# Examples of order-disorder in macromolecular crystals

Andrey Lebedev

York Structural Biology Laboratory
University of York

One example in detail

OD-structures

OD-twins, allotwins

Partial disorder

Inverted stacking vectors

One example in detail

OD-structures

OD-twins, allotwins

Partial disorder

Inverted stacking vectors

# One example in detail

Rye et al. (2007).
*Acta Cryst.* D**63**, 926–930.

Molecule: L-2-haloacid dehalogenase from *Sulfolobus tokodaii*

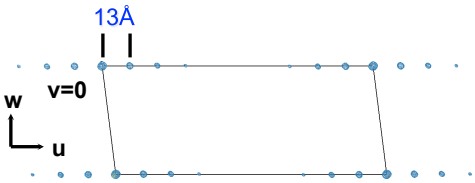
Spacegroup: C2 Unit cell: a = 127.6, b = 58.1, c = 51.2 Å, β = 97.2˚

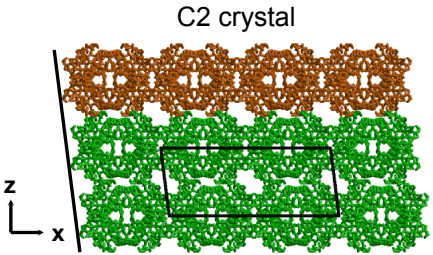
Resolution = 1.9 Å R-sym = 9.5% (16%)

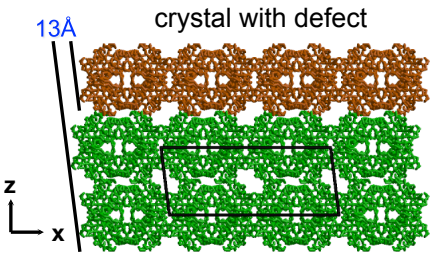
Asymmetric unit: dimer Method: MR

Initial refinement: R = 0.21 R-free = 0.27

The electron density for protein atoms is defined, but it is poor for solvent structure.




## Inspection of the Patterson map

There are series of non-origin peaks.
Their heights are up to 0.2 of the origin peak height.

## Analysis of crystal packing

There is no NCS by translation

However:

C2 crystal is formed by layers with diperiodic symmetry C22(2).

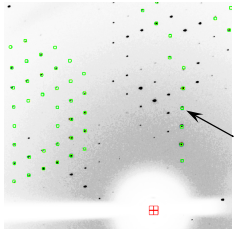
Defects preserving bilayers (i.e. contacts between layers) are possible.

Defective layer would be shifted from the "correct" position by 13Å. This is the distance between neighbouring peaks in the Patterson map.

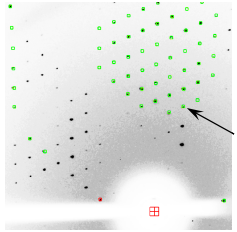
The crystal must have been a twinned crystal.

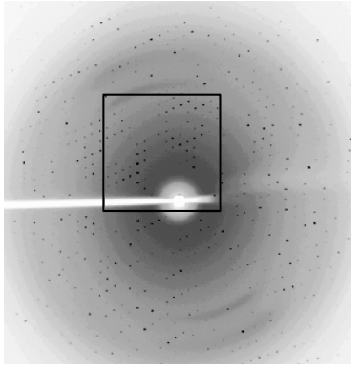
# Analysis of diffraction images

Indexing in C2



Indexing in C2





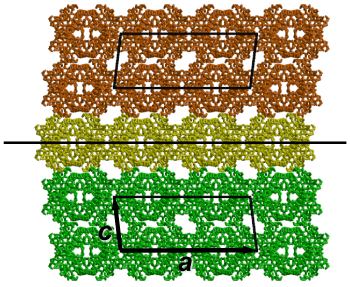
A typical story: the data were not analysed carefully until problem occurred with the structure solution or refinement.

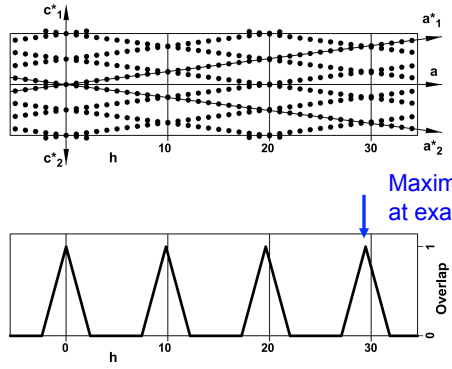
The diffraction images can be indexed in C2 with two different orientation of the crystal.

The two lattices partially overlap.

The presence of layers of overlapping reflections must be the reason of non-origin peaks in the Patterson map.

# Relation between two lattices



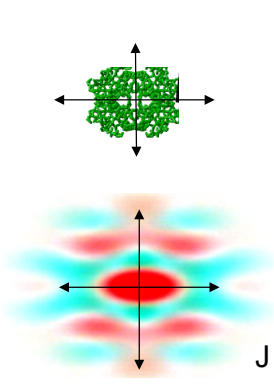

The twin geometry follows from the atomic model of an individual crystal.

In particular:

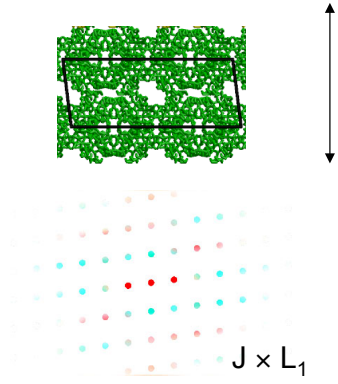
twinning by reticylar merohedry with twin index 10 and obliquity 0.1°

The overlap of spots is independent on *k* and *l*.
The overlap is a periodic function of *h* (in the first approximation).

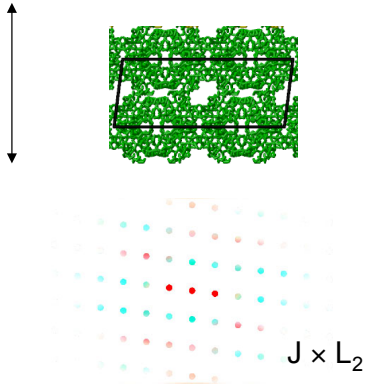
# Intensities of the overlapping reflections




Fourier transform
of the tetramer




Diffraction pattern of domain 1




Diffraction pattern of domain 2

Tetramers in different twin domains are in the same orientation

Therefore, if reflections of the two lattices overlap, they have close intensities.
The stronger the overlap, the closer the intensities are.

# Detwinning

$I_{T1} = (1 - \alpha)\, I_1 + \alpha\, q_{12}\, I_2$

$I_{T1}$ - an observed intensity from twinned crystal;
$I_1$ and $I_2$ - intensities from different domains;
$q_{12}$ - overlap between twin related reflections;
$\alpha$ - the twinning fraction.

$q_{12} = q(h)$

In addition, we know that only reflections with the same $h$ can overlap and that the overlap is independent on $k$ and $l$.

$I_1 = I_2$

In addition, we know that overlapping reflections have close intensities.

$I_{T1} = (1 - \alpha + \alpha\, q(h))\, I_1$

Altogether: de-twinning is reduced to demodulation; the modulation is defined by one-dimensional periodic function.

$I_{T1} = q'(h)\, I_1$

Concise form: there is no need to determine the twinning fraction $\alpha$.

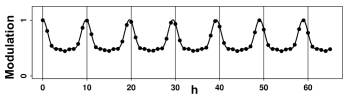
# Demodulation

$I_{T1} = q'(h)\ I_1$

$q'(h) = p_0 + p_1 \cos(2\pi th) + p_2 \cos(4\pi th) + \ldots$

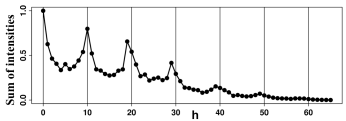
$p_i$ and $t$ were refined

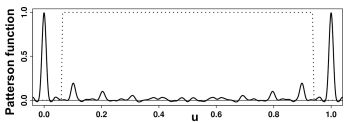
$q'(h)$




R / R-free

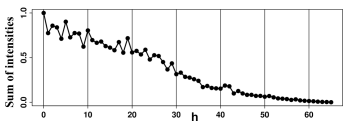
Original data

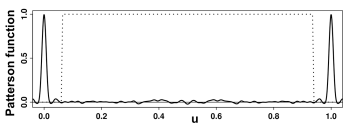



0.21 / 0.27

Demodulated data



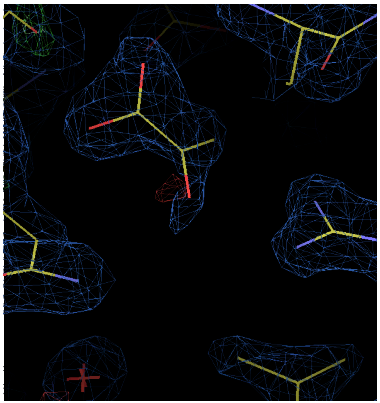

0.162 / 0.225

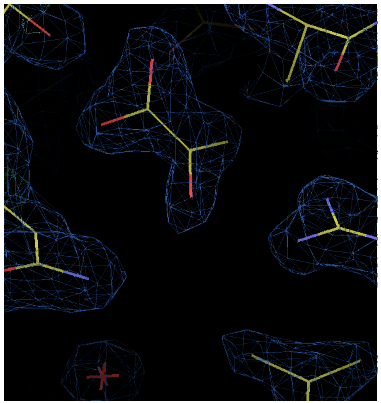
v = w = 0

# Improvement in the electron density

Visual improvement occurred only for the electron density for solvent molecules
(Poor density for solvent was the original reason for data revision)

The electron density maps (2-1 at 1.5σ and 1-1 at 3σ)
around the pyruvate molecule before and after demodulation





One example in detail

OD-structures

OD-twins, allotwins

Partial disorder

Inverted stacking vectors

*Acta Cryst.* (1961). **14**, 167

# On the Theory of Order-Disorder (OD) Structures

By K. Dornberger-Schiff and H. Grell-Niemann

*Institut für Strukturforschung der Deutschen Akademie der Wissenschaften zu Berlin, Germany*

(*Received* 4 *January* 1960 *and in revised form* 17 *March* 1960)

**OD**-structures consisting of equivalent layers are first characterized as having pairs of adjacent layers which are all equivalent. Then a slightly more general condition—the *'vicinity condition'*—is formulated which is satisfied not only by all ordered structures but also by all **OD**-structures. Partial operations (**PO**s) are seen to be of fundamental importance for characterizing the symmetry properties of **OD**-structures and the set of **PO**s of a certain structure is called an **OD**-groupoid. **OD**-structures of the same substance, built of the same kind of layers with the same kinds of pairs of adjacent layers are said to belong to the same family, the corresponding **OD**-groupoids to the same **OD**-groupoid family. Twins of one particular type are described as special members of families of **OD**-structures. A report on the deduction of a complete list of **OD**-groupoid families is given, and the resulting numbers of such families with different symmetry characteristics are listed in tables. There are 333 in all.

## 1. Introduction

In earlier papers (Dornberger-Schiff, 1956, 1957, 1959*c*) one of us has described some examples of what we propose to call **OD**-*structures*. In such structures equivalent parts lie in equivalent vicinities but there need not be perfect long-range order.

repeating operation under consideration may be described by a **PO** or by combinations of **PO**s.

A **PO** is fully characterized by

(*a*) the transformation of space, and
(*b*) the layer which is to be transformed.

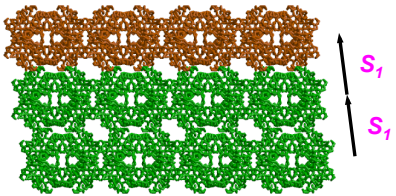
# OD-structures

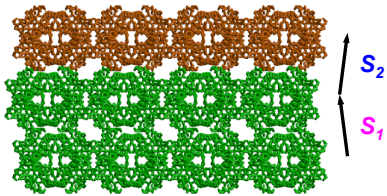
OD-structure

An OD-structure is composed of geometrically identical layers.
All pairs of contacting layers are equivalent, but triplets may differ.

Stacking vector

Stacking vector relates two neighbouring layers. There are two possible stacking vectors in the discussed example, $S_1$ and $S_2$. (In general case operation relating neighbouring layers includes rotation or reflection)




OD-family

All (putative or real) structures built of the same layers with the same interfaces form a family of OD-structures.

Single crystal

Regular sequences of stacking vectors: $S_1S_1S_1$... or $S_1S_1S_2S_1S_1S_2$... etc. represent single crystals.

OD-twin

$S_1S_1S_1S_1S_2S_2S_2S_2$... Two individuals have the same symmetry.

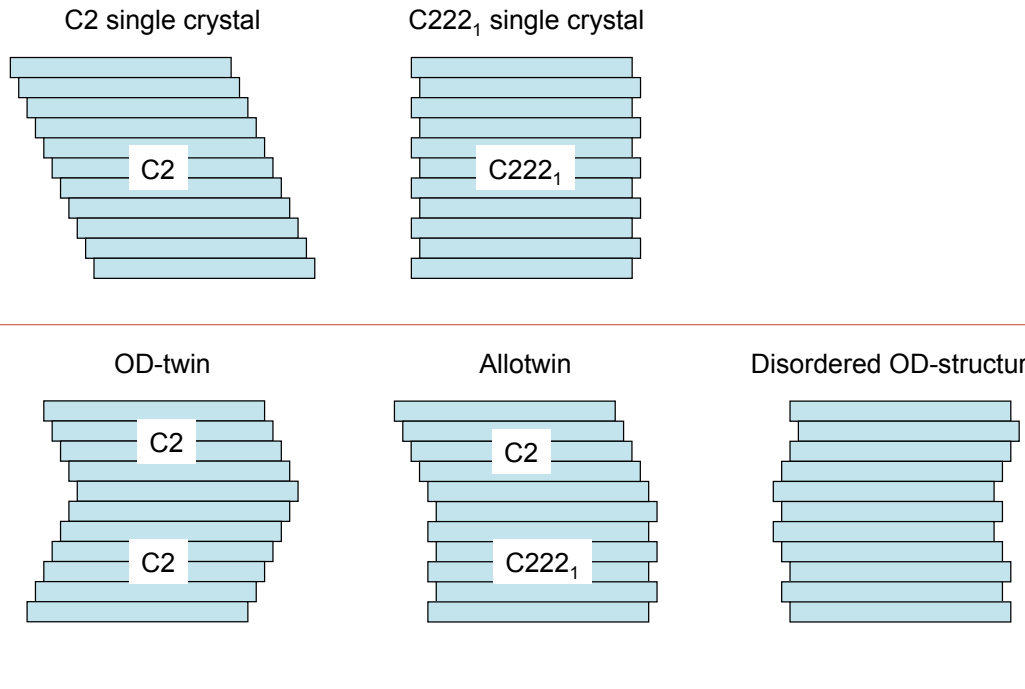
Allotwin

$S_1S_1S_1S_1S_2S_1S_2S_1S_2S_1S_2$... Two individuals have different symmetries.

Disordered OD-structure

$S_1S_1S_2S_1S_1S_2S_2S_2S_1$... Irregular sequence of stacking vectors.

# Long-range order in OD-structures



# Classification: OD-structures *vs.* twins

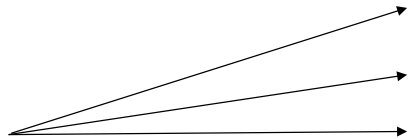
OD-structures:

Single crystals

allotwin

OD-twin

(partially) disordered OD-structure



Twinning:

by merohedry

by pseudomerohedry

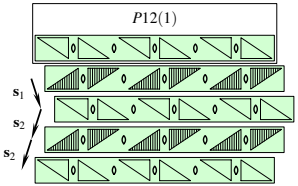
by reticular (pseudo)merohedry

...

This is structure based classification of a specific class of structures

This is geometry based classification accounting for crystal and lattice symmetries.

# Symbols for groupoid symmetry




$P \quad 1 \quad 2 \quad (1)$

$\{ \quad 2_P \quad 1 \quad (2_2) \quad \}$

P is a non-integer subscript.

Special values of P correspond to space group symmetry or specialised groupoid symmetry

The following types are possible
(I) two surfaces of a single layer are identical;
(II) two surfaces of a single layer are different and contacts are made by different surfaces.
(III) two surfaces of a single layer are different but contacts are made by identical surfaces.

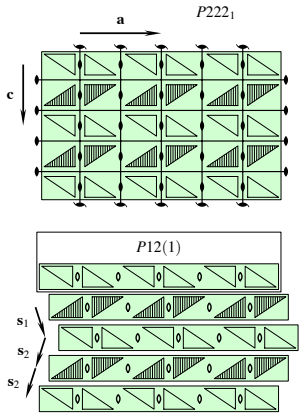
An example of symbol for groupoid of type (III):

$P \quad 1 \quad 1 \quad (4) \quad 1 \quad 1$

$\{ \quad 2_P \quad 2_Q \quad (1) \quad 2_U \quad 2_V \quad \}$

$\{ \quad 2_{P'} \quad 2_{Q'} \quad (1) \quad 2_{U'} \quad 2_{V'} \quad \}$

# Fully ordered reference structure




Fully ordered structure is uniquely determined by the structure of a single layer and relative arrangement of a pair of adjacent layers

$P \quad 1 \quad 2 \quad (1)$

$\{ \quad 2_P \quad 1 \quad (2_2) \quad \}$

Twins by pseudomerohedry can be specified by the composition point group and crystal point group. By analogy,

a groupoid symmetry can be specified by
- the space group symmetry of the reference fully ordered structure $P222_1$
- the plane space group of the OD-layer $P12(1)$

Easier to write and to grasp, but
(i) the reference s.g. is not necessarily uniquely defined
(ii) special values of subscripts cannot be indicated

# Practical implications

The crystallographic contacts in protein crystals may have substantially different strengths dependent on the shape of molecules and exposed residues. One consequence of such variation is a frequent occurrence of protein OD-structures.

From practical point of view:

- The OD-crystals are more vulnerable to twinning or disorder because there are two or more possibilities for adjacent layers to pack and form the same contacts.

- Fortunately, because of long range interactions, only few of all the OD-crystals are actually twinned or disordered.

- Fortunately, the disorder if occurred is only in one direction and can be handled.

- Any representative of the OD family gives full information on the covalent bonding and intermolecular contacts in all other members. Therefore even if the twinning or disorder occurs, these can be modeled based on the structure of a single crystal from the same OD-family.

One example in detail

OD-structures

OD-twins, allotwins

Partial disorder

Inverted stacking vectors

# An OD-twin with zero obliquity

Uppenberg *et al*. (1995).
*Biochemistry* **34**, 16838-51.

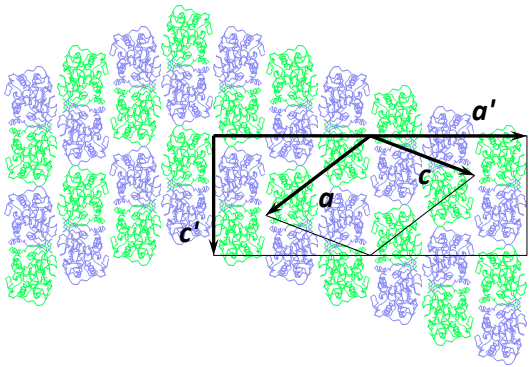
Molecule: Lipase B from *Candida antarctica*

PDB code 1lbs

Space group: C2

a = 95.9 Å, b = 95.6 Å, c = 81.8 Å
β =122.2°

OD layer: $P(2)2_12_1$
The closest
fully ordered structure: I222




- The data were processed in C2 but in the twin lattice (twin index = 3)

  a'=229.5 Å, c'=86.8Å, β =90°

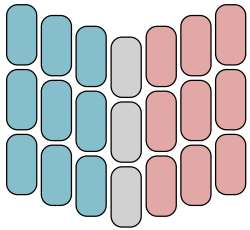
- non-overlapping reflections from the minor twin component were removed

- overlapping reflections were detwinned

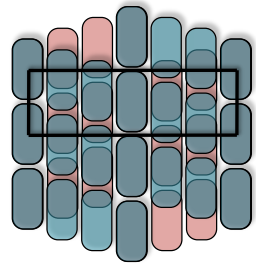
# An OD-twin with zero obliquity

Uppenberg *et al*. (1995).
*Biochemistry* **34**, 16838-51.

This packing could be assumed by similarity with the previous example



This packing is more likely to occur as it explains the exactly orthorhombic twin lattice

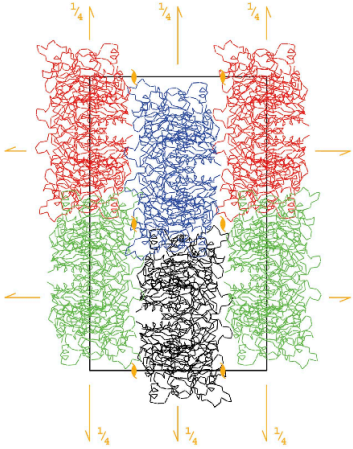


The previous example: twin index 10 obliquity 0.1°
This example: twin index 3 obliquity 0°

In general, protein OD-twins frequently have zero obliquity (twins by metric merohedry)

# Example of allotwin

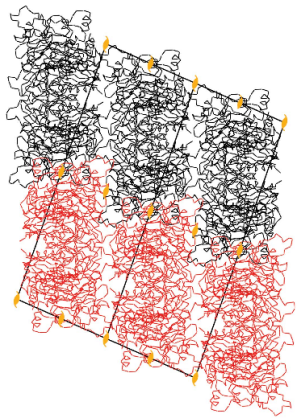
PDB code 1z0t



$P2_12_12_1$

a = 86.3 Å
b = 90.6 Å
c = 148.0 Å

PDB code 1z0v



$P2_1$

a = 48.5 Å
b = 86.3 Å
c = 138.0 Å
β = 92.3°

Dauter *et al.* (2005).
*Acta Cryst.* D**61**, 967-975.

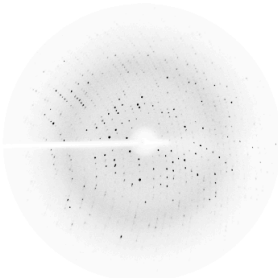
Crystal of Lon protease
Resolution 3Å

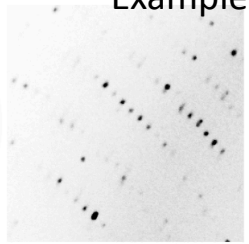
Detailed analysis of the case

Clear demonstration
of the presence of two lattices
with different symmetries

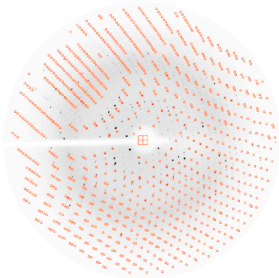
OD layer: $P2_12_1(2)$
The closest
fully ordered structure: $P2_12_12$

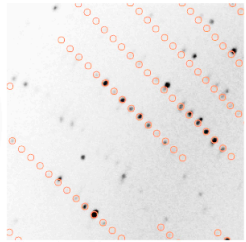
# Example of allotwin



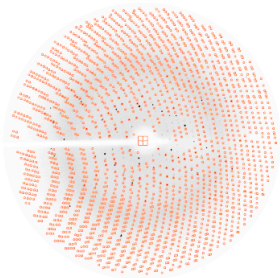


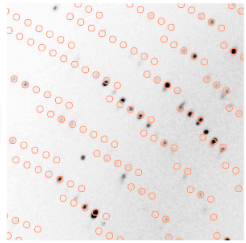
Dauter *et al.* (2005).
*Acta Cryst.* D**61**, 967-975.





$P2_1$ R / R-free = 0.21 / 0.31





$P2_12_12_1$ R / R-free = 0.19 / 0.35

# An example from the PDB

Muraki et al. (2002).
*Biochim. Biophys. Acta* 1569, 10-20.

Complexes of wheat-germ agglutinin

PDB codes: 1k7u, 1k7v

Resolution: 2.2 Å

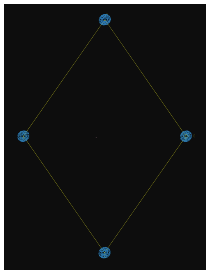
Space group: $P2_1$

$a = c = 44.9$ Å, $b = 91.8$ Å, $\beta = 110.3°$
Twinning by pseudomerohedry

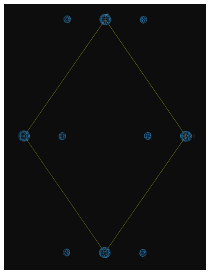
OD layer: C12(1)
The closest
fully ordered structure: $C222_1$

Patterson maps at v=0

1k7u



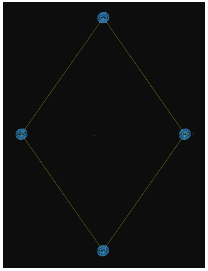
1k7v



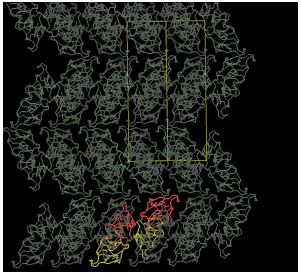
# An example from the PDB

Patterson maps at w=0

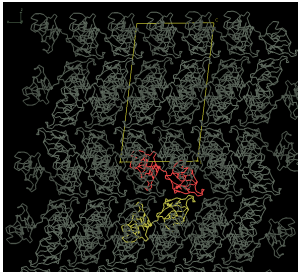
1k7u



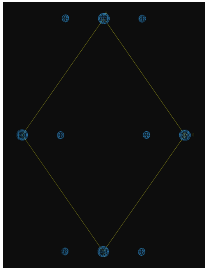
$P2_1$ structure (1k7u, 1k7v)



Putative C2 structure



1k7v



Interpretation of the Patterson map for 1k7v: allotwin

- $P2_1$ (orientation 1) lattice 1
- $P2_1$ (orientation 2) lattice 1

} Data processed in $P2_1$

- C2 (orientation 1) lattice 2 (peaks in the Patterson map)
- C2 (orientation 2) lattice 3 (peaks in the Patterson map)

# An example from the PDB

Another possible explanation of non-origin Patterson peaks is a partial disorder:

- both 1k7u and 1k7v represent twinned crystals with $P2_1$ domains
- but with different sizes of domains

Significant total interference term between small domains in 1k7v would generate non-origin Patterson peaks via modulations of intensities.

Small domains means their sizes < coherence radius of X-rays.
(In general, the difference between twin and partial disorder is the size of individual crystals.)

In structural terms the two interpretations are very similar
because in the second interpretation C2 domains are still there
but are represented by three adjacent layers at the domain interfaces.

1k7u: R = 24.2 R-free = 30.6
1k7v: R = 23.2 R-free = 31.0

} Are not informative, analysis of X-ray images is required.

One example in detail

OD-structures

OD-twins, allotwins

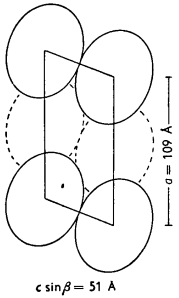
Partial disorder

Inverted stacking vectors

# Statistically orthorhombic crystals

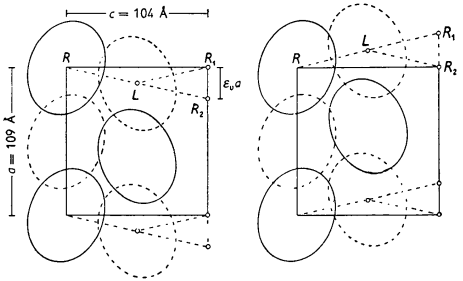
Bragg and Howells (1954).
*Acta Cryst*. **7**, 409-411.

Horse methaemoglobin
Monoclinic crystal




The unit cell parameters of this crystal helped build the model of disorder for the orthorhombic form

Imidazole methaemoglobin
"Statistically orthorhombic" crystal



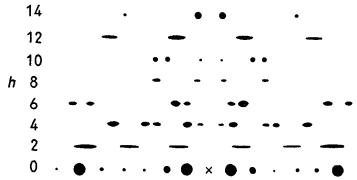

OD layer: C12(1)
Fully ordered structure: $C222_1$

(can be modeled using the monoclinic form, PDB code 2mhb)

# Statistically orthorhombic crystals

The strength and diffuseness of the (h0l) reflections.

Bragg and Howells (1954). *Acta Cryst*. **7**, 409-411.

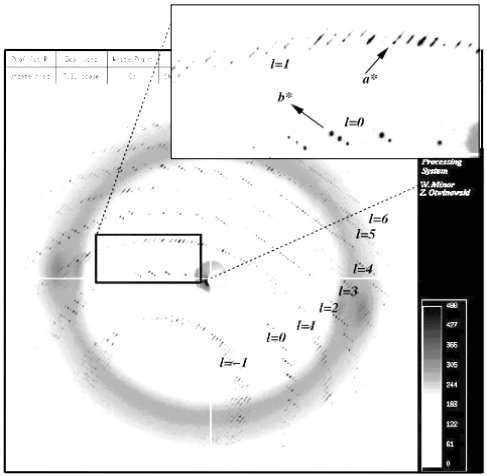



- Diffraction from a disordered OD crystal has the symmetry of corresponding fully ordered structure

- Reflection are diffuse and elongated in the direction of missing translation symmetry; the diffuseness varies in the plane parallel to the OD layers.

- Diffraction from a disordered OD-crystal can be modeled (Cochran & Howells (1954). *Acta Cryst*. 7, 412-415.)
This requires a statistical model for the interference terms between the OD-layers.

# Partially disordered OD-structure

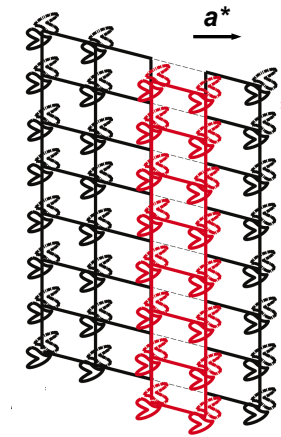



The translation symmetry is perturbed in the direction ***a****.

The diffraction pattern is characterised by the presence of the diffuse streaks along ***a**** for odd *l*.

Wang *et al.* (2005). *Acta Cryst.* D**61**, 67-74.

- Crystals of DNA polymerase from phage ϕ29
- Resolution 2.2Å
- Refinement against corrected data: R=0.28




Proposed model

# Partially disordered OD-structure

Wang *et al.* (2005). *Acta Cryst.* D**61**, 67-74.

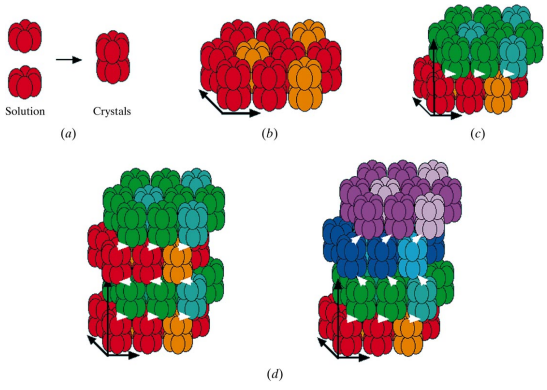
- 300 crystals tested. Observed:
  - partially disordered monoclinic form (was there the second lattice?)
  - statistically orthorhombic form as in the case by Bragg and Howells

- Demodulation involving interference term between the OD-layers

- The structure was solved using experimental phasing (combination of MIR/MAD) and demodulated data

# Partial disorder with several stacking vectors

Trame, C. B. & McKay, D. B. (2001).
*ActaCryst. D***57**, 1079–1090.




model of $P222_1$ single crystal

model of disordered crystal

Heat-shock locus U protein from *Haemophilus influenzae* and its complexes

Several crystal forms,
all partially disordered OD
belonging to different OD-families.

Data:

| | |
|---|---|
| Resolution | 2.3Å |
| Processed in | P622 |

a = 110.6, c = 335.8

The native crystal is shown

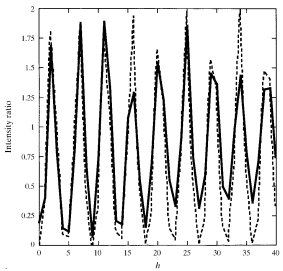
| | |
|---|---|
| OD layer: | P(6)22 |
| Fully ordered structure: | P622 |

# Partial disorder with several stacking vectors

Trame, C. B. & McKay, D. B. (2001).
*ActaCryst. D***57**, 1079–1090.

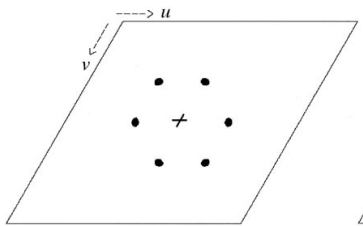
Interference term can be excluded by demodulation of data to get data corresponding to P622 crystal with one dodecamer per unit cell

mean *l*-odd to mean *l*-even intensity ratio
solid: experimental data
dashed: interference term

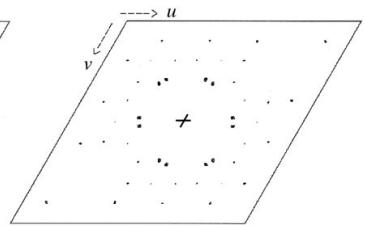



Patterson map, w=0.5

native data




demodulated data



One example in detail

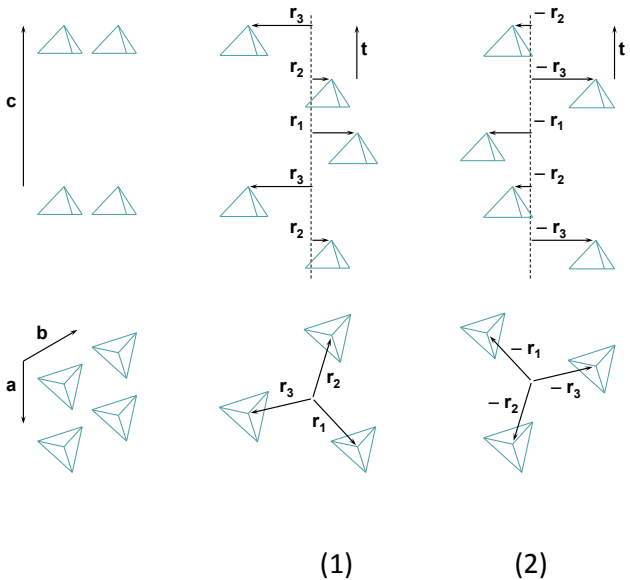
OD-structures

OD-twins, allotwins

Partial disorder

Inverted stacking vectors

# Enantiomorphic stacking vectors




(1) (2)

Structures (1) and (2)

- belong to different space groups:

  (1) $P3_1$ (2) $P3_2$

- are not necessarily related by inversion
- but have the same structure amplitudes:

  $F(1) = F(2)$

- and belong to the same OD family

# Enantiomorphic stacking vectors

Gulbis et al. (1996). Structure of the C-terminal region of p21WAF1/CIP1 complexed with human PCNA. *Cell* **87**, 297–306.

PDB code 1axc

Space group: $P3_221$
a = 83.5 Å, c = 233.9Å

OD layer: P(3)21
Fully ordered structures: R32

| Structure: | from PDB | generated |
|---|---|---|
| Spacegroup: | $P3_221$ | $P3_121$ |
| R (%): | 22.09 | 22.35 |
| R-free (%): | 29.15 | 30.02 |

Asymmetry of OD layer is within 0.2Å, but it helps choosing the right space group

# Conclusions

Presented examples of OD-structures show

- that the X-ray data from OD-twins and partially disordered structures can be corrected by demodulation or directly used for refinement and even for structure solution

- that in the required geometric information can be derived from the structure of an individual crystal or from the lattice parameters and Patterson map.

An automated refinement of partially disordered structures would be useful.
Twin refinement can be adapted for this purpose.
"Extended" twin refinement needs to account for
- interference term between reflections related by "twin" operation,
- incomplete integration of the diffuse spots,
- non-zero obliquity

# Acknowledgements

All the authors of cited papers

Zbigniew Dauter
Michail Isupov
Garib Murshudov

BBSRC

# OD-twins: twinning tests

Twin operation in the OD twin belongs to the point group of the OD-layer.
Consequently, the twin related complex structure factors correlate.

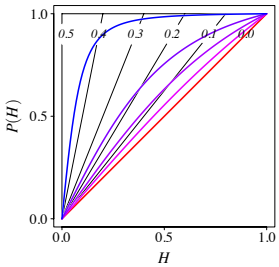
The model of random atoms is no longer valid
and reference curves for twinning tests need to be adjusted.

Partial twinning test (H-test),
theoretical curves:

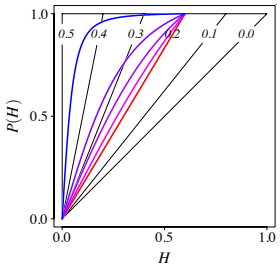
Red: uncorrelated structure factors

Blue: correlation = 0.99

α=0.0




α=0.2




In the perfect twinning test,
the apparent twinning fraction is smaller if the structure factors correlate

# Monoclinic OD-twin

Au et al. (2006).
*Acta Cryst*. D**62**, 1267-1275.

Molecule: Ferrochelatase-1 (HemH)
from *B. anthracis*

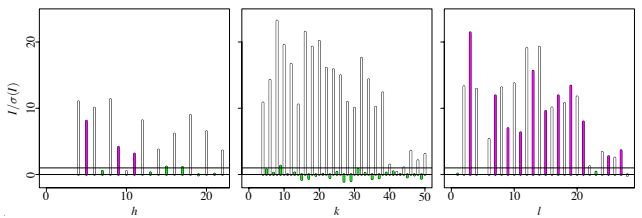
PDB code 2c8j

Space group: $P2_1$
Resolution 2.2Å

a = 49.9, b = 109.9, c = 59.4 Å
α = β = γ = 90°

OD layer: P2(1)1
The closest
fully ordered structure: $P2_12_12$

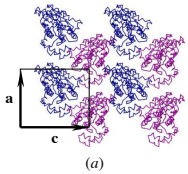
Only one reflection with h = 2n, k=l=0 has I/sig(I) > 3

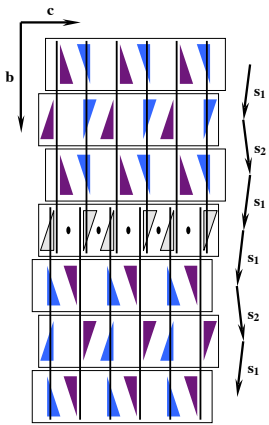


# Monoclinic OD-twin

$P2_1$ true structure

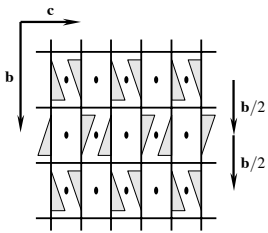
The lattice is perfectly orthorhombic (twin by metric merohedry)




$P2_12_12$ false structure

molecules shifted along **c** by 2.5Å




Twinning was suspected only after several unsuccessful attempts at solving structure in an orthorhombic space group

# Conserved one-dimensional substructure

Another important special case is the crystal consists of one-dimensional substructures with strong internal and weak external contacts.

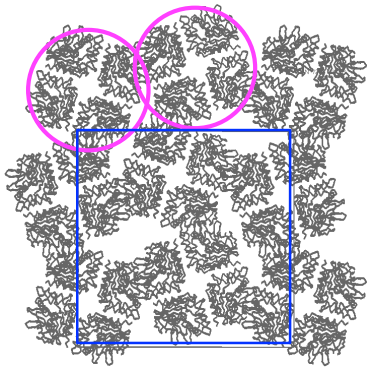
There is an extra degree of freedom compared to OD-structures meaning that

- these structures are more likely to be globally disordered
  (hence less frequently used for structure determination)

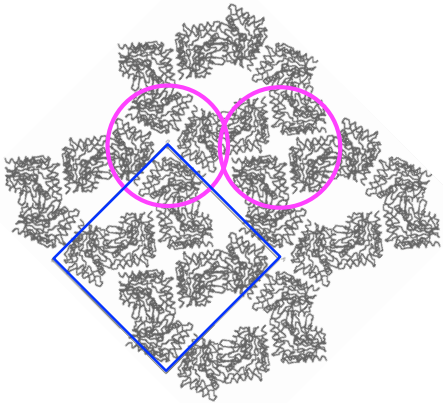
- various crystal forms and phase transitions are possible

# Conserved one-dimensional substructure

Roberto Steiner, Kings college, University of London



crystal soaking



twinned orthorhombic crystal

twinned tetragonal crystal