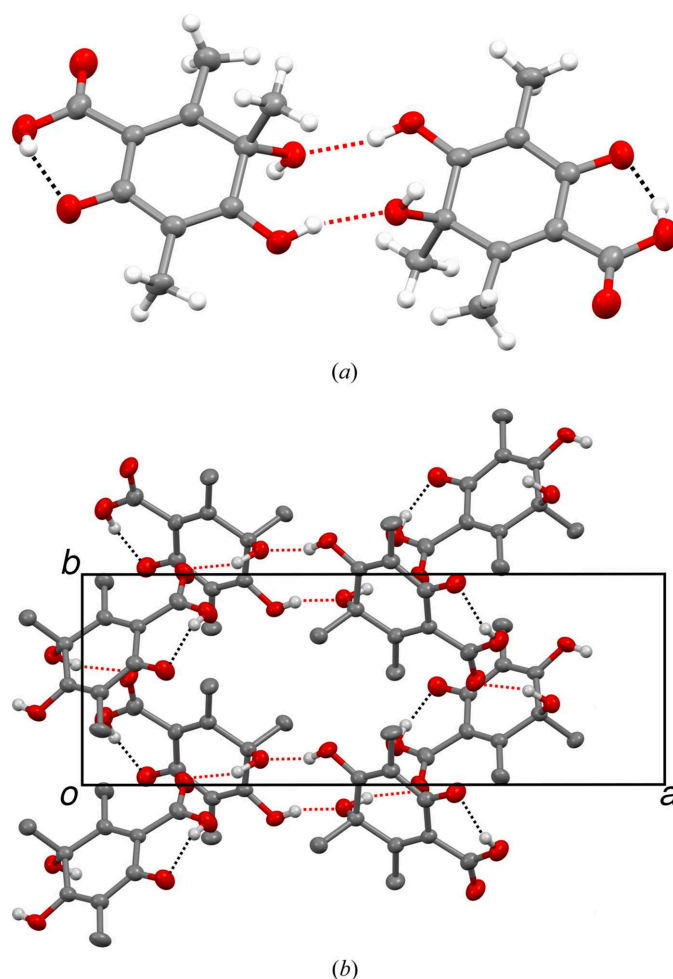(iv) Additional diffraction peaks of alternative polytypes.

(*a*)

(*b*)

Figure 2
(*a*) Pair of **1** molecules connected by hydrogen bonds and (*b*) diperiodic hydrogen-bonding network extending parallel to the (001) plane. Atoms as in Fig. 1. Inter- and intramolecular hydrogen bonds are represented by red and black dashed lines, respectively.

None of these prove the OD character of the structure [see Schöbinger & Stöger (2024) for a counter-example], but they are a good indication.

The first step in an OD interpretation is identification of the OD layers, their (idealized) symmetry, and the operations relating layers. These operations need not be symmetry operations of the space group. They are *partial (symmetry) operations* (POs), because they apply only to distinct layers. The layers in **1** should be parallel to (001) in accordance with the observed diffuse scattering. Apart from the translations in the (001) plane, only the $a_{[010]}$ glide reflection of the *Pna2*$_1$ space group (and the compositions with the translations) maps potential layers parallel to (001) onto themselves. This however means that the symmetry groups of the OD layers must be supergroups of $p1a1$, where we use the notation of the *International Tables for Crystallography*, Vol. E (Kopsky & Litvin, 2006), which assumes layers extending parallel to the (001) plane.

In the refined *Pna2*$_1$ polytype, there are two non-equivalent molecules [marked by distinct colors in Fig. 3(*a*)]. Note that the stacking direction [001] in Fig. 3 and in the two-dimen-

sional reciprocal space reconstructions below is pointing up, as is common in the OD literature and which is implied by the expression 'layer stacking'. Since the molecules are 'asymmetric', or more precisely only symmetric by the trivial point group 1 (C_1 after Schönfließ), any additional layer symmetry element must be located between molecules. In Fig. 3(b) a simplified view is given by replacing the molecules with arrows pointing from the 6-oxy group to the opposite 3-hydroxy group, whereby red and yellow colors give the orientation with respect to the viewing direction [010]. The symmetry elements of the $Pna2_1$ space group are indicated by black symbols.

The two molecules at the bottom left of Fig. 3(b) feature opposite orientation in [100] and [001] and the same orientation in [010] direction. Their orientations are therefore related by a $2_{[010]}$ rotation. Observe that when discussing
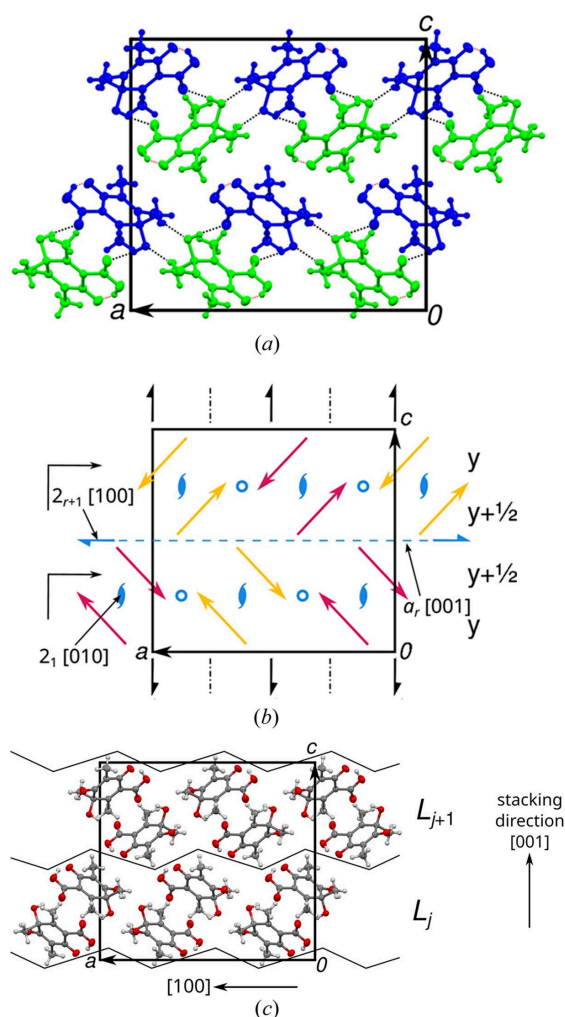

(a)

(b)

(c)

Figure 3
The $Pna2_1$ polytype of **1** viewed down [010] with (a) molecules colored according to symmetry equivalence, (b) molecules represented by red (pointing toward the reader) and yellow (pointing away) arrows and (c) layer interfaces indicated by a black polygonal chain. In (b) symmetry elements of the space group (black) and non-space group symmetry elements (blue) are given by the common symbols. Additional translation along [010] is given to the right. Inter- and intramolecular hydrogen bonds are represented in (a) by black and red dashed lines, respectively. The geometric elements discussed in the text are marked by their printed symbol in (b). Atom colors and ellipsoids in (c) as in Fig. 2.

orientation relations of molecules, we only consider the linear part of the operations relating the molecules (the 'point operation'). An additional translational component of $\frac{1}{2}\mathbf{b}$ means that the molecules are related by a 2_1 screw rotation as indicated by the blue symbol in Fig. 3(b), whereby the blue color indicates that the operation is not a space group operation of the $Pna2_1$ polytype.

Double application of this screw rotation gives a $\mathbf{b}$ translation, which means that the 2_1 screw rotation is a symmetry element of a layer formed by two half-layers. Combination of the 2_1 screw rotation and the $a_{[010]}$ glide reflection results in inversions, as indicated by blue circles in Fig. 3(b). Pairs of molecules related by inversion correspond to the hydrogen-bond connected pair in Fig. 2. Combining these symmetries, the overall layer group symmetry is $p12_1/a1$.

Additional POs are identified by considering pairs of red (or yellow) arrows at the center of Fig. 3(b). Here, the orientations are related by an $m_{[001]}$ reflection. The intrinsic translational component of the operation relating the molecules is $ca\ \frac{1}{4}\mathbf{a}$, resulting in a *glide reflection*. In the OD literature, a generalization of Hermann Mauguin symbols is used and if the translational component were precisely $\frac{1}{4}\mathbf{a}$, the glide reflection would be written as $a_{\frac{1}{2}}$. Owing to a small deviation, it is instead written as a_r, indicating an intrinsic translation of $\frac{r}{2}\mathbf{a}$. Observe that reflections are the composition of a twofold rotation and an inversion and therefore the subscript r is interpreted as in the case of a twofold screw rotation. r is a metric parameter of the structure (see below). Double application of the a_r operation results in a translation by $r\mathbf{a} \approx \frac{1}{2}\mathbf{a}$, which is *not* a (pseudo) translation of the crystal lattice (the lattice formed by all translations of the crystal). Thus, this glide reflection is *not* a symmetry operation of a layer, since it does not map a potential layer onto itself. However it plays a crucial role in the OD description, as it relates two distinct layers. By composition with the $a_{[010]}$ operation of the $Pna2_1$ space group, one obtains a screw rotation with intrinsic translation $ca\ \frac{3}{4}\mathbf{a}$ in the [100] direction, with the symbol 2_{r+1}. Both operations are indicated in Fig. 3(b) by the graphical symbols for regular a and 2_1 operations.

In summary, the structure of **1** can be considered as a structure of layers parallel to (001) with $p12_1/a1$ symmetry. The layer boundaries are indicated in Fig. 3(c). These layers corresponds to hydrogen bonded sheets described in the previous section, though this is purely coincidental. OD layer boundaries may even 'slice' through molecules as for example described by Stöger *et al.* (2013).

Note that the layer boundary is not a plane. The crucial point from an OD perspective is that the layers possess a substantial thickness at every point so that there are only weak interactions across a full layer width. The layers will be designated as A_j, j being a sequential integer.

3.3. OD groupoid family

In a proper OD structure, the layers may contact in different, yet geometrical equivalent ways, resulting in an infinity of stacking possibilities built according to the same

Table 2
Intrinsic translation components of the σ-POs in the OD groupoid symbol of equation (1).

Symbol	Direction	Type	Intrinsic translation
2_{r+1}	[100]	Screw rotation	$(\frac{r}{2} + \frac{1}{2})\mathbf{a}$
$n_{s+1,2}$	[100]	Glide reflection	$(\frac{s}{2} + \frac{1}{2})\mathbf{b} + \mathbf{c}_0$
2_2	[001]	Screw rotation	$\mathbf{c}_0$
$n_{r,s}$	[001]	Glide reflection	$\frac{r}{2}\mathbf{a} + \frac{s}{2}\mathbf{b}$

symmetry principle. The set of these stacking arrangements form an OD family. The full symmetry of each family member is given by the POs mapping individual layers. These do not form a group, but a groupoid, because an operation that maps layer A_j onto layer A_k can only be composed with an operation that maps layer A_k onto a layer A_l (Ito & Sadanaga, 1976).

To classify the symmetry of OD families, the groupoids are grouped into OD groupoid families which, in analogy to space group types, abstract from metric parameters (Fichtner, 1977*b*). For the special case of OD structures of one kind of layer with all the same translation lattice there are 400 kinds of OD groupoid families (Fichtner, 1977*a*). The OD groupoid family is derived from the layer symmetry, called λ-POs in the OD literature (λ for layer), and one operation that relates two adjacent layers, called σ-PO (σ for space). The full set of σ-POs is generated by composing the λ-POs with the σ-PO (in mathematical terms, this can be regarded as forming a coset of the layer group). To do so, we use a computer program that we plan to publish in the foreseeable future.

The layer symmetry and a σ-PO were derived in the prior section, leading to the OD groupoid family

$$\begin{array}{ccccc} p & 1 & 2_1/a & 1 \\ \{ & 2_{r+1}/n_{s+1,2} & - & 2_2/n_{r,s} & \} \end{array} \quad (1)$$

according to a notation inspired by Dornberger-Schiff & Grell-Niemann (1961). The first line indicates the layer symmetry, the second line the σ-POs using the usual positional Hermann–Mauguin symbols. The meaning of the individual symbols is given in Table 2, where $\mathbf{c}_0$ is the vector perpendicular to the layer planes with the length of one layer width. r and s are metric parameters. As noted above, the subscripts of the glide reflections symbols are based on those of twofold rotations. In particular, $n_{r,s}$ in [001] direction standing for a glide with intrinsic translation $\frac{r}{2}\mathbf{a} + \frac{s}{2}\mathbf{b}$ is an unfortunate quirk of the notation, because n is defined as a glide reflection with intrinsic translation half the diagonal of a base of the unit cell. The intrinsic translations of symmetry group operations are restricted because n-fold application of an operation with linear part of order n must result in a lattice translation. In particular, the intrinsic translations of glide reflections and twofold screw rotations (higher order screw rotations cannot exist in layers) must be a full or half of a lattice vector. In contrast, the intrinsic translation of σ-POs is not restricted and, therefore, r and s may adopt any value, integral, rational or irrational. As will be shown below, these parameters and their deviations from rational values with low denominators

have profound implications on the observed diffraction patterns.

3.4. Possible layer arrangements and metric parameters

As noted before, the OD groupoid family abstracts from the metric parameters, namely the layer's cell parameters ($a = |\mathbf{a}|$, $b = |\mathbf{b}|$), the layer spacing ($c_0 = |\mathbf{c}_0|$) and the relative positions of adjacent layers with respect to $\mathbf{a}$ and $\mathbf{b}$ (r, s). The latter determine the number of distinct ways in which adjacent layers can be placed and they therefore need to be determined.

There are two ways of generating an ambiguity in the stacking arrangement. Firstly, application of a symmetry operation of layer A_j that does not invert the orientation of $\mathbf{c}_0$ and is not a symmetry operation of A_{j+1} to the layer pair (A_j, A_{j+1}) generates a a new pair (A_j, A'_{j+1}), which has the same interatomic distances as the original pair. In **1**, $p1a1$ is the subgroup of operations of A_j that does not invert $\mathbf{c}_0$ (called λ-τ-POs in the OD literature). The 2_1 operation is ignored, as it inverts the orientation of $\mathbf{c}_0$ and therefore maps A_{j+1} onto the space occupied by A_{j-1}. It consequently does not generate a new (A_j, A'_{j+1}) pair. For general s, the a-glide planes of A_j and A_{j+1} do not overlap and only the operations of the subgroup $p1 < p1a1$ also apply to A_{j+1}. Given A_j, there are thus $[p1a1:p1] = 2$ possible ways of placing A_{j+1} ($[\mathcal{G} : \mathcal{H}]$ is the index of the subgroup $\mathcal{H}$ in $\mathcal{G}$).

Secondly, if there is a σ-PO that inverts $\mathbf{c}_0$ (called σ-ρ-PO in the OD literature) and the pair (A_j, A_{j+1}) is polar, *i.e.* none of the σ-ρ-POs map A_j onto A_{j+1} *and* A_{j+1} onto A_j, then application of the inverse of a σ-ρ-PO creates a new pair (A_j, A'_{j+1}): A_{j+1} is mapped onto A_j, but A_j is not mapped onto any A_{j+1} derived by coset decomposition above. Thus the number of stacking possibilities is doubled. In **1**, for general r, the 2_{r+1} and $n_{r,s}$ operations are of this kind and therefore given A_j, there are $Z = 2[p1a1:p1] = 4$ ways of placing the adjacent layers. This is known as the NFZ relationship in the OD literature (Ďurovič, 1997).

For special values of r and s, the number of stacking possibilities decreases. Again, we use custom software routines, which we plan to publish in due course, to enumerate these special cases. For the OD groupoid family of **1**, there are the following sets of special parameters:
(i) $s = n$, $n \in \mathbb{Z}$: $Z = 2[p1a1:p1a1] = 2$
(ii) $r = n$, $n \in \mathbb{Z}$: $Z = [p1a1:p1] = 2$
(iii) $r = n$, $s = m$, $n, m \in \mathbb{Z}$: $Z = [p1a1:p1a1] = 1$

In the first case, the a glide planes of adjacent layers overlap. In the second case, the 2_{r+1} operation becomes a symmetry operation of the (A_j, A_{j+1}) layer pair, which therefore ceases being polar. In the third case, both conditions are true and the structure is fully ordered, *i.e.* there is no stacking ambiguity. Its space group depends on the parity of r and s.

Since the major polytype of **1** possesses $Pna2_1$ symmetry, the a glide planes of adjacent layers overlap and we can conclude that s is an integer. For even s, adjacent A layers are related by an $a_{[001]}$, for odd s, by a $n_{[001]}$ glide reflection. In **1**,

Table 3
Polytypes with stacking sequences of primitive period with lengths $n = 1 \ldots 6$.

The stacking sequences and σ-sequences are given by a primitive period.

n	Stacking sequence	σ-sequence	$\mathbf{c}_0$	Symmetry
1	1	$+ +$	$2r\mathbf{a} + 2\mathbf{c}_0$	$P2_1/a$ (MDO$_1$)
1	2	$+ -$	$2\mathbf{c}_0$	$Pna2_1$ (MDO$_2$)
2	12	$+ + - -$	$4\mathbf{c}_0$	$P2_1/a$
3	112	$+ + + - - -$	$6\mathbf{c}_0$	$Pna2_1$
3	122	$+ + -$	$2r\mathbf{a} + 6\mathbf{c}_0$	$P2_1/a$
4	1112	$+ + + + - - - -$	$8\mathbf{c}_0$	$P2_1/a$
4	1122	$+ + + -$	$2r\mathbf{a} + 4\mathbf{c}_0$	Pa
4	1222	$+ + - + - - + -$	$8\mathbf{c}_0$	$P2_1/a$
5	11112	$+ + + + + - - - - -$	$10\mathbf{c}_0$	$Pna2_1$
5	11122	$+ + + + -$	$6r\mathbf{a} + 10\mathbf{c}_0$	$P2_1/a$
5	11212	$+ + + - -$	$2r\mathbf{a} + 10\mathbf{c}_0$	$P2_1/a$
5	11222	$+ + + - + - - - + -$	$10\mathbf{c}_0$	$Pna2_1$
5	12122	$+ + - - + - - + + -$	$10\mathbf{c}_0$	$Pna2_1$
5	12222	$+ + - + -$	$2r\mathbf{a} + 10\mathbf{c}_0$	$P2_1/a$
6	111112	$+ + + + + - - - - - -$	$12\mathbf{c}_0$	$P2_1/a$
6	111122	$+ + + + + -$	$4r\mathbf{a} + 6\mathbf{c}_0$	Pa
6	111212	$+ + + + - -$	$2r\mathbf{a} + 6\mathbf{c}_0$	$P2_1/a$
6	111222	$+ + + + - + - - - - + -$	$12\mathbf{c}_0$	$P2_1/a$
6	112122	$+ + + - - + - - - - + + -$	$12\mathbf{c}_0$	Pa
6	112222	$+ + + - + -$	$2r\mathbf{a} + 6\mathbf{c}_0$	Pa
6	121222	$+ + - - + -$	$6\mathbf{c}_0$	Pa
6	122222	$+ + - + - + - - + - + -$	$12\mathbf{c}_0$	$P2_1/a$

the former is the case and we conclude that $s = 0$. This can be expressed by the (nonstandard) symbol

$$
\begin{matrix}
p & 1 & 2_1/a & (1) \\
\{ & 2_{1+r}/n_{1,2} & - & (2_2/a_r) \quad \}
\end{matrix} \tag{2}
$$

The remaining metric parameter r can be derived from the refined coordinates of the **1** molecules. Averaging the atomic coordinates (excluding H atoms) of two independent molecules gives the inversion center of an A_j layer, from which r can be inferred. We obtain a clearly non-integral value of $r = 0.510$ which means that the intrinsic translational component of the a_r glide is $\frac{r}{2}\mathbf{a} = 0.255\mathbf{a}$.

Thus, given a layer A_j, the adjacent layer A_{j+1} can be placed in two ways, namely application of an a_r or an a_{-r} operation.

3.5. Maximum degree of order (MDO) polytypes

The stacking ambiguity leads to an infinite number of potential polytypes, which are all locally equivalent. Pairs of adjacent layers are always equivalent. However, there are two kinds of triples (A_j, A_{j+1}, A_{j+2}): application of twice a_r (or a_{-r}) and application of a_r followed by a_{-r} (or *vice versa*).

The two polytypes consisting of only either of the triples are said to be of an MDO (Dornberger-Schiff, 1982). They are particularly simple in the sense that all other polytypes can be decomposed into fragments of MDO polytypes. Moreover, ordered MDO polytypes need the least amount of contextual information during growth. The two MDO polytypes are:

(i) MDO$_1$: $P2_1/a$, $\mathbf{c} = 2\mathbf{c}_0 + r\mathbf{a}$, generated by repeated application of a_r (or a_{-r} for different orientation).

(ii) MDO$_2$: $Pna2_1$, $\mathbf{c} = 2\mathbf{c}_0$, generated by alternating application of a_r and a_{-r}.

The MDO$_2$ polytype corresponds to the refined $Pna2_1$ structure, fragments of MDO$_1$ were observed by 'phantom' molecules and additional diffraction peaks.

The point group of the OD groupoid family (the group generated by the linear part of all POs) is mmm. The point groups of the MDO polytypes are proper subgroups of mmm and they may therefore crystallize as twins if a stacking fault (a fragment of the other MDO polytype) appears.

The expected twin operations are the operations of the mmm point group of the OD groupoid family that do not appear in the point group of the polytype. The orientation states are determined formally by a coset decomposition: for the MDO$_1$ ($P2_1/a$) polytype there are [mmm:2] = 2 orientation states, related by $2_{[100]}$, $m_{[100]}$, $2_{[001]}$ or $m_{[001]}$. If r is idealized to $\frac{1}{2}$, the translation lattice of MDO$_1$ would be orthorhombic B-centered (oB). Since the twin operations are point operations of this lattice (the point group of an oB lattice is mmm), the twinning would be by *metric merohedry*, which means that the direct and reciprocal lattices of both twin domains coincide perfectly (Grimmer & Nespolo, 2006). Since r deviates from the ideal value, the twinning can be considered as being by *pseudo-merohedry*, with the twin obliquity depending on the deviation from $r = \frac{1}{2}$.

For MDO$_2$ ($Pna2_1$), there are [mmm:$mm2$] = 2 expected twin domains. This is a twin by inversion and, since the inversion is always a symmetry operation of any lattice, a twin by *merohedry*. It can't be directly showed in our experiment, since the short wavelength of 0.56 Å was far from the absorption edges of the C and O elements, which means that the deviation from Friedel's 'law' (stating that diffraction patterns are centrosymmetric) is too small to determine the volume fraction of the two twin domains (the Flack parameter). In fact, the estimated standard uncertainty of the Flack parameter was higher than 0.5, if refined.

Given the observed diffuse scattering, which implies an intimate stochastic mix of MDO$_1$ and MDO$_2$ fragments, one can assume that the MDO$_2$ fragments appear in both orientations to equal fractions, as per the law of big numbers. The structure refinement was therefore performed with a fixed 1:1 volume ratio.

3.6. Other polytypes

Potentially, there are an infinity of polytypes, which can be built of fragments of MDO$_1$ and MDO$_2$. In this section an in-depth discussion of these polytypes and their symmetries is given. Readers not interested in these details may proceed to §3.7.

We will designate as a *translational period* of a polytype the smallest n-tuple of adjacent layers from which the polytype can be constructed by translation. It must contain an even number of layers because A_j with even and odd j possess different orientations with respect to [100].

Since two adjacent layers are related either by a_{+r} or a_{-r} σ-POs, a classical approach of enumerating polytypes would be generating bi-infinite periodic strings over the binary alphabet $\{+, -\}$. Symmetry equivalents (*e.g.* those obtained by

Table 4
Number and fractions of polytype symmetries with stacking sequences of primitive period with length $n = 0 \ldots 30$.

The three rightmost columns give the fractions of all stacking sequences of period $\leq n$.

n	All	$P2_1/a$	$Pna2_1$	Pa	Running fractions (%)		
					$P2_1/a$	$Pna2_1$	Pa
1	2	1 (50.0%)	1 (50.0%)	0 (0.0%)	50.0	50.0	0
2	1	1 (100.0%)	0 (0.0%)	0 (0.0%)	66.7	33.3	0
3	2	1 (50.0%)	1 (50.0%)	0 (0.0%)	60.0	40.0	0
4	3	2 (66.7%)	0 (0.0%)	1 (33.3%)	62.5	25.0	12.5
5	6	3 (50.0%)	3 (50.0%)	0 (0.0%)	57.1	35.7	7.1
6	8	4 (50.0%)	0 (0.0%)	4 (50.0%)	54.5	22.7	22.7
7	16	7 (43.8%)	8 (50.0%)	1 (6.2%)	50.0	34.2	15.8
8	24	10 (41.7%)	0 (0.0%)	14 (58.3%)	46.8	21.0	32.3
9	42	14 (33.3%)	21 (50.0%)	7 (16.7%)	41.3	32.7	26.0
10	69	21 (30.4%)	0 (0.0%)	48 (69.6%)	37.0	19.7	43.4
11	124	31 (25.0%)	62 (50.0%)	31 (25.0%)	32.0	32.3	35.7
12	208	42 (20.2%)	0 (0.0%)	166 (79.8%)	27.1	19.0	53.9
13	378	63 (16.7%)	189 (50.0%)	126 (33.3%)	22.7	32.3	45.1
14	668	91 (13.6%)	0 (0.0%)	577 (86.4%)	18.8	18.4	62.9
15	1214	123 (10.1%)	607 (50.0%)	484 (39.9%)	15.0	32.3	52.8
16	2220	184 (8.3%)	0 (0.0%)	2036 (91.7%)	12.0	17.9	70.1
17	4110	255 (6.2%)	2055 (50.0%)	1800 (43.8%)	9.4	32.4	58.2
18	7630	371 (4.9%)	0 (0.0%)	7259 (95.1%)	7.3	17.6	75.1
19	14308	511 (3.6%)	7154 (50.0%)	6643 (46.4%)	5.6	32.5	61.9
20	26931	750 (2.8%)	0 (0.0%)	26181 (97.2%)	4.3	17.4	78.3
21	50944	1015 (2.0%)	25472 (50.0%)	24457 (48.0%)	3.2	32.7	64.1
22	96782	1519 (1.6%)	0 (0.0%)	95263 (98.4%)	2.4	17.3	80.3
23	184408	2047 (1.1%)	92204 (50.0%)	90157 (48.9%)	1.8	32.8	65.4
24	352450	3030 (0.9%)	0 (0.0%)	349420 (99.1%)	1.4	17.2	81.4
25	675180	4092 (0.6%)	337590 (50.0%)	333498 (49.4%)	1.0	32.8	66.2
26	1296477	6111 (0.5%)	0 (0.0%)	1290366 (99.5%)	0.7	17.1	82.1
27	2493680	8176 (0.3%)	1246840 (50.0%)	1238664 (49.7%)	0.5	32.9	66.6
28	4805388	12222 (0.3%)	0 (0.0%)	4793166 (99.7%)	0.4	17.1	82.5
29	9272778	16383 (0.2%)	4636389 (50.0%)	4620006 (49.8%)	0.3	32.9	66.8
30	17919558	24486 (0.1%)	0 (0.0%)	17895072 (99.9%)	0.2	17.1	82.7

switching + and −) are then removed. We will call these strings σ-sequences, as they list the σ-POs. The primitive period of a periodic bi-infinite word is the smallest fragment from which the word can be generated by repeated juxtaposition. Even-length primitive periods correspond to a translation period. For odd lengths, the primitive period must be doubled to obtain a translation period.

Here we propose a different approach, which respects the OD description: consider (periodic) polytypes as a periodic sequence of MDO_1 and MDO_2 triples (sharing a layer), represented by a periodic bi-infinite string over the alphabet {1, 2} (the number stands for the kind of MDO triple). These strings will be called stacking sequences. The primitive period of these strings correspond to an asymmetric unit of the polytype and better represents the complexity of a polytype than the σ-sequence. Moreover, weeding out symmetry equivalents is simpler, as only reverse strings have to be removed.

Periodic strings can be enumerated very efficiently according to Duval (1983), however, for primitive periods of lengths $n \geq 6$, pairs of non-palindromic reflected strings appear (*e.g.* strings with the primitive period 112122 and 221211), where one instance has to be removed. An efficient algorithm to directly avoid these pairs was described by Sawada (2001).

The polytypes with stacking sequences up to a primitive period length $n = 6$ are compiled in Table 3. For odd n, the translation period is doubled, owing to the different orientations of the A_j and A_{j+n} layers. For even n with an odd number of MDO_2 triples in the primitive period, the translation period is likewise doubled. The σ-sequence (third column) and **c** vector (fourth column) are trivially derived by observing that every MDO_2 triple changes direction ($a_{+r} \leftrightarrow a_{-r}$).

Since the $a_{[010]}$ glide reflection applies to all layers, it is a total operation of every polytype, *i.e* it is element of the symmetry group of the polytype. *Periodic* polytypes possess space group symmetry and their symmetry must therefore be a supergroup of a space group of type Pa. Other total symmetry operations can either derive from layer operations (λ-POs) or operations relating layers (σ-POs). A polytype is $P2_1/a$ symmetric, if the 2_1 operation of a layer is a global operation. This is only possible inside MDO_1 fragments and the remaining part of the primitive period must be palindromic. Thus, to check for $P2_1/a$ symmetry, locate all odd-length 1 . . . 1 subsequences in the primitive period of the stacking sequence and check if the rest is palindromic. An example is the primitive period 111121112, where 111 is a block of an odd number of MDO_1 fragments and the remainder 211112 is palindromic.[1]

---

[1] This example was chosen such that the central MDO_1 block is not at the beginning in the lexicographically sorted representative, which are the strings output from Duval's algorithm.

A polytype is $Pna2_1$ symmetric, if all layers A_j, A_{j+k}, where k is half the translation period, are related by an $n_{[100]}$ glide reflection. This is only possible if k is odd (because only then A_j and A_{j+k} have opposite orientation with respect to [100]), which means that the stacking sequence has an odd primitive period. In this case it is sufficient that there is an odd number of MDO_2 triples in the stacking sequence, because then the second half of the primitive period of the σ-sequence is precisely the reflected version of the first half. Consider for example the stacking sequence with the primitive period 12122 resulting in the σ-sequence with primitive period $+$ $+ - - + - - + + -$. Note that this condition precludes the conditions for $P2_1/a$ symmetry, because for odd n the latter requires and even number of MDO_2 triples in the palindromic part of even length.

All other polytypes possess only Pa symmetry. It is interesting to investigate the asymptotic behavior for stacking sequences with large periods. The number of polytypes whose stacking sequence has a primitive period of length n corresponds to the integer sequence A001371 of the *On-line Encyclopedia of Integer Sequences* (OEIS Foundation Inc., 2025). Those with $P2_1/a$ symmetry are given by the sequence A056513. In the $n \to \infty$ limit, the density of $P2_1/a$ sequences approaches zero. In contrast, $Pna2_1$ polytypes appear for precisely half of the stacking sequences with odd periods. Thus, in the limit, 25% of stacking sequences are $Pna2_1$ and the remaining 75% are Pa, whereas for small n, $P2_1/a$ dominates. Table 4 lists the numbers for $n \leq 30$.

3.7. Diffraction pattern

The diffraction pattern of **1** is complex, featuring more or less diffuse peaks that can be attributed to different domains on top of pronounced diffuse scattering (Fig. 4).
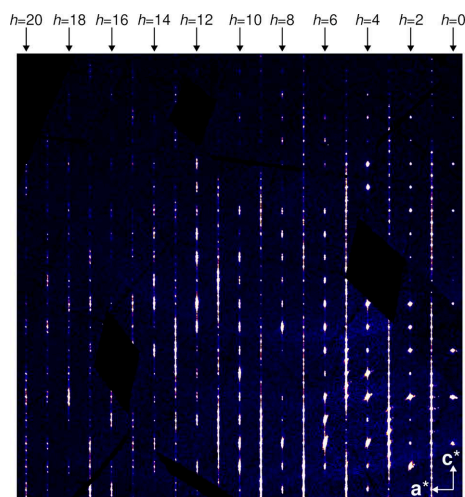

Figure 4
Reconstruction of the $k = 2$ plane from diffraction data. $k = 2$ was chosen arbitrarily. Other k constant sections are qualitatively similar, except $k = 0$, which is affected by the systematic absences of the a glide reflection common to all polytypes.

Describing the diffraction pattern requires a reciprocal basis $(\mathbf{a}^*, \mathbf{b}^*, \mathbf{c}^*)^T$ (the superscript T stays for the transpose operation, as reciprocal bases are written by convention as columns). Since any periodic polytype must possess a multiple of two layers in the primitive translation period (see above), we will use the dual basis of $(\mathbf{a}, \mathbf{b}, 2\mathbf{c}_0)$. Since $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}_0$ are mutually orthogonal, $\mathbf{a}^*$ is the vector parallel to $\mathbf{a}$ with the length $1/a$, *etc.*

The coordinate triples in reciprocal space will be designated as $hk\nu$ ($\mathbf{s} = h\mathbf{a}^* + k\mathbf{b}^* + \nu\mathbf{c}^*$) to highlight that diffraction can occur anywhere on rods parallel to $\mathbf{c}^*$ with integral h and k.

A crucial point in understanding diffraction patterns of polytypes is identifying reflections that are common to all domains. Thereto, the *family structure* is determined, which is informally an equal overlay of all possible polytypes. The symmetry group of the family structure is generated by extending the POs to global operations. For the OD family of **1**, let us first concentrate on the group of translations. Layers A_j and A_{j+2n} ($n \in \mathbb{Z}$) are related by a translation. The vector connecting the origins of both layers can be written as $2mr\mathbf{a} + 2n\mathbf{c}_0$, $-n \leq m \leq n$, whereby it is useful to reduce the $\mathbf{a}$ component to its fractional part frac$(2mr)$. The metric parameter r is not close to a rational with small denominator and can be assumed as being irrational. This means that with growing n, frac$(2mr)$ can be arbitrarily close to any number in the 0–1 range. Overall, the translation group of the family structure is dense in the $\mathbf{a}$ direction, and has periodicities $\mathbf{b}$ and $2\mathbf{c}_0$ in the other directions. Accordingly, the symmetry of the family structure is *not* a space group, but the smallest supergroup of the space groups of all possible polytypes. Owing to the dense translation vectors in the $\mathbf{a}$ direction, the $\mathbf{a}$-component of glides and screw rotations can likewise be arbitrary small. In the context of Euclidean normalizers (Koch *et al.*, 2016), similar groups exist, with a technical difference: for normalizers, the set of translations is uncountably infinite, so in a sense the set of translations is even 'more dense'.[2] If, nevertheless, the same notation is used, the symmetry group of the family structure can be stated as being of type P^1bmm with the 'basis' $(\varepsilon\mathbf{a}, \mathbf{b}, 2\mathbf{c}_0)$.

While this discussion may appear academical, a crucial consequence is that diffraction intensities may appear anywhere on rods h and k integral (since any A_j layer can only produce intensities on such rods), except for $h = 0$ rods. There, sharp reflections are observed at positions $k + l$ even owing to the b-glides in the P^1bmm group. A statistical stacking without any correlation between layers would produce unstructured diffuse scattering along $\mathbf{c}^*$ on all rods $h \neq 0$.

These considerations are only perfectly true if there is no desymmetrization, *i.e.* all layers and all layer contacts are perfectly equivalent. This assumption is confirmed by the diffraction pattern, as indeed for $h = 0$ it features sharp peaks and no diffuse scattering.

---

[2] Here, the translations in [100] direction form a $\mathbb{Z}$-module, for Euclidean normalizers they are isomorphic to $\mathbb{R}$.

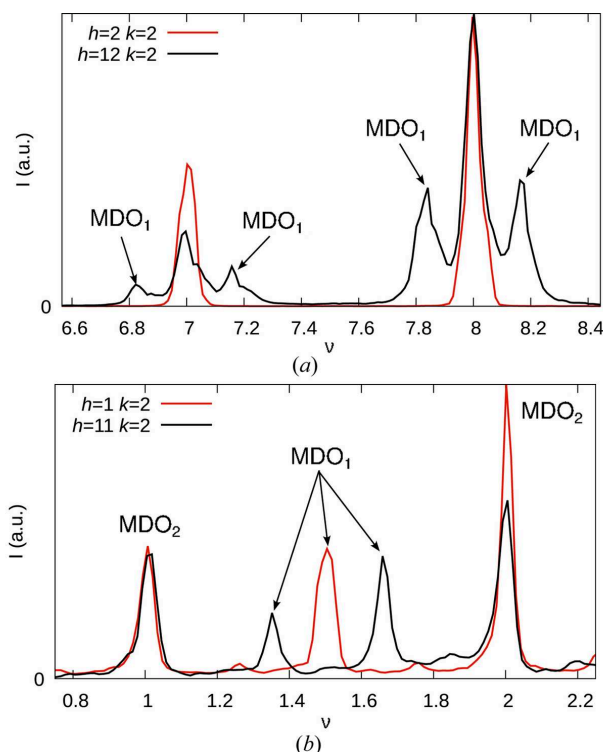## 3.8. Qualitative interpretation of the diffraction pattern

Figure 5
Comparison of experimental reflection profiles on (*a*) even *h* and (*b*) odd *h* rods with low (2, 1) and high (12, 11) *h* values, showing splitting of the MDO$_1$ 'reflections' for high *h* values resulting in characteristic triple and double peaks, respectively. In all cases, *k* = 2. The high-*h* intensities were scaled by a factor of 3 to bring them to a similar scale as the low-*h* intensities.

For $h \neq 0$, diffraction peaks are broader and placed on top of diffuse scattering (Fig. 4), which means that the crystal is not periodic (or quasi-periodic) and has to be described using disorder models. Nevertheless, it is useful to describe the peak positions qualitatively at first. Peaks corresponding to both MDO polytypes are observed. The reciprocal bases are:

(i) MDO$_1$ (first orientation): $(\mathbf{a}^* - r\mathbf{c}^*, \mathbf{b}^*, \mathbf{c}^*)^T$
(ii) MDO$_1$ (second orientation): $(\mathbf{a}^* + r\mathbf{c}^*, -\mathbf{b}^*, -\mathbf{c}^*)^T$
(iii) MDO$_2$ (both orientations): $(\mathbf{a}^*, \mathbf{b}^*, \mathbf{c}^*)^T$

The reflections of MDO$_2$ are all located at integral ν. Since $r \approx \frac{1}{2}$, for small even $|h|$ MDO$_1$ reflection are likewise expected close to integral ν. With increasing $|h|$, their positions deviate, resulting in distinct 'triplets' of MDO$_1$ and MDO$_2$ peaks [Fig. 5(*a*)]. For small odd $|h|$, MDO$_1$ reflections are close to half-integer ν. Deviations for increasing $|h|$ lead to a distinct splitting into 'doublets' [Fig. 5(*b*)].

Close inspection of the rods with low odd *h* values reveal additional peaks at quarter points of ν, marked by black arrows in Fig. 6. These peaks are significantly weaker than the MDO$_1$/MDO$_2$ peaks. For $|h| = 5$ they are barely observed, for $|h| = 7$ they are practically absent. It is unknown whether they belong to distinct domains or are part of the disordered MDO$_1$/MDO$_2$ domains.
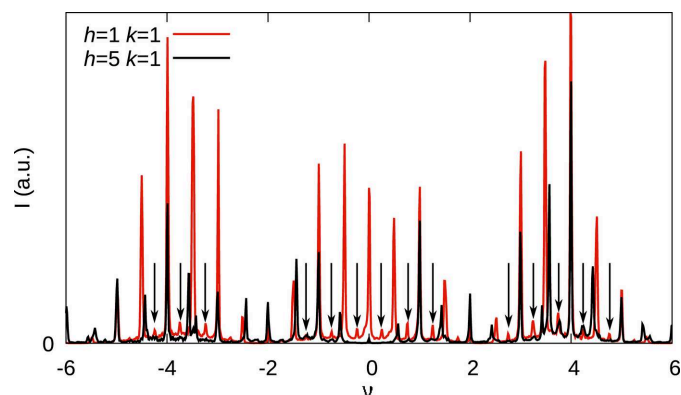

Figure 6
Comparison of reflection profiles on $(11\nu)^*$ (red) and $(51\nu)^*$ (black) rods showing additional peaks at quarter points of ν.

Since there is no broadening, splitting or deviation from $\nu = \frac{1}{4}n$ ($n \in \mathbb{Z}$) observed up to $|h| = 5$, we assume a polytype with a **c** vector without $r\mathbf{a}$ component. The potential polytypes with $\mathbf{c} = 8\mathbf{c}_0$ are (see §3.6 and Table 3):

1112, $+ + + + - - - -$, $P2_1/a$
1222, $+ + - + - - + -$, $P2_1/a$
11211222, $+ + + - - - + -$, Pa
11212212, $+ + + - - + - -$, Pa
12122222, $+ + - - + - + -$, Pa

Applying Occam's razor principle we can suppose that the first two, the simplest, polytypes are the most probable. The absence of these peaks on *h* even rods is a case of systematic non-space group absences and is due to $r \approx \frac{1}{4}$, which are characteristic of OD structures where layers are translationally equivalent. For derivation, see *e.g.*, Ferraris *et al.* (2008).

If the diffraction peaks are interpreted as Bragg peaks (*i.e.* their shape is only determined by experimental artifacts), then the crystal would be considered as an allotwin composed mostly of MDO$_1$ and MDO$_2$ with minor domains of an unknown non-MDO polytype, showed by very minor peaks at quarter points of ν, see above. Yet, owing to broad peaks and diffuse scattering, we will categorize the crystal under investigation as an allotwin-like disordered OD polytype. The aforementioned contributions of the non-MDO polytype are so weak that we will ignore them in the upcoming modeling of the diffraction pattern.

3.9. Growth models

Here we assume that the crystal is one homogeneous disordered domain, in contrast to distinct independently scattering domains. There are different ways to model such domains. For layer-stacking disorder we find that the most intuitive approach is growth models, which means that the position/orientation of the layer A_j is given in a probabilistic way depending on the *s* prior layers $A_{j-s} \ldots A_{j-1}$, where *s* is called the range of interaction or *reichweite* (Treacy *et al.*, 1991). It should be stressed that such a description does not describe the actual physico-chemical process of crystal growth, *i.e.* it does not imply that the crystal is grown layer-by-layer. It

is merely a mathematical model of expressing probabilities of finding specific layer n-tuples. Alternative equivalent descriptions are Ising models (Welberry, 2010), which have the advantage of being applicable to two-dimensional disorder (*i.e.* arrangements of rods).

The range of interaction s is a measure of complexity as is the number of parameters used to describe the model. We will require that the model follows the symmetry of the OD description, because we assume no external influences that could favor one orientation of a layer n-tuple over another.

The simplest model is described by a single parameter p, the probability of encountering an MDO_1 fragment (the probability of an MDO_2 fragment then is $1 - p$). Expressed as a growth model, it is a range of interaction $s = 1$ model represented by the graph

$$\text{(3)}$$
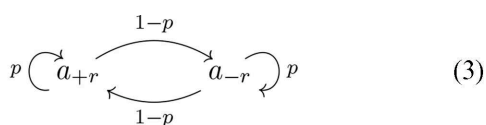
since every step depends only on the prior step. The nodes in the graph represent the σ-PO relating adjacent layers and the edges are marked with the transition probabilities. Strictly speaking, $p = \frac{1}{2}$ would represent an $s = 0$ model, as every step is independent. Such a model would result in unstructured diffuse scattering. For $p > \frac{1}{2}$ broad peaks at the MDO_1 positions appear, converging to the Bragg peaks of a 1:1 MDO_1 twin for $p \to 1$. The shape of the peaks is well understood and given for example in Hendricks & Teller (1942). In return, for $p \to 0$ the diffraction pattern converges to the 1:1 MDO_2 inversion twin. However, there exists no probability p for which peaks at the MDO_1 *and* MDO_2 positions are observed, *i.e.* such a simple model cannot describe the observed diffraction pattern.

The simplest model able to describe an allotwin-like disordered structure possesses two parameters p_1 and p_2 giving the probability of MDO_1 after MDO_1 and MDO_2 after MDO_2, respectively. It corresponds to the range of interaction $s = 2$ model represented by the graph

$$\text{(4)}$$
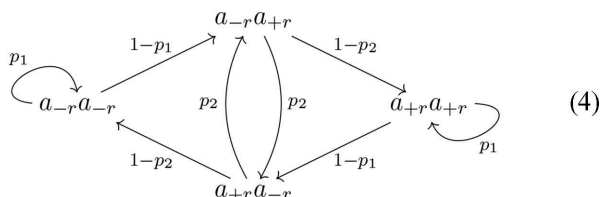
where each node represents the two prior steps. For $p = p_1 = 1 - p_2$ the model degenerates to the simple model of equation (3). The $s = 2$ model has two additional extreme cases:

$p_1, p_2 \to 0$: the non-MDO polytype with stacking sequence of primitive period 12, respectively σ-sequence of primitive period $+ + - - + + - -$ (see §3.6)

$p_1, p_2 \to 1$: an allotwin of MDO_1 and MDO_2

Thus, in a sense an allotwin and the simplest non-MDO polytype possess the same complexity: both require an $s = 2$ range of interaction, in contrast to MDO polytypes ($s = 1$).

Note that only probabilistic statements can be given on such random processes. In particular, in probability theory, statements with a probability *1* should be prefixed with the expression *almost surely*. Here, we will omit this expression for brevity. Solving the steady state of the random process in equation (4) shows that the fraction of MDO_1 and MDO_2 triples are

$$\frac{1 - p_2}{2 - p_1 - p_2} \tag{5}$$

and

$$\frac{1 - p_1}{2 - p_1 - p_2}, \tag{6}$$

respectively. Another interesting property is the average length of MDO_1 and MDO_2 fragments. It is a standard result that these are

$$1/(1 - p_1) \tag{7}$$

and

$$1/(1 - p_2), \tag{8}$$

respectively.

3.10. Diffuse scattering

Even though modeling diffuse scattering is often considered as being tricky, from a mathematical point of view it is easier to handle than diffraction of periodic structures. The latter is plagued by the issue of sums of regular density functions converging to discrete distributions. The convergence behavior of diffuse scattering is distinctly 'nicer' (uniform *versus* distributional convergence) and very reasonable approximations can be obtained with early series terminations. Though general software such as *DIFFaX* (Treacy *et al.*, 1991) and its derivatives can simulate diffuse scattering of layer structures in the general case, it is instructive to analyze the concrete case of **1** in detail. Readers not interested in structure factor calculations may proceed to §3.11.

Let us assume that the diffraction intensity $|F|^2$ of a polytype is the sum of the pair correlations between any two pairs of layers (Welberry, 2010):

$$|F|^2 = \lim_{N \to \infty} \frac{1}{2N + 1} \sum_{j_1 = -N}^{N} \sum_{j_2 = -N}^{N} F_{j_1} \overline{F}_{j_2}, \tag{9}$$

where F_j is the Fourier transform of the layer A_j and an overline indicates the complex conjugate. The overall sum is real, because for every $F_{j_1} \overline{F}_{j_2}$ term, it contains the complex

conjugate $F_{j_2}\overline{F}_{j_1}$. This can be made more explicit by summing over all unordered pairs of layers:

$$|F|^2 = \lim_{N\to\infty}\frac{1}{2N+1}\left[\left(\sum_{j=-N}^{N}F_j\overline{F}_j\right)\right.$$
$$\left.+\left(\sum_{j_1=-N}^{N-1}\sum_{j_2=j_1+1}^{N}F_{j_1}\overline{F}_{j_2}+F_{j_2}\overline{F}_{j_1}\right)\right] \quad (10)$$
$$=\lim_{N\to\infty}\frac{1}{2N+1}\left[\left(\sum_{j=-N}^{N}F_j\overline{F}_j\right)\right.$$
$$\left.+\left(\sum_{j=-N}^{N-1}\sum_{\Delta j=1}^{N-j}F_j\overline{F}_{j+\Delta j}+F_{j+\Delta j}\overline{F}_j\right)\right]. \quad (11)$$

Here, all terms are real (though potentially negative). The second line conveniently groups pair-correlations of layers spaced by Δj.

Note that all even j and all odd j layers are translationally equivalent and call F^+ and F^- the Fourier transforms of the respective layers when placed at the origin $[F^+(hk\nu)=F^-(\overline{hk}\nu)]$. Then we have:

$$F_j = \begin{cases} F^+\exp\left(2\pi i\mathbf{h}\cdot\mathbf{o}_j\right) & j \text{ even} \\ F^-\exp\left(2\pi i\mathbf{h}\cdot\mathbf{o}_j\right) & j \text{ odd} \end{cases} \quad (12)$$

where $\mathbf{o}_j$ is the origin of the layer A_j and the terms in equation (16) calculate as

$$F_j\overline{F}_j = \begin{cases} F^+\overline{F}^+ & j \text{ even} \\ F^-\overline{F}^- & j \text{ odd} \end{cases} \quad (13)$$

and

$$F_j\overline{F}_{j+\Delta j}+F_{j+\Delta j}\overline{F}_j$$
$$= \begin{cases} 2F^+\overline{F}^+\cos\left[2\pi\mathbf{h}\cdot\left(\mathbf{o}_j-\mathbf{o}_{j+\Delta j}\right)\right] & j\text{ even}, \Delta j\text{ even} \\ 2F^-\overline{F}^-\cos\left[2\pi\mathbf{h}\cdot\left(\mathbf{o}_j-\mathbf{o}_{j+\Delta j}\right)\right] & j\text{ odd}, \Delta j\text{ even} \\ 2\Re\left\{F^+\overline{F}^-\exp\left[2\pi i\mathbf{h}\cdot\left(\mathbf{o}_j-\mathbf{o}_{j+\Delta j}\right)\right]\right\} & j\text{ even}, \Delta j\text{ odd} \\ 2\Re\left\{F^-\overline{F}^+\exp\left[2\pi i\mathbf{h}\cdot\left(\mathbf{o}_j-\mathbf{o}_{j+\Delta j}\right)\right]\right\} & j\text{ odd}, \Delta j\text{ odd} \end{cases}$$
$$(14)$$

where $\Re$ designates the real value.

It is easy to construct layer sequences where the limit of equation (9) does not converge. However stochastic processes such as the one given in equation (4) typically possess very 'nice' (exponential) convergence behavior. The expected densities of $\mathbf{o}_j-\mathbf{o}_{j+\Delta j}$ are well defined and easily calculated.

For a given $\Delta j\geq 1$, $\mathbf{o}_j-\mathbf{o}_{j+\Delta j}$ may adopt the form $m r\mathbf{a}/2+\Delta j\mathbf{c}_0$ with m being of the same parity as Δj and in the range $-\Delta j\to\Delta j$. In the model of equation (9), the expected density of $\mathbf{o}_j-\mathbf{o}_{j+\Delta j}$ is independent of the parity of j and will be designated as $P_{\Delta j}(m)$, which accordingly must fulfill

$$\sum_{m=-\Delta j}^{\Delta j}P_{\Delta j}(m)=1. \quad (15)$$

To clarify, let us consider a few examples:

(i) In the MDO$_1$ polytype (repeated application of a_r), the origins of layers L_j and $L_{j+\Delta j}$ are connected by $\Delta jr\mathbf{a}/2+\Delta j\mathbf{c}_0$. Thus, $P_{\Delta j}(m)=1$ if $m=\Delta j$ and 0 otherwise.

(ii) In the other orientation of the MDO$_1$ polytype (repeated application of a_{-r}), $P_{\Delta j}(m)=1$ if $m=-\Delta j$ and 0 otherwise.

(iii) In a twin of both MDO$_1$ orientations with twin volume fraction $\frac{1}{2}$, $P_0(0)=1$ and for $\Delta j\neq 0$: $P_{\Delta j}(m)=\frac{1}{2}$ if $|m|=|\Delta j|$ and 0 otherwise.

(iv) In the MDO$_2$ polytype (alternating application of a_r and a_{-r}), $P_{\Delta j}(0)=1$ for even Δj and $P_{\Delta j}(-1)=P_{\Delta j}(+1)=\frac{1}{2}$ for odd Δj. $P_{\Delta j}(m)=0$ for all other combinations. The same is true for twins of both MDO$_2$ orientations, because, by definition, the twin interfaces are infinitely sparse compared to the bulk of the twin domains.

(v) In a purely random stacking, where a_r and a_{-r} are chosen for each layer independently with a probability of $\frac{1}{2}$, for Δj and m of same parity,

$$P_{\Delta j}(m)=\binom{|\Delta j|}{\frac{m+|\Delta j|}{2}}/2^{|\Delta j|},$$

where $\binom{n}{k}$ is the binomial coefficient, leading to a Pascal's triangle-like construction.

For a $p_1\neq 0,1$ and $p_2\neq 0,1$ model, the expected densities of the vectors connecting the origins of the layers L_j and $L_{j+\Delta j}$ are independent of the parity of j. Thus, for a given $m r\mathbf{a}/2+\Delta j\mathbf{c}_0$, the terms in the top two and bottom two cases of equation (14) possess the same expected density of $P_{\Delta j}(m)/2$ and we can further simplify the Δj odd case to

$$\frac{P_{\Delta j}(m)}{2}\left(2\Re\left\{F^+\overline{F}^-\exp\left[2\pi i\mathbf{h}\cdot(m r\mathbf{a}/2+\Delta j\mathbf{c}_0)\right]\right\}\right.$$
$$\left.+2\Re\left\{F^-\overline{F}^+\exp\left[2\pi i\mathbf{h}\cdot(m r\mathbf{a}/2+\Delta j\mathbf{c}_0)\right]\right\}\right)$$
$$=P_{\Delta j}(m)(F^+\overline{F}^-+F^-\overline{F}^+)\cos\left[2\pi\mathbf{h}\cdot(m r\mathbf{a}/2+\Delta j\mathbf{c}_0)\right].$$
$$(16)$$

We can now resolve the limits of equation (11). The auto-correlation of layers is

$$\lim_{N\to\infty}\frac{1}{2N+1}\sum_{j=-N}^{N}F_j\overline{F}_j=\frac{F^+\overline{F}^++F^-\overline{F}^-}{2} \quad (17)$$

and the cross-correlation of layers spaced by Δj:

$$\lim_{N \to \infty} \frac{1}{2N+1} \sum_{j=-N}^{N-1} \sum_{\Delta j=1}^{N-j} F_j \overline{F}_{j+\Delta j} + F_{j+\Delta j} \overline{F}_j$$

$$= (F^+ \overline{F}^+ + F^- \overline{F}^-)$$

$$\times \left\{ \sum_{\Delta j \in 2\mathbb{N}^{>0}} \sum_{m=-\Delta j/2}^{\Delta j/2} P_{\Delta j}(2m) \cos\left[2\pi \mathbf{h} \cdot (2m r \mathbf{a}/2 + \Delta j \mathbf{c}_0)\right] \right\}$$

$$+ (F^+ \overline{F}^- + F^- \overline{F}^+)$$

$$\times \sum_{\Delta j \in 2/\mathbb{N}+1} \left\{ \sum_{m=-(\Delta j-1)/2}^{(\Delta j-1)/2} P_{\Delta j}(2m+1) \right.$$

$$\left. \times \cos\left[2\pi \mathbf{h} \cdot \left((2m+1) r \mathbf{a}/2 + \Delta j \mathbf{c}_0\right)\right] \right\}, \qquad (18)$$

where $2\mathbb{N}^{>0}$ and $2\mathbb{N}+1$ are the sets of positive even and odd numbers, respectively. The scalar product $\mathbf{h} \cdot (m r \mathbf{a}/2 + \Delta j \mathbf{c}_0)$ can be written in $hk\nu$ coordinates as $(mrh + \Delta j\nu)/2$ and thus the diffraction intensity at $hk\nu$ is given as

$$|F|^2 = (F^+ \overline{F}^+ + F^- \overline{F}^-)$$

$$\times \left\{ 1 + \sum_{\Delta j \in 2\mathbb{N}^{>0}} \sum_{m=-\Delta j/2}^{\Delta j/2} P_{\Delta j}(2m) \cos\left[\pi(2mrh + \Delta j\nu)\right] \right\}$$

$$+ (F^+ \overline{F}^- + F^- \overline{F}^+)$$

$$\times \left\{ \sum_{\Delta j \in 2\mathbb{N}+1} \sum_{m=-(\Delta j-1)/2}^{(\Delta j-1)/2} P_{\Delta j}(2m+1) \right.$$

$$\left. \times \cos\left[\pi((2m+1)rh + \Delta j\nu)\right] \right\}. \qquad (19)$$

Note that the factors $(F^+ \overline{F}^+ + F^- \overline{F}^-)$ and $(F^+ \overline{F}^- + F^- \overline{F}^+)$ are real and can be pre-calculated for every h, k rod with the experimental resolution of ν. The $P_{\Delta j}(m)$ are easily calculated by step-wise development of the stochastic process with the starting probabilities derived from equations (5, 6), For p_1, p_2 not close to 0 or 1, the $P_{\Delta j}$ converge quickly and the series in equation (19) can be truncated at low Δj values. In our simulations, we used $\Delta j_{\max} = 40$ for adequate results with very short calculation times in the microsecond range.

3.11. Properties of the diffuse scattering

Before analyzing the actual diffraction pattern, let us examine a few of its properties. When analyzing diffuse scattering, homometry has to be taken into account (Welberry, 2010), that is distinct models resulting in the same diffraction pattern. As we will show in an upcoming publication, for $r = \frac{1}{2}$, there are homometric pairs, in particular models with $p_1 = p_2 = p$ and $p_1 = p_2 = 1 - p$ describe distinct structures resulting in the same diffraction pattern. Here, this affects only rods with low $|h|$ values, because the hr factors in equation (19) are close to $\frac{1}{2}$. Consider for example the allotwin-like $p_1 = p_2 = 0.9$ and the non-MDO-like $p_1 = p_2 = 0.1$ structures. Both are composed of an equal amount of MDO_1 and MDO_2 triples. However in the former, the average length of MDO_1 and MDO_2 fragments is 10, in the latter $10/9 \approx 1.11$, *i.e.* the structures are fundamentally different. Yet, the $h = 1$ rods of both structures are
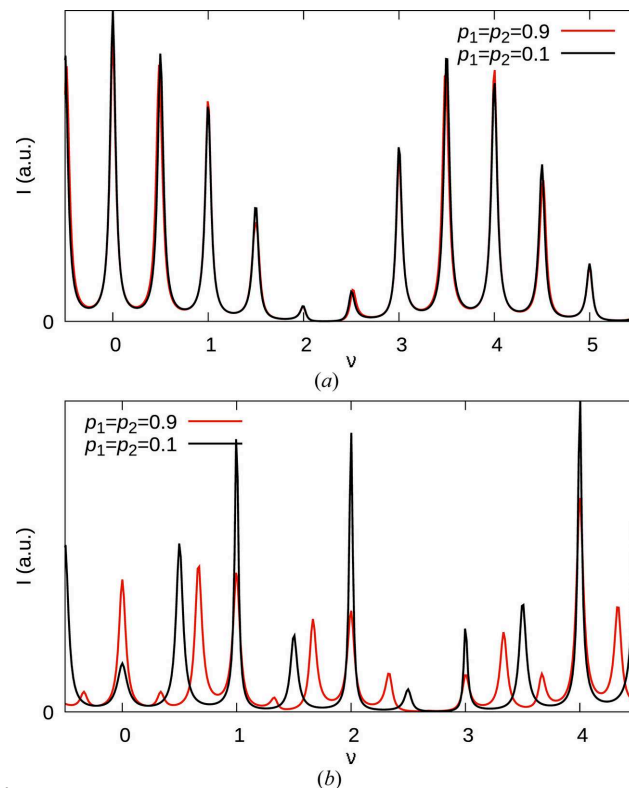

Figure 7
Simulations of the (a) $(11\nu)^*$ and (b) $(\underline{1}11\nu)^*$ rods with the parameters $p_1 = p_2 = 0.9$ (red) and $p_1 = p_2 = 0.1$ (black) and $r = 0.51$.

practically indistinguishable [Fig. 7(a)]. There are only very subtle changes in the peak heights and positions. In contrast, the $h = 11$ rods differ substantially [Fig. 7(b)]. In particular, the peaks at half-integer ν are split in the allotwin-like phase, but remain unsplit in the non-MDO-like phase, as expected. Note that homometric mates also exist for pairs $p_1 \neq p_2$ in the case of $r = \frac{1}{2}$, though the deviation of their growth-model probabilities is non-trivial and will be discussed in a future publication. In these cases too, differences in the diffraction pattern increase with $|h|$.
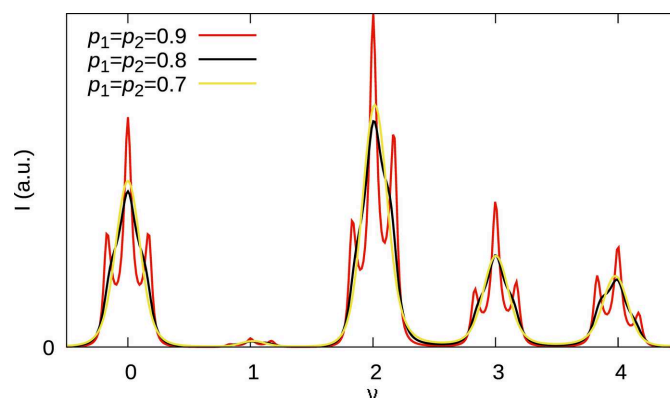

Figure 8
Simulations of the $(\underline{1}21\nu)^*$ rod with the parameters $p_1 = p_2 = 0.9$ (red), $p_1 = p_2 = 0.8$ (black) and $p_1 = p_2 = 0.7$ (yellow) and $r = 0.51$. With decreasing h, the broadening / splitting is even less pronounced.

Thus, in the present case, it is most productive to analyze rods with high h values. Moreover, the splitting of the peaks in the experimental data [Fig 5(b)] clearly shows that only models with $p_1, p_2 > \frac{1}{2}$ have to be considered.

Let us further evaluate the effect of the parameters p_1, p_2 on characteristic triplet peaks at high even h values [see experimental data in Fig. 5(b)]. Fig. 8 shows the $(\underline{1}21\nu)^*$ rod for $p_1 = p_2 = 0.9, 0.8, 0.7$ models. These correspond to allotwin-like disordered structures with equal amounts of MDO_1 and MDO_2 triples and average MDO_1 and MDO_2 fragment sizes of 10, 5 and $3\frac{1}{3}$, respectively. A clear separation of MDO_1 and MDO_2 peaks is only observed for the $p_1 = p_2 = 0.9$ case. For $p_1 = p_2 = 0.8$ there is only a single peak with slight shoulders. A qualitative interpretation in this case would suggest an MDO_2 disordered phase, even though the structure is built of equal amounts of MDO_1 and MDO_2 fragments, which proves that such interpretations are treacherous. The diffraction intensities of the $p_1 = p_2 = 0.8$ and $p_1 = p_2 = 0.7$ models are barely distinguishable though the average size of the fragments decreased by 33%. This shows that determining the actual domain size is very difficult in highly disordered polytypes. Given the clear splitting into triplets in the experimental pattern [compare Fig. 5(b)] one can expect values of p_1 and p_2 close to 1.

3.12. Semi-quantitative interpretation of the diffuse scattering

Despite the expected difficulties mentioned in the last section, here we will try to give a semi-quantitative evaluation of the diffuse scattering.

One issue to consider is that the shape of the diffuse scattering is convolved with an experimental profile. Even though we minimized artifacts by using synchrotron radiation featuring high monochromaticity and low beam divergence, some broadening due to still not fully perfect beam, detector resolution, and most notably crystal size and mosaicity cannot be prevented. Moreover, the probabilities p_1, p_2 may change throughout the macroscopic crystal. As noted in §3.7, the $h = 0$ peaks can be considered being mostly unaffected by diffuse scattering. Therefore we used these sharp peaks to estimate
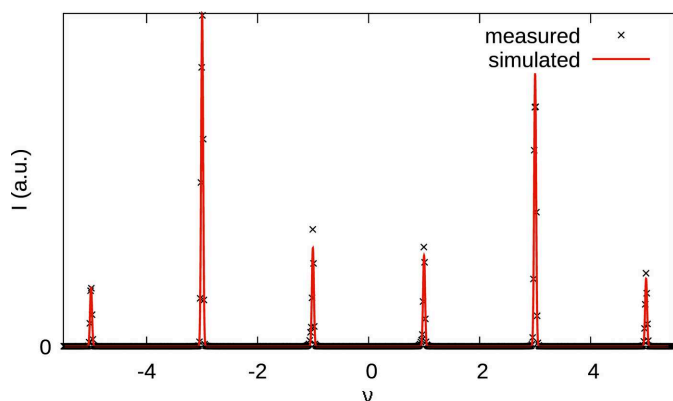
the experimental broadening. By fitting a sum of normal distributions on the peaks $01l$, $l = -5 \ldots 5$ extracted from two-dimensional reconstructed reciprocal space sections, we obtained a variance of $\sigma^2 = 1.32352\,(12)$ in pixels (Fig. 9). The peak profile is well described by a Gauss function. A small anisotropy at the base of some peaks is due to parasitic diffraction of additional domains, respectively mosaicity. By using high quality synchrotron radiation, the broadening is tiny, but as we will show below not negligible.

We optimized the model based on two sections of two high-h rods ($\underline{1}12\nu$ and $\underline{1}22\nu$) using the MCS algorithm (Huyer & Neumaier, 1999), which performs a global search. In contrast to classical local searches, a global search finds all local minima, which might be necessary in the case of (pseudo-) homometry. Here, owing to high-h values this was not the case.

Owing to inconsistencies in the reciprocal space reconstructions, which are due to different crystal positions with different incoherent scattering and absorption effects, we only optimized short sections of the rods. The refined parameters were: center in pixels, length of $\mathbf{c}^*$ in pixels, r, p_1 and p_2. We performed optimizations with and without convolution of the experimental normal distribution. The loss function was the unweighted R_p (the average difference between calculated and measured intensity, p for profile), as determined by a linear regression. Fig. 10 shows the measured and simulated intensities of the optimized model. Table 5 summarizes the refined parameters.
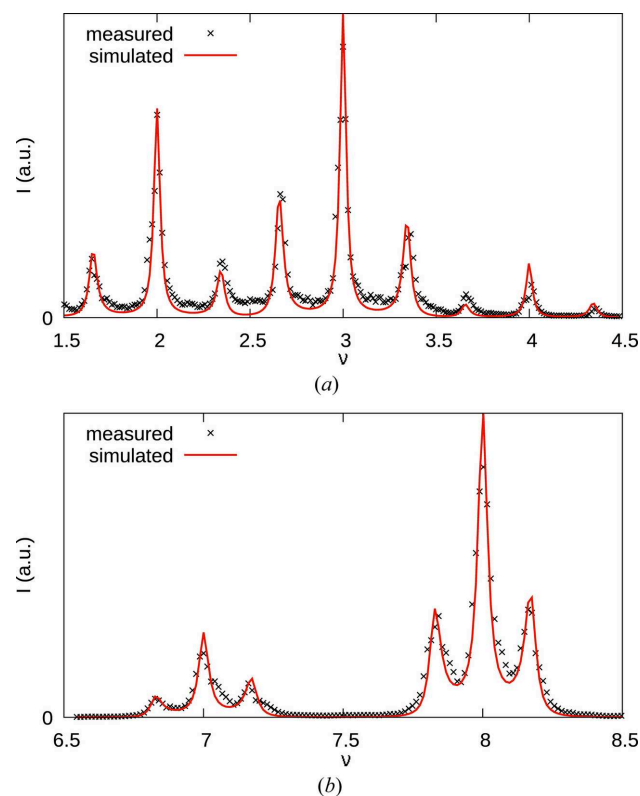

Figure 9
Experimental profile on the $(01\nu)^*$ rod (black crosses) and fitted sum of normal distributions (red line) with common variance.

Figure 10
Optimization results on the (a) $(\underline{1}12\nu)^*$ and (b) $(\underline{1}22\nu)^*$ rods. Simulated and measured intensity profiles are given by red lines and black crosses, respectively.

Table 5
Optimization results with and without convolution with experimental peak profile.

	With convolution		Without convolution	
	$(112\nu)^*$	$(122\nu)^*$	$(112\nu)^*$	$(122\nu)^*$
Center (pixels)	1567.68	1570.00	1567.77	1570.00
$|\mathbf{c}^*|$ (pixels)	74.123	74.854	74.103	74.851
r	0.5141	0.5146	0.5140	0.5149
p_1	0.9273	0.9301	0.9036	0.9052
p_2	0.9428	0.9399	0.9226	0.9173
R_p	0.0522	0.0081	0.0389	0.0087

Generally we found simulations of h even rods much more robust than of h odd rods, which is reflected by significantly improved R_p values. As an example of a mediocre fit, see for example the valley at $\nu = 2.5$ in Fig. 10(a). In a different structure (publication in preparation) we found these errors to be due to unaccounted desymmetrization: In MDO_1 fragments and MDO_2 fragments one can expect the molecules to adopt slightly different orientations. There, better fits of the valleys were obtained by using different structure factors for distinct fragments. Thereto, structures of alternative polytypes have to be determined either experimentally by growing crystals of distinct polytypes or by theoretical calculations, based on quantum-chemistry or molecular mechanics.

Despite these issues, as well as parasitic diffraction resulting in shoulders, the evaluation of the parameters p_1 and p_2 seems to be rather reproducible. They are determined by the width of the peaks, but also the relative heights of MDO_1 versus MDO_2 peaks. When convolving with the experimental profile one finds $p_1 \approx 0.93$ and $p_2 \approx 0.94$, which would correspond to an allotwin-like structure with ca 47% MDO_1 and 53% MDO_2 fragments. Though, given the inherent imprecision, one probably should consider the structure as being composed of approximately equal amounts of MDO_1 and MDO_2. The average length of the fragments is ca 14 (MDO_1) and 17 (MDO_2).

By not taking into account the experimental peak broadening, the values of p_1 and p_2 decrease to $p_1 \approx 0.91$ and $p_2 \approx 0.92$, which proves that even with low-divergence synchrotron radiation the experimental peak profile cannot by neglected completely.

The refined metric parameter $r \approx 0.515$ is within 1% of r derived from the structural refinements ($r = 0.51034$, see §3.4). The latter is certainly the more precise value, though it has to be noted that again r may be affected by desymmetrization and change for different local environments in the structure.

4. Conclusion and outlook

The title compound is further proof of the fascinating complexity that often arises from polytypism, with a single crystal showing peaks that can be attributed to three distinct polytypes. It also confirms the pervasiveness of OD structures and the generality of OD theory. However, software support such as automatic determination of partial pseudo-symmetry is still lacking and needs to be improved.

Our analysis shows that a quantitative interpretation of the diffraction pattern of disordered crystals can be very treacherous. The occurrence of fragments of minor polytypes are easily underestimated. Pronounced changes in the stacking probabilities result in only mild changes of the peak profiles, which are however affected by experimental artifacts. The robustness of classical single-crystal diffraction is due to the ability of ignoring real structure effects leading to peak broadening by reducing intensities to discrete values and a luxurious data to parameter ratio. Both are lost when analyzing diffuse scattering. On the flip side, crucial structural information is lost by averaging over all unit cells in the crystal, as is done in a classical structure refinement.

Notably, the analysis of the diffuse scattering provides a complexity measure of the structure, namely the range of interaction. To properly describe the observed diffraction pattern, it was necessary to use a range of interaction $s = 2$ model, which allows the modeling of an allotwin-like disordered structure. The same parametrization can describe the simplest of the non-MDO polytype. The crystal under investigation also featured minor amounts of a more complex non-MDO polytype, which requires a longer range of interaction.

References

Aksenov, S. M., Charkin, D. O., Banaru, A. M., Banaru, D. A., Volkov, S. N., Deineko, D. V., Kuznetsov, A. N., Rastsvetaeva, R. K., Chukanov, N. V., Shkurskii, B. B. & Yamnova, N. A. (2023). *J. Struct. Chem.* **64**, 1797–2028.
Dornberger-Schiff, K. (1982). *Acta Cryst.* A**38**, 483–491.
Dornberger-Schiff, K. & Grell-Niemann, H. (1961). *Acta Cryst.* **14**, 167–177.
Ďurovič, S. (1997). *EMU Notes Mineral.* **1**, 3–28.
Duval, J. P. (1983). *J. Algorithms* **4**, 363–381.
Ferraris, G., Makovicky, E. & Merlino, S. (2008). *Crystallography of Modular Materials. IUCr Monographs on Crystallography* 15, Oxford: Oxford University Press.
Fichtner, K. (1977*a*). *Cryst. Res. Technol.* **12**, 1263–1267.
Fichtner, K. (1977*b*). *Beitr. Algebra Geom.* **6**, 71–99.
Grimmer, H. & Nespolo, M. (2006). *Z. Kristallogr.* **221**, 28–50.
Hendricks, S. & Teller, E. (1942). *J. Chem. Phys.* **10**, 147–167.
Huyer, W. & Neumaier, A. (1999). *J. Glob. Optim.* **14**, 331–355.
Ito, T. & Sadanaga, R. (1976). *Proc. Japan Acad.* **52**, 119–121.
Koch, E., Fischer, W. & Müller, U. (2016). In *International Tables For Crystallography*, Vol. A, *Space-group symmetry*, ch. 3.5, pp. 826–851. Chester: IUCr.
Kopsky, V. & Litvin, D. B. (2006). Editors. *International Tables for Crystallography*, Vol. E, *Subperiodic groups*. Chester: IUCr.
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**, 266–235.
Nespolo, M. & Ferraris, G. (2004). *Acta Cryst.* A**60**, 89–95.
Nespolo, M., Kogure, T. & Ferraris, G. (1999). *Z. Kristallogr.* **214**, 5–8.

OEIS Foundation Inc. (2025). *The On-Line Encyclopedia of Integer Sequences.* https://oeis.org.

Powers, Z., Scharf, A., Cheng, A., Yang, F., Himmelbauer, M., Mitsuhashi, T., Barra, L., Taniguchi, Y., Kikuchi, T., Fujita, M., Abe, I. & Porco, J. A. (2019). *Angew. Chem. Int. Ed.* **58**, 16141–16146.

Rigaku Oxford Diffraction, (2022). *CrysAlisPro Software System* (version 1.171.42.69a). Rigaku Oxford Diffraction, Yarnton, England.

Sawada, J. (2001). *SIAM J. Comput.* **31**, 259–268.

Schöbinger, M. & Stöger, B. (2024). *Z. Kristallogr.* **239**, 331–338.

Sheldrick, G. M. (2015a). *Acta Cryst.* C**71**, 3–8.

Sheldrick, G. M. (2015b). *Acta Cryst.* A**71**, 3–8.

Stöger, B., Pokorny, B., Lumpi, D., Zobetz, E. & Fröhlich, J. (2013). *Z. Kristallogr.* **228**, 106–112.

Treacy, M. M. J., Newsam, J. M. & Deem, M. W. (1991). *Proc. R. Soc. A* **433**, 499–520.

Welberry, T. R. (2010). *Diffuse X-ray Scattering and Models of Disorder*, Vol. 16 of *IUCr Monographs on Crystallography*. Oxford: Oxford University Press.

Wondratschek, H. & Jeitschko, W. (1976). *Acta Cryst.* A**32**, 664–666.