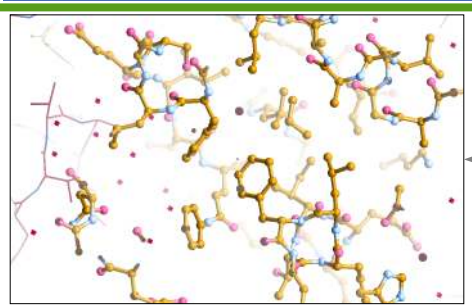
Experimental Phasing with SHARP/autoSHARP

Clemens Vonrhein
Global Phasing Ltd.

DLS/CCP4, 2021

GΦL
Global Phasing Limited

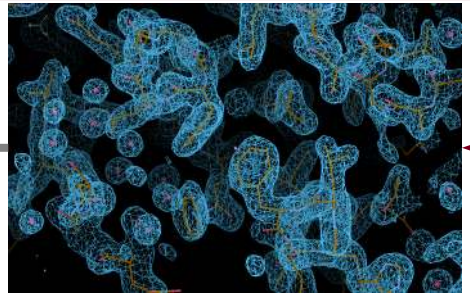
Where are we in the process?

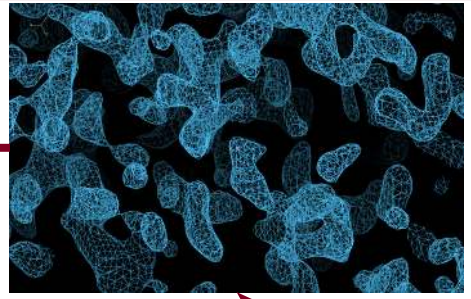


BUSTER

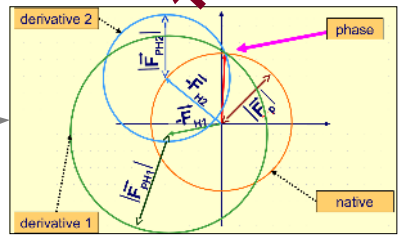
REFMAC Phenix

Crank2
hkl2map
Solve
shelxc/d/e
Phaser
...

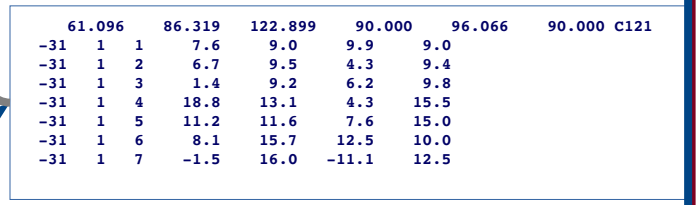




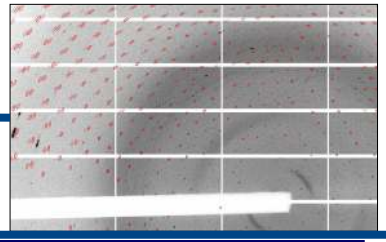
SHARP/autoSHARP



XDS
Dials
HKL
Xia2
iMOSFLM
fast_dp
...



	61.096		86.319	122.899	90.000	96.066	90.000 C121
-31	1	1	7.6	9.0	9.9	9.0	
-31	1	2	6.7	9.5	4.3	9.4	
-31	1	3	1.4	9.2	6.2	9.8	
-31	1	4	18.8	13.1	4.3	15.5	
-31	1	5	11.2	11.6	7.6	15.0	
-31	1	6	8.1	15.7	12.5	10.0	
-31	1	7	-1.5	16.0	-11.1	12.5	



autoPROC

GΦL
Global Phasing Limited

SHARP/autoSHARP Design

SHARP:

- getting **best phases** possible
- applicable to large variety of **phasing scenarios**: MAD+native, SAD+SIR, RIP, ... and variations
- *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.

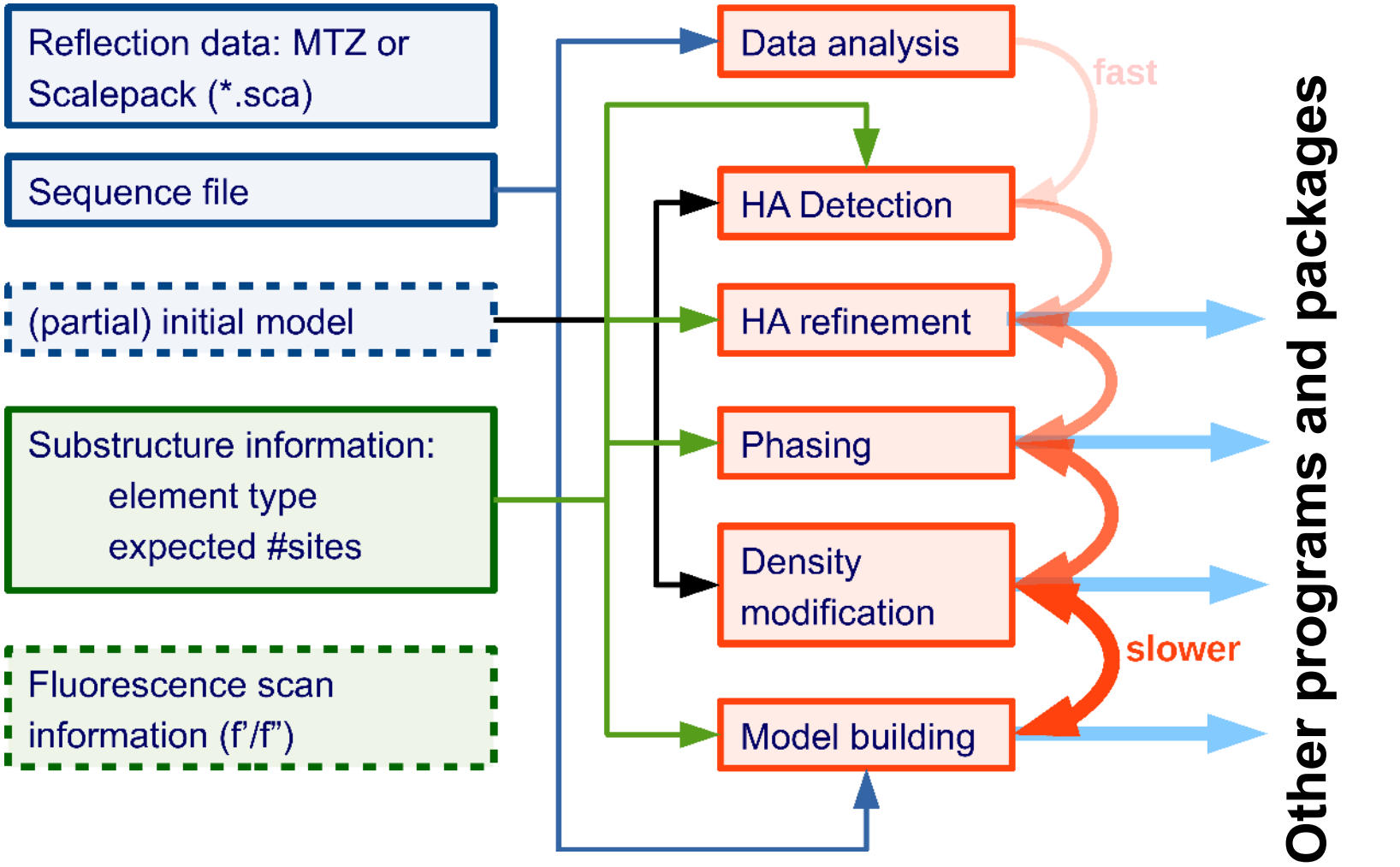
autoSHARP:

- 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.

GΦL
Global Phasing Limited

What do I need as input and what do I get on output?



Global Phasing Limited

Command-line interface

run_autoSHARP.sh -h

writes HTML output (view in browser)

creates scripts ready to start in 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
```

https://www.globalphasing.com/sharp/wiki/index.cgi?TutorialRapiData2021

GΦL
Global Phasing Limited

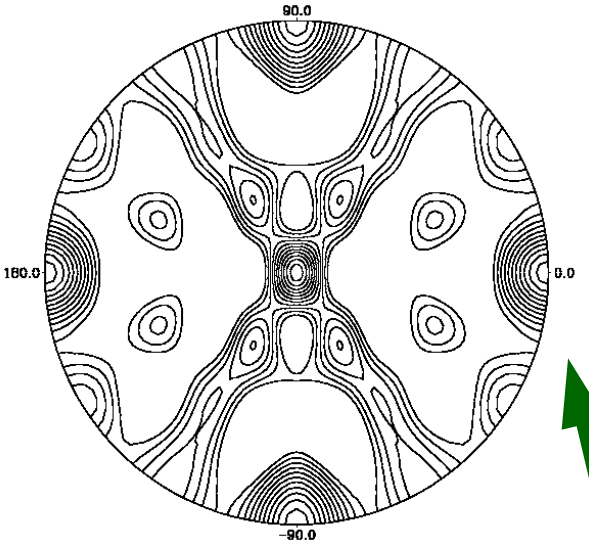
autoSHARP - Analysis

- Consistency (more than 1 dataset):
  - cell (significant changes)
  - spacegroup (screw axes)
  - indexing (alternate possibilities)
  - sequence (Se-Met, S-SAD)
- Data quality:
  - CC(Δano), R-values, completeness
  - outliers (E-values)

SIGNAL

NOISE

- self-rotation function



- native Patterson
- Matthews parameters (probability)

?/ASU

autoSHARP – HA detection

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

2.1.1 SHELXC (details)

105261 Reflections read from SAD file 1.sca

54117 Unique reflections, highest resolution 1.849 Angstroms
154.7 Friedel pairs used on average for local scaling

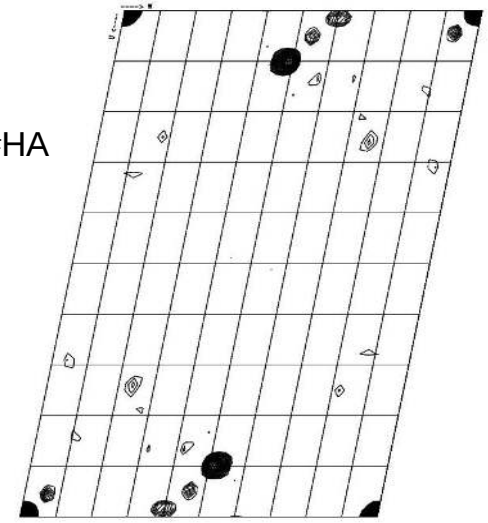
Resl.	Inf.	9.91	6.10	4.59	3.76	3.21	2.83	2.54	2.31	2.13	1.98	1.85
N(data)	588	1899	3290	4269	5741	7698	8074	10398	7161	4070	929	
<I/sig>	**51.3**	**46.0**	**42.0**	**39.8**	26.9	16.0	7.7	4.2	2.6	1.7	1.4	
%Complete	96.2	99.4	98.2	90.2	90.4	99.7	86.1	93.6	58.0	29.0	5.7	
Multipl.	1.8	1.9	1.9	1.9	1.9	2.0	1.9	2.0	1.9	1.9	1.9	
R(pim)%	2.78	3.62	3.45	2.83	3.81	6.04	12.86	24.07	33.92	44.73	56.94	
Ranom%	5.55	7.24	6.90	5.66	7.63	12.08	25.72	48.15	67.84	89.45	113.9	
<d"/sig>	**2.43**	**2.50**	**2.09**	**1.55**	**1.33**	1.11	0.92	0.83	0.75	0.69	0.73	

For zero signal <d'/sig> and <d"/sig> should be about 0.80

NOTE : suggesting high-resolution cut-off of 3.21 Å

Harker section

tricky with large #HA

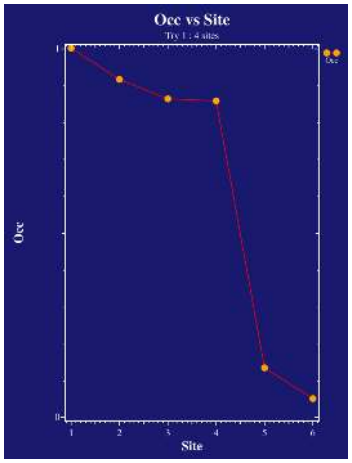


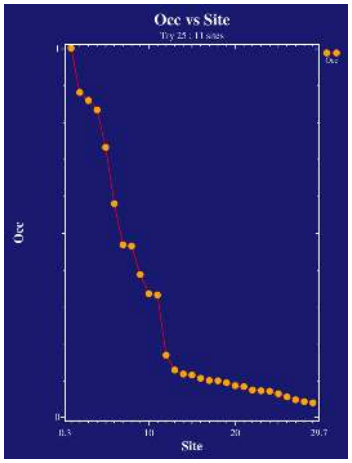
- **SHELXC/SHELXD** [1] :
  - stop if significant solution found
  - present results graphical
  - try detecting ‘jump’ in occupancy

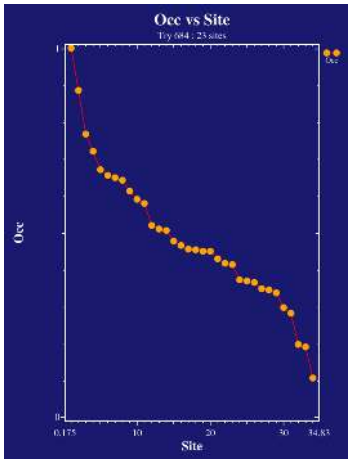
Alternative: give your own set of initial sites to autoSHARP

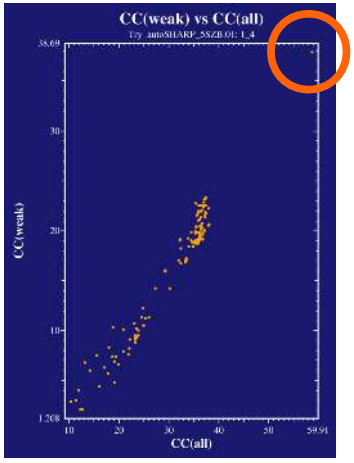
1) Schneider, T. R. & Sheldrick, G. M. (2002). Acta Cryst. D58, 1772-1779.

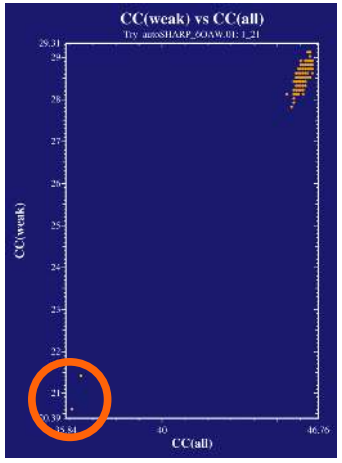
HA detection - does it look promising?

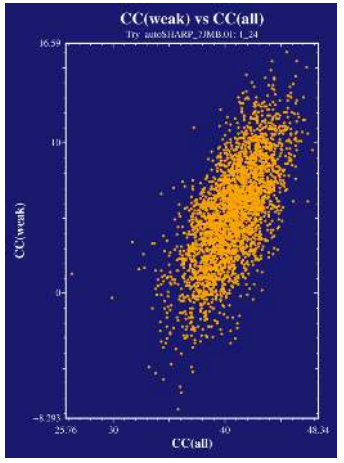












GΦL
Global Phasing Limited

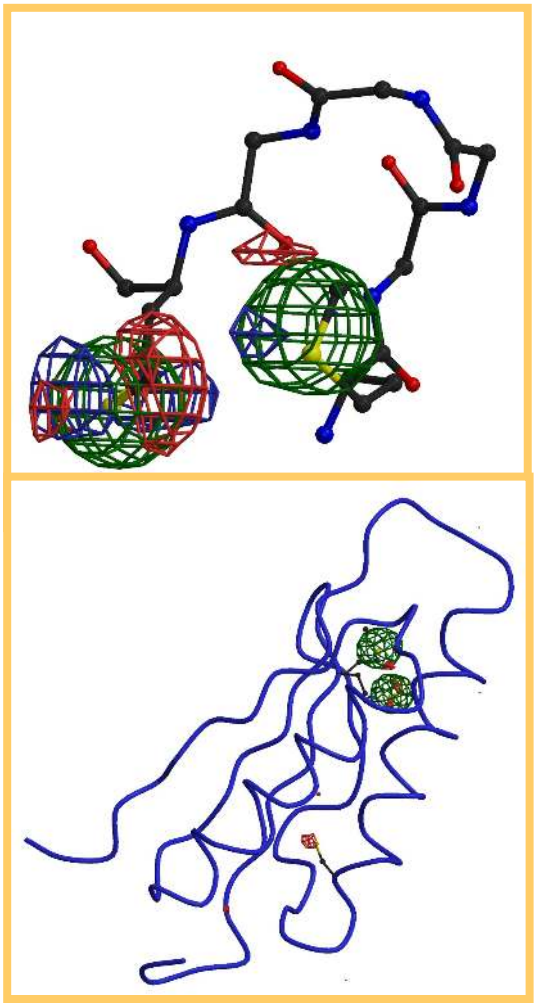
SHARP

- HA refinement and phasing
- overall anisotropic scaling
- analysis of log-likelihood gradient 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)



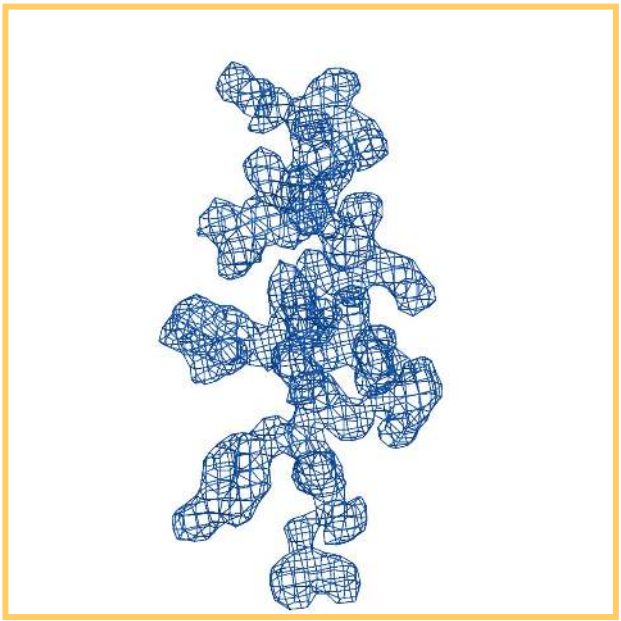
Goal: best possible phases

GΦL
Global Phasing Limited

autoSHARP – density modification

- **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 no. of molecules/ASU

Similar to MR (test enantiomorphs)



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.sh -h`*)

- 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

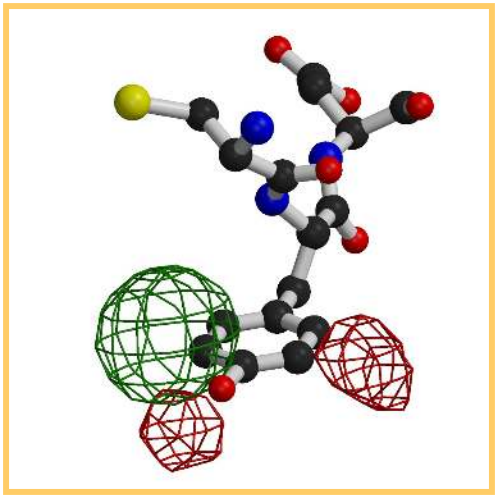
You could use "**-nowarp**" or "**-nobuild**" option to **`run_autoSHARP.sh`**.

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

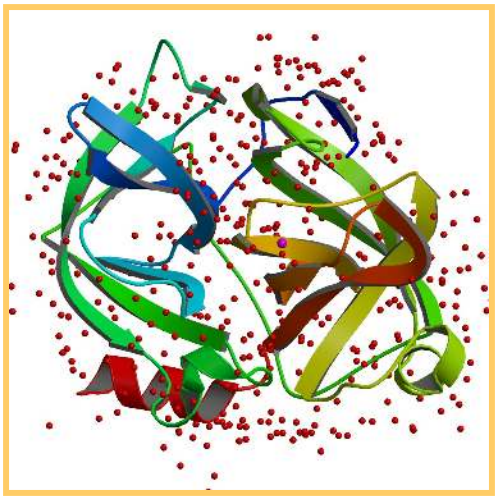
GΦL
Global Phasing Limited

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 (→ e.g. ACORN)



- **SHARP** phases (→ e.g. SHELXE building)
- LLG (residual) maps



- density modified phases/maps (→ e.g. NCS-detection via GETAX)
- initial model(s) (→ completion, refinement etc)

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**[2])
  - 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

"web-based" interface (local installation)

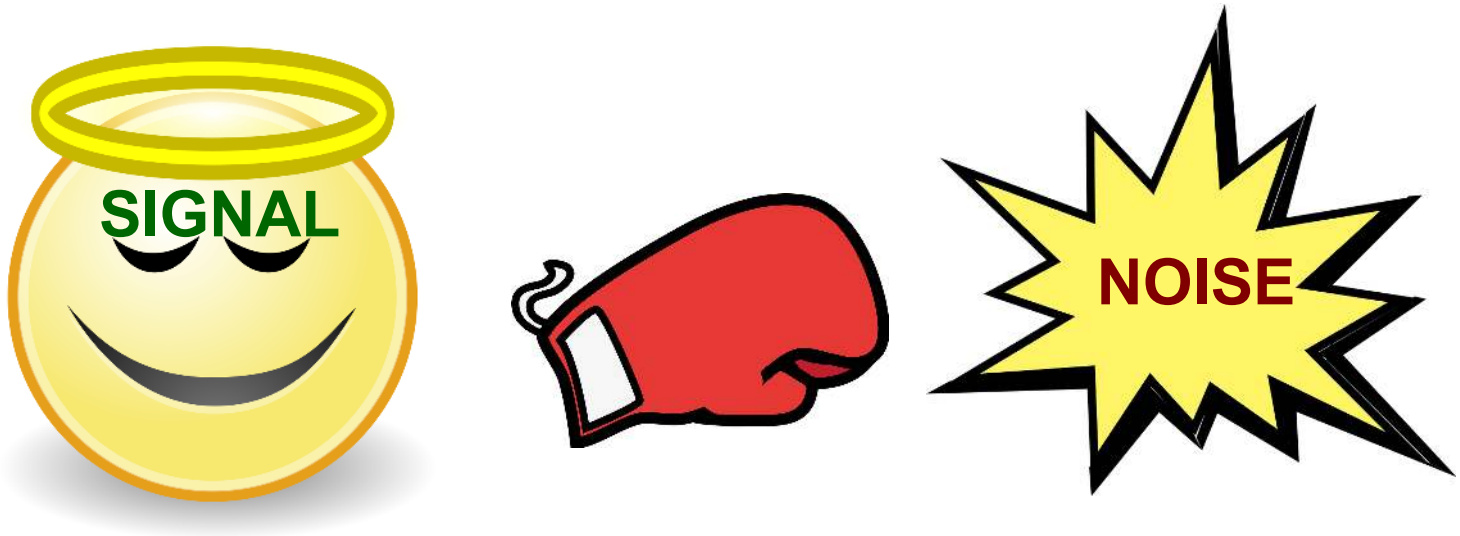
1) C. Vonrhein and G. E. Schulz, Acta Cryst., D55, 225 - 229 (1999)

2) K. Cowtan (1994), Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography, 31, p34-38.

GΦL
Global Phasing Limited

How to improve on first results

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



GΦL
Global Phasing Limited

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).
```

```
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** *

GΦL
Global Phasing Limited

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.
```

Overloads?

```
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 *
```

Beamstop masking?

GΦL
Global Phasing Limited

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** *

GΦL
Global Phasing Limited

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 *

Ice-rings?

GΦL
Global Phasing Limited

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.

GΦL
Global Phasing Limited

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
```

Beamstop?
Outlier rejection?

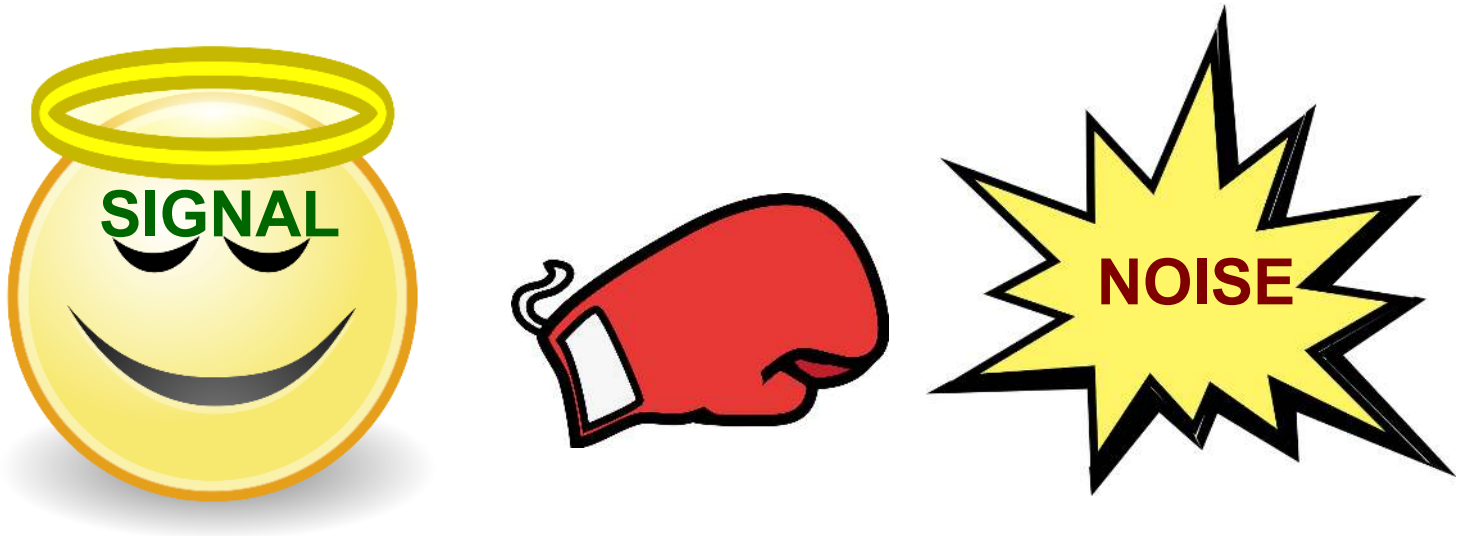
```
WARNING : the two hands don't show a significant difference in the two scores!
          This might be due to a very weak phasing signal, wrong heavy atom
          sites or something else (single site in polar spacegroups?). Please
          check your heavy atom sites solution carefully (does it 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.
```

Not good ... no solution?

GΦL
Global Phasing Limited

How to improve on first results

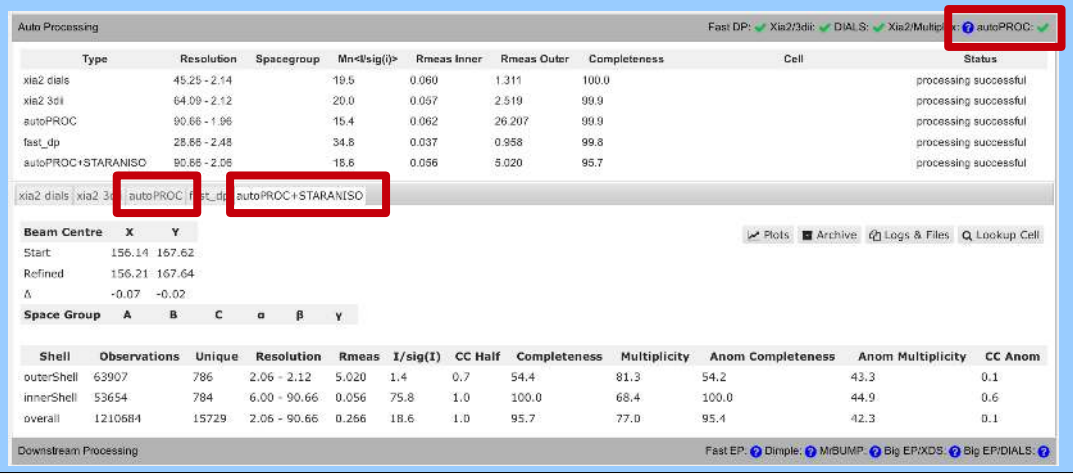
- after **autoSHARP**:
  - **warning messages** point back to data processing/scaling problems (beamstop, ice-rings, decay, …)

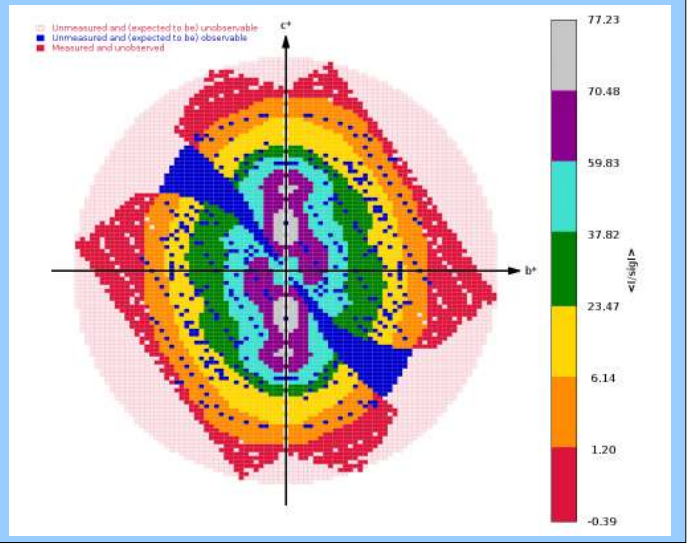


GΦL
Global Phasing Limited

How to improve on first results

- after **autoSHARP**:
  - **warning messages** point back to data processing/scaling problems (beamstop, ice-rings, decay, ...)
- **autoPROC (2005-)**: *www.globalphasing.com/autoproc/*
  - expert system (using XDS, POINTLESS, AIMLESS, XSCALE)
  - multiple lattices, multi-wavelength, multi-sweep, multi-orientation
  - anisotropy (STARANISO)





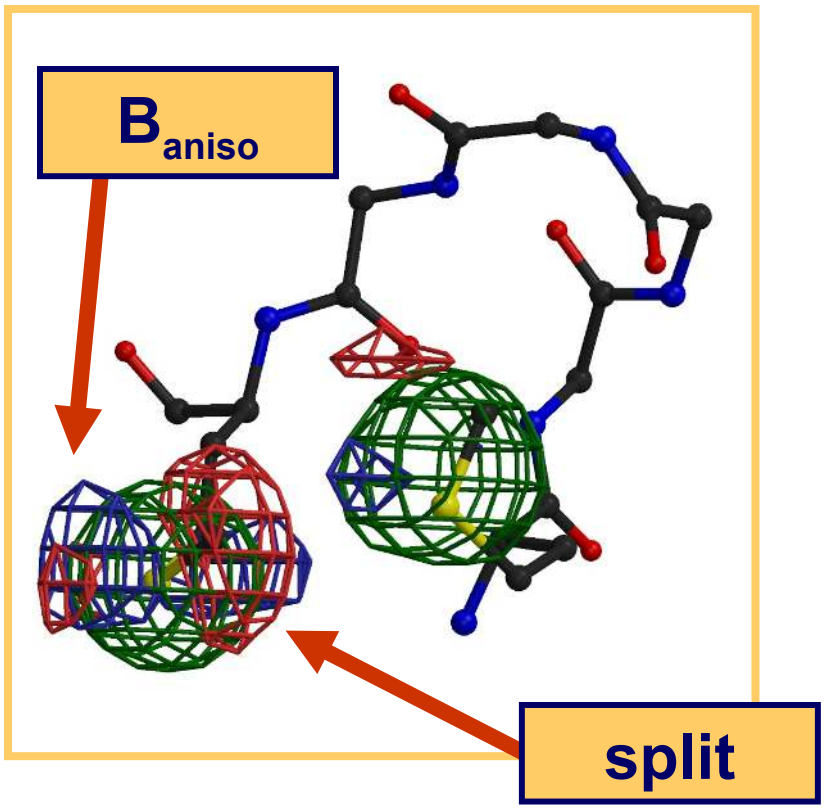
GΦL
Global Phasing Limited

How to improve on first results

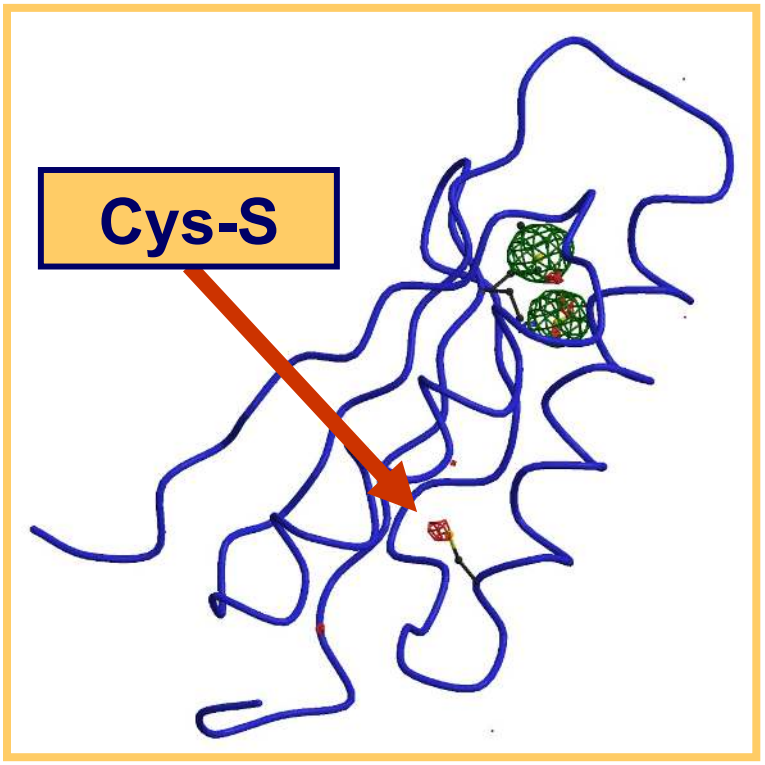
- after **autoSHARP**:
  - **warning messages** point back to data processing/scaling problems (beamstop, ice-rings, decay, ...)
  - analyse **log-likelihood gradient** (residual) maps: additional sites, f'/f'' refinement, anisotropic B-factor for sites, ...

GΦL
Global Phasing Limited

Example (LLG maps)



Se sites (LLG peak)



Anomalous LLG

GΦL
Global Phasing Limited

How to improve on first results

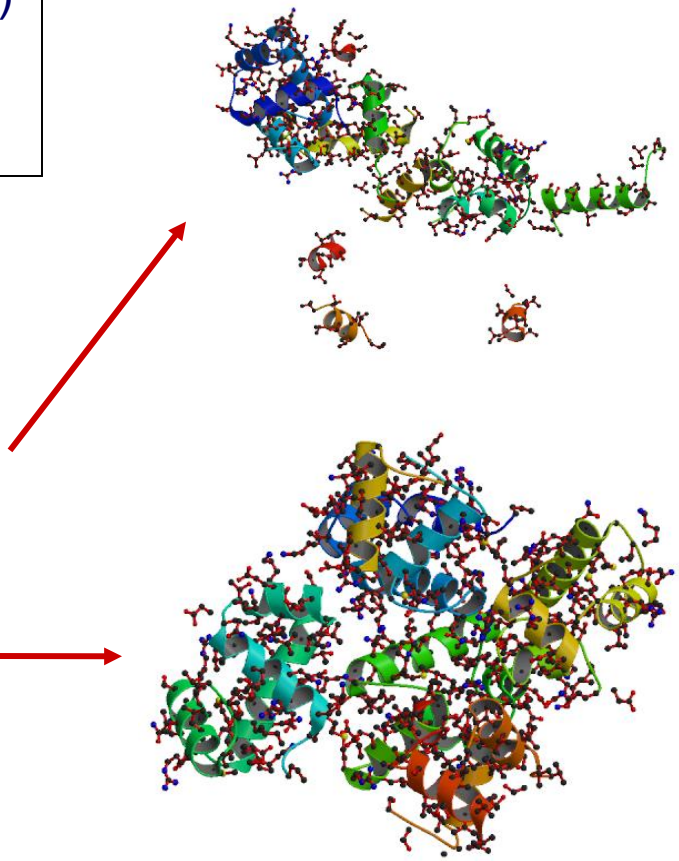
- after **autoSHARP**:
  - **warning messages** point back to data processing/scaling problems (beamstop, ice-rings, decay, ...)
  - analyse **log-likelihood gradient** (residual) maps: additional sites, f'/f'' refinement, anisotropic B-factor for sites, ...
  - significant difference between score of **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**

GΦL
Global Phasing Limited

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

GΦL
Global Phasing Limited

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 log-likelihood gradient maps → HA detection
  - Overcoming non-isomorphism issues

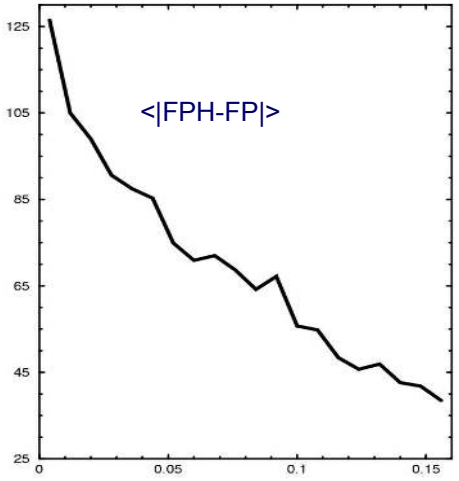
"MR-SAD"

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

GΦL
Global Phasing Limited

Why not MIR?

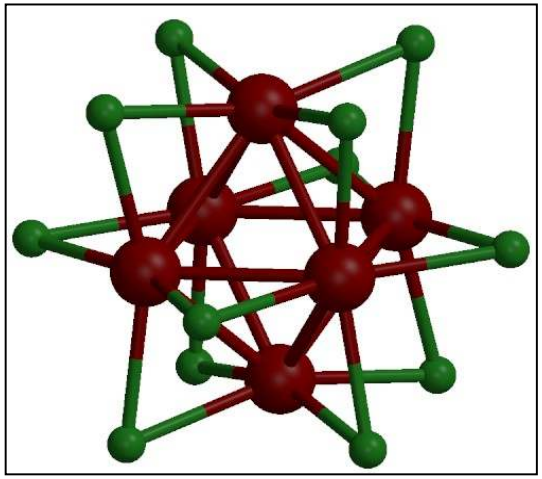
- **Non-isomorphism**
  - **datasets from different crystals**
  - **agressive soaks, co-crystallization**
- **no. of sites (and occupancy) not well known**
- **'old-fashioned'**



- **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}



GΦL
Global Phasing Limited

"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)

RESEARCH PAPERS

Acta Cryst. (2004). D60, 1024-1031
https://doi.org/10.1107/S0907444904006377

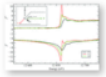
Phasing in the presence of severe site-specific radiation damage through dose-dependent modelling of heavy atoms

M. Schiltz, P. Dumas, E. Ennifar, C. Flensburg, W. Paciorek, C. Vonrhein and G. Bricogne

The case of a brominated RNA crystal structure determination in which standard three-wavelength MAD phasing was unsuccessful because of fast X-ray-induced debromination was reinvestigated [Ennifar *et al.* (2002), *Acta Cryst.* D58, 1262-1268]. It was found that if the data are kept unmerged and if a dose-stamp is associated with each reflection measurement, dose-dependent occupancies can be refined for the Br atoms. Such a parametrization has been implemented in the macromolecular phasing program *SHARP*. Refining such dose-dependent occupancies on an unmerged data set gave a dramatic improvement, even for SAD phases from only the first wavelength (peak), and resulted in a good electron-density map after solvent flattening. The adverse effect of radiation damage has been turned into a beneficial one. The crucial difference is made by the use of unmerged data: phasing power is generated through the intensity differences of symmetry-related reflections recorded at different doses, *i.e.* corresponding to different states of the X-ray-induced debromination. This approach should prove useful in all situations of experimental phasing where site-specific radiation damage occurs unavoidably and undesirably and not only in cases in which radiation damage is purposely being created in order to demonstrate its potential usefulness.

RESEARCH PAPERS

Acta Cryst. (2008). D64, 711-729
https://doi.org/10.1107/S0907444908010202



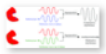
Exploiting the anisotropy of anomalous scattering boosts the phasing power of SAD and MAD experiments

M. Schiltz and G. Bricogne

The X-ray polarization anisotropy of anomalous scattering in crystals of brominated nucleic acids and selenated proteins is shown to have significant effects on the diffraction data collected at an absorption edge. For conventionally collected single- or multi-wavelength anomalous diffraction data, the main manifestation of the anisotropy of anomalous scattering is the breakage of the equivalence between symmetry-related reflections, inducing intensity differences between them that can be exploited to yield extra phase information in the structure-solution process. A new formalism for describing the anisotropy of anomalous scattering which allows these effects to be incorporated into the general scheme of experimental phasing methods using an extended Harker construction is introduced. This requires a paradigm shift in the data-processing strategy, since the usual separation of the data-merging and phasing steps is abandoned. The data are kept unmerged down to the Harker construction, where the symmetry-breaking is explicitly modelled and refined and becomes a source of supplementary phase information. These ideas have been implemented in the phasing program *SHARP*. Refinements using actual data show that exploitation of the anisotropy of anomalous scattering can deliver substantial extra phasing power compared with conventional approaches using the same raw data. Examples are given that show improvements in the phases which are typically of the same order of magnitude as those obtained in a conventional approach by adding a second-wavelength data set to a SAD experiment. It is argued that such gains, which come essentially for free, *i.e.* without the collection of new data, are highly significant, since radiation damage can frequently preclude the collection of a second-wavelength data set. Finally, further developments in synchrotron instrumentation and in the design of data-collection strategies that could help to maximize these gains are outlined.

RESEARCH PAPERS

Acta Cryst. (2010). D66, 447-457
https://doi.org/10.1107/S0907444909053578



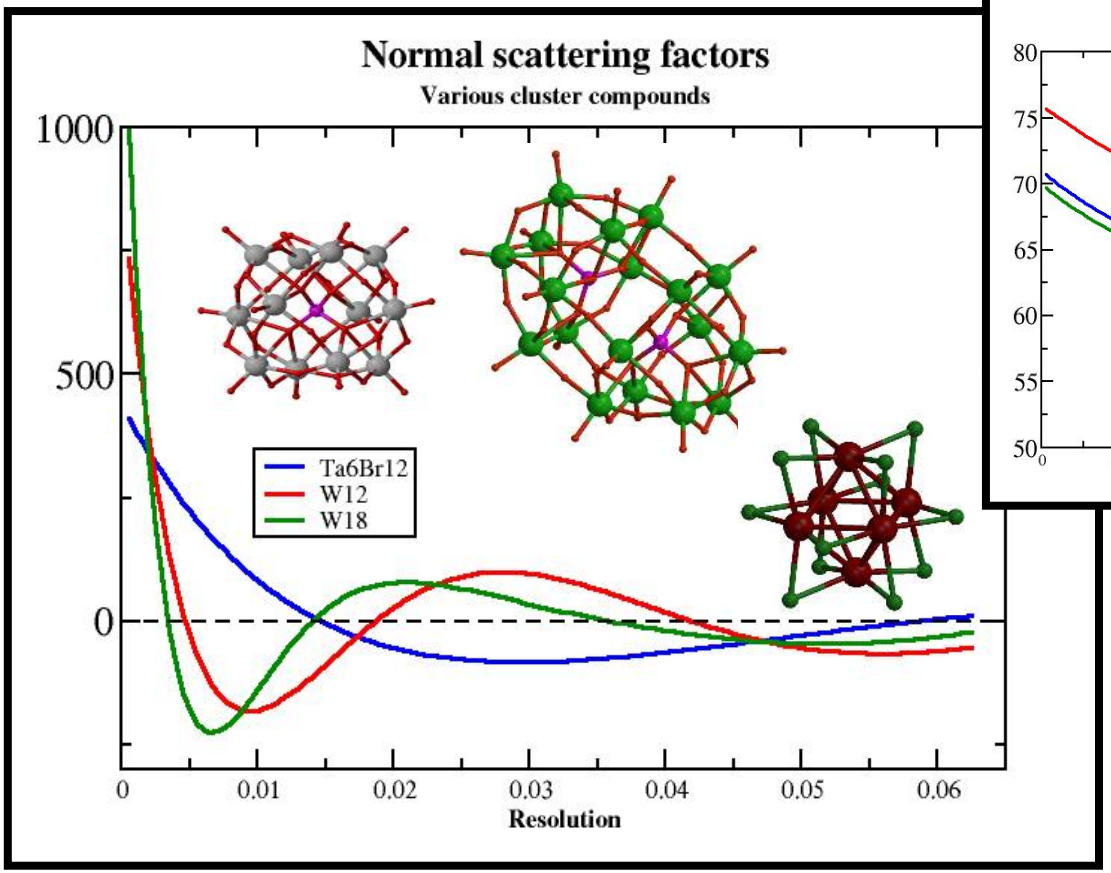
`Broken symmetries' in macromolecular crystallography: phasing from unmerged data

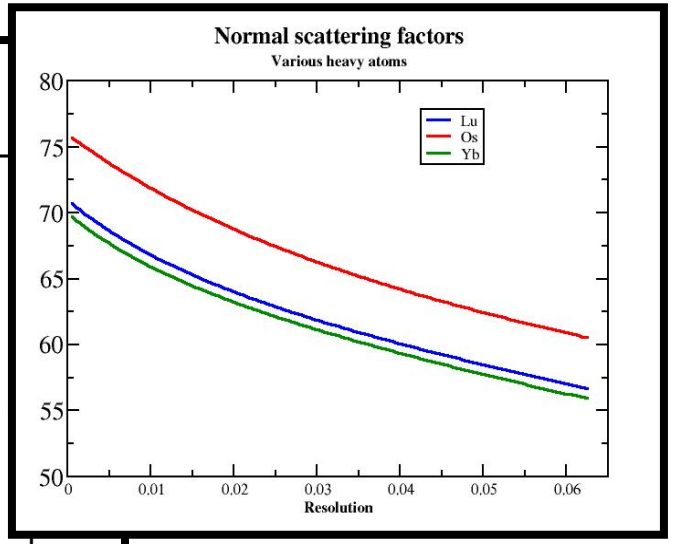
M. Schiltz and G. Bricogne

The space-group symmetry of a crystal structure imposes a point-group symmetry on its diffraction pattern, giving rise to so-called symmetry-equivalent reflections. Instances in macromolecular crystallography are discussed in which the symmetry in reciprocal space is broken, *i.e.* where symmetry-related reflections are no longer equivalent. Such a situation occurs when the sample suffers from site-specific radiation damage during the X-ray measurements. Another example of broken symmetry arises from the polarization anisotropy of anomalous scattering. In these cases, the genuine intensity differences between symmetry-related reflections can be exploited to yield phase information in the structure-solution process. In this approach, the usual separation of the data merging and phasing steps is abandoned. The data are kept unmerged down to the Harker construction, where the symmetry-breaking effects are explicitly modelled and refined and become a source of supplementary phase information.

GΦL
Global Phasing Limited

Spherical averaged cluster compounds





Ribosome: 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.

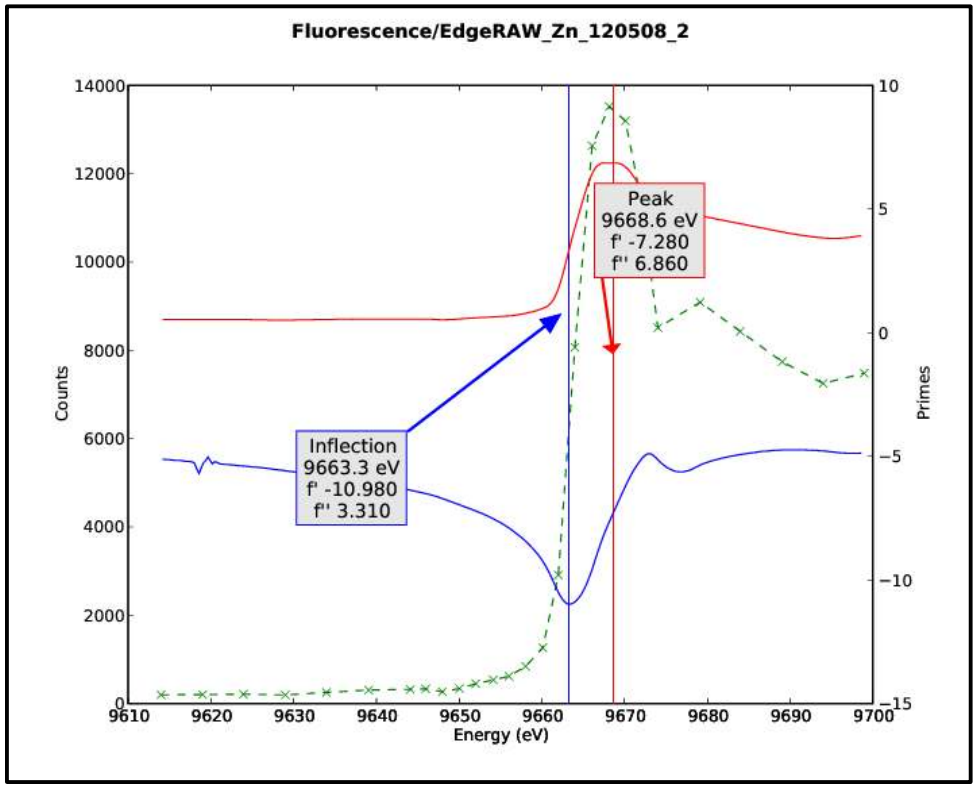
SHARP: special **SPHCLUSTER** keyword – also in autoSHARP

GΦL
Global Phasing Limited

SAD or 2-wvl MAD?

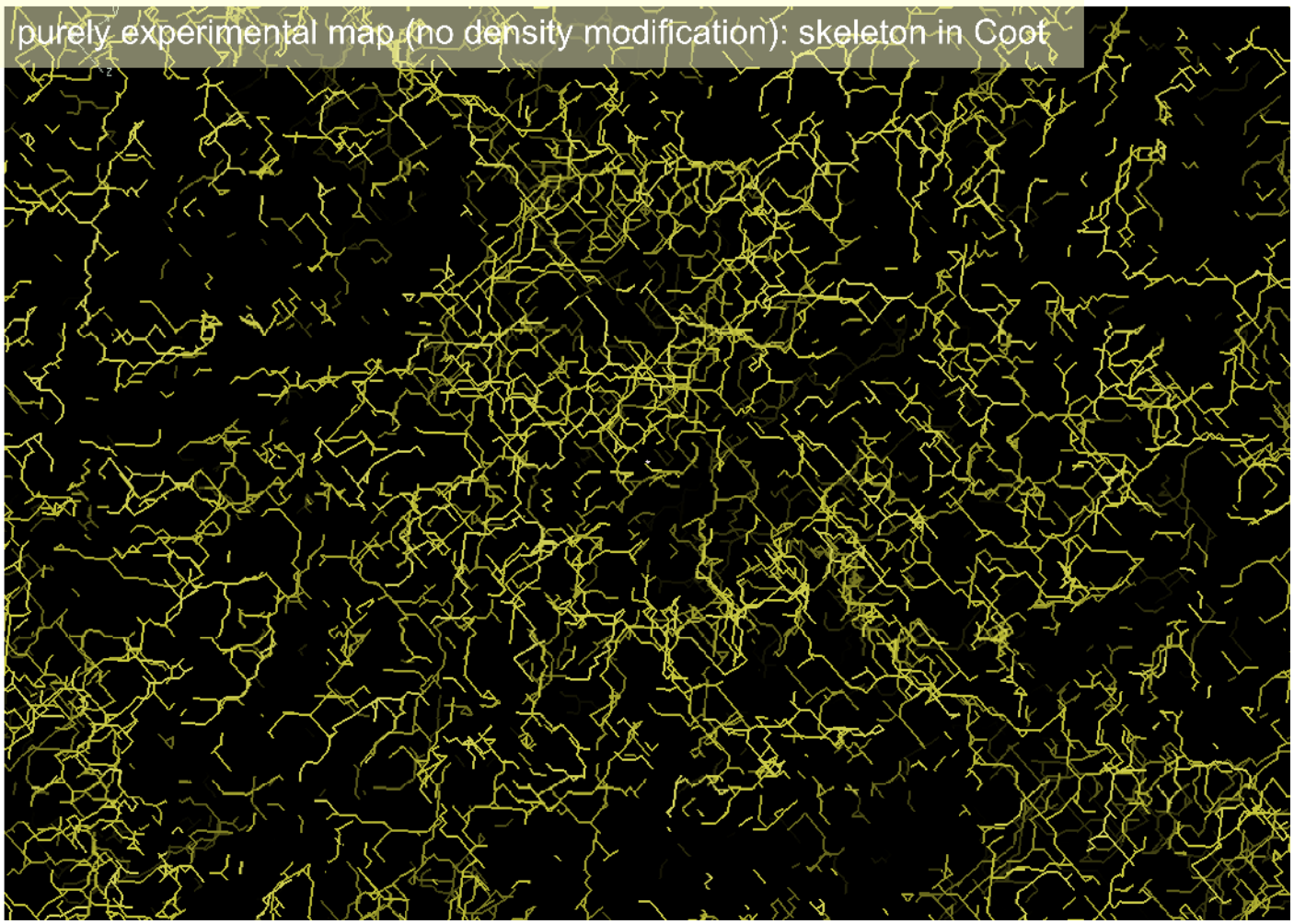
A. González (2003). Optimizing data collection for structure determination. Acta Cryst. D59, 1935-1942

- 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)



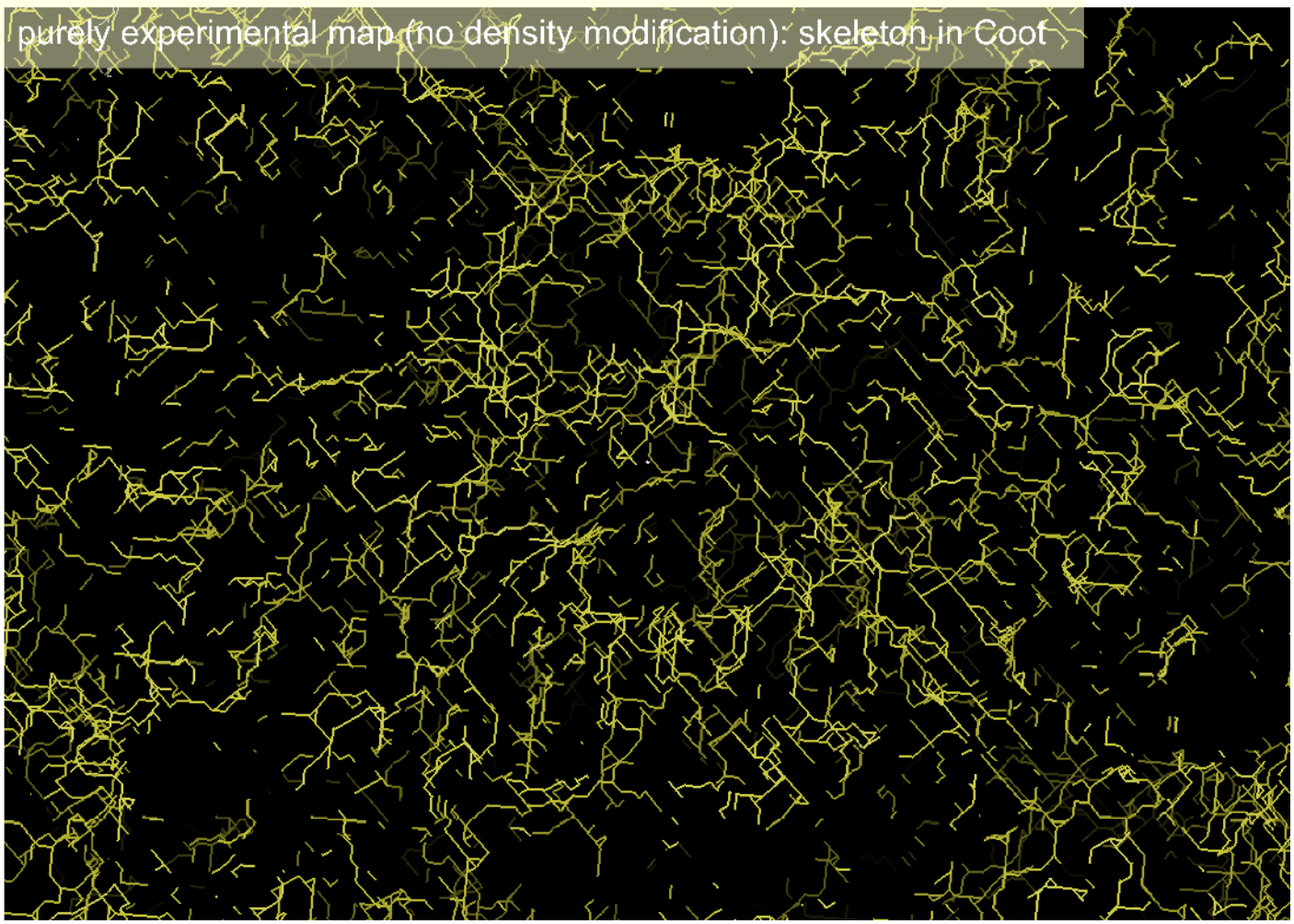
GΦL
Global Phasing Limited

SAD (PEAK, 3600 images, multiplicity 25)



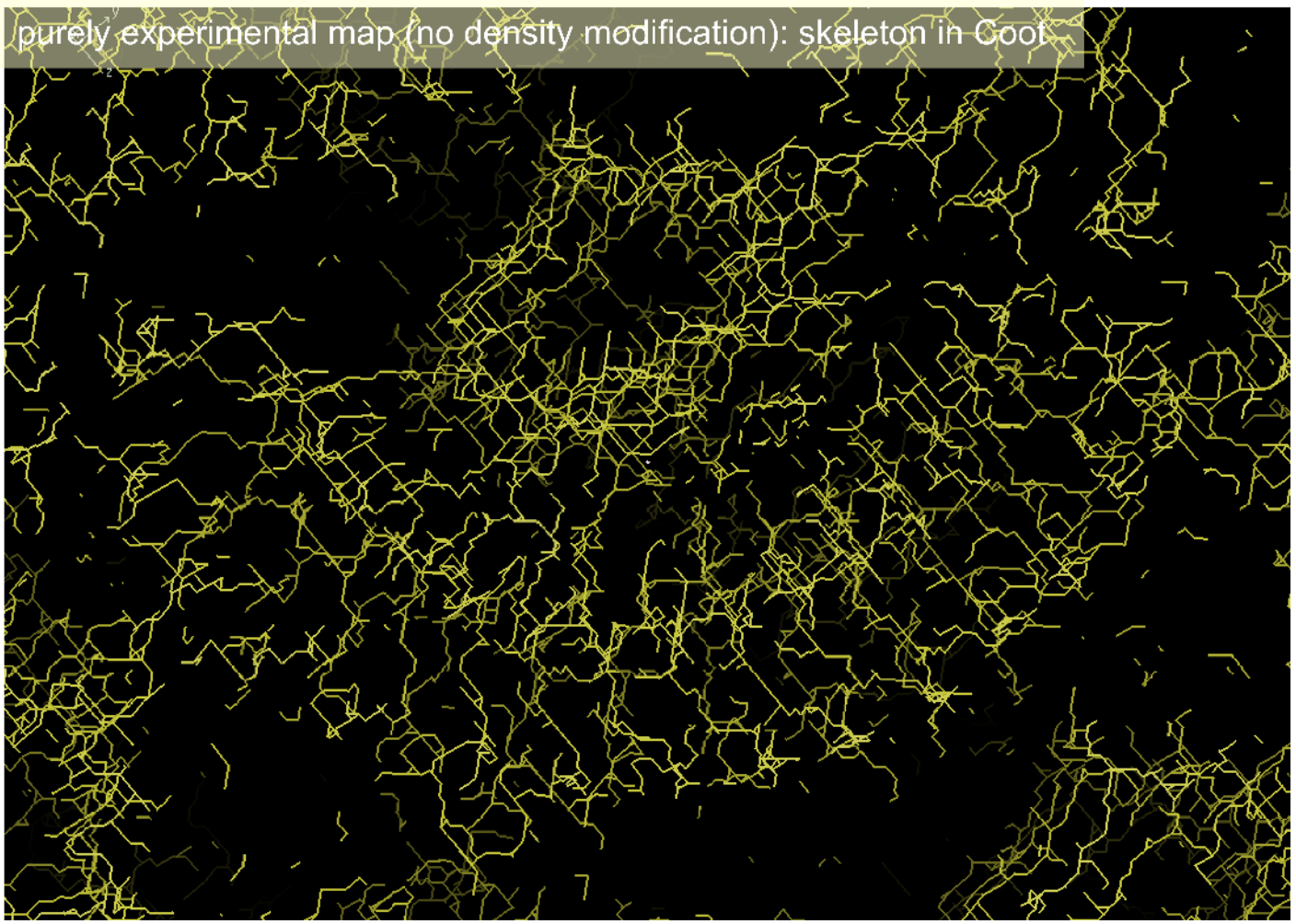
GΦL
Global Phasing Limited

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



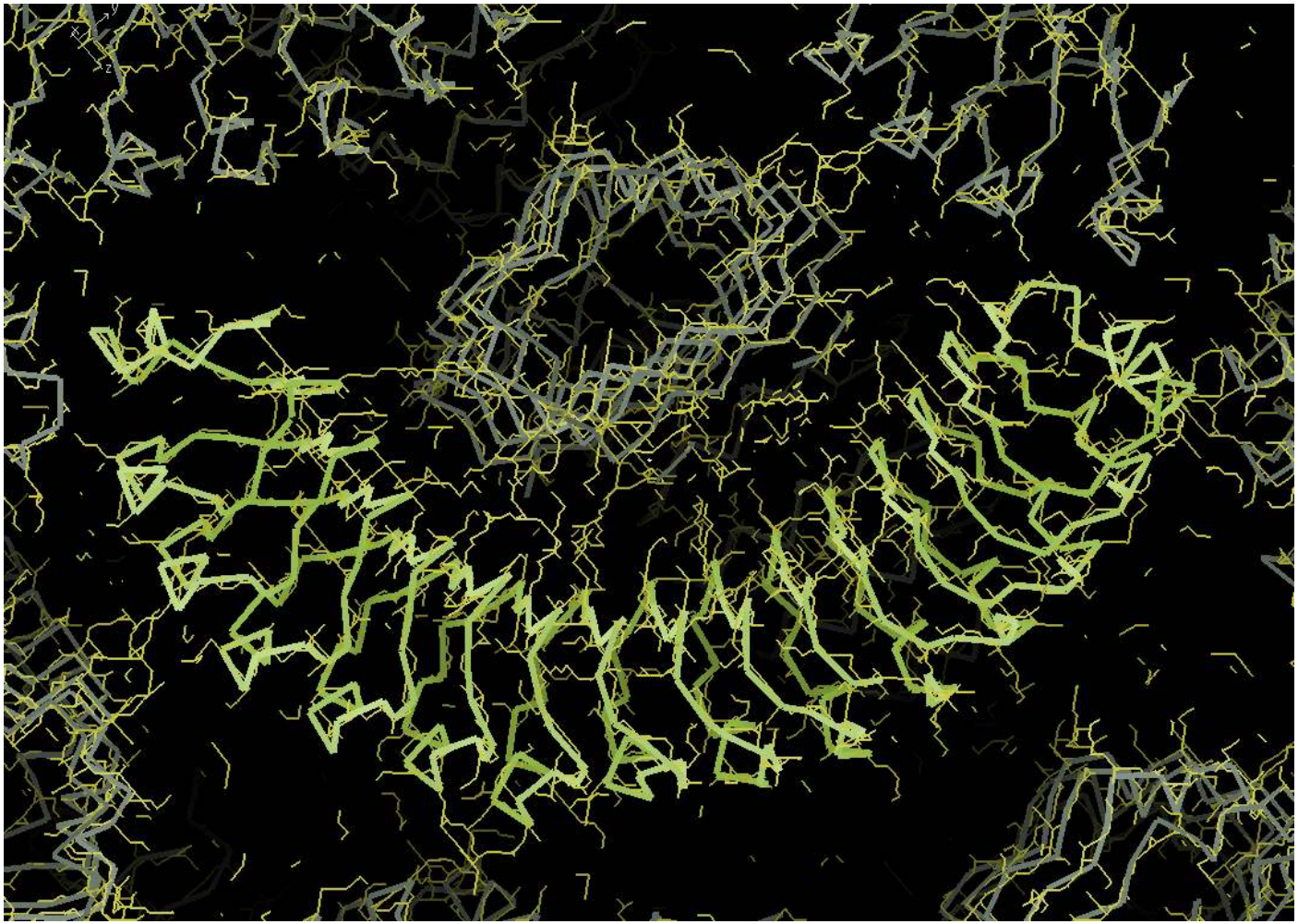
GΦL
Global Phasing Limited

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



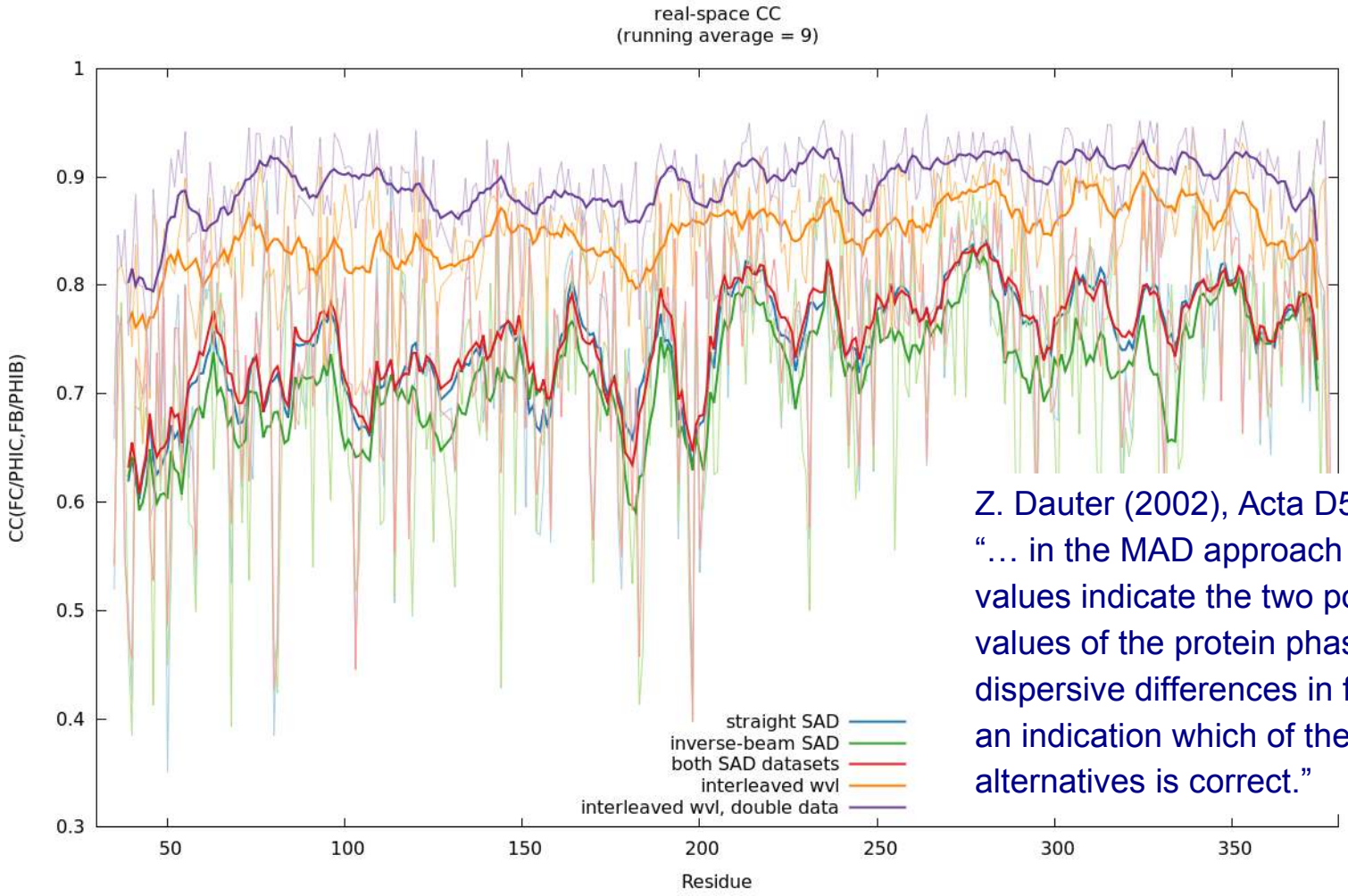
GΦL
Global Phasing Limited

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



GΦL
Global Phasing Limited

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."

GΦL
Global Phasing Limited

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.

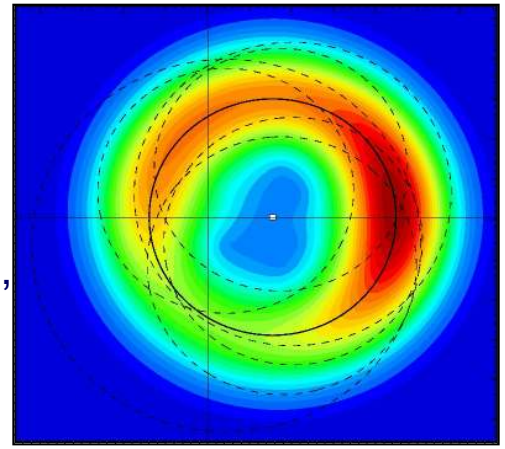
- Think about **differences** (and how to measure, scale, analyse and describe them): anomalous, dispersive, isomorphous, radiation damage, MAD+native etc.

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

- High multiplicity (good): maybe split between different wavelengths instead of staying with SAD

- Nomenclature can be confusing between programs

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



2D phase probability in SHARP

GΦL
Global Phasing Limited

Acknowledgement

Global Phasing, Cambridge (UK):

- Gérard Bricogne, Claus Flensburg, Andrew Sharff, Wlodek Pacziorek, Peter Keller, Rasmus Fogh, Ian Tickle, Leigh Carter, Marcin Wojdyr
- Previous group-members: M. Schiltz, E. Blanc, P. Roversi, G. Evans, R. Morris, T. Womack, Oliver Smart
- Global Phasing Industrial Consortum (funding, discussions, reports, examples)

Licence/download = free to academics

https://www.globalphasing.com/ **(also: BUSTER, Grade, Rhofit, Pipedream, autoPROC)**
https://staraniso.globalphasing.org/ **(STARANISO & PDBpeep server)**
https://grade.globalphasing.org/ **(restraints dictionary generation)**
https://www.globalphasing.com/autoproc/wiki/
https://www.globalphasing.com/sharp/wiki/

Tutorials, FAQ, documentation, examples, ...

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

Test data: JCSG, SBGrid, Global Phasing Consortium members, SLS/PSI, Soleil, DLS, ESRF, IMCA-CAT, SSRL, APS, ALS, PF, LNLS, NSLS, … many others and users