Twinning and other pathologies

Andrey Lebedev

CCP4
OD-structures

Twinning by (pseudo)merohedry

Statistics of one observation

Statistics of two observations

Twinning tests summary

Space group validation
- identical layers
- identical interfaces between the layers
- but: two or more ways of packing three adjacent layers

*) MX: "identical" means Ca r.m.s.d. < 1 Å

*) \( S_1 \) and \( S_2 \) are called stacking vectors

- two-dimensional periodicity
- a potential for disorder in the third dimension
Example 1: OD-twin

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

Some reflections from two lattices overlap.

L-2-haloacid dehalogenase from *Sulfolobus tokodaii*

Real and reciprocal lattices

Maximum overlap is not at integer $h$.

Twinning by reticular merohedry with twin index 10 and obliquity 0.1°

Integration of a single lattice: in effect, twinning coefficient depends on $h$
Intensities of the overlapping reflections

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.
Demodulation

Original data: \( R / R\text{-free} = 0.21 / 0.27 \)

Corrected data: \( R / R\text{-free} = 0.16 / 0.23 \)

Modulation function

\[ q'(h) = p_0 + p_1 \cos(2\pi h) + p_2 \cos(4\pi h) + \ldots \]
Indexed lattice

The second lattice

Non-origin peaks in the Patterson map:
• contribution from the second lattice
• because of the overlapping spots
Improvement in the electron density

Visually, 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 $1.5\sigma$ and 1-1 at $3\sigma$) around the pyruvate molecule before and after demodulation
OD-structures

Single crystal

OD-twin

Examples 1 & 2

Single crystal

Allotwin

Example 3

Partially disordered OD-structure

Examples 4 & 5
Classification: OD-structures vs. twins

**OD-structures:**

- Single crystals
- allotwin
- OD-twin
- (partially) disordered OD-structure

This is structure based classification of a specific class of structures

**Twinning:**

- by (pseudo)merohedry
- by reticular (pseudo)merohedry = non-merohedral
- ...

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

\[
P \quad 1 \quad 2 \quad (1) \\
\{ \quad 2_p \quad 1 \quad (2_2) \quad \}
\]

In $2_p$, $P$ is a non-integer subscript.

Special values of $P$ correspond to space group symmetry or specialised groupoid symmetry.
Example 2: OD-twin with zero obliquity

Space group: \( C2 \)
\[ a = 95.9 \, \text{Å}, \quad b = 95.6 \, \text{Å}, \quad c = 81.8 \, \text{Å} \]
\( \beta = 122.2^\circ \)

OD layer: \( P(2)2_12_1 \)

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

\[ a' = 229.5 \, \text{Å}, \quad c' = 86.8 \, \text{Å}, \quad \beta = 90^\circ \]

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


Molecule: Lipase B from *Candida antarctica*

PDB code 1lbs
Example 2: OD-twin with zero obliquity

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°
The previous example: twin index 3 obliquity 0°

In general, protein OD-twins frequently have zero obliquity (twins by metric merohedry)
Example 3: allotwin

Crystals of Lon protease Resolution 3Å

Dauter et al. (2005).

P2₁

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

P2₁2₁2₁

a = 86.3 Å  
b = 90.6 Å  
c = 148.0 Å
Example 3: allotwin

Crystals of Lon protease
Resolution 3Å

Dauter et al. (2005).
*Acta Cryst.* D61, 967-975.

Structures of both crystal forms were solved

<table>
<thead>
<tr>
<th>PDB code 1z0t</th>
<th>PDB code 1z0v</th>
</tr>
</thead>
<tbody>
<tr>
<td>R / R-free</td>
<td>0.19 / 0.35</td>
</tr>
<tr>
<td></td>
<td>0.21 / 0.31</td>
</tr>
</tbody>
</table>
Crystal disorder

Twinning, partial disorder: Missing global periodicity

- Single crystal: (Single ordered domain)
- Twinned crystal: (Two or more ordered domains)
- Partially disordered crystal: (Many ordered domains)

size of ordered domains

Coherence length of X-rays
Example 4: partially disordered OD-structure

Crystals of Phi29 DNA polymerase
Resolution 2.2Å

The translation symmetry is not global in the direction $a^*$.

The diffraction pattern is characterized by the presence of the diffuse streaks along $a^*$.

The structure was solved using demodulated data and experimental phasing

Refinement against corrected data: $R=0.28$
Example 5: Partial disorder with several stacking vectors


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

OD layer: P(6)22

- Model is the model of largest ordered domains
- Space group is the space group of the model
Example 6: Enantiomorphomorphic stacking vectors

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

PDB code 1axc

**Structure:**
- **Space group:** P3$_{21}$
- **a = 83.5 Å, c = 233.9Å**

**OD layer:**
- **Space group:** P(3)$_{21}$

**Structure:**
- from PDB
- generated

**Space group:**
- P3$_{21}$
- 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
OD-structures

Twinning by (pseudo)merohedry

Statistics of one observation

Statistics of two observations

Twinning tests summary

Space group validation
Important special case: twinning by (pseudo)merohedry

- All spots overlap with related spots from another individual crystal
- Detection requires analysis of intensity statistics
- More significant effect on model if ignored
- Point group and, consequently, space group determination may be a problem
Twinning by (pseudo)merohedry

\[
\begin{align*}
\text{lattice in orientation 1} & = \text{lattice in orientation 2} \\
\text{crystal in orientation 1} & \neq \text{crystal in orientation 2} \\
\text{crystal in orientation 1} & + \text{crystal in orientation 2} = \text{Twinned crystal}
\end{align*}
\]

Example: P2, \( \beta = 90 \)
## Twinning by (pseudo)merohedry

<table>
<thead>
<tr>
<th>lattice in orientation 1</th>
<th>=</th>
<th>lattice in orientation 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>crystal in orientation 1</td>
<td>≠</td>
<td>crystal in orientation 2</td>
</tr>
<tr>
<td>crystal in orientation 1</td>
<td>+</td>
<td>crystal in orientation 2</td>
</tr>
</tbody>
</table>

### P121, \( \beta = 90 \)

| Intensities from individual crystal 1 | Intensities from individual crystal 2 |

**Perfect twin**

- (individual crystal of **equal** size)

**Partial twin**

- (individual crystal of **different** size)

Some weak reflections vanish

- Twin two-fold axis
- Crystallographic two-fold axis
Twinning by (pseudo)merohedry

- higher lattice symmetry is determined by crystal symmetry

Twinning by pseudomeroedry
- specialised unit cell parameters
• Crystal content analysis *may* help recognising twinned crystal
• Deterministic case: structure cannot be "solved" in wrong space group

Not always that easy
Monoclinic OD-twin (twin by pseudomerohedry)

Au et al. (2006).
*Acta Cryst.* D62, 1267-1275.

Ferrochelatase-1 from *B. anthracis*

PDB code 2c8j

Space group: \( P2_1 \)

Resolution 2.2\(\AA \)

\( a = 49.9, \ b = 109.9, \ c = 59.4 \ \text{Å} \)

\( \alpha = \beta = \gamma = 90^\circ \)

OD layer: \( P2(1)1 \)

The only reflection with \( h = 2n, \ k=l=0 \) and \( l > 3 \text{ sig}(l) \)
Monoclinic OD-twin (twin by pseudomomerohedry)

P2\textsubscript{1}2\textsubscript{1}2 symmetrised structure

Molecules shifted along \textbf{c} by 2.5Å
R-free = 40%

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

P2\textsubscript{1} true structure

The lattice is exactly orthorhombic
R-free = 27%
Examples of crystal pathologies

Twinning by (pseudo)merohedry

Statistics of one observation

Statistics of two observations

Twinning tests summary

Space group validation
Resolution bins (resolution shells)

\[ s < s(h,k,l) < s + ds \]

\[ \langle I(s) \rangle = \text{mean}( I(h,k,l) ) \]

\[ \langle I(s) \rangle \approx C \cdot \exp(-2B \cdot s^2) \]

B - Overall temperature factor

Wilson Plot:

\[ \log(\langle I(s) \rangle) \text{ against } s \]
\[
Z(hkl) = \frac{I(hkl)}{\langle I(s) \rangle} \quad \langle Z(s) \rangle = 1 \quad \text{(by definition of } Z) \]

<table>
<thead>
<tr>
<th>(Z(hkl))</th>
<th>(Z^2(hkl))</th>
<th>(\langle Z(s) \rangle)</th>
<th>(\langle Z^2(s) \rangle)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 1 1 1</td>
<td>1 1 1 1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0 0 2 2</td>
<td>0 0 4 4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0 0 0 4</td>
<td>0 0 0 16</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

many weak reflections

\[\langle Z^2(s) \rangle\]

a few weak reflections

08/11/2014
KEK-CCP4 Workshop
$$\{ Z(hkl) \} \rightarrow \text{sort} \rightarrow \{ Z_i \}$$

$$Z_1 < Z_2 < Z_3 < \ldots < Z_N$$

**Normalised intensity: Cumulative distribution**

- a few weak reflections
- many weak reflections
Theoretical distribution of intensities

Single crystal

Partial twin

Partial twin

Perfect twin

\[ P(Z) \]

\[ P(Z) \]

\[ P(Z) \]

\[ P(Z) \]

\[ \langle Z^2 \rangle \]

\[ \langle Z^2 \rangle \]

\[ \langle Z^2 \rangle \]

\[ \langle Z^2 \rangle \]

(works for incomplete data set)

Acentric reflections

Centric reflections
Perfect twinning tests: proper behaviour

PDB entry 1i1j
single crystal

C-terminal domain of gp2 protein from phage SPP1
perfect twin
Perfect twinning tests: resolution cutoff

PDB code 1l2h
Partial twin
Perfect twinning tests: resolution cutoff

human deoxycytidine kinase single crystal
### Pseudosymmetry

**twin axis**

Twinning + Generic NCS

Twinning axis || NCS axis

Single crystal
Theoretical distribution of intensities

Single crystal

Partial twin

Partial twin

Perfect twin

\[ P(Z) \]

\[ <Z^2> \]

\( <Z> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)

\( <Z^2> \)

\( <Z> \)

\( <Z> \)

\( <Z^2> \)
Theoretical distribution of intensities

Single crystal

Partial twin

OR

Perfect twin + pseudosymmetry axis || twin axis

Perfect twin

Acentric reflections
Centric reflections
Figure 3.12. OD-twin of HemH.

(a, b) An individual P1211 crystal of the twin: (a) as single OD-layer and (b) three adjacent OD-layers shown as Cα traces with symmetry related molecules in the same colour.

(c) Schematic view of a single OD-layer with P2111 planes pseudo symmetry. The pseudo-symmetry axes of the OD-layer are shown by dashed black lines.

(d) Schematic view of the OD-twin with two individual crystals at the top and bottom with the interface OD-layer in the middle. Crystallographic axes are shown by black lines, the pseudo-symmetry axes of the interface OD-layer are shown by black ovals and the molecules related by these symmetry elements are shown in the same colour. The stacking vectors s1 and s2 relating the origins of the adjacent layers are shown by black arrows at the right margin.

(e) The symmetrised P21212 structure that would occur if s1 and s2 in (d) were equal to b/2.

Monoclinic OD-twin (twin by pseudomerohedry)

P21212 symmetrised structure

Molecules shifted along c by 2.5Å

R-free = 40%

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

P21 true structure

The lattice is exactly orthorhombic

R-free = 27%
OD-structures

Twinning by (pseudo)merohedry

Statistics of one observation

Statistics of two observations

Twinning tests summary

Space group validation
H-test and L-test

\[ H = \frac{|J_1 - J_2|}{|J_1 + J_2|} \]

\[ L = \frac{|J_1 - J_2|}{|J_1 + J_2|} \]

Originally, L-test was designed to deal with pseudotranslation generating sublattices with strong and weak reflections alternating.
H-test and L-test

\[ H = \frac{|J_1 - J_2|}{(J_1 + J_2)} \]

\[ L = \frac{|J_1 - J_2|}{(J_1 + J_2)} \]

Originally, L-test was designed to deal with pseudotranslation generating sublattices with strong and weak reflections alternating.
Theoretical distribution of H

What is the tested operation?

- not a twin operation
- twin operation (partial twinning)
- twin operation (perfect twinning)

or operation of the crystal point group
Distribution of H can be perturbed by NCS and weak observations

Blue:
ideal distribution for partial twin

Green:
blue + effect of NCS axis || twin axis

Red:
green + effect of intensities with small I/ \text{sig}(I)
Examples of experimental $P(H)$

An almost ideal case

+ effect of NCS axis || twin axis

+ effect of intensities with small I/ sig(I)
H-test and L-test

H = |J1 – J2| / (J1 + J2)

L = |J1 – J2| / (J1 + J2)

Originally, L-test was designed to deal with pseudotranslation generating sublattices with strong and weak reflections alternating.
Theoretical distribution of $L$

Single crystal

Partial twin

Perfect twin

\[
P(L)
\]
Distribution of L can be strongly perturbed by weak observations

Cell: 64.2 109.2 100.2 90 93.8 90
Space group: P2₁
No pseudo symmetry
Pseudomeroehedral twinning is impossible

All data: as if a perfect twin
Data below 3A: untwinned

Nevertheless the L-test is very useful when performed with right resolution range (or with several ranges)
Pseudotranslation

Crystallographic translation

Pseudotranslation

Translation \( C \)

Pseudotranslation \( C \)

Translation \( C' = 2C \)

Two times larger reciprocal lattice spacing

- Weak reflections appear (planes \( 2L+1 \))
Statistics of **one** intensity are strongly affected by pseudotranslation

1jfk: Pseudotranslation results in alteration of strong and weak reflections

> Acentric moments of E for k=1,3,4
> 4th moments of E ...
L-test and H-test are not affected by pseudotranslation

> L test for twinning
> cumulative distribution function for |L|

> H test for twinning (operator ...)
> cumulative distribution function for |H|
OD-structures

Twinning by (pseudo)merohedry

Statistics of one observation

Statistics of two observations

Twinning tests summary

Space group validation
### Why so many tests?

<table>
<thead>
<tr>
<th>Statistics of one observation</th>
<th>Statistics of two observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(Z)</td>
<td>H-test</td>
</tr>
<tr>
<td>Specific for a given resolution shell</td>
<td>No</td>
</tr>
<tr>
<td>Specific for a given twin operation</td>
<td>No</td>
</tr>
<tr>
<td>distinguishes perfect twinning vs. crystal symmetry</td>
<td>+</td>
</tr>
<tr>
<td>Insensitive to pseudotranslation</td>
<td>–</td>
</tr>
<tr>
<td><em>Insensitive to anisotropy</em></td>
<td>–</td>
</tr>
<tr>
<td>Insensitive to weak reflections at high resolution</td>
<td>–</td>
</tr>
</tbody>
</table>
Symmetry, Scale, Merge (Aimless)

Analysis with ctruncate
Run of ctruncate

L test for twinning

Acentric moments of I

Cumulative intensity distribution
OD-structures

Twinning by (pseudo)merohedry

Statistics of one observation

Statistics of two observations

Twinning tests summary

Space group validation
Are these tests always sufficient?

How to handle the cases with strong pseudosymmetry?

Validation of crystallographic symmetry:
- refinement in space groups compatible with
  - unit cell
  - current model (considered as at least approximately correct)
Collaborative Computational Project No. 4
Software for Macromolecular X-Ray Crystallography

Welcome to CCP4 online

Login

Other Options - Register, Forgotten Password, Change Password

Runnable programs
The following programs are available:

<table>
<thead>
<tr>
<th>Program</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balbes</td>
<td>An automated Molecular Replacement (MR) pipeline - Balbes integrates into one system all the components necessary for solving a crystal structure by Molecular Replacement</td>
</tr>
<tr>
<td>MrBUMP</td>
<td>An automated Molecular Replacement (MR) pipeline - Given a target sequence and experimental structure factors, it will search for homologous structures, create a set of suitable search models from the template structures, do molecular replacement, and test the solutions with some rounds of restrained refinement.</td>
</tr>
<tr>
<td>Zenuda</td>
<td>Space group and crystallographic origin validation</td>
</tr>
<tr>
<td>jsPISA</td>
<td>Calculation and analysis of macromolecular surfaces and interfaces</td>
</tr>
</tbody>
</table>

Zanuda
CCP4I > Validation & Deposition > Validate space group

• Modes
  » REFINE all transformed models and save the best model
  » SAVE all transformed models and data without refinement

• Option
  » SYMMETRYSE input model before further refinements
Pseudotranslation: what else can go wrong?

Cell and H-M symbol are the same.

Crystallographic and pseudosymmetry axes are confused.

Molecular Replacement:
- Two structures are globally very similar (e.g. rmsd = 0.5Å)
- MR can in some cases pick up a wrong solution
An example of symmetry correction

PDB code: 1yup

space group (PDB): P1  8 molecules per a.u.

space group (true): P2₁  4 molecules per a.u.

Pseudo-symmetry space group: C2  2 molecules per a.u.
Monoclinic structures related to 1yup

Positions of molecules

Crystallographic axes
Pseudosymmetry axes

Space group and its relation to the structure 1yup

C2  Pseudo-symmetry space group

P2  False space group

P2_{1}  True space group
Structure solution and symmetry validation

Data processing (2/m)

Molecular replacement (P2)

Refinement (P2)
R-free ≈ 0.37

Data processing (-1)

Molecular replacement (P1)

Refinement (P1)
R / R-free = 0.24 / 0.31

PDB: 1yup (P1)

PDB: 1yup

Zanuda

(P2₁)
R-free = 0.33
Assumptions:

• The pseudosymmetry is very strong (r.m.s.d. from exact symmetry ≈ 1Å)
• The structures of individual molecules are almost correct
  – although they might have been refined / rebuilt in an incorrect space group

If assumptions are not satisfied, the results will likely to be wrong.
Four alternative solutions in two space groups

GAF (N-terminal) domain of CodY protein from Bacillus subtilis

True $P_{4\bar{3}}22$  True $P_{4\bar{3}}22$  False $P_{4\bar{3}}22$  False $P_{4\bar{1}}22$

MR solution

$z = 1/8$  $z = 3/8$  $z = 1/8$  $z = 3/8$

$z = 0$  $z = 1/4$  $z = 0$  $z = 1/4$
Twinned crystal with pseudo-symmetric substructure

Human macrophage receptor CLEC5A for dengue virus

- Substructure (A) is common for twin individuals
- Substructure (B) is not even approximately symmetric relative to ▲ and ▲
  - The choice of correct origin was essential for structure completion
Acknowledgements

<table>
<thead>
<tr>
<th>Eleanor Dodson</th>
<th>University of York</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roberto Steiner</td>
<td>King's College, London</td>
</tr>
<tr>
<td>Zbigniew Dauter</td>
<td>APS, Chicago</td>
</tr>
<tr>
<td>Gleb Bournkov</td>
<td>EMBL-Hamburg</td>
</tr>
<tr>
<td>Michail Isupov</td>
<td>University of Exeter</td>
</tr>
<tr>
<td>Garib Murshudov</td>
<td>MRC LMB Cambridge</td>
</tr>
</tbody>
</table>

All the authors of cited papers