Computing School 2011 Twinning Andrea Thorn









Image property of Rob Lavinsky, iRocks.com – CC-BY-SA-3.0

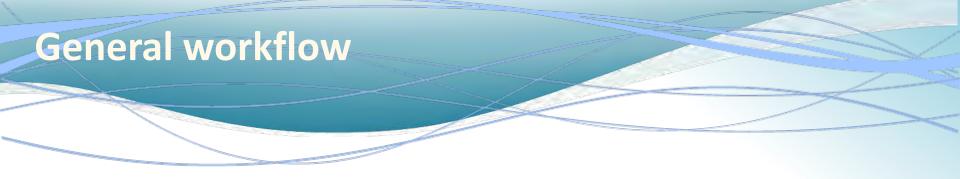


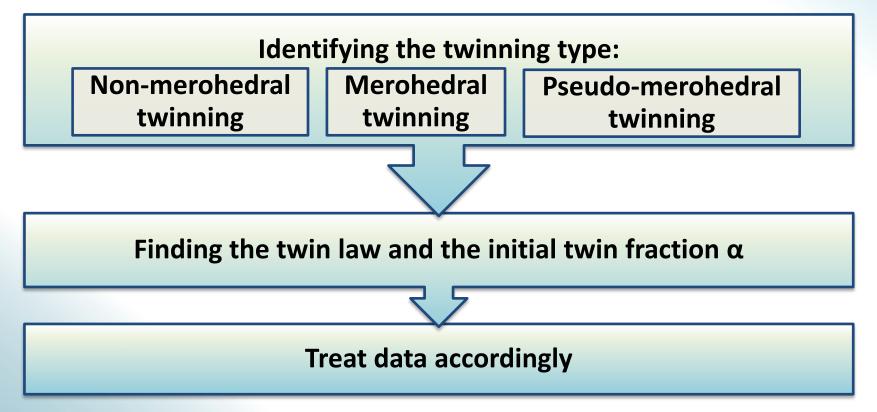
Twin: Two or more crystals of the same species are joined together in different orientation

α is the twin fraction: $I_{1+2} = αI_1 + (1-α)I_2$

The twin law (twin operator) is the operator between the cojoined crystals.



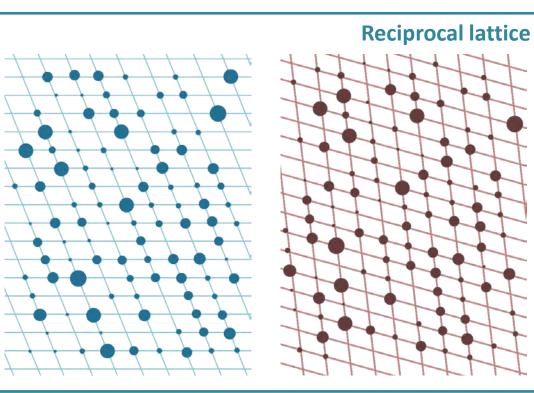


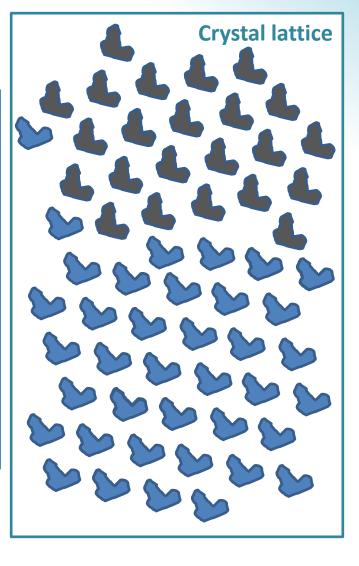




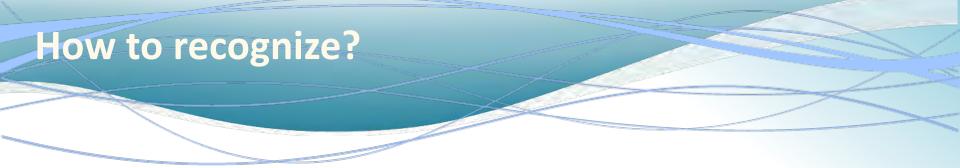
What is a non-merohedral twin?

Twin law is arbitrary.









- From detector frames
- By reciprocal lattice viewers: RLATT (*proprietary*) or RLAT4XDS (*available at <u>http://www.cb-huebschle.de</u>*)
- No suitable cell for all reflections (many outliers)
- Cell refinement difficult
- An unusual long cell axis
- Some reflections sharp, others split



Courtsey of Madhumati Sevvana

Example #1: Glucose Isomerase

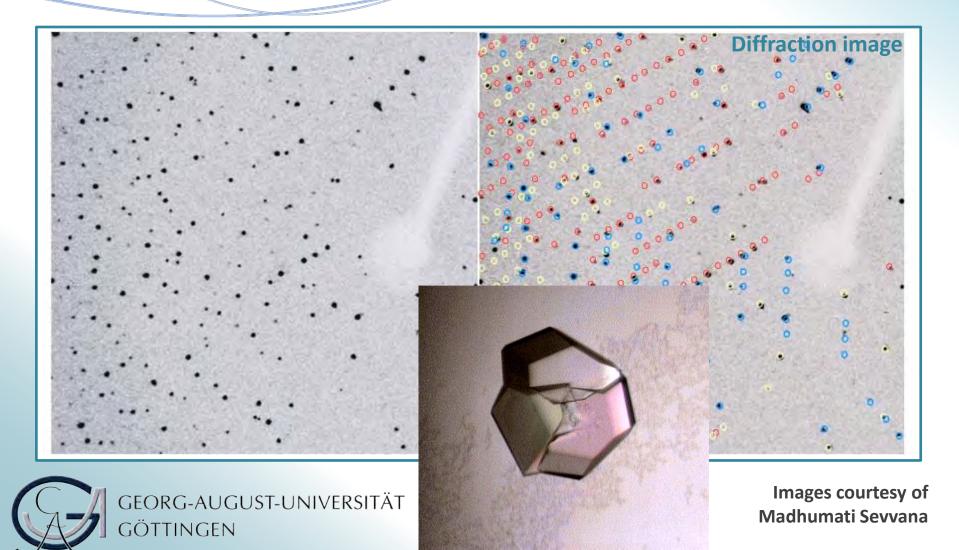


- "Drilling"
- three separate centers
- diffraction pattern consists of
 - single reflections
 - overlapping spots

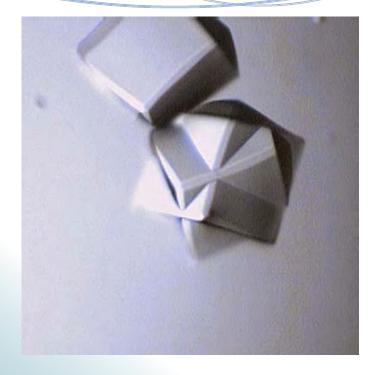


Images courtesy of Madhumati Sevvana

Example #1: Glucose Isomerase



Example #2: Bovine insulin

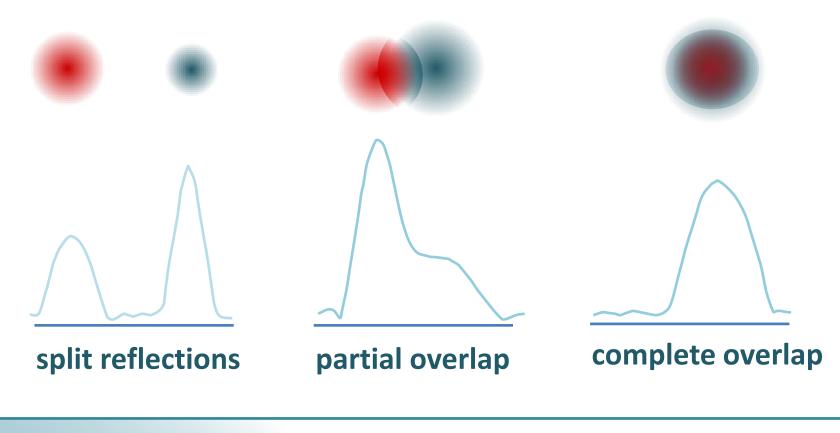


- 51 amino acids
- Resolution to 1.60 Å
- Cubic (I2₁3)
- interpenetrant twins with approximately the same center

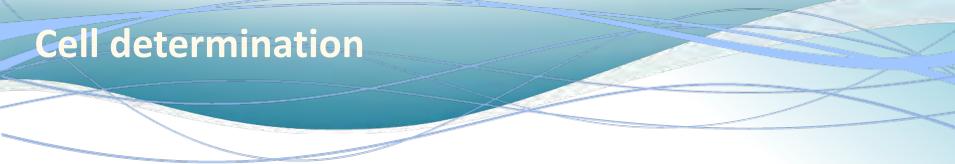


Images courtesy of Madhumati Sevvana









- For indexing, leave out partial overlaps at first.
- Find a cell that fits a reasonable fraction of spots.
- Use the not-yet-indexed reflections to find an alternative orientation of the same cell.
- If necessary, repeat this step to find more twin domain orientations.



CELL_NOW: Determining the twin law and cell

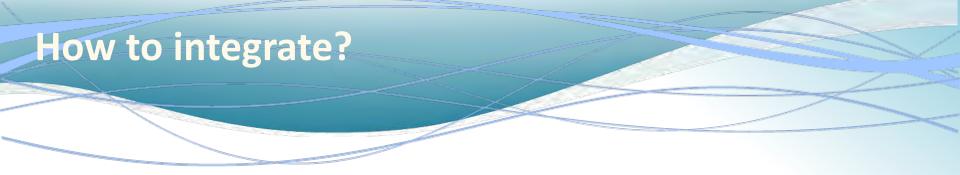
Cell for domain	1: 78.04	40 77.	986	78.024	89.99	89.94	90.01
Figure of merit:	0.560	%(0.1):	51.6	%(0.2):	55.2	%(0.3):	62.6

4072 reflections within 0.250 of an integer index assigned to domain 1 Cell for domain 2: 78.040 77.986 78.024 89.99 89.94 90.01 Figure of merit: 0.910 %(0.1): 91.4 %(0.2): 93.6 %(0.3): 94.5

Rotated from first domain by 89.2 degrees	about	reciprocal	axis
0.928 0.207 1.000 and real axis 0.927	0.208	1.000	
Twin law to convert hkl from first to	0.459	-0.625	0.631
this domain (SHELXL TWIN matrix):	0.824	0.036	-0.565
	0.330	0.780	0.532

3564 reflections within 0.250 of an integer index assigned to domain 2, 2751 of them exclusively; 184 reflections not yet assigned to a domain

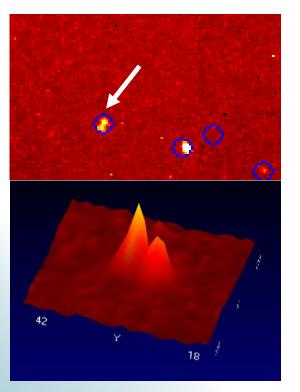




- Ignoring the twinning: No special software needed, but poor results (if any).
- Omission of all overlapped reflections: Possible in *many integration programs, but low data completeness.*
- Omission of all partially overlapping reflections
- Integrating both domains by profile fitting the overlapping reflections:



How to integrate?

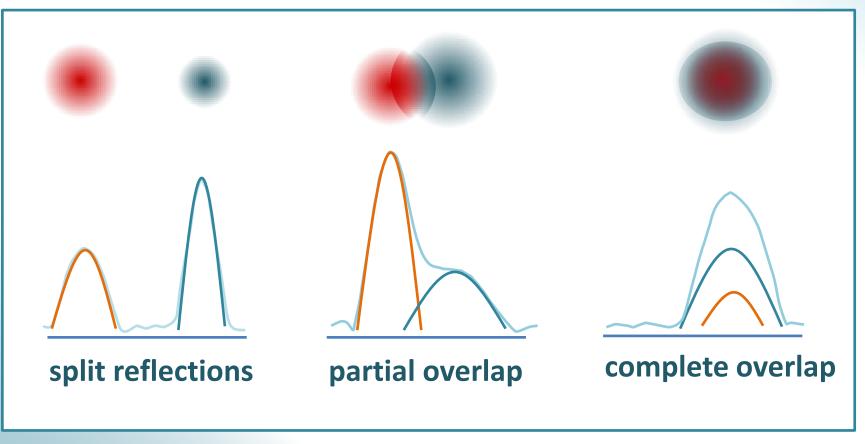


Split reflections

- The total intensity of a group of overlapping reflections can be determined precisely!
- Partitioning can be determined by 3D profile fit.

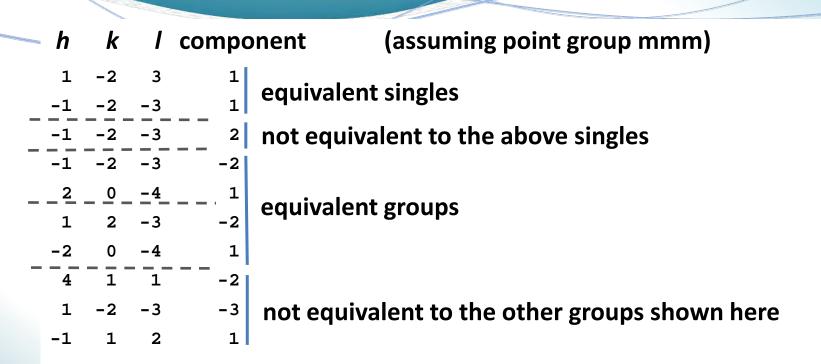








HKLF5 format definition



In HKLF5 format, a group of overlapping reflections is defined by negative component numbers for all but the last reflection. For scaling the numbers **must** match.



Approximation:

 $I_{\rm c} = I_{\rm o} \cdot S(n) \cdot P(u,v,w)$

S(n) scale factor for frame *n* (incident beam I only) *P* absorption factor (diffracted beam D only) *u,v,w* direction cosines relative to a*, b* and c*

S and **P** are refined alternately to minimize $\Sigma w (\langle I_c \rangle - I_c)^2$, where $\langle I_c \rangle$ is the mean of a group of equivalents. Scaling requires a high redundancy to work well.



The scale factors **S** correct for:

- absorption of primary beam by crystal and support
- crystal decomposition
- intensity variation of the primary beam
- changes in the volume irradiated. (twin components)
- beam inhomogeniety

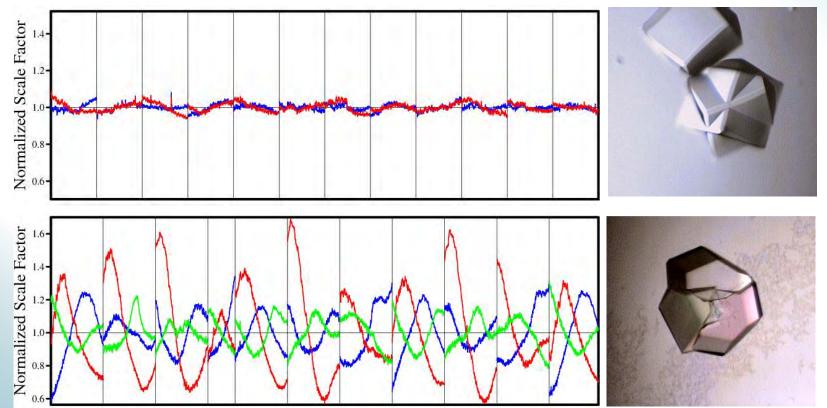
The diffracted beam absorption factor P(u,v,w) is a sum of *spherical harmonics* for the direction cosines u, vand w as suggested by Blessing (1995). Only one refinement iteration is required (linear function).



- Fitting individual intensities to the mean corrected intensities (averaged over equivalents).
- However, for non-merohedral twins the equivalents may be single reflections or groups of overlapping reflections with the same pattern of contributors.
- Outliers are only downweighted.
- After parameter establishment, outliers are deleted.



In TWINABS:





GEORG-AUGUST-UNIVERSITÄT Göttingen The smallest crystal (red) was furthest from the center.

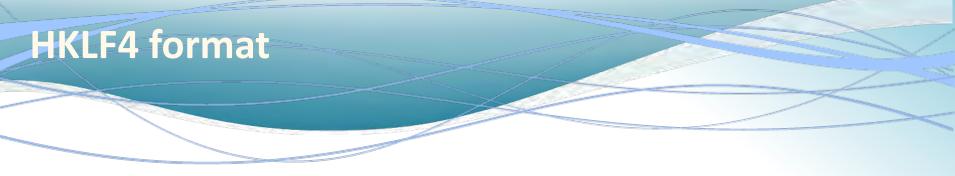


For HKLF4, the best list of unique reflections using all the available information has to be generated by **detwinning.** We solve a set of equations in the form:

$$I_{\rm m} = k_1 I_{\rm h} + k_2 I_{\rm h'} + \dots + k_{\rm n} I_{\rm h''}$$

*I*_m measured total intensities of single or composite reflections
 *k*_n relative twin masses
 *I*_h intensities of the unique reflections





- From the profile fitting of split reflections in refinement, we know the approximate partitioning of the reflections.
- These are used as **weak restraints** for the solution of the equations by least-squares.
- This algorithm is robust, converges fast and can process several million reflections in a few seconds (by using sparse matrix techniques).
- The resulting twin ratios agree well with those from the HKLF5 refinement in SHELXL.





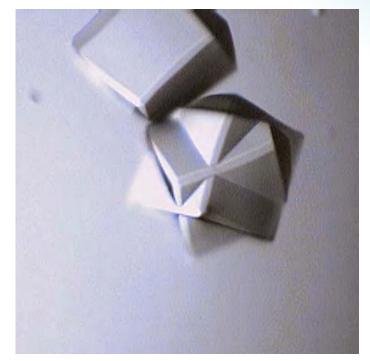
But beware:

- If Laue symmetry is lower than the metric symmetry of the lattice, the components may be indexed inconsistently!
- This is similar to the generation of a merohedral twin
- The scaling is not affected by this inconsistent indexing, but the resulting data is as problematic as a merohedral twin would be.
- Warning sign: High R_{int} for deconvolution!





- This inconsistent indexing actually happened in the case of this bovine insulin crystal.
- For both glucose isomerase and insulin HKLF4 files were created.
- SAD structure solution with SHELXC/D/E proved to be straightforward for both!

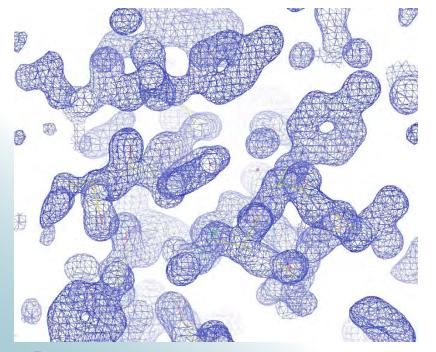




Structure solution

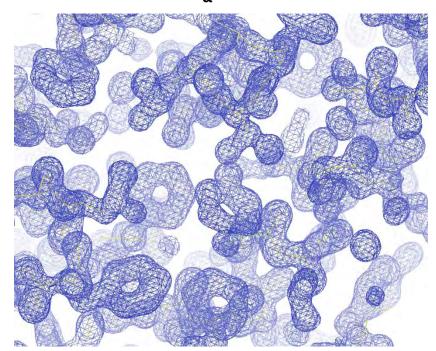
Cubic insulin

SHELXD: CC 50.5, CC_{weak} 29.1 for 6 S SHELXE: 50 out of 51 residues traced; MPE 21°, 94% of C_{α} atoms within 0.5 Å



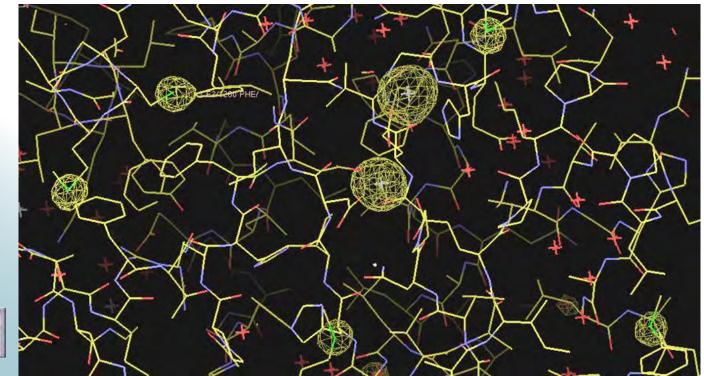


GEORG-AUGUST-UNIVERSITÄT Göttingen Glucose isomerase SHELXD: CC 31.4, CC_{weak} 20.4 for 2 Mn SHELXE: 383 out of 388 residues traced MPE 21°, 95% of C_{α} atoms within 0.5 Å



Non-merohedral twinning

Glucose isomerase was phased using the two Mn atoms (f" = 2.81), all sulfurs (f" = 0.56) are well defined. The mean anomalous densities were 68.3σ for Mn, 43.4 for Mn/Mg, 11.1 for S and less than 1.0 for all other atoms [ANODE].





- Compare the R1 value after 'merging for Fourier' at the end of the HKLF5 refinement so that the same total number of reflections are used as in HKLF4 refinement.
- Free R reflections have currently to be put in a separate file for HKLF5 refinement with SHELXL.
- It may be difficult to select a suitable free R set of reflections for HKLF5 (for HKLF4, they can be chosen in the usual way).





	Insulin	Glucose isomerase
HKLF 4:	R1 _{work} 18.9	R1 _{work} 18.3
	R1 _{free} 23.3	R1 _{free} 22.1
HKLF 5:	R1 _{work} 14.6	R1 _{work} 13.0
	R1 _{free} 22.3	R1 _{free} 22.8

- Free *R*-values are similar for both HKLF4 and HKLF5 formats and are in the range that would be expected for isotropic refinements without H-atoms and TLS.
- The R_{work} values for HKLF 5 format are suspiciously low! This might be another reason for preferring the HKLF4 data.





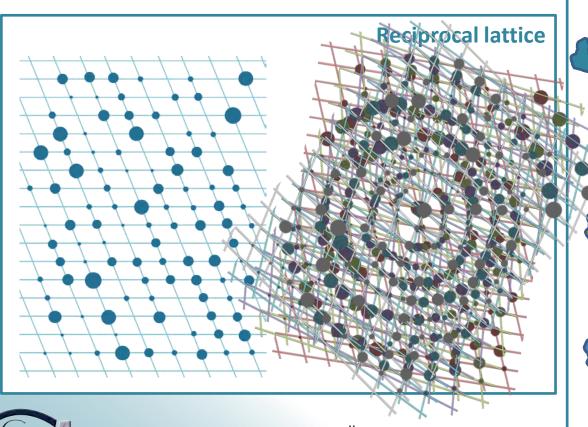
Structure solution and refinement

- The **HKLF4** file contains merged reflections and is suitable for structure solution and initial refinement using any of the usual programs.
- The HKLF5 data contains unmerged data with a total intensity for overlapping reflections. HKLF5 data are not suitable for structure solution, but – especially for small molecules – usually give slightly better R values in the final refinement.

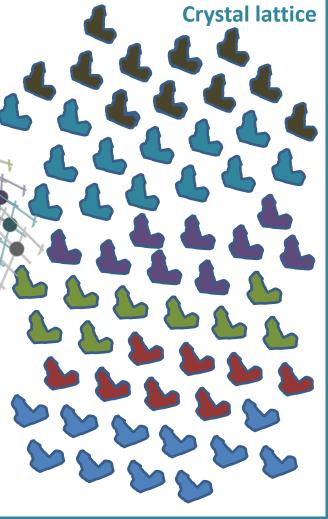




This is NOT a twin!



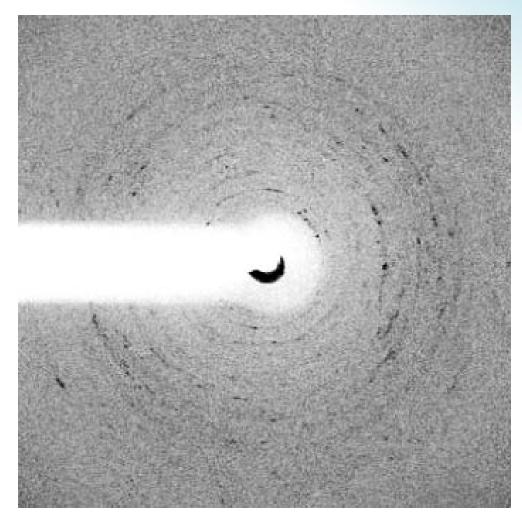






This is NOT a twin!

- Twin law near to unity
- Indexing gets better with box being bigger
- Bad data quality



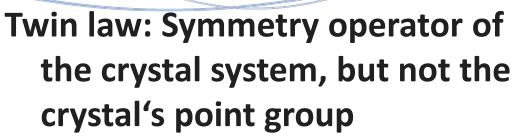


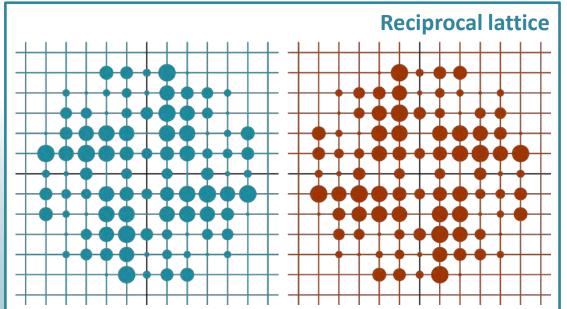


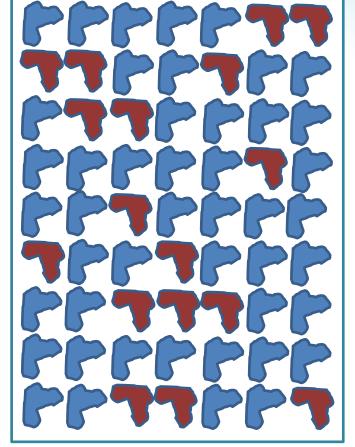
- Non-merohedral twin domains are related by an arbitrary twin law.
- To index all reflections, the cell has to be applied in different orientations.
- Partially overlapping reflections can be integrated by profile fitting.
- Scaling is applied to single reflections and to equivalents groups of overlapping reflections in HKLF5 format.
- HKLF4 format files are used for structure solution. For macromolecules deconvolution has many advantages.



Merohedral twinning







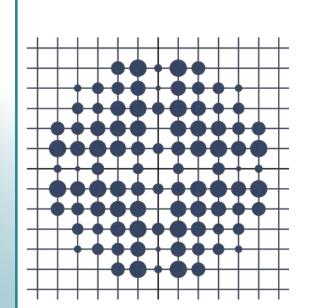
PPPPP

Crystal lattice



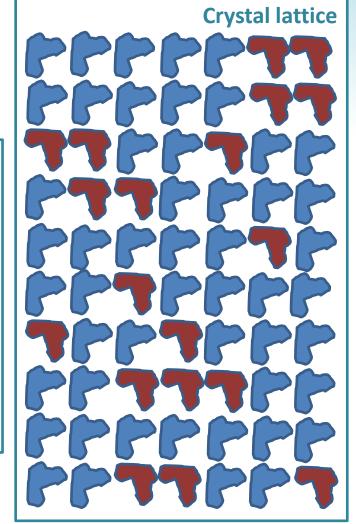
Merohedral twinning

Twin law: Symmetry operator of the crystal system, but not the crystal's point group



Reciprocal lattice







Twin law: Symmetry operator of the crystal system, but not the crystal's point group

- Only in tetragonal, trigonal, hexagonal and cubic space groups possible.
- Exact overlap of reciprocal lattices, but different intensity distribution
- Correct space group determination and phasing difficult



Merohedral twinning

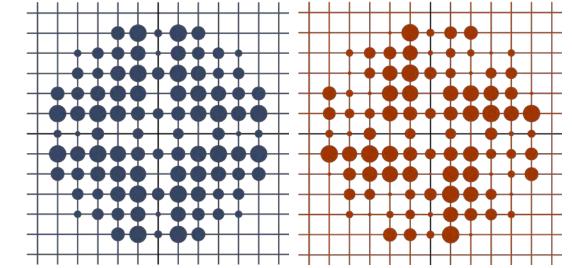
How to recognize?

- Lower symmetry point group of the trigonal, hexagonal, tetragonal or cubic system
- Symmetry looks possibly higher than it really is
- Changed intensity distribution
- R_{int} for the higher Laue groups
- Typical space group
- No structure solution



Merohedral twinning

Depending on the twin fraction α: The **intensity distribution** has been changed by the twinning. Also, symmetry looks higher!



The $\langle E^2 - 1 \rangle$ could be too low for twins (below 0.736).





Identification of a merohedral twin, the twin law and determination of the twin fraction α ?

- PLATON
- XPREP
- Yeates-Test and Yeates-Padilla-Test
 - with CTRUNCATE and DETWIN
- Merohedral twinning occurs more frequently than commonly recognized!
 Integration in the lower Laue group



Merohedral twinning

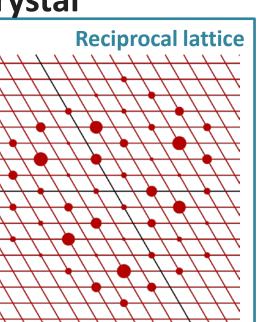
Refinement

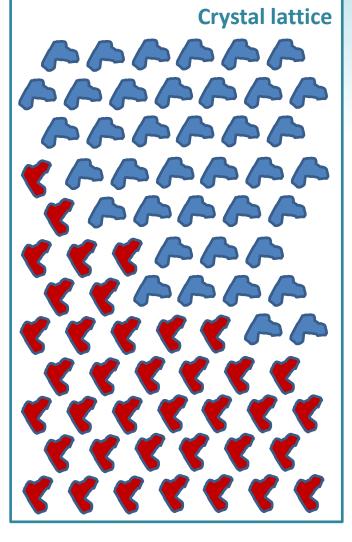
- Several programs can refine merohedral twins.
- Difference density might have fewer features.
- R factors may be lower than in single crystals.
 (Random R value goes down from 58.5% to 50%!)
- R_{free} set should include all twin-related reflections.
- Do not use a merohedral twin refinement on data which is not twinned. It will lower the R value possibly, but it is not a valid treatment!



Reticular merohedry

Most common case: Obverse/reverse twinning in a rhombohedral crystal









The typical case (in proteins)

ls:	Looks like:
R32	P3 ₁ 21
R3	P3 ₁

1/3 of all reflections are missing.

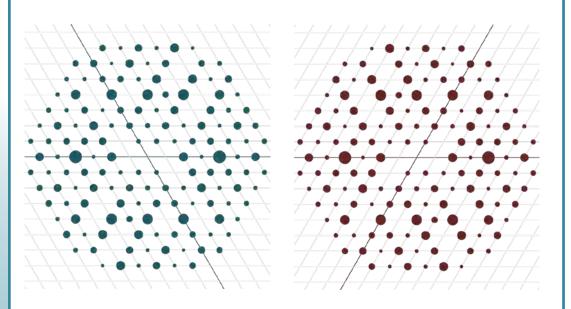
The missing reflections form a funny pattern, which is inconsistent with any systematic absence.



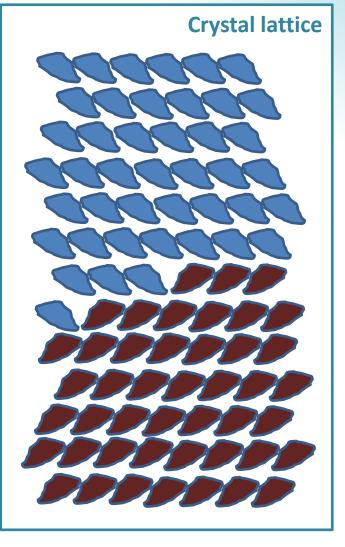
Pseudo-merohedral twinning

Twin law: Belongs to a higher crystal system than the structure.

Reciprocal lattice







Pseudo-merohedral twinning

How to recognize?

- Like a merohedral twin. The real space group belongs to another crystal system than the observed one.
- The overlap of the lattices might not be perfect for all reflections.
- All hints for merohedral twinning might also work for pseudo-merohedral ones.
- R_{int} behaves like in merohedral twins.





How to treat?

- Treatment is very similar to the one for merohedral twins.
- Most programs that can process merohedral data will also process pseudo-merohedral one.
- Be careful to choose the right (lower) crystal system.





All data and crystal images shown here were part of **Madhumati Sevvana**'s PhD thesis about non-merohedral twins:

Crystallographic Analysis of Pathological Crystals, Periplasmic Domain of Ligand-free CitA Sensor Kinase and PDI-related Chaperones

http://webdoc.sub.gwdg.de/diss/2006/sevvana





ACKNOWLEDGEMENTS

I am grateful to **George Sheldrick** and **Regine Herbst-Irmer**, who freely shared with me their knowledge, ideas and material.

I also want to thank the **Sheldrick group** for all the support in the final months of my own thesis!

Martin Lutz kindly provided the P4P file for the RLATT demonstration.



