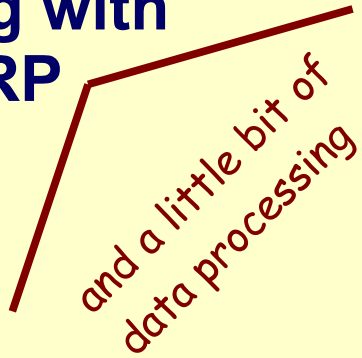
Experimental Phasing with SHARP/autoSHARP



Clemens Vonrhein
Global Phasing Ltd.

DLS/CCP4 workshop
06.12.2015

Overview

SHARP/autoSHARP:

- principles
- the “pipelines”
- tools around it

•**Examples:**

- Radiation damage
- Improve on first results
- SAD, MAD, MIR … alphabet soup

Non-automated structure solution:

- Why bother?
- **SHARP** – advanced phasing
- (Some) tips ...

Tutorials:

http://www.globalphasing.com/sharp/wiki/
http://www.globalphasing.com/autoproc/wiki/
http://www.globalphasing.com/buster/wiki/

SHARP/autoSHARP Design

autoSHARP:

- centred around **SHARP**
- supports common phasing scenarios: **MAD**, **SAD**, **SIR(AS)**, **MIR(AS)**
- *minimum* input
- explanation of results
- online help

Vonrhein, C, Blanc, E., Roversi, P and Bricogne, G. (2007). Automated structure solution with autoSHARP. Methods Mol Biol 364, 215-30.

SHARP:

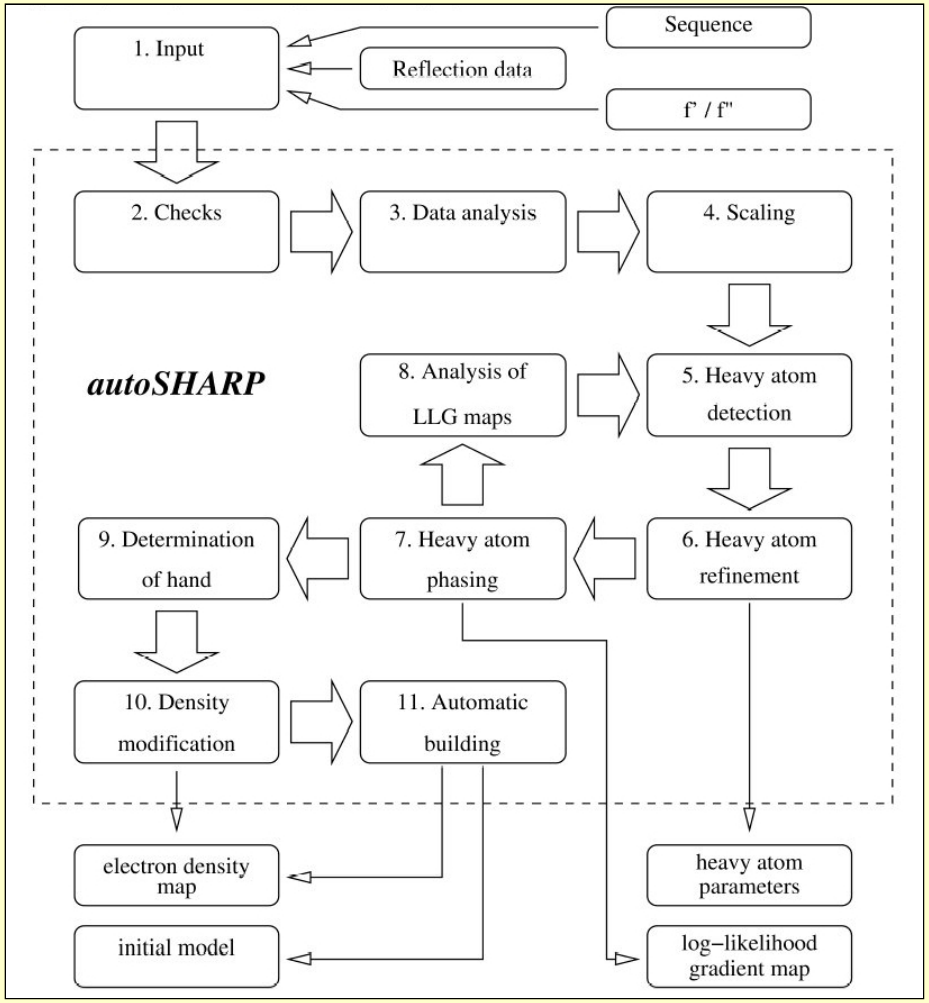
- getting **best phases** possible
- applicable to large variety of **phasing scenarios**: MAD+native, SAD+SIR, ...
- *extensive* user control
- interfaced to downstream phase improvement
- online help

Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. (2003). Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. Acta Cryst. D59, 2023-2030.

SHARP Control Panel

1. Typically start with autoSHARP
2. Optimise with SHARP

autoSHARP “pipeline”



minimum amount of input

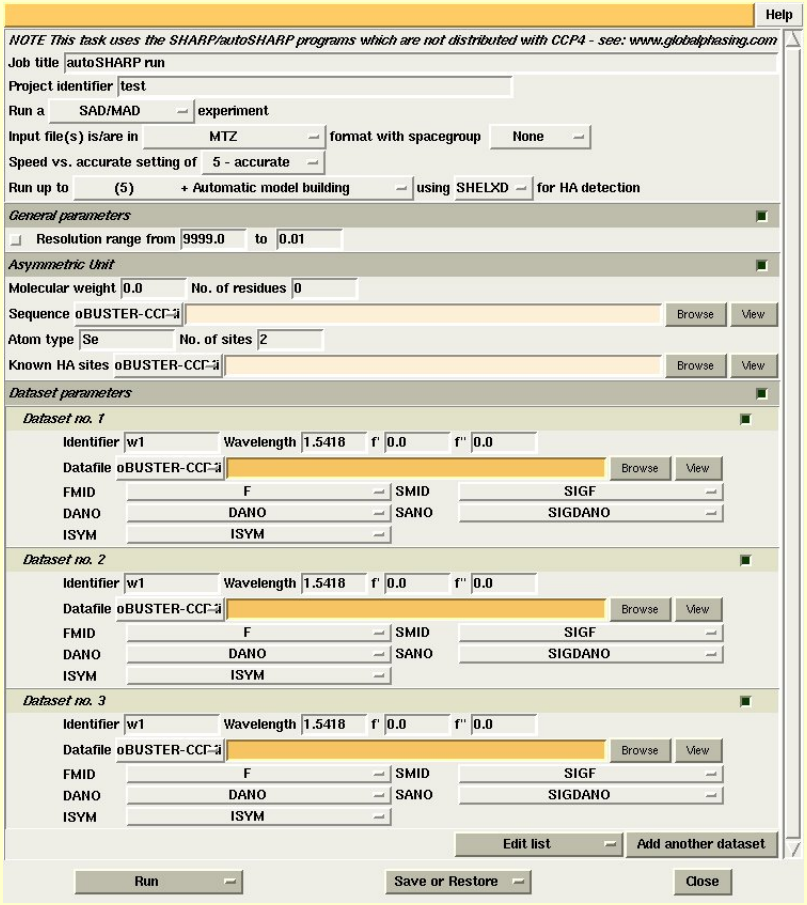
feedback:

internally and to the user

several levels of results

autoSHARP - Input

- experiment type: **SAD, MAD, SIR(AS), MIR(AS)**
- **data files**: MTZ, Scalepack (merged or unmerged)
- additional information:
  - **no. of HA sites** (estimate)
  - **type of HA** (predominant one)
- optional information:
  - sequence
  - wavelength, f'/f''
  - space group



CCP4i interface

Command-line interface

`run_autoSHARP.sh -h`

writes HTML output (view in browser)

creates scripts ready to start Coot

SAD:

```
run_autoSHARP.sh -seq 1o22.pir -ha "Se" \
  -wvl 0.9778 peak -7 5 -sca 1o22_peak.sca
```

MAD:

```
run_autoSHARP.sh -seq 3isy.pir -ha "Se" \
  -wvl 0.97934 infl -11 3.3  -sca 3isy_aimless_0.97934A.sca \
  -wvl 0.91162 hrem -1.8 3.3 -sca 3isy_aimless_0.91162A.sca
```

SIR(AS):

```
run_autoSHARP.sh -seq 1GXT.pir -nat -mtz 1GXT_nat.mtz \
  -ha "Hg" -nsit 2 \
  -wvl 0.99970 peak -16 10 -mtz 1GXT_hg.mtz
```

MIR(AS):

```
run_autoSHARP.sh -seq 3zft.pir -nat -mtz 3zft_nat.mtz \
  -ha "Hg" -nsit 1 -wvl 1.54179 -mtz 3zfq_Hg.mtz \
  -ha "Ir" -nsit 2 -wvl 1.54179 -mtz 3zfr_Ir.mtz
```

partial model:

```
run_autoSHARP.sh -seq 3get.pir -ha "Se" \
  -pdb 3ffh_ala_MR.pdb \
  -wvl 0.9789 peak -8 4 -sca 3get.sca
```

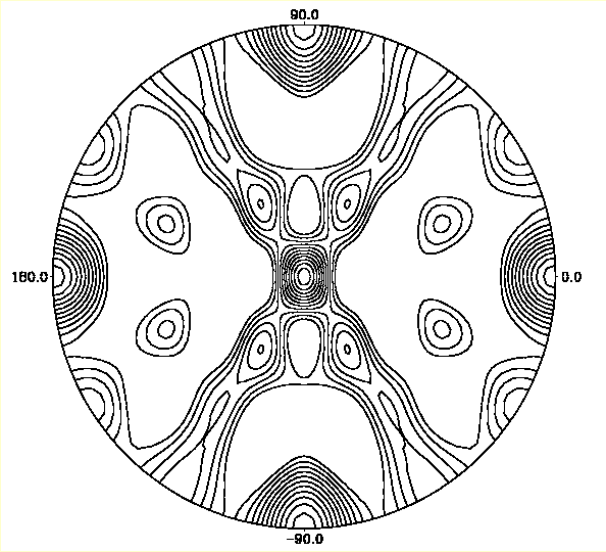
SAD with Ta6Br12 cluster:

```
run_autoSHARP.sh -seq 4cv5.pir -ha "Ta6Br12:Ta" \
  -wvl 1.25472 peak -mtz 4cv5.mtz
```

autoSHARP - Analysis

- consistency:
  - cell (significant changes)
  - spacegroup (screw axes)
  - indexing (alternate possibilities)
  - sequence (Se-Met, S-SAD)
- data quality:
  - CC(Δano), R-factor, completeness
  - outliers (E-values)

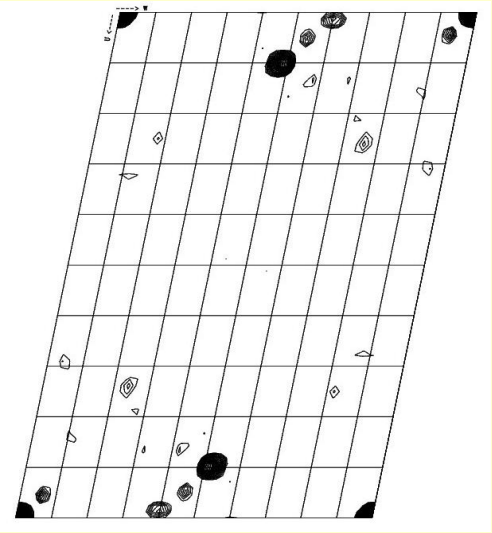
- self-rotation function
- native Patterson
- Matthews parameters (probability, Rupp)

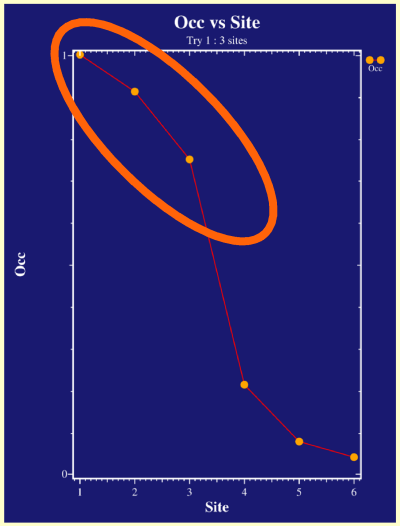


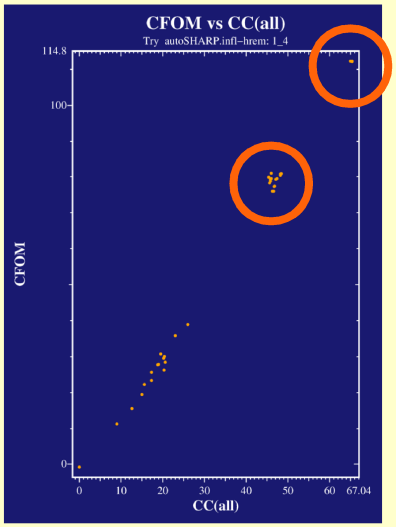
autoSHARP - Detection

- only data with significant signal (from *Analysis* step)
- try different scenaria (MAD, peak-SAD, 2-wvl MAD, ...)
- adjust number of HA sites (e.g. if exact number not known)
- use solution with best score

- **SHELXC/SHELXD** [1] :
  - stop if significant solution found
  - present results graphical
  - try detecting ‘jump’ in occupancy







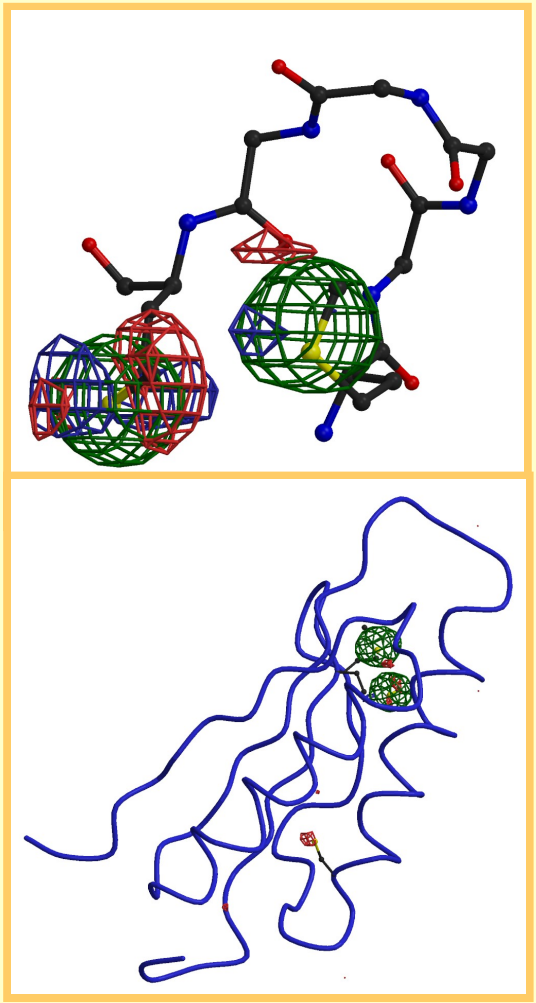
1) Schneider, T. R. & Sheldrick, G. M. (2002). Acta Cryst. D58, 1772-1779.

autoSHARP - SHARP

- HA refinement and phasing
- overall anisotropic scaling
- analysis of LLG (residual) maps:

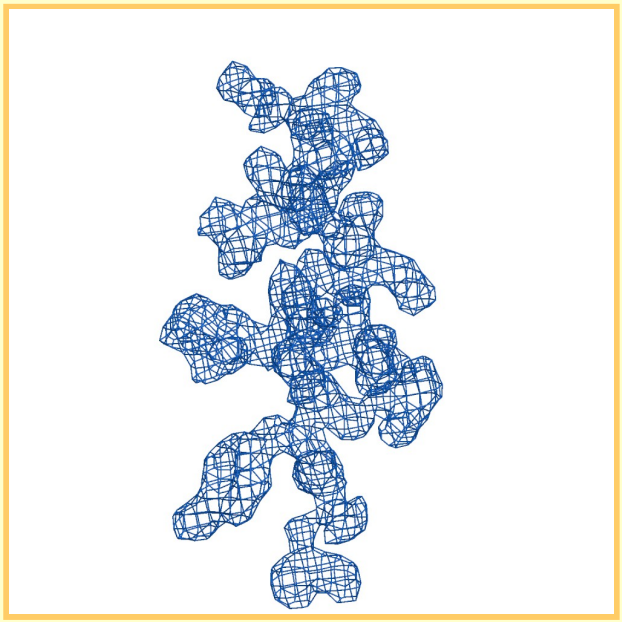
G. Bricogne (1992). A Statistical Formulation of the Molecular Replacement and Molecular Averaging Methods.In: "The Molecular Replacement Method, Proceedings of the CCP4 Study Weekend 31 January - 1st February 1992" (W. Wolf, E.J. Dodson & S. Gover, eds.) pp. 62-75. Daresbury Laboratory.

  - remove spurious sites
  - add additional sites
  - switch on refinement of f' and/or f''
  - feed-back into **SHARP**
- phases in two hands (enantiomorphs)



autoSHARP - Denmod

- **SOLOMON** [1] program
- start from resolution with significant phase information
- take heavy-atoms into account
- detect correct hand (based on CC(E**2) and “contrast”)
- optimize solvent content
- adjust correct no. of molecules



1) J. P. Abrahams and A. G. W. Leslie (1996). Acta Cryst. D52, 30-42.

autoSHARP – automatic building

- **ARP/wARP** [1)] suite of programs (latest ARP/wARP as distributed via CCP4)
- **PARROT/BUCCANEER** [2)] iterative density-modification and automatic building tool (“Long John Silver”)

- Results from heavy atom phasing (SHARP) plus density modification (SOLOMON) can easily be feed into other building programs:
  - Good experience with BUCCANEER at medium/low resolution and for quick/initial building
  - taking advantage of NCS information (if possible): very powerfull

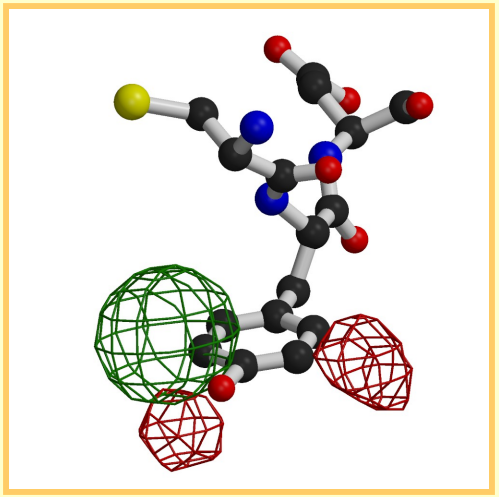
1) Perrakis, A., Morris, R. & Lamzin, V.S. (1999). Nat. Struct. Biol. 6, 458-63.
2) K. Cowtan (2006) Acta Cryst. D62, 1002-1011.

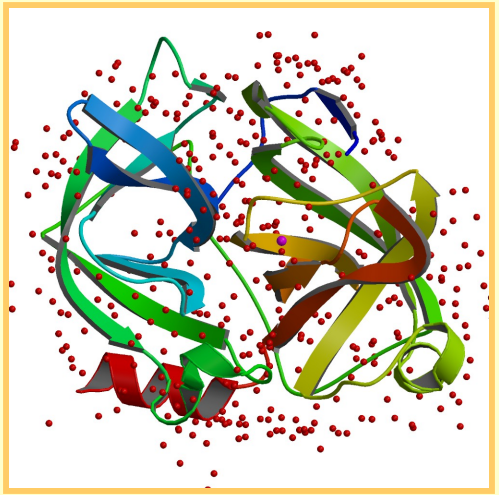
autoSHARP - Output

- self-rotation plots, native Patterson, difference Patterson (Harker sections), most meaningful/important analysis plots/tables
- scaled datasets (MAD, SIR, MIR)
- initial HA sites
- refined (completed) HA sites

- **SHARP** phases
- LLG (residual) maps

- density modified phases/maps
- initial model





Tools around SHARP/autoSHARP

- Any phasing scenario with **SHARP**, e.g.:
  - MAD+native
  - MAD+derivative+native
  - radiation damage (refining exponential occupancy decay)
  - cluster compounds (spherically averaged)
  - complete control
- separate density modification control panel:
  - inclusion of partial model
  - NCS detection (**GETAX**[1]) and averaging (**DM**)
  - extraction of HLs for **SHARP** (from PDB or MTZ)
  - difference Fourier maps (detection of additional sites)
  - **ARP/wARP** control panel
  - view results (maps, model, skeleton, ...)
- LLG (residual) map panel: viewing, analysis, peak-picking

1) C. Vonrhein and G. E. Schulz, Acta Cryst., D55, 225 - 229 (1999)

How to improve on first results

- after **autoSHARP**:
  - **warning messages** point back to data processing/scaling problems (beamstop, ice-rings, decay, ...)

Warning messages - 1

```
WARNING : No f'/f'' values defined for atom of type "I" in wavelength 1:
          automatically set to calculated values for wavelength
          0.8900 Å (-0.4308,2.6934).
```

Take fluorescence scan home (and do one in the first place)

```
WARNING : completeness of only 65.00% (below 90%) for native dataset in the
          range 18.47 - 10.0 Å. This might give serious problems later!
          Maybe you haven't included your low-resolution data into this dataset?
          Or there were a lot of overloads in the reflection patter? Anyway, it
          is never good to miss this many low resolution reflections.
```

```
WARNING : there are serious differences between 2 amplitudes from different
          datasets (as judged by analysing E values). If these appear only in
          specific resolution ranges or shells you might be able to improve
          results by restricting e.g. low resolution. Here is a list of the
          reflections that look suspicious:

     H    K    L   Reso       infl       hrem (all)
   ---------------------------------------------
     1    0    2  42.02     425.45       3.15 *
     1    1    1  45.00     539.98       3.82 *
```

Warning messages - 2

WARNING : there are serious differences between 49 amplitudes from different datasets (as judged by analysing E values). If these appear only in specific resolution ranges or shells you might be able to improve results by restricting e.g. low resolution. Here is a list of the reflections that look **suspicious:**

H	K	L	Reso	natder1_Hg_peak (all)	
5	2	9	7.30	**252.60**	**613.66** *
8	4	-62	2.22	**727.50**	**3.86** *
9	9	51	2.22	**684.00**	**3.80** *
10	0	-11	4.74	**80.20**	**332.09** *
10	2	-61	2.22	**453.10**	**4.00** *
10	6	55	2.22	**817.30**	**4.03** *
10	10	45	2.22	**1732.80**	**6.44** *
11	0	-61	2.23	**919.30**	**4.14** *
11	8	48	2.22	**1050.60**	**5.26** *
11	8	-48	2.22	**530.20**	**3.86** *
12	7	44	2.30	**35.40**	**108.02** *
12	10	38	2.22	**1086.80**	**5.63** *
13	0	13	3.68	**1516.30**	**16.43** *
13	4	51	2.23	**1306.20**	**11.32** *
13	6	-47	2.22	**1448.40**	**3.94** *
13	8	41	2.22	**1032.70**	**5.22** *
13	10	33	2.22	**1070.80**	**5.81** *
13	11	-28	2.22	**1466.30**	**8.43** *
13	13	9	2.22	**1339.00**	**9.46** *

Warning messages - 3

NOTE : 4 reflexions have large anomalous differences:

H	K	L	FMID	SMID	DANO	SANO	ABS(DANO)/FMID
-3	1	12	203.27	5.43	398.98	7.68	1.96 *
-2	2	12	79.28	2.86	-150.88	4.05	1.90 *
-2	6	17	177.12	8.44	-259.31	11.94	1.46
-1	1	2	28.28	2.17	-43.53	3.07	1.54

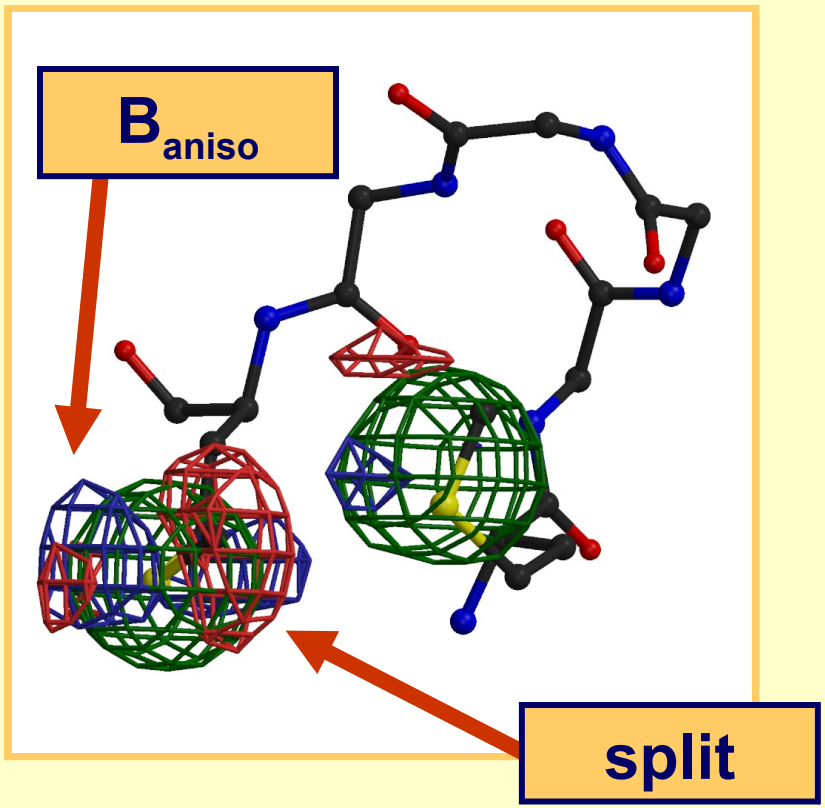
WARNING : We will remove 2 reflexions with a ABS(DANO)/FMID ratio > 1.9

WARNING : the two hands don't show a significant difference in the two scores! This might be due to a very weak phasing signal, wrong heavy atom sites or something else (single site in polar spacegroups?). Please check your heavy atom sites solution carefully (does it look convincing?) as well as the LLG/residual maps from SHARP (for minor sites or something pathologically going on). However, if you start seeing meaningful features in the electron density map after density modification (e.g. secondary structures) this might not be a cause for concern.

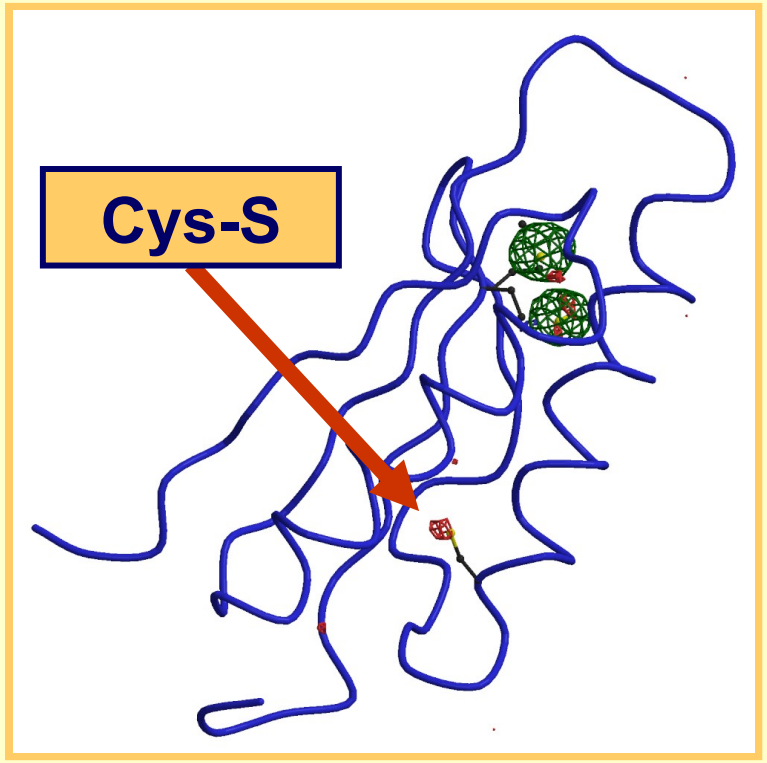
How to improve on first results

- after **autoSHARP**:
  - **warning messages** point back to data processing/scaling problems (beamstop, ice-rings, decay, ...)
  - analyse **LLG** (residual) maps: additional sites, f'/f'' refinement, anisotropic B-factor for sites, ...

Example (LLG maps)



Se sites (LLG peak)



Anomalous LLG

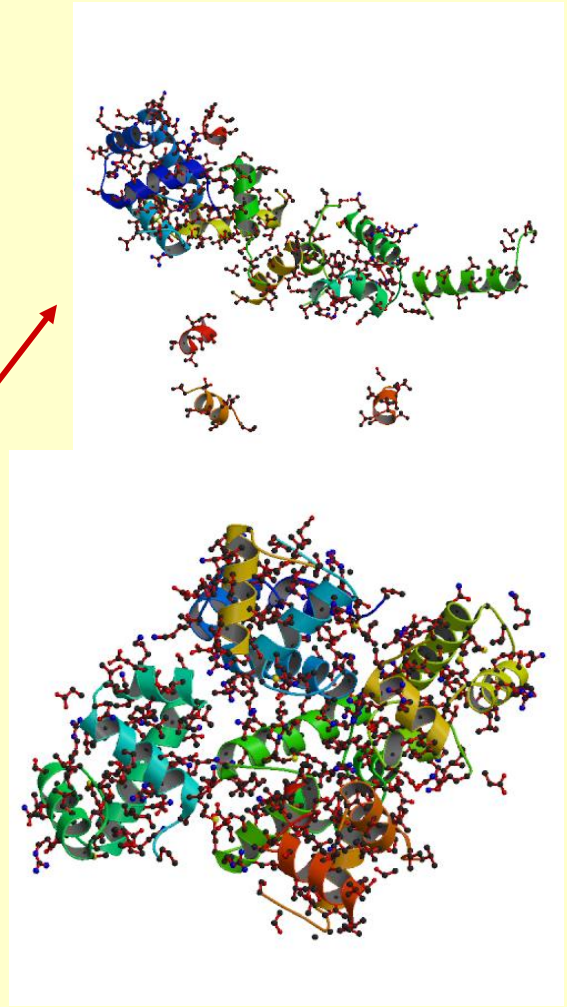
How to improve on first results

- after **autoSHARP**:
  - **warning messages** point back to data processing/scaling problems (beamstop, ice-rings, decay, ...)
  - analyse **LLG** (residual) maps: additional sites, f'/f'' refinement, anisotropic B-factor for sites, ...
  - significant difference between score of **original/inverted** hand?
  - feed-back of density modified phases into **SHARP** (improving LLG maps)
  - feed-back of **partial model** into density modification and/or **SHARP**

Example (feed-back of partial model)

- GerE (3wvl-MAD, 2.66 Å, 444 aa, 6 mol)
- cycling ARP/wARP with DenMod:
  using model in first DenMod cycle(s)

Cycle	DenMod	ARP/wARP		
		built	docked	R / Rfree
0	33.8%	251 / 13	103 / 2	0.249 / 0.513
1	40.3%	262 / 9	128 / 2	0.251 / 0.465
2	37.3%	310 / 10	239 / 4	0.236 / 0.432
3	36.5%	315 / 11	246 / 5	0.232 / 0.432



Morris, R.J., Zwart, P.H., Cohen, S., Fernandez, F.J., Kakaris, M., Kirillova, O., Vonrhein, C., Perrakis, A. & Lamzin, V.S. (2004) Breaking good resolutions with ARP/wARP. J. Synchr. Rad. 11, 56-59

External Phase Information

- **SHARP** allows use of additional phase information in likelihood function

E de la Fortelle & G Bricogne (1997) "Maximum-Likelihood Heavy-Atom Parameter Refinement for Multiple Isomorphous Replacement and Multiwavelength Anomalous Diffraction Methods". Methods in Enzymology 276, 472-494.

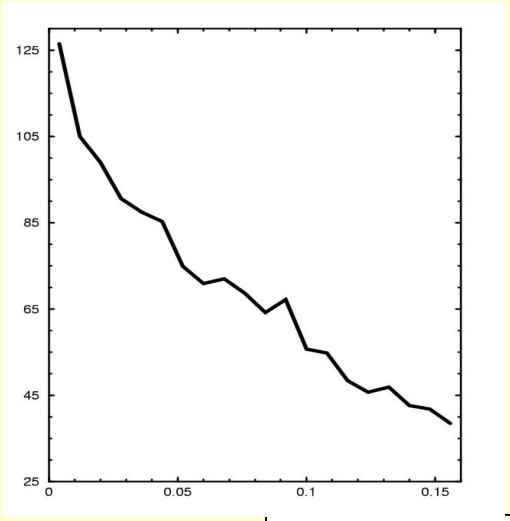
- often encoded as HL-coefficients
- from density-modification, non-isomorphous crystal, model, etc
- Interface (SHARP and autoSHARP) allows easy use of this feature
- Beneficial for:
  - Stabilising parameter refinement
  - Sensitivity of LLG residual maps → HA detection
  - Overcoming non-iomorphism issues

"MR-SAD"

- **autoSHARP** can do: **"ExtPhas-SAD"** **"ExtPhas-SIR"** **"ExtPhas-MIRAS"** **"ExtPhas-MAD"** **"ExtPhas-MIR"** **"ExtPhas-SIRAS"**
- **SHARP** can also do: **"ExtPhas-MAD-NATIVE"** **"ExtPhas-XYZ"**

Why not MIR?

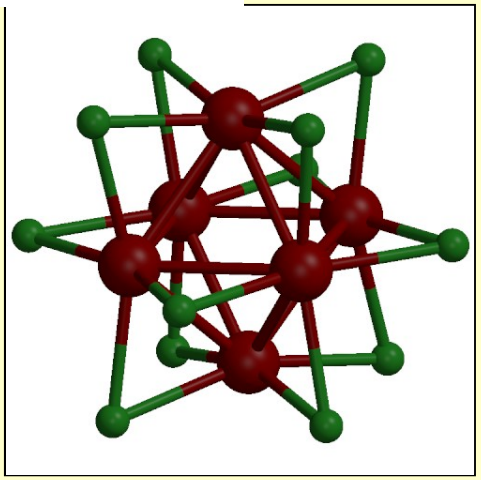
- **Non-isomorphism**
  - **datasets from different crystals**
  - **agressive soaks, co-crystallization**
- **no. of sites (and occupancy) not well known**
- **'old-fashioned'**



- **BUT:**
  - *spectacular **signal***
  - *soaking can be very fast*
  - *no need for Se-Met expression*
  - *soaks can be used for SAD/MAD too*
  - *quick soaks (Dauter) more gentle to crystals*
  - *good statistics:*

    <|FPH-FP|>

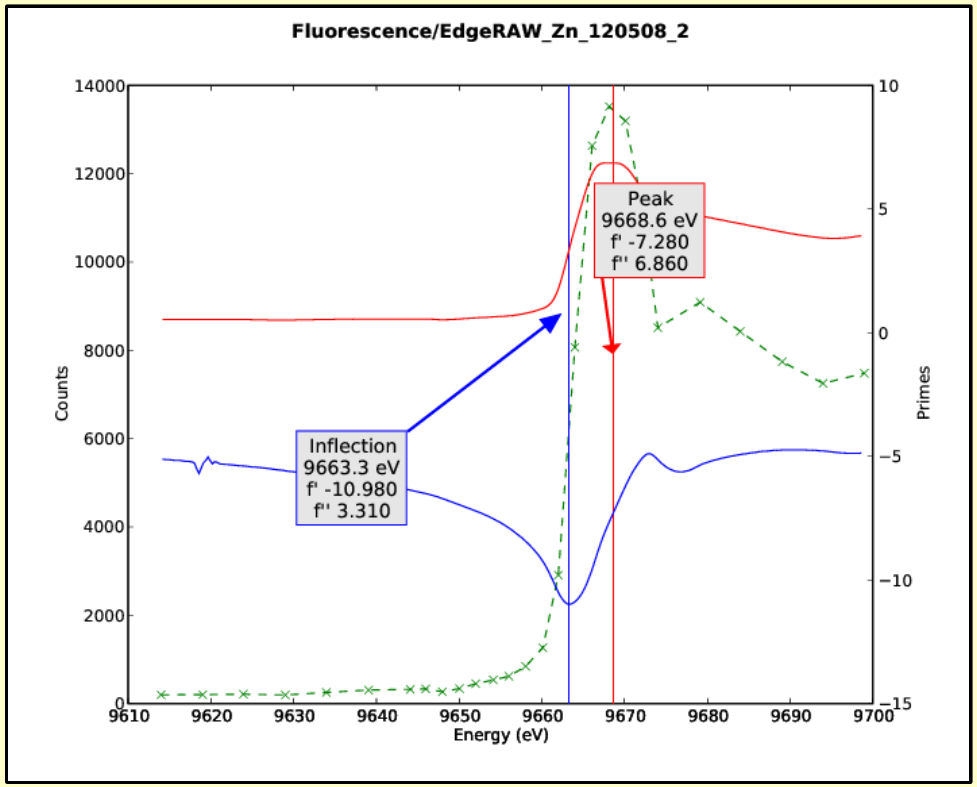
Ta_6Br_{12}



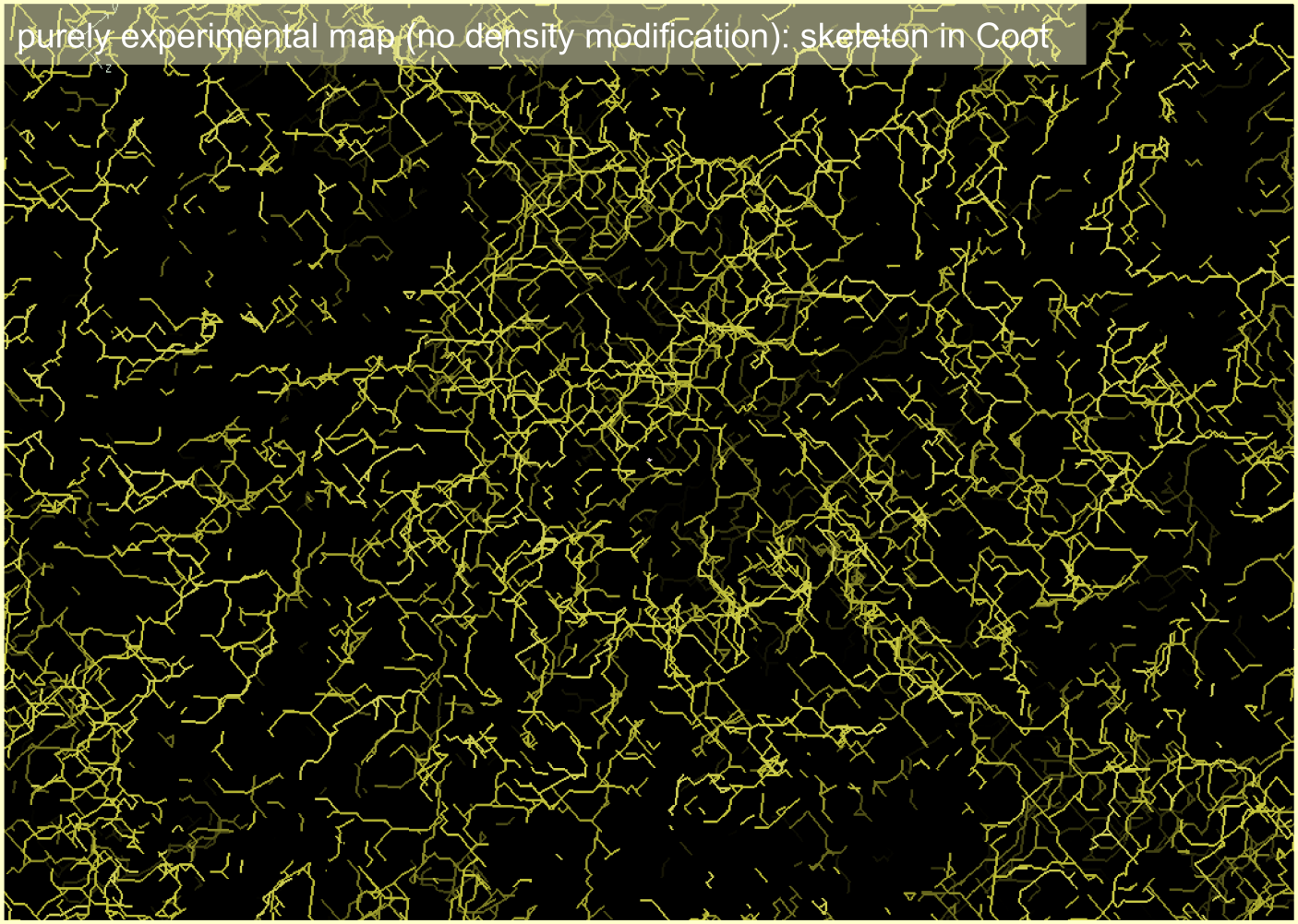
SAD or 2-wvl MAD?

A. González (2003). Optimizing data collection for structure determination. Acta Cryst. D59, 1935-1942

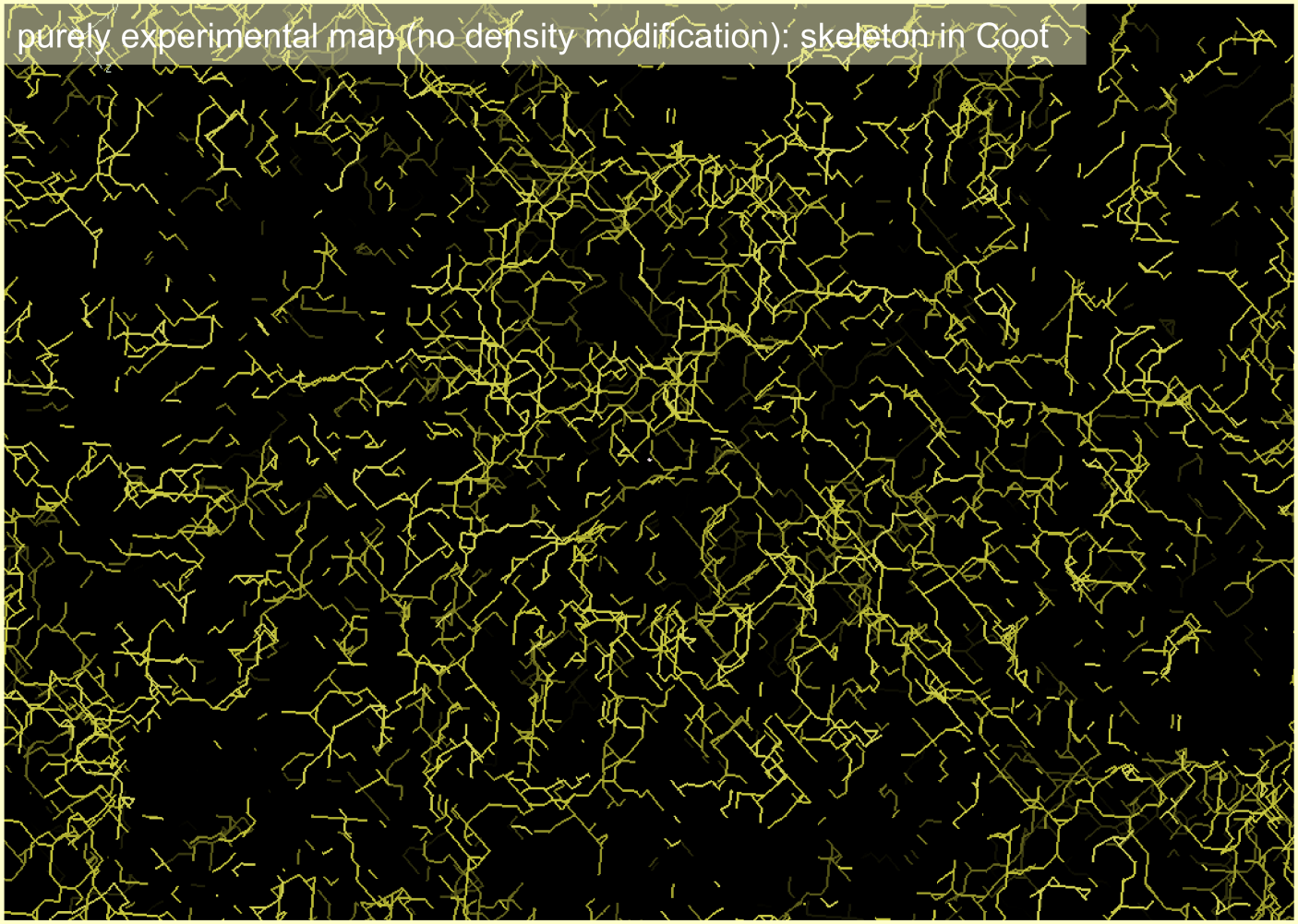
- SLS (M. Wang, V. Olieric), Soleil (A. Thompson, P. Legrand), May 2012
- Zn-containing protein
- tetragonal SG
- Data collected at very low dose (Pilatus 2M)



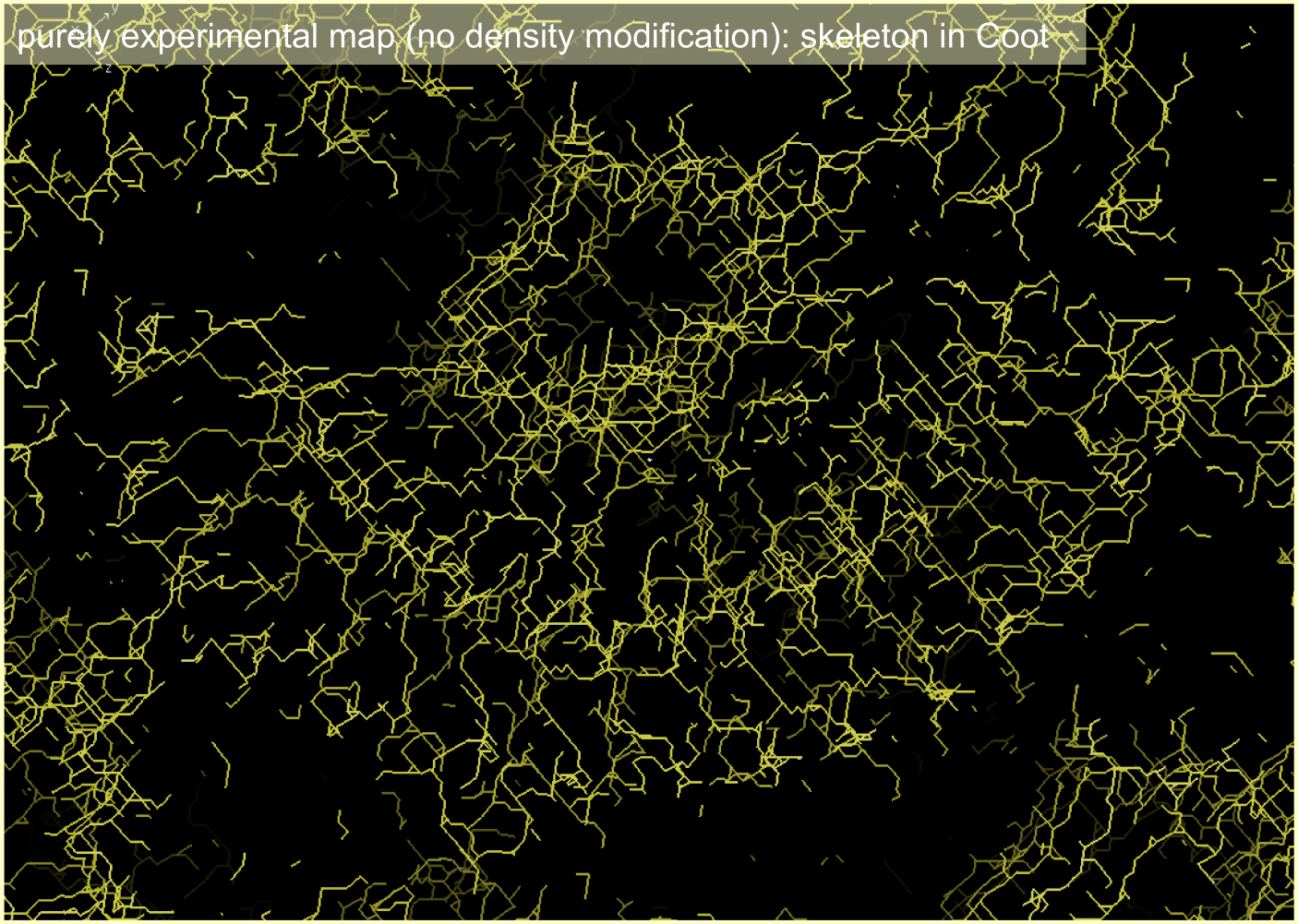
SAD (PEAK, 3600 images, multiplicity 25)



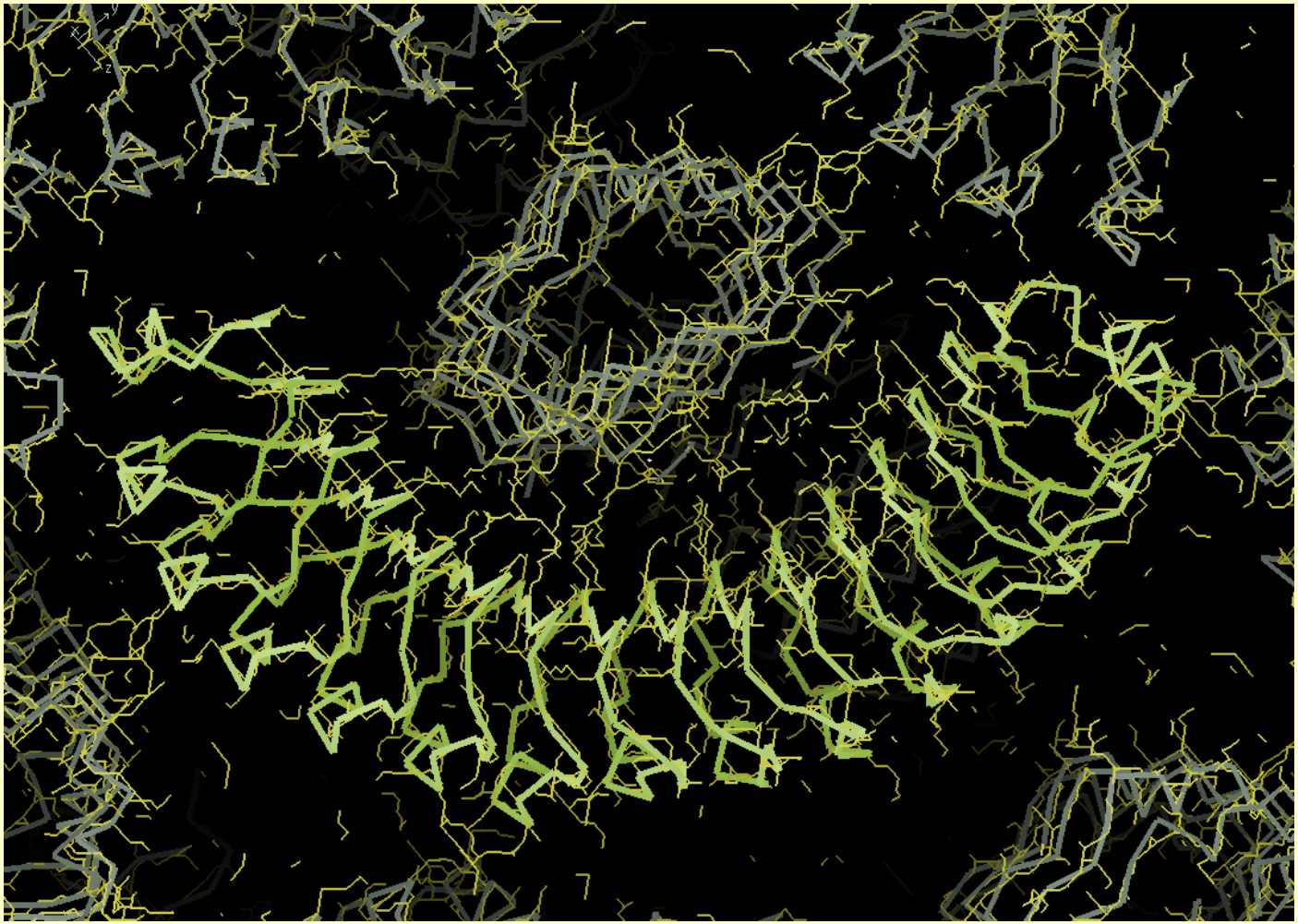
SAD, inverse-beam (1800+1800 images, multiplicity = 25)



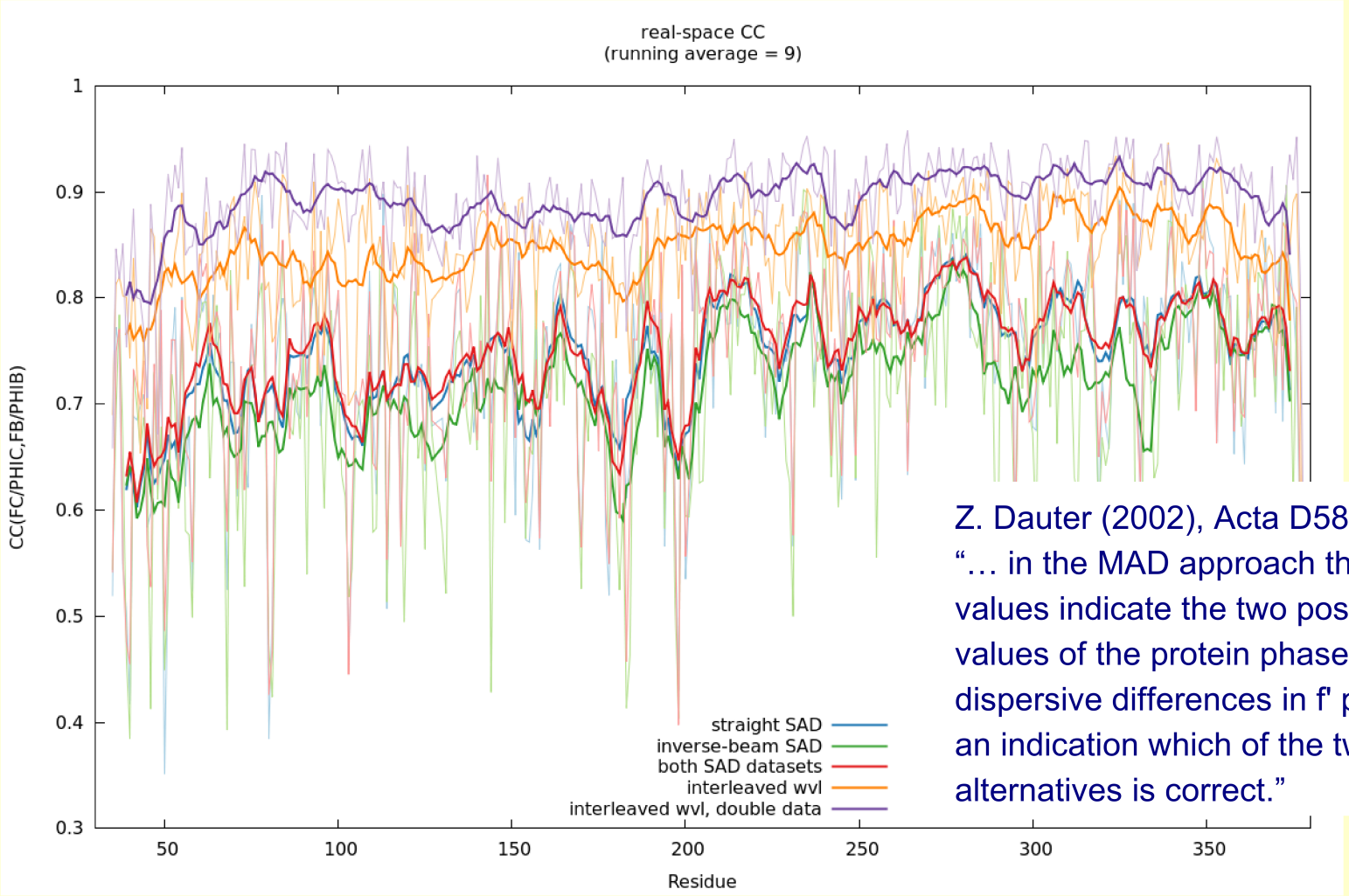
Double-inflection, interleaved-wvl (1800+1800 images, multiplicity = 12.5 + 12.5)



MAD, unaligned crystal, 2 x 180 degree (interleaved-wvl)



Phase ambiguity - SAD-vs-MAD

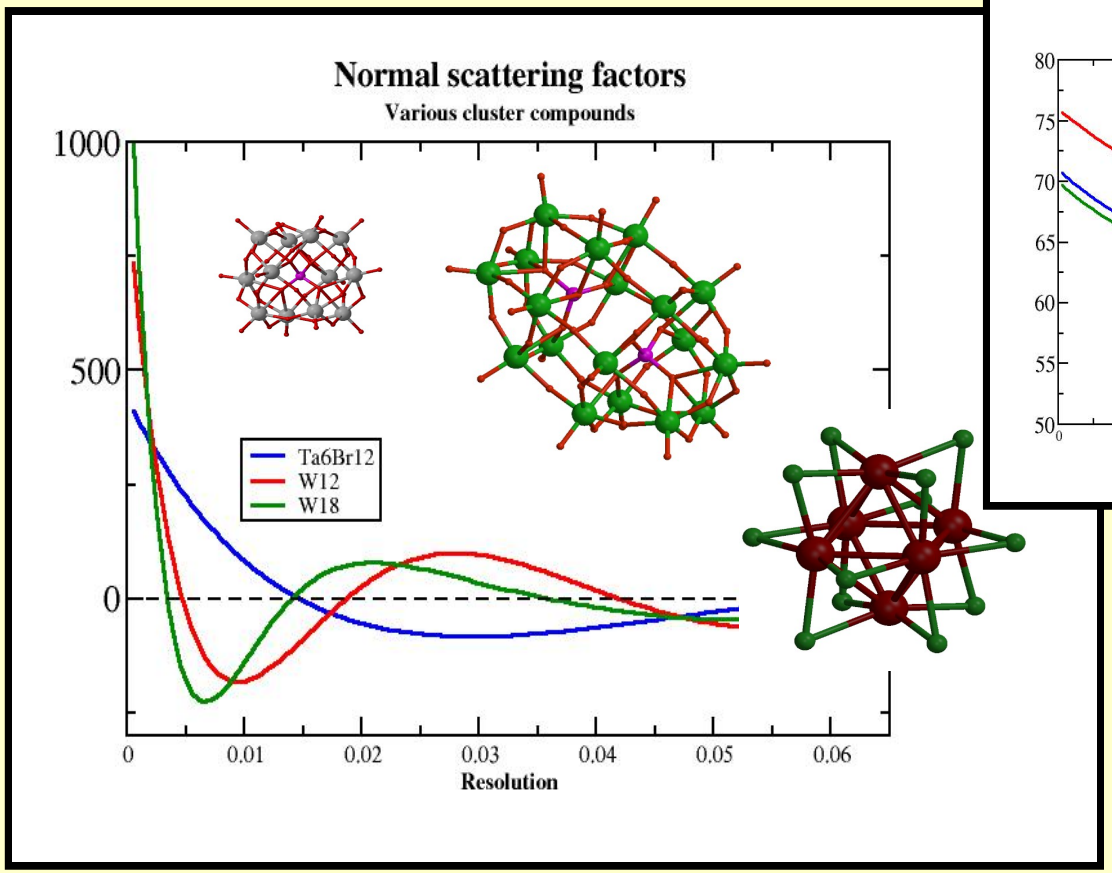


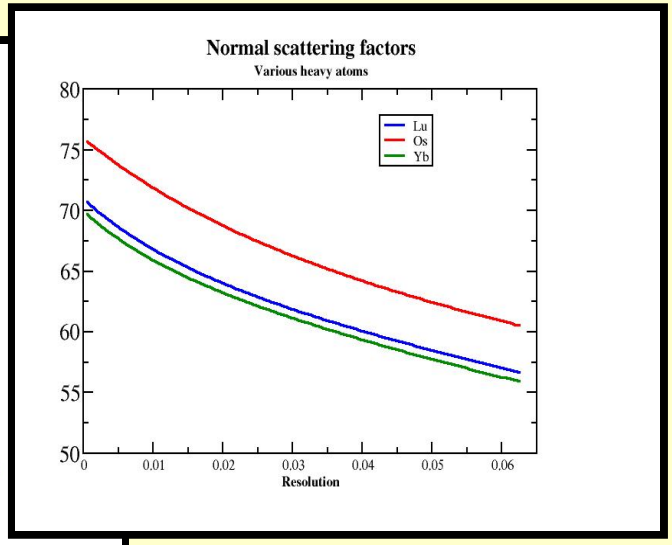
Z. Dauter (2002), Acta D58, 1958:
"… in the MAD approach the f" values indicate the two possible values of the protein phase and the dispersive differences in f' provide an indication which of the two alternatives is correct."

“Advanced” experimental phasing

- **Spherical average clusters** (getting a hold on very large complexes)
- **Exponential occupancy decay** (turning site-specific radiation damage into something positive)
- **Anisotropy of anomalous scattering** (expect the unexpected...)
- **Breaking of symmetries** (… and the even more unexpected)

Spherical averaged cluster compounds





Wimberly, B.T., Brodersen, D.E., Clemons, W.M. Jr., Carter, A.P., Morgan-Warren, R.J., Vonrhein, C., Hartsch, T. and Ramakrishnan, V. (2000). Nature 407: 327-339.

SHARP: special **SPHCLUSTER** keyword – also in autoSHARP

Session example datasets

- 2vb1
  - High resolution
  - P1 Lysozyme
- 2qvo
  - Medium resolution
  - S-SAD phasing
- 3csl
  - Low resolution
  - Se-MET
  - Difficult spots

- 1y13
  - Medium resolution
  - 2-wvl MAD
  - Small phasing signal

http://www.globalphasing.com/autoproc/wiki/index.cgi?ACA2011Tutorial
http://www.globalphasing.com/sharp/wiki/index.cgi?ACA2011Tutorial

Session example datasets

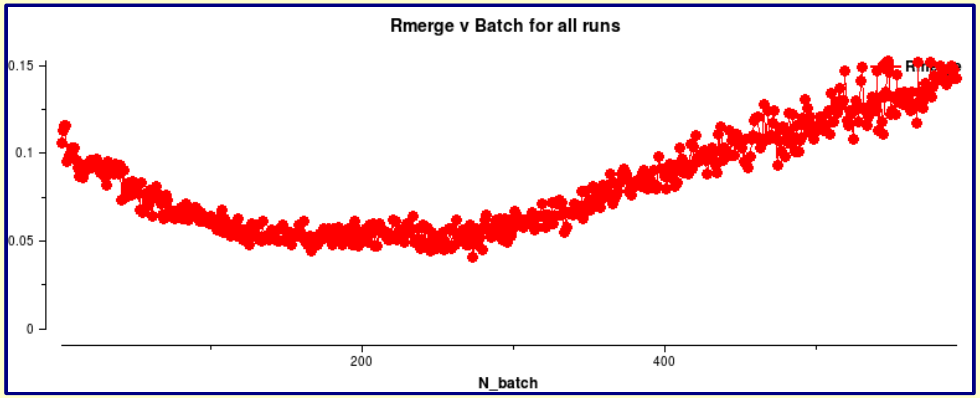
- 2vb1
  - High resolution
  - P1 Lysozyme
- 2qvo
  - Medium resolution
  - S-SAD phasing
- 3csl
  - Low resolution
  - Se-MET
  - Difficult spots

- 1y13
  - Medium resolution
  - 2-wvl MAD
  - Small phasing signal

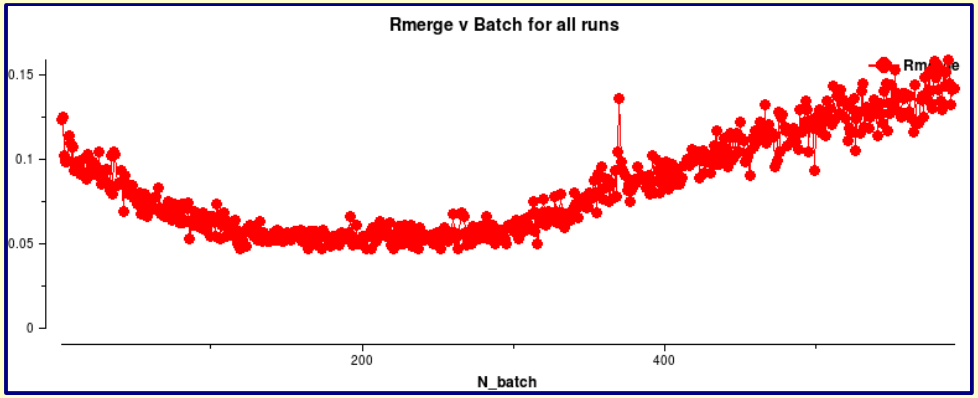
http://www.globalphasing.com/autoproc/wiki/index.cgi?ACA2011Tutorial
http://www.globalphasing.com/sharp/wiki/index.cgi?ACA2011Tutorial

1Y13 - autoPROC

infl



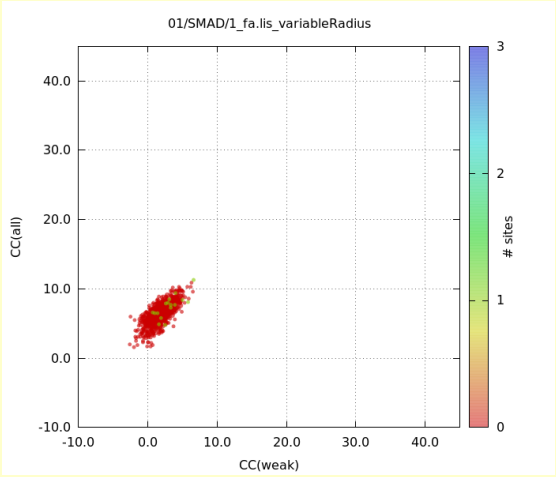
hrem



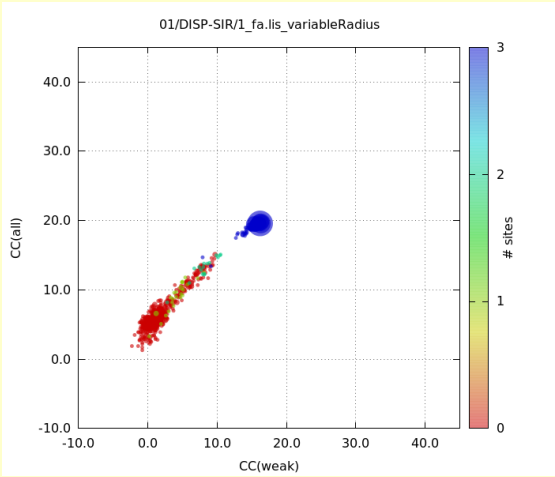
Radiation damage?

1Y13 – anomalous/dispersive differences

Anomalous 2.8Å

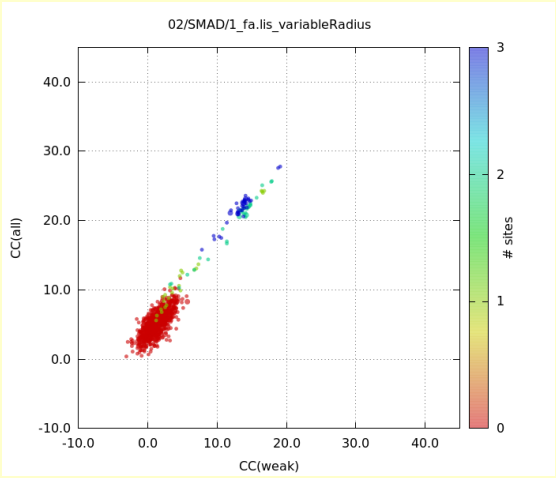


Dispersive 2.8Å

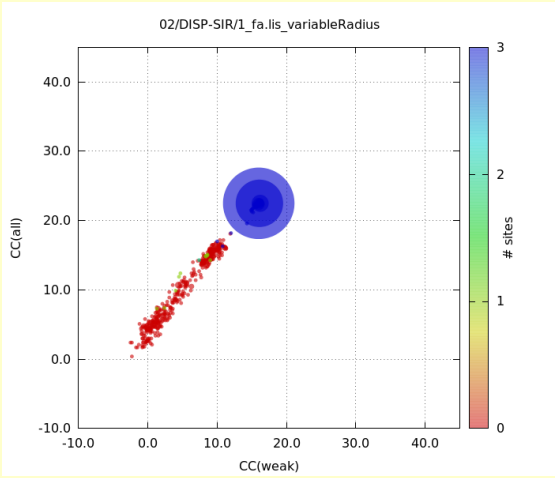


No 0dose

Anomalous 2.8Å



Dispersive 2.8Å

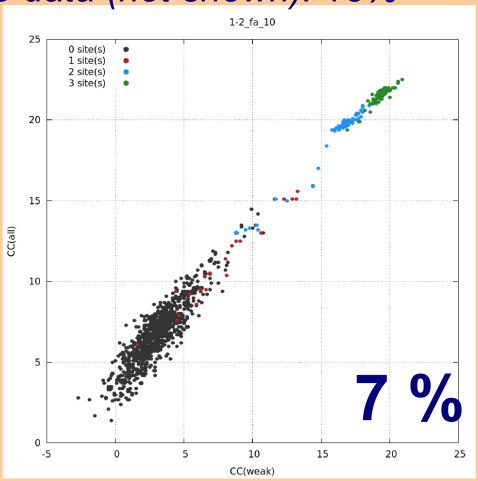


0dose

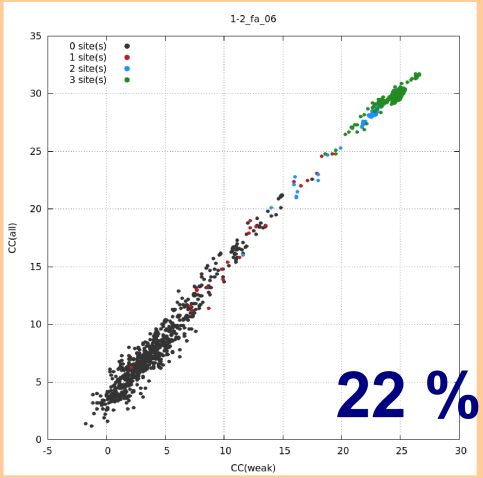
1Y13 - SHELXD (MAD) success rate

AIMLESS data (not shown): 13%

XSCALE no 0-dose

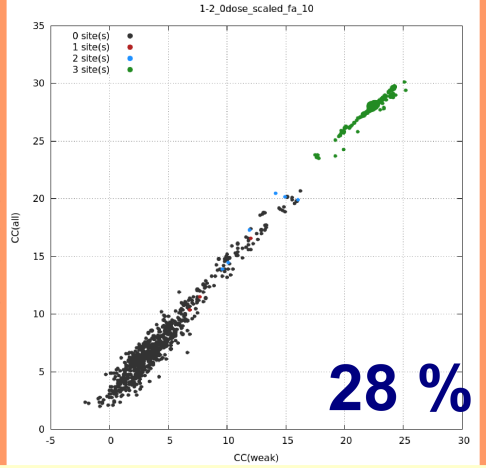


XSCALE 0-dose

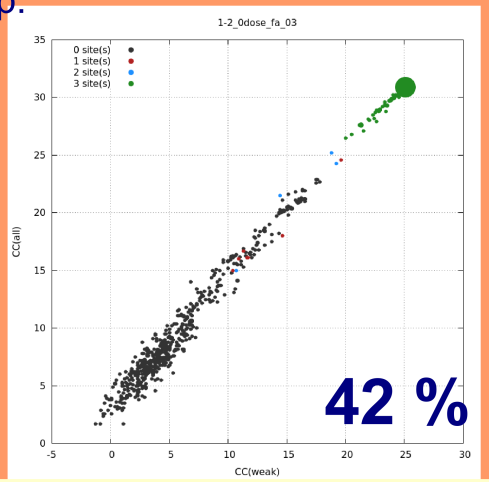


single 0dose on combined datasets with correct dose-stamp:

XSCALE* 0-dose



AIMLESS-XSCALE* 0-dose



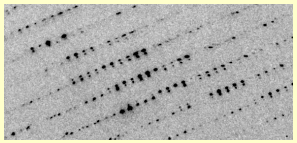
Session example datasets

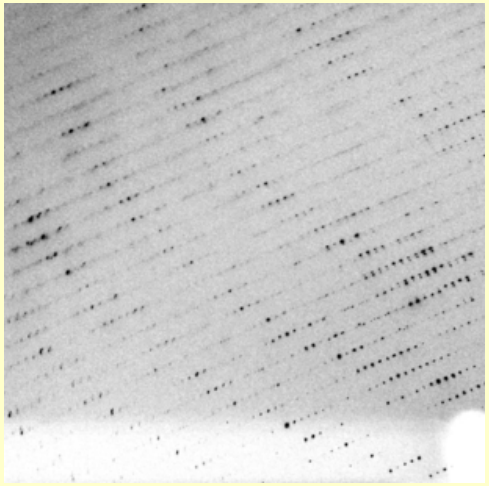
- 2vb1
  - High resolution
  - P1 Lysozyme
- 2qvo
  - Medium resolution
  - S-SAD phasing
- 3csl
  - Low resolution
  - Se-MET
  - Difficult spots
- 1y13
  - Medium resolution
  - 2-wvl MAD
  - Small phasing signal

http://www.globalphasing.com/autoproc/wiki/index.cgi?ACA2011Tutorial
http://www.globalphasing.com/sharp/wiki/index.cgi?ACA2011Tutorial

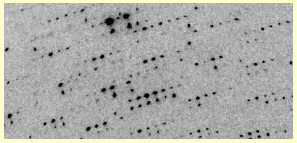
3CSL - autoPROC

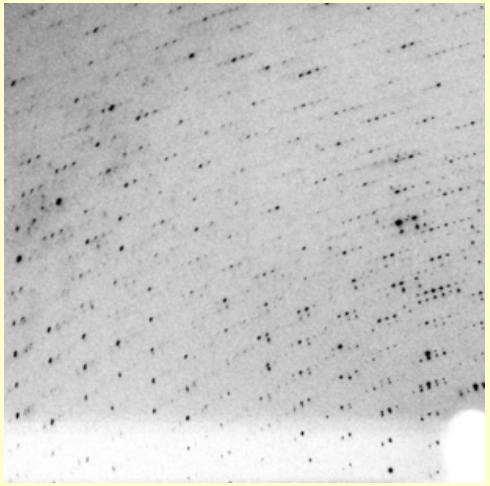
SK3_2_001.img



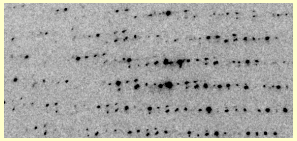


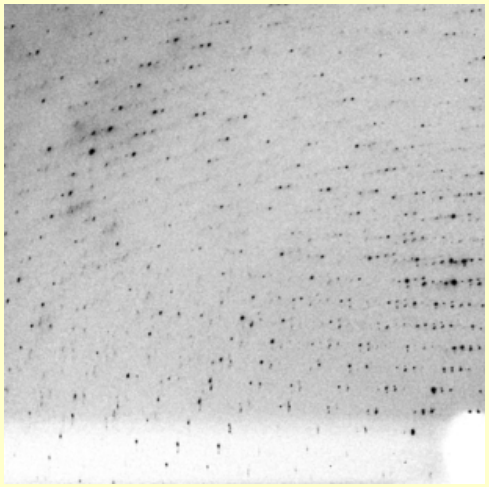
SK3_2_061.img



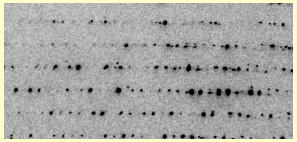


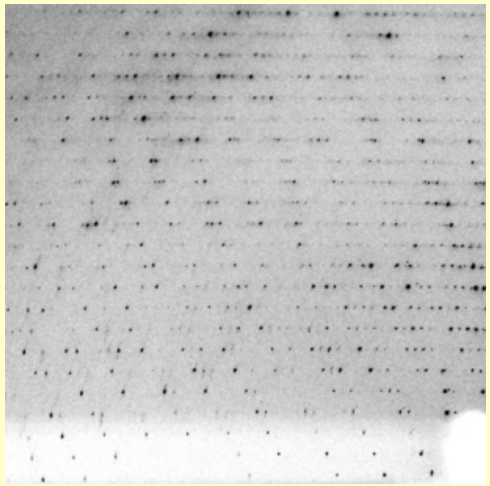
SK3_2_121.img





SK3_2_181.img





- Simple processing with autoPROC:

```
% process -d 01 | tee 01.lis
```

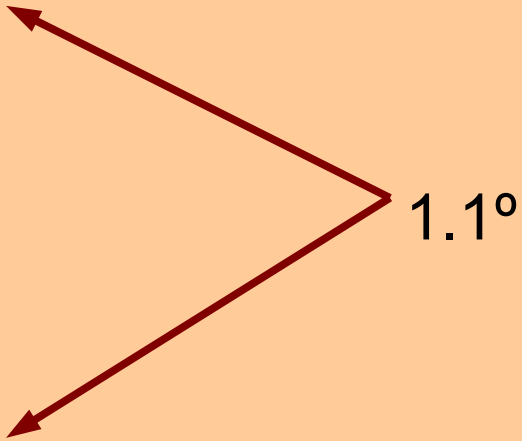
	Overall			InnerShell			OuterShell		
Low resolution limit	68.142	68.142	68.142	68.142	68.142	68.142	3.200	3.494	3.340
High resolution limit	**3.189**	**3.483**	**3.330**	14.752	14.752	14.752	3.189	3.483	3.330
Rmerge	**0.103**	**0.126**	**0.148**	0.034	0.061	0.046	0.296	0.222	0.312
Ranom	0.090	0.111	0.137	0.025	0.051	0.039	0.268	0.202	0.308
Rmeas (within I+/I-)	0.112	0.132	0.160	0.032	0.060	0.046	0.347	0.242	0.377
Rmeas (all I+ & I-)	0.114	0.137	0.160	0.038	0.067	0.051	0.345	0.242	0.347
Rpim (within I+/I-)	0.065	0.070	0.083	0.019	0.032	0.025	0.217	0.132	0.215
Rpim (all I+ & I-)	**0.049**	**0.053**	**0.061**	0.017	0.028	0.021	0.172	0.096	0.148
Total number of observations	332783	309564	363672	1878	4125	4079	1459	698	461
Total number unique	62862	47265	54321	408	697	696	425	111	92
Mean(I)/sd(I)	**14.7**	**13.9**	**13.3**	37.0	29.7	36.1	3.8	8.4	5.6
Completeness	**98.1**	**95.7**	**96.4**	56.7	96.9	96.8	68.2	22.7	17.4
Multiplicity	5.3	6.5	6.7	4.6	5.9	5.9	3.4	6.3	5.0
Anomalous completeness	**95.6**	**93.4**	**93.8**	55.3	95.3	93.2	57.8	22.6	16.7
Anomalous multiplicity	2.8	3.4	3.5	2.7	3.5	3.5	1.9	3.2	2.6

hrem peak infl

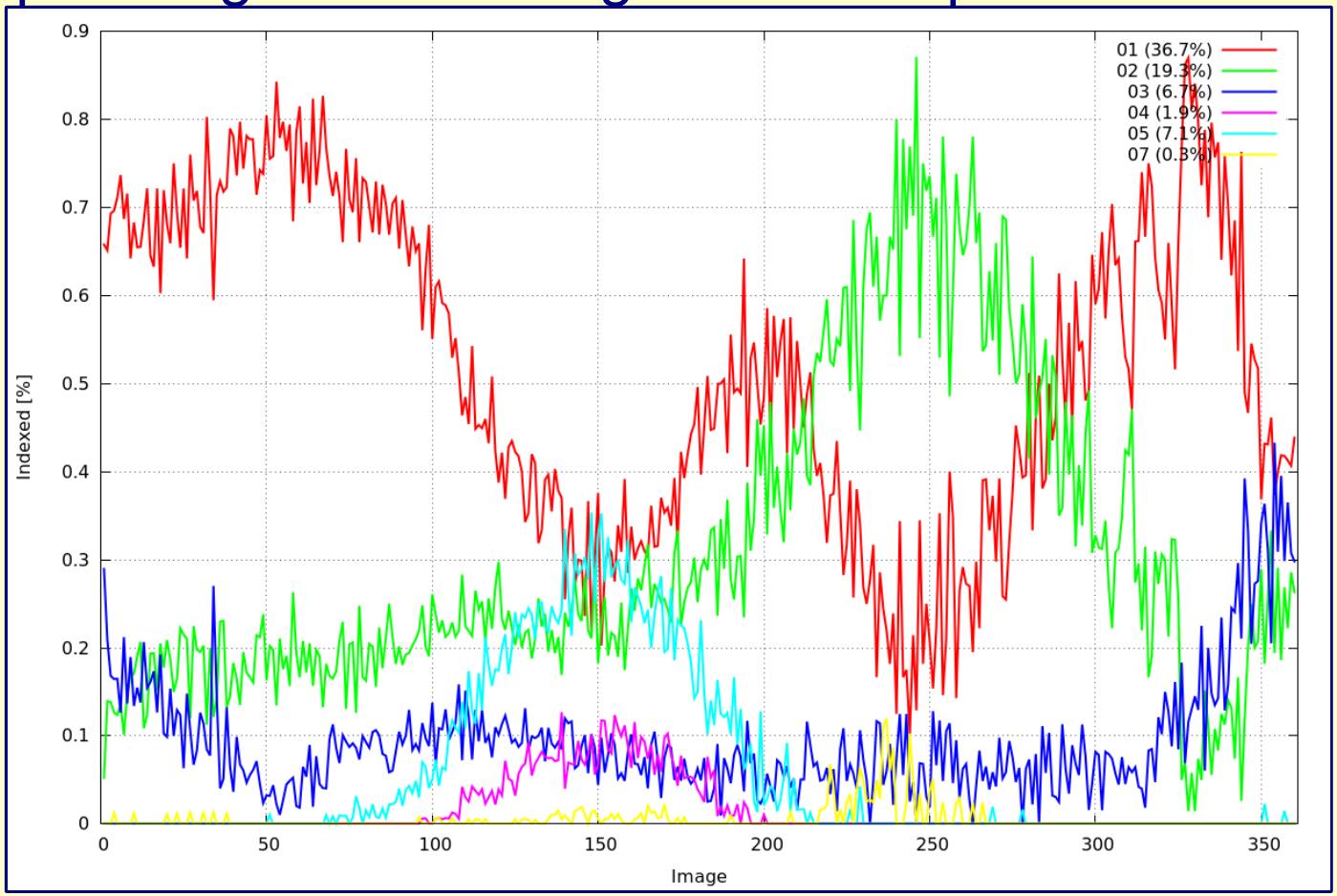
3CSL - autoPROC

- Dealing with **multiple lattices**: *iterative indexing* (automatic or user-defined)

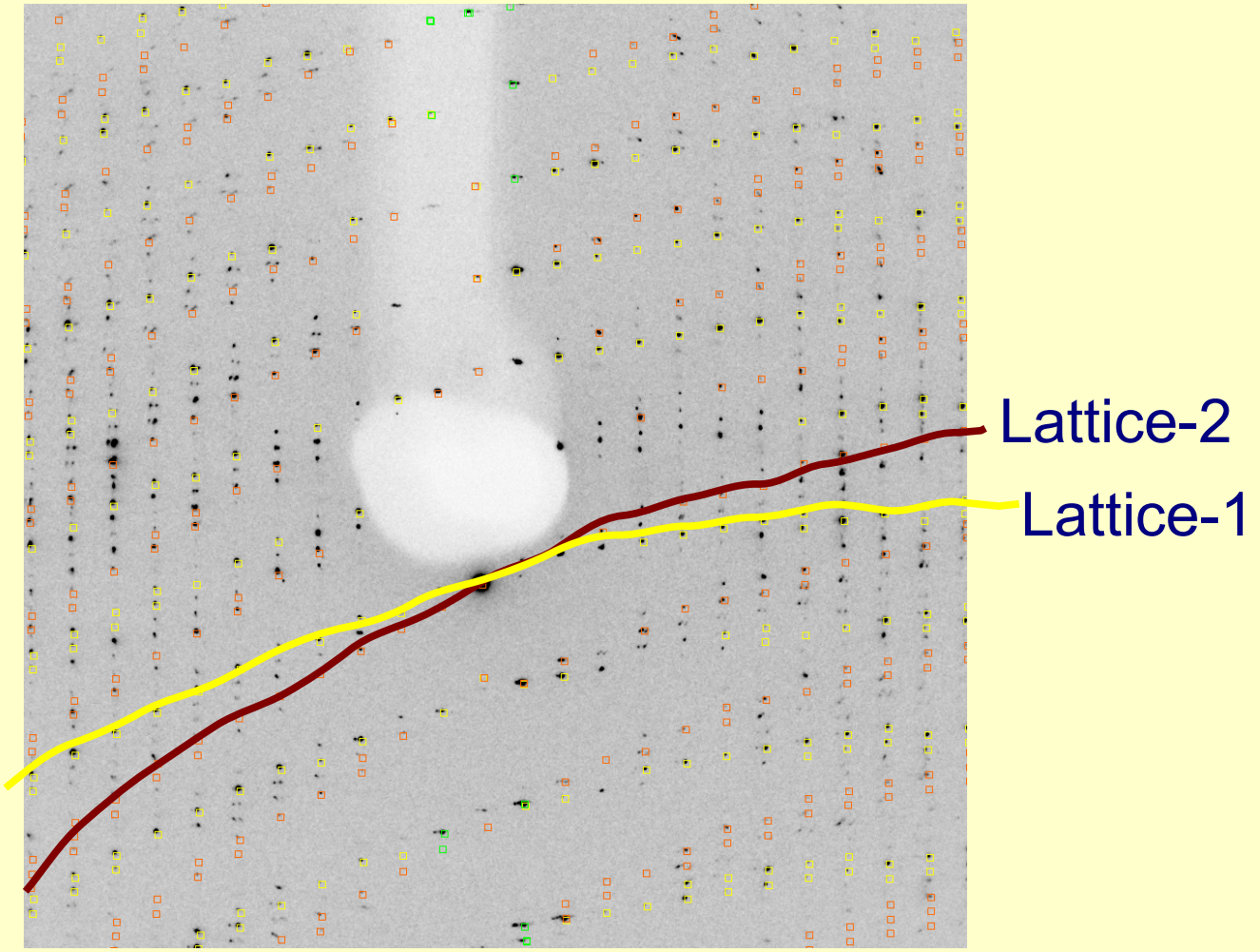
```
starting from a total of 47287 spots!

############################### 1 ################################
  indexing: 25913/47287 (54.8%)  24815/25913 (52.5%)  24624/24815 (52.1%)  24561/24624 (51.9%)  done

     24561 spots: <spot strength> = 110.5  with  sd = 814.9
REFINED VALUES OF DIFFRACTION PARAMETERS DERIVED FROM  24535 INDEXED SPOTS
STANDARD DEVIATION OF SPOT    POSITION (PIXELS)     1.27
STANDARD DEVIATION OF SPINDLE POSITION (DEGREES)    0.60
DETECTOR COORDINATES (PIXELS) OF DIRECT BEAM    1536.63   1519.09
DETECTOR ORIGIN (PIXELS) AT                     1526.35   1536.20
CRYSTAL TO DETECTOR DISTANCE (mm)       370.29
COORDINATES OF UNIT CELL A-AXIS    41.594   128.226   -80.319
COORDINATES OF UNIT CELL B-AXIS    45.885    72.403   139.350
COORDINATES OF UNIT CELL C-AXIS   550.349  -220.326   -66.740
UNIT CELL PARAMETERS    156.917   163.603   596.558  90.000  90.000  90.000

############################### 2 ################################
  indexing: 11070/21374 (23.4%)  10995/11070 (23.3%)  10982/10995 (23.2%)  done

     10982 spots: <spot strength> = 66.6  with  sd = 99.4
REFINED VALUES OF DIFFRACTION PARAMETERS DERIVED FROM  10955 INDEXED SPOTS
STANDARD DEVIATION OF SPOT    POSITION (PIXELS)     1.37
STANDARD DEVIATION OF SPINDLE POSITION (DEGREES)    0.47
DETECTOR COORDINATES (PIXELS) OF DIRECT BEAM    1536.44   1519.02
DETECTOR ORIGIN (PIXELS) AT                     1523.24   1535.93
CRYSTAL TO DETECTOR DISTANCE (mm)       372.26
COORDINATES OF UNIT CELL A-AXIS    41.654   132.712   -86.420
COORDINATES OF UNIT CELL B-AXIS    46.008    72.124   132.933
COORDINATES OF UNIT CELL C-AXIS   553.123  -220.399   -71.857
UNIT CELL PARAMETERS    163.755   158.082   599.737  90.000  90.000  90.000
```



Spot-usage versus image number - peak

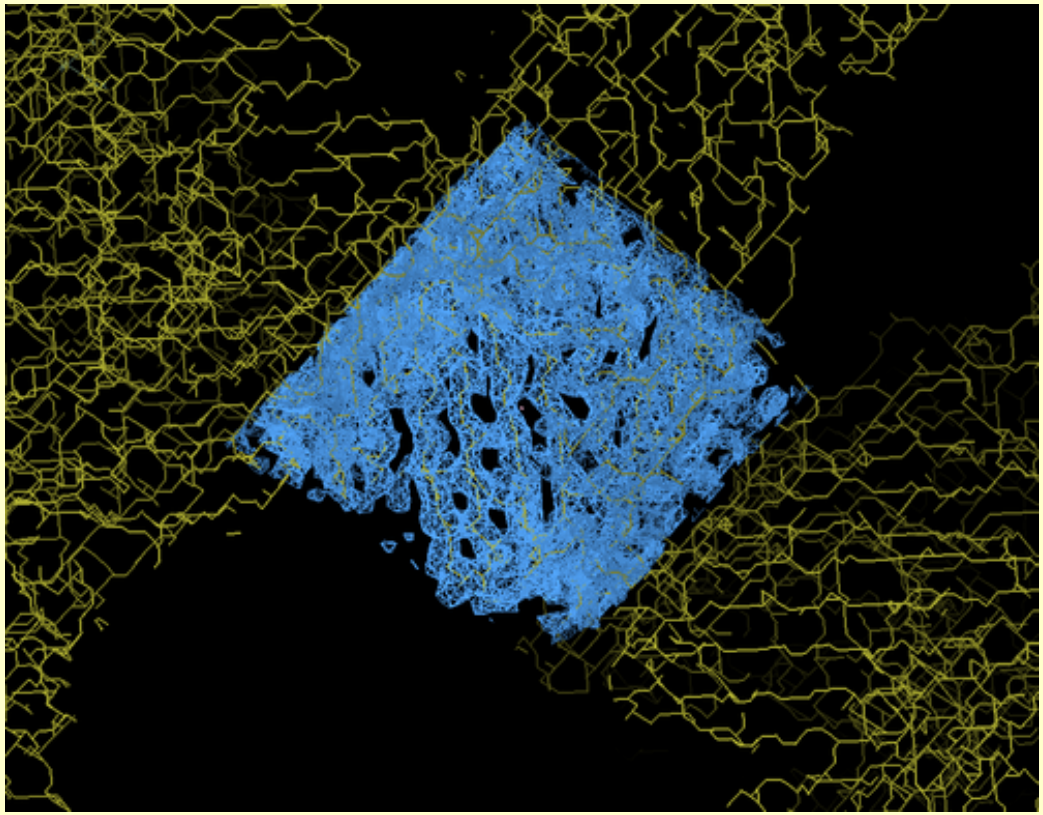


3CSL - autoPROC



3CSL - autoSHARP

- As a result we get a map to look at:

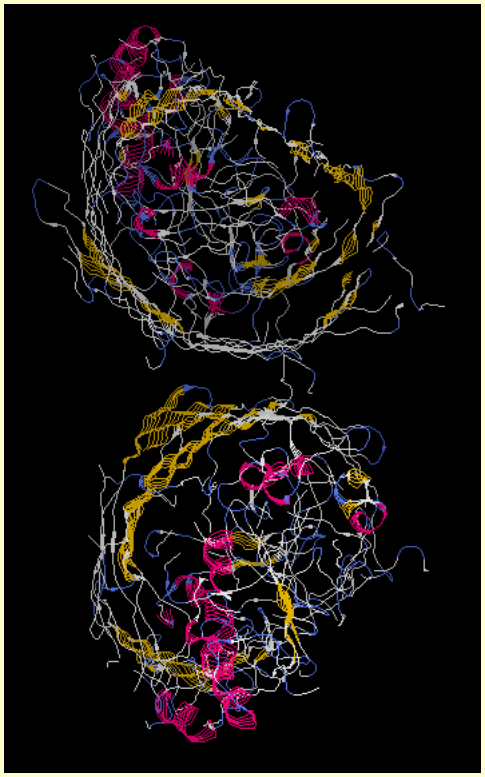


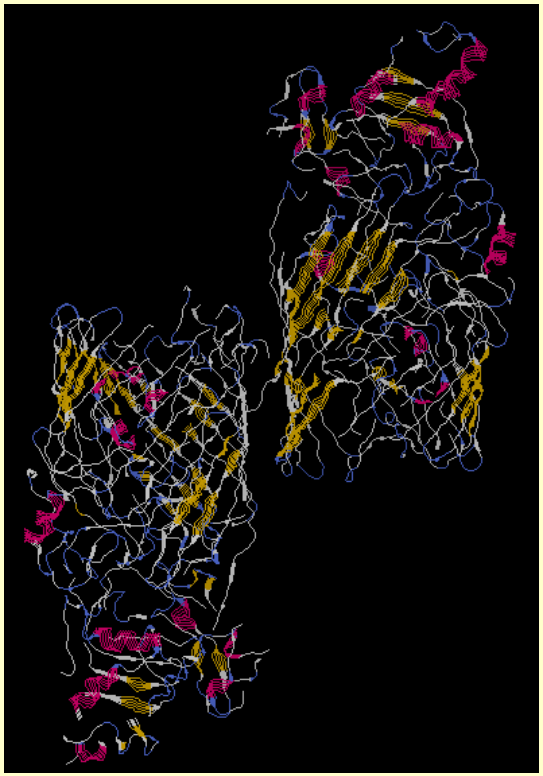
- **Can we built into it?**

3CSL - building

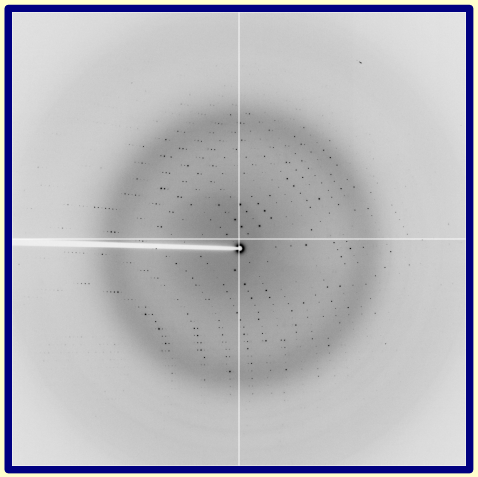
After 30 minutes using Pirate/Buccaneer (K. Cowtan) via autoSHARP tool:

- 1838 residues with 902 docked (out of 2142 according to HasR and HasA sequence)
- PDB entry contains 1832 residues.

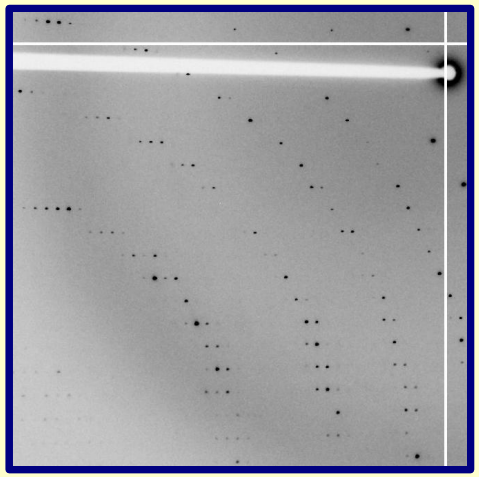




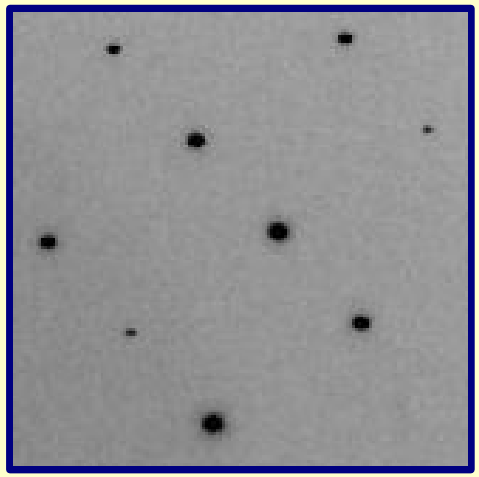
Expectations



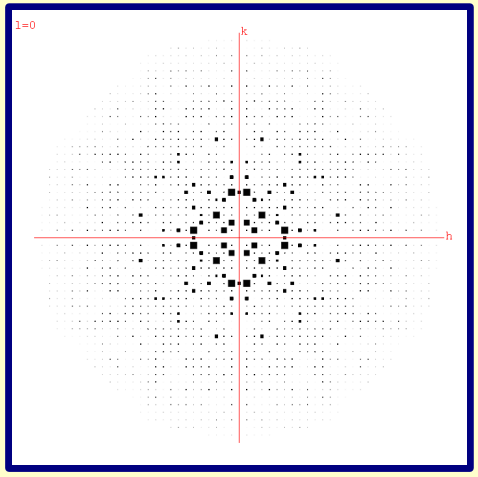
nice diffraction



separated lunes



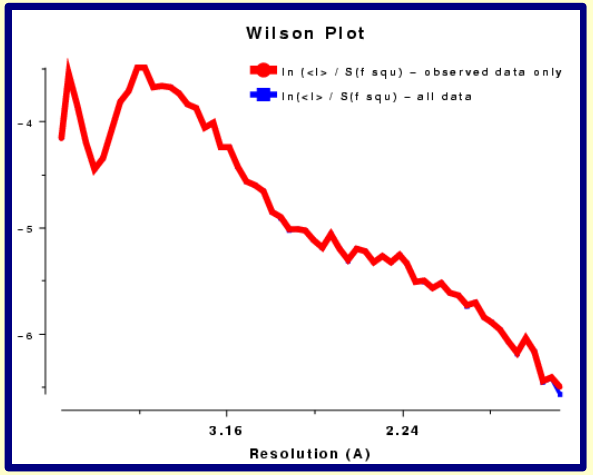
perfect spots



completeness

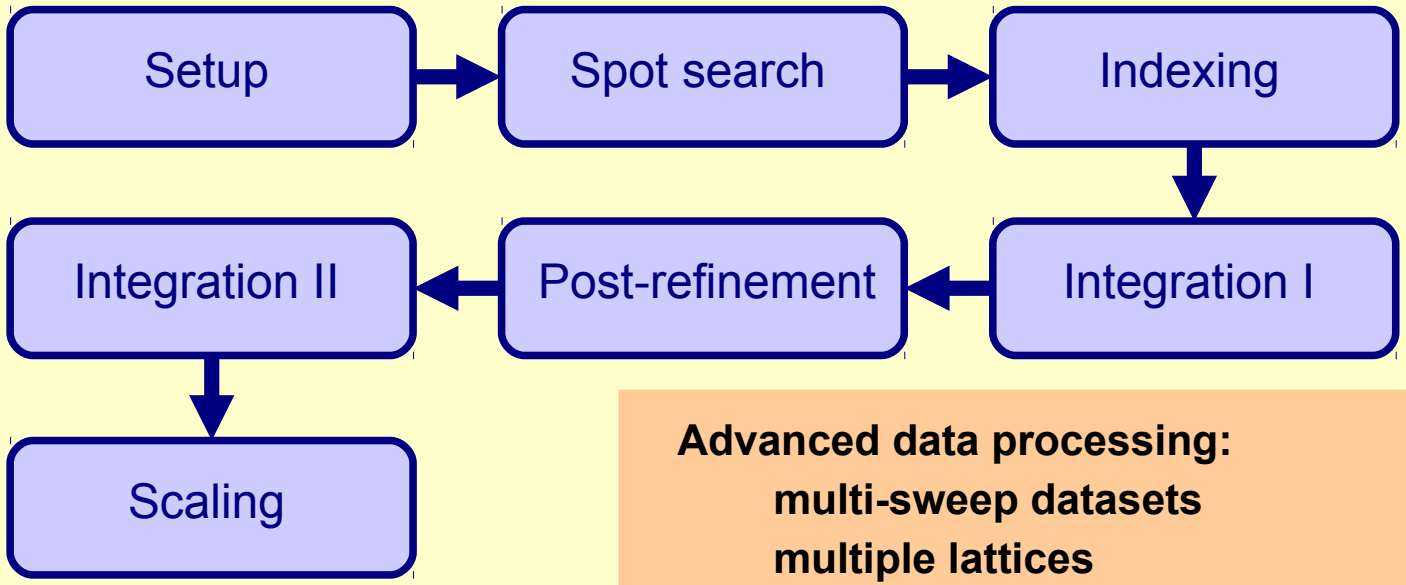
ideal world

well behaved statistics



autoPROC

Typical steps for a single set of consecutive images:



autoPROC (1) adds expert knowledge at and between each step and (2) combines separate sweeps/datasets

Advanced data processing:

- **multi-sweep datasets**
- **multiple lattices**
- **multi-axis goniostats**
- **ice-rings**
- **resolution limits, cut-offs etc**
- **reporting (HTML, XML)**

autoPROC - behind the scenes

*Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T. & Bricogne, G. (2011). Data processing and analysis with the **autoPROC** toolbox. Acta Cryst. D67, 293-302.*

- Uses external programs
  - **XDS/XSCALE**: *Kabsch, W. XDS. Acta Cryst. D66, 125-132 (2010).*
  - **POINTLESS, AIMLESS**: Evans, P. (2006). Scaling and assessment of data quality. Acta Cryst. D62, 72-82.
  - **CCP4**: *Collaborative Computational Project, Number 4. 1994. "The CCP4 Suite: Programs for Protein Crystallography". Acta Cryst. D50, 760-763.*

Summary

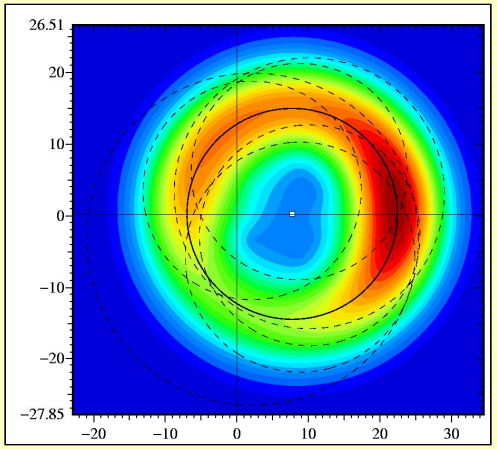
- Automation is great ... if it works
- Meaningful feedback important!
- Improving on initial results often involves going back to scaling and/or data processing and/or collection strategy.

- Think about **differences** (and how to measure, scale, analyse and describe them): anomalous, dispersive, isomorphous, radiation damage, MAD+native etc.

- It's all about **signal-over-noise**: HA signal, noise in data … MIR still very useful … so is MAD

- Nomenclature can be confusing between programs

- Goal: best set of phases from well measured and processed data



2D phase probability in SHARP

Acknowledgement

Global Phasing, Cambridge (UK):

- Gérard Bricogne, Claus Flensburg, Andrew Sharff, Wlodek Pacziorek, Peter Keller
- Previous group-members: M. Schiltz, E. Blanc, P. Roversi, G. Evans, R. Morris, T. Womack, Oliver Smart

http://www.globalphasing.com/ (also: BUSTER)
http://www.globalphasing.com/autoproc/wiki/
http://www.globalphasing.com/sharp/wiki/

Tutorials, FAQ, documentation, examples, ...

ARP/WARP (Tassos Perrakis, Victor Lamzin, Richard Morris)
PARROT/BUCCANEER (Kevin Cowtan)
SHELXC/SHELXD: (George Sheldrick)
CCP4
MOSFLM (Andrew Leslie & Harry Powell)
SCALA/POINTLESS/AIMLESS (Phil Evans)
XDS (Wolfgang Kabsch & Kay Diederichs)

Test data: C. Yang/J. Pflugrath (Rigaku/MSC), Z. Dauter, CCP4/AutoStruct, **JCSG**, Global Phasing Consortium members, SLS/PSI (CH), Soleil/Proxima1 (F), DLS (UK), ESRF (F), many users