

ACA 2009

Handling Twinning in Macromolecular Crystallography

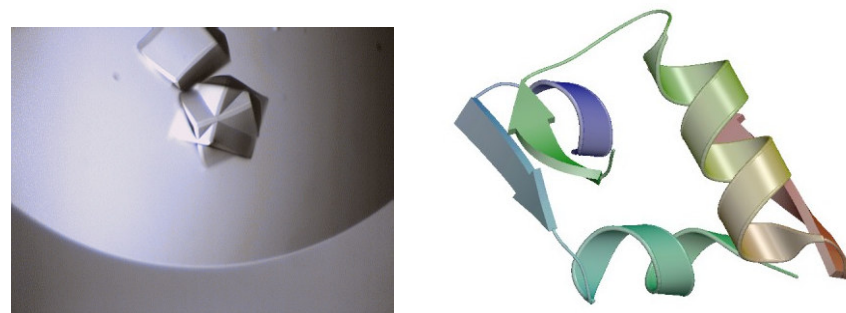
Toronto, July 25th, 2009

Non-merohedral Twinning

rherbst@shelx.uni-ac.gwdg.de

<http://shelx.uni-ac.gwdg.de/~rherbst/twin.html>

Bovine Insulin



51 amino acids in the asymmetric unit

CELL_NOW - Input

Full name of .p4p, .spin or .drx file to read: **test.spin**
253 reflections read in

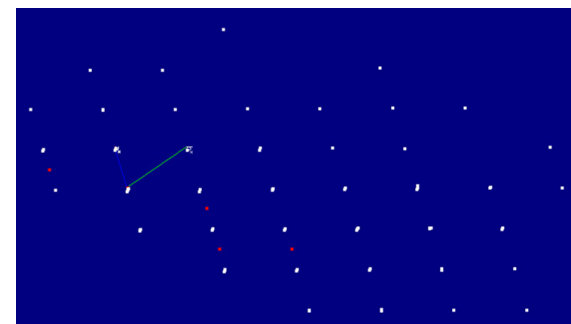
Listing file [test._cn]:

Initial search (<Enter>) or specified cell search (S):

Superlattice threshold: an axis will be rejected if less than this percentage of reflections has indices not equal to $2n$ or $3n$ resp. [10]:

Minimum and maximum allowed values for cell edge [5 45]:

RLATT - Superlattice threshold



CELL_NOW - Input

Full name of .p4p, .spin or .drx file to read: **test.spin**
253 reflections read in

Listing file [test._cn]:

Initial search (<Enter>) or specified cell search (S):

Superlattice threshold: an axis will be rejected if less than this percentage of reflections has indices not equal to 2n or 3n resp. [10]:

Minimum and maximum allowed values for cell edge [5 45]:
60 80

CELL_NOW - Output (I)

The following cells would appear to be plausible, but should be checked using XPREP because they are not necessarily the conventional cells.

FOM, % within 0.2, a..gamma, volume and lattice type for potential unit-cells:

1	1.000	53.7	77.986	78.024	78.040	89.94	89.99	90.01	474854.1	I
2	0.380	55.4	67.532	67.549	67.577	109.49	109.44	109.47	237348.4	P
3	0.372	55.3	67.532	67.577	67.595	109.51	109.46	109.44	237464.6	P
4	0.372	55.3	67.532	67.549	67.595	109.45	109.46	109.47	237455.8	P
5	0.369	55.2	67.549	67.577	67.595	109.51	109.45	109.49	237407.3	P
6	0.294	46.1	77.988	78.033	78.036	90.03	89.99	90.02	474899.7	I

CELL_NOW - Output (II)

Cell for domain 1: **77.986 78.024 78.040 89.94 89.99 90.01**

Figure of merit: 0.560 % (0.1): 51.6 % (0.2): 55.2 % (0.3): 62.6

Orientation matrix:

0.00347148	-0.01005931	-0.00713157
-0.00035888	-0.00749939	0.01039335
-0.01233875	-0.00261444	-0.00230569

Maximum deviation from integer index. If this is given as negative, the tolerance is based on the longest axis and applied to all three [0.25]:

Percentages of reflections in this domain not consistent with lattice types:
A: 49.3, B: 50.2, C: 49.9, **I: 5.0**, F: 74.7, O: 66.1 and R: 66.1%

Percentages of reflections in this domain that do not have:
h=2n: 49.2, k=2n: 50.0, l=2n: 50.0, h=3n: 66.1, k=3n: 65.5, l=3n: 66.8%

CELL_NOW - Output (II)

New cell from list (number), reorientate (R), accept (A) or quit (Q) [A]:

.p4p or .spin file to write domain to: **d1.spin**

RLATT color-coding employed in file: d1.spin
White: indexed for first domain
Red: not yet indexed

4072 reflections within tolerance assigned to domain 1,
4072 of them exclusively; 2935 reflections not yet assigned to a domain

Re-refine initial cell (R), search for next domain (S), quit (Q) or choose new cell from list (enter number) [S]:

CELL_NOW - Output (III)

Cell for domain 2: 77.986 78.024 78.040 89.94 89.99 90.01

Figure of merit: 0.872 %(0.1): 90.4 %(0.2): 93.6 %(0.3): 94.4

Orientation matrix: 0.00029610 0.00743809 -0.01043920
-0.01281940 0.00016027 -0.00024474
-0.00001133 0.01043623 0.00742700

Rotated from first domain by 129.1 degrees about
reciprocal axis -0.830 0.038 1.000 and real axis -0.831 0.037 1.000

Twin law to convert hkl from first to 0.035 0.567 -0.823
this domain (SHELXL TWIN matrix): -0.628 -0.628 -0.459
-0.779 0.532 0.333

RLATT color-coding employed in file: d2.spin

White: indexed for first domain

Green: current domain (but not in a previous domain)

Red: not yet indexed

3574 reflections within tolerance assigned to domain 2,
2756 of them exclusively; 179 reflections not yet assigned to a domain

d2.spin

```
...
CELL 77.9253 78.0153 78.0891 90.0900 90.0029 90.0224 474731.500
CELLSD 0.0156 0.0156 0.0156 0.0300 0.0300 0.0300 237.366
ORT1 0.003470643 0.007125409 -0.010055407
ORT2 -0.000361324 -0.010403471 -0.007488692
ORT3 -0.012349292 0.002301708 -0.002607543
ZEROS 0.0000000 0.0000000 0.0000000 0.0000 0.0000 0.0000
ADCOR 0.7900 0.1500 -0.1400 0.0000 0.0000 0.0000
CELL2 77.9253 78.0153 78.0891 90.0900 90.0029 90.0224 474731.500
CELLSD2 0.0156 0.0156 0.0156 0.0300 0.0300 0.0300 237.366
ORT12 0.007453182 0.010432143 -0.000315667
ORT22 0.000184934 0.000277006 0.012802006
ORT32 0.010444943 -0.007442796 -0.000000610
ZEROS2 0.0000000 0.0000000 0.0000000 0.0000 0.0000 0.0000
ADCOR2 0.7900 0.1500 -0.1400 0.0000 0.0000 0.0000
```

TWINABS

Rint = 0.4347 for all 435389 observations and

Rint = 0.4276 for all 295354 observations with $I > 3\sigma(I)$

Rint is based on agreement between observed single and composite intensities and those calculated from refined unique intensities and twin fractions.

**** Warning: components may be inconsistently indexed, try reindex option.**

After reindexing the second domain with

0 1 0 1 0 0 0 0 -1:

Rint = 0.0351 for all 435389 observations and

Rint = 0.0307 for all 295354 observations with $I > 3\sigma(I)$

"Twin Pairing Errors"

Most Disagreeable Reflections

h	k	l	F_o^2	F_c^2	$\Delta(F^2)/\sigma$	F_o/F_{cmax}	Res.(Å)
4	5	10	1403.88	325.9	16.81	0.056	1.15
6	12	3	466.90	132.48	6.55	0.036	1.10
9	10	0	391.68	40.15	6.16	0.020	1.09
4	1	1	3603.49	2038.18	5.90	0.140	2.94

...

.hkl file:

7	5	7	170.12	5.38	-2	9	10	0	40.05	4.96	1
4	5	10	170.12	5.38	1	-2	10	11	128.97	6.65	-2
4	5	10	143.53	14.20	1	9	10	0	128.97	6.65	1
1	12	8	100.26	5.44	-2	0	1	5	548.82	10.10	-2
6	12	3	100.26	5.44	1	4	1	1	548.82	10.10	1
6	12	3	47.74	4.01	1	4	1	1	367.60	22.30	1

TWINABS

Generate HKLF 4 format file (4) with crude averaged intensities for structure solution or more accurate HKLF 5 format file (5) for refinement, reindex (I), repeat (R) or quit (Q) [5]:

Average equivalent reflections (Y or N) [Y]:

Enter name of output file [twin5.hkl]:

Make file using domain N only, -N to use domains 1..N or 0 to use all [1]:

Average Friedel opposites in output file (Y or N) [Y]:

Leave out single reflections that also occur in composites (Y or N) [Y]:

Structure Solution

- Resolution to 1.60 Å
- Cubic symmetry (space group $I2_13$): high redundancy
- 6 sulfur atoms : structure solution with single wavelength anomalous scattering?

Sad Phasing:

- Redundant data
- Complete data
- Precise data (small signal lost easily in the noise)
- Avoid systematic noise (ice rings, pin in the beam, ...)

Twinning ???

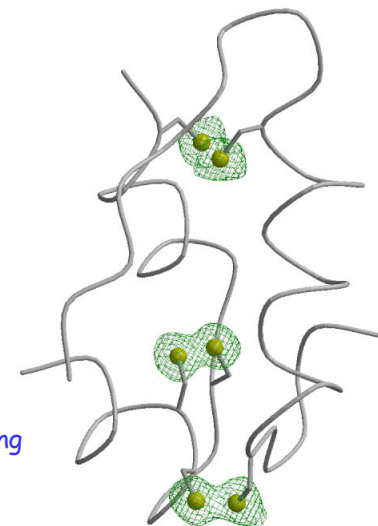
<http://erice2005.docking.org/vcourse/21sat/1130-Sevana/Sevana.pdf>

Substructure Solution with SHELXD

Data truncated to 1.90 Å

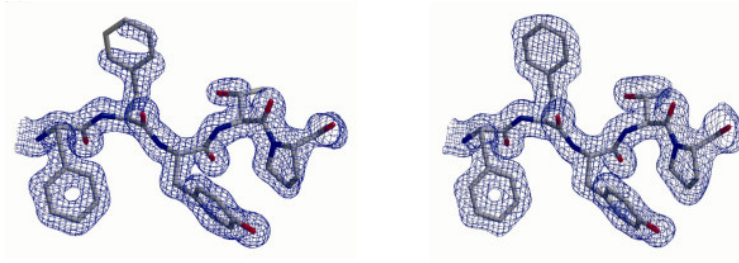
CC 50.5%, weakCC 28.1%

(could also be solved with wrong indexing of the second domain)



Bovine Insulin - Structure Solution

- Phasing and Density Modification with SHELXE
- SHELXE map was traced using ARP/wARP



SHELXE map

Final map

contoured at 1σ

Final Refinement

	HKLF 5	HKLF 4	HKLF 5	HKLF5P
	wrongly indexed			
R1 ($F_o > 4\sigma(F_o)$)	0.2004	0.1487	0.1150	0.1050
R1(free)	-	0.1814	-	-
wR2 (all data)	0.4841	0.3876	0.3314	0.2856
Data	15167	10583	22468	15167
Twin fraction	0.212	0.417	0.433	0.427
R1 (after merging for Fourier)	0.1876	0.1503	0.1414	0.1175
Unique data	10582	10583	10582	10582

Glucose Isomerase

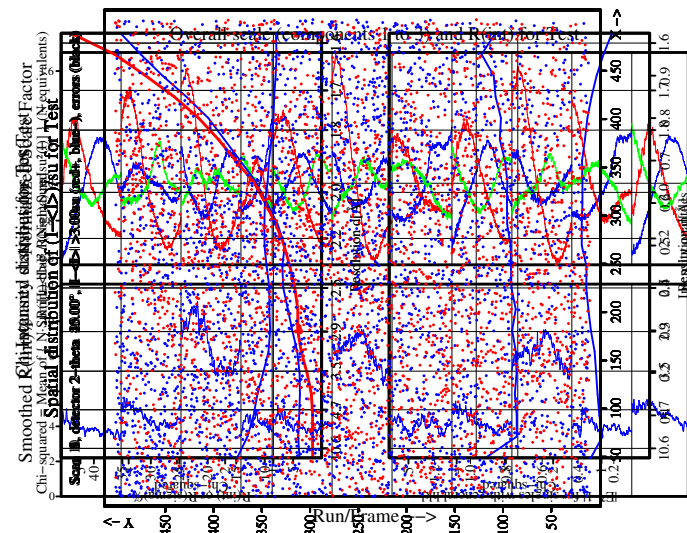


space group I222
cell: 94.34 99.36 103.64
three twin domains
rotated by 120°



<http://erice2005.docking.org/vcourse/21sat/1130-Sevvana/Sevvana.pdf>

TWINABS



Solution

SAD:

SHELXD found 2 cations

with data truncated to 2.10 Å

CC 33.1% and weak CC 21.2

Excellent experimental map traced with ARP/wARP

Final Refinement

	HKLF 4	HKLF 5	HKLF5P
R1 ($F_o > 4\sigma(F_o)$)	0.1638	0.1298	0.1199
R1(free)	0.1924	-	-
wR2 (all data)	0.4261	0.3580	0.3223
Data	58841	106442	75109
Twin fraction 1	0.184	0.192	0.187
Twin fraction 2	0.400	0.416	0.411
R1 (after merging for Fourier)	0.1807	0.1571	0.1368
Unique data	58841	49100	49100

Summary

- Non-merohedrally twinned structures can be solved and refined to satisfying results - even for macromolecules
- More integration programs for twins are needed
- Refinement against the detwinned HKLF4 data is sufficient for routine structures refinement and in first stages of refinement
- Refinement against HKLF5 data leads to slightly better data

Acknowledgements

Madhumati Sevana, University of Erlangen

George Sheldrick, University of Göttingen