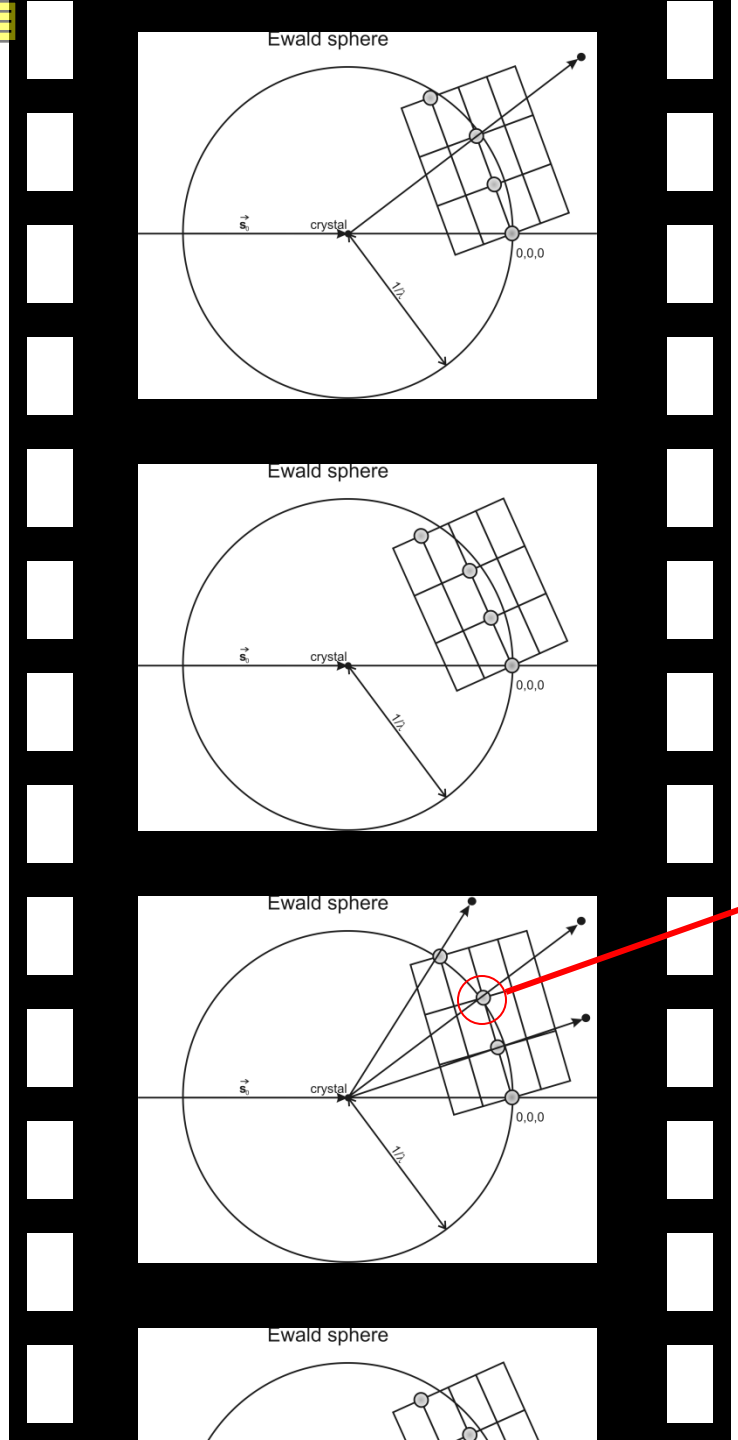
Rapid Data Quality Feedback and Online Monitoring using CASS

Karol Nass

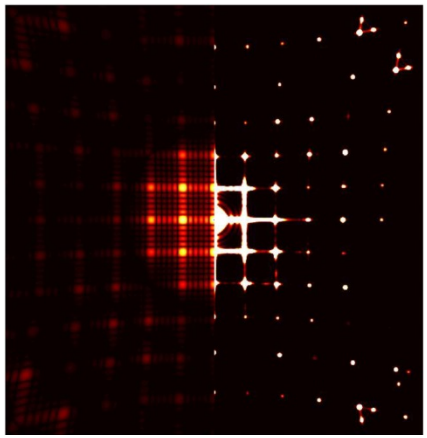
MPI Medical Research Heidelberg



Features of serial femtosecond crystallography:

- Every exposed crystal is destroyed
- Numerous shots of different crystals with possibly different sizes
- No control over crystal orientation in the jet
- Crystals stand still during 40 fs exposure
- Only part of reflection intersects Ewald sphere ("partial" reflection)
- Fringes rather than neat spots

6x6x6 unit cells



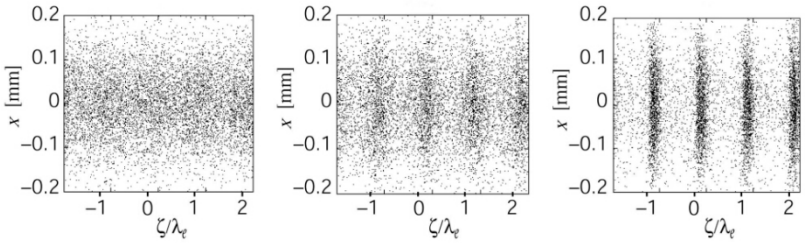
200x200x200 unit cells

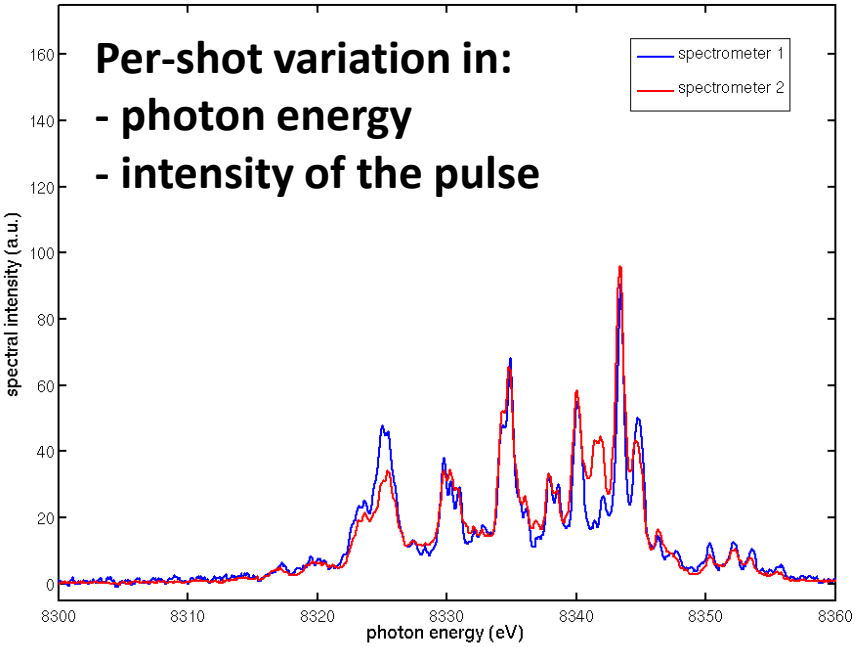
(Simulation software by Wolfgang Kabsch)

- Crystals may be hit by center of the beam, or just grazed
- Beam intensity and spectrum varies from shot to shot

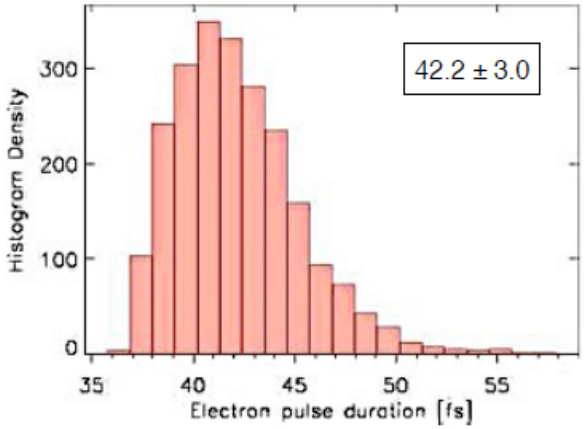
X-ray pulse fluctuations

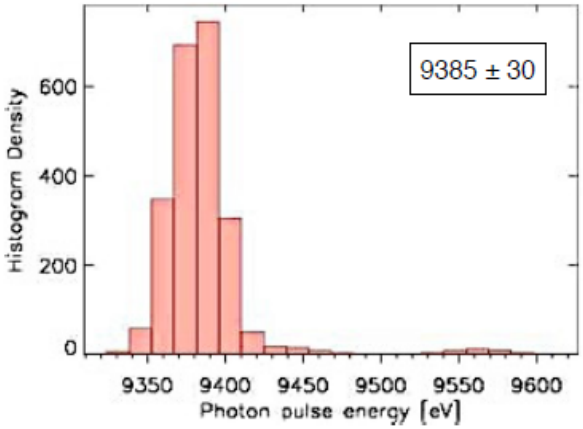
Self amplified spontaneous emission (SASE)

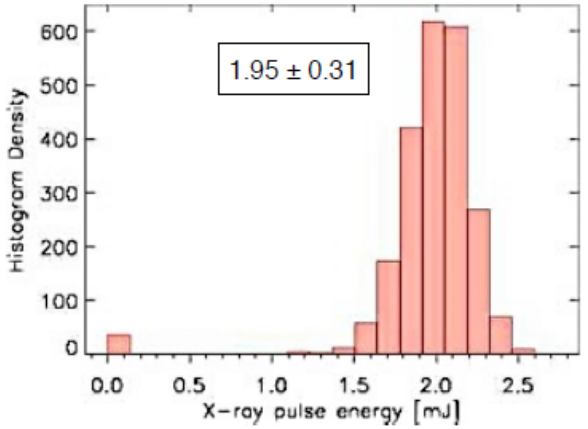




Histograms from few thousands shots for:
electron bunch duration [fs], X-ray pulse energy [eV], pulse energy [mJ]:

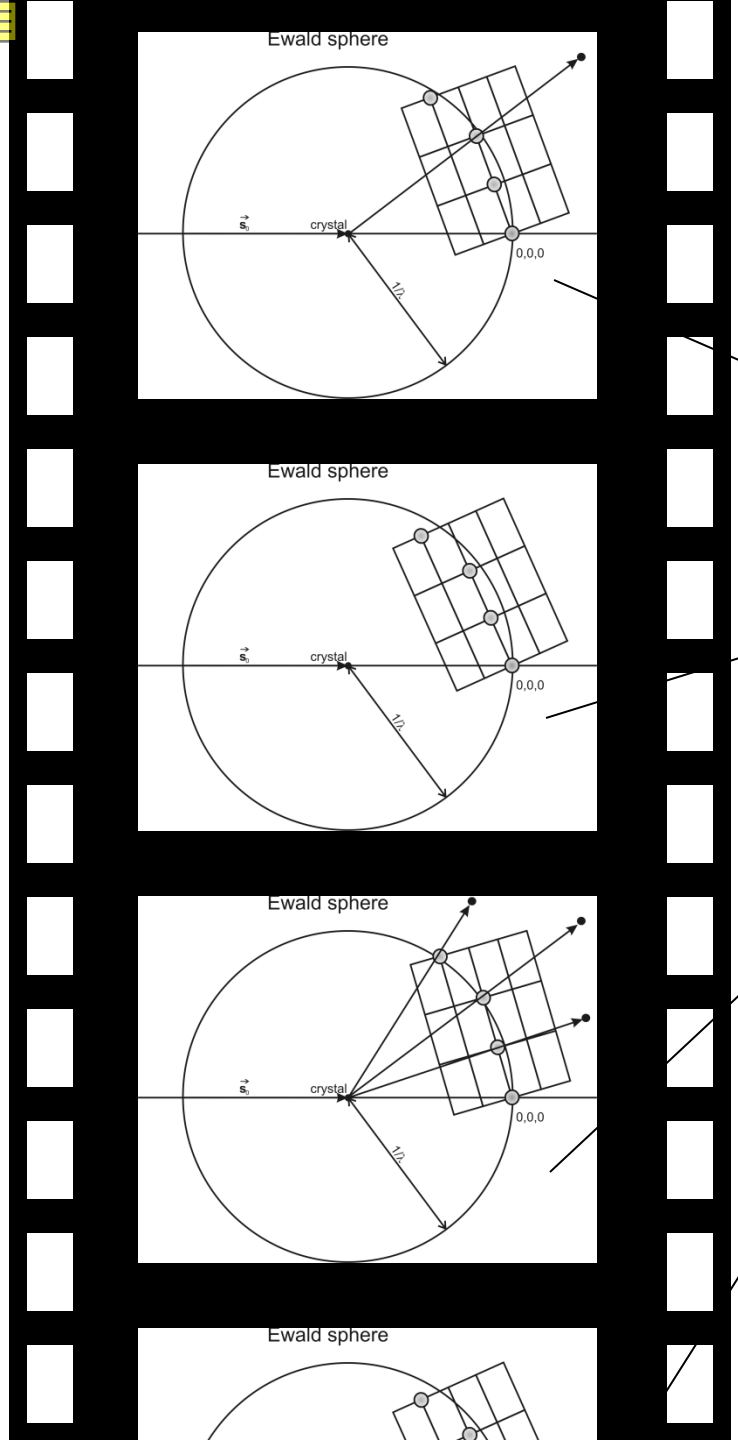


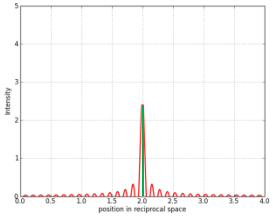


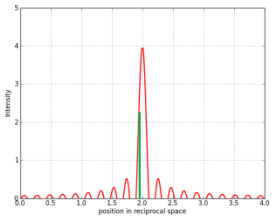


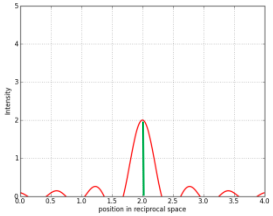
It is possible to do a "*Monte Carlo*" integration over multiple *indexed* femtosecond images and obtain a dataset of fully integrated reflections

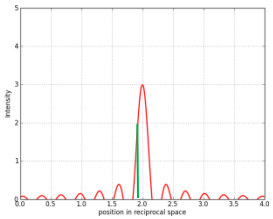
Kirian *et al* (2010), *Optics Express*,**18**, 5713-5723:

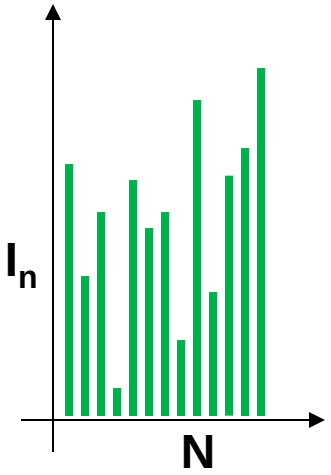








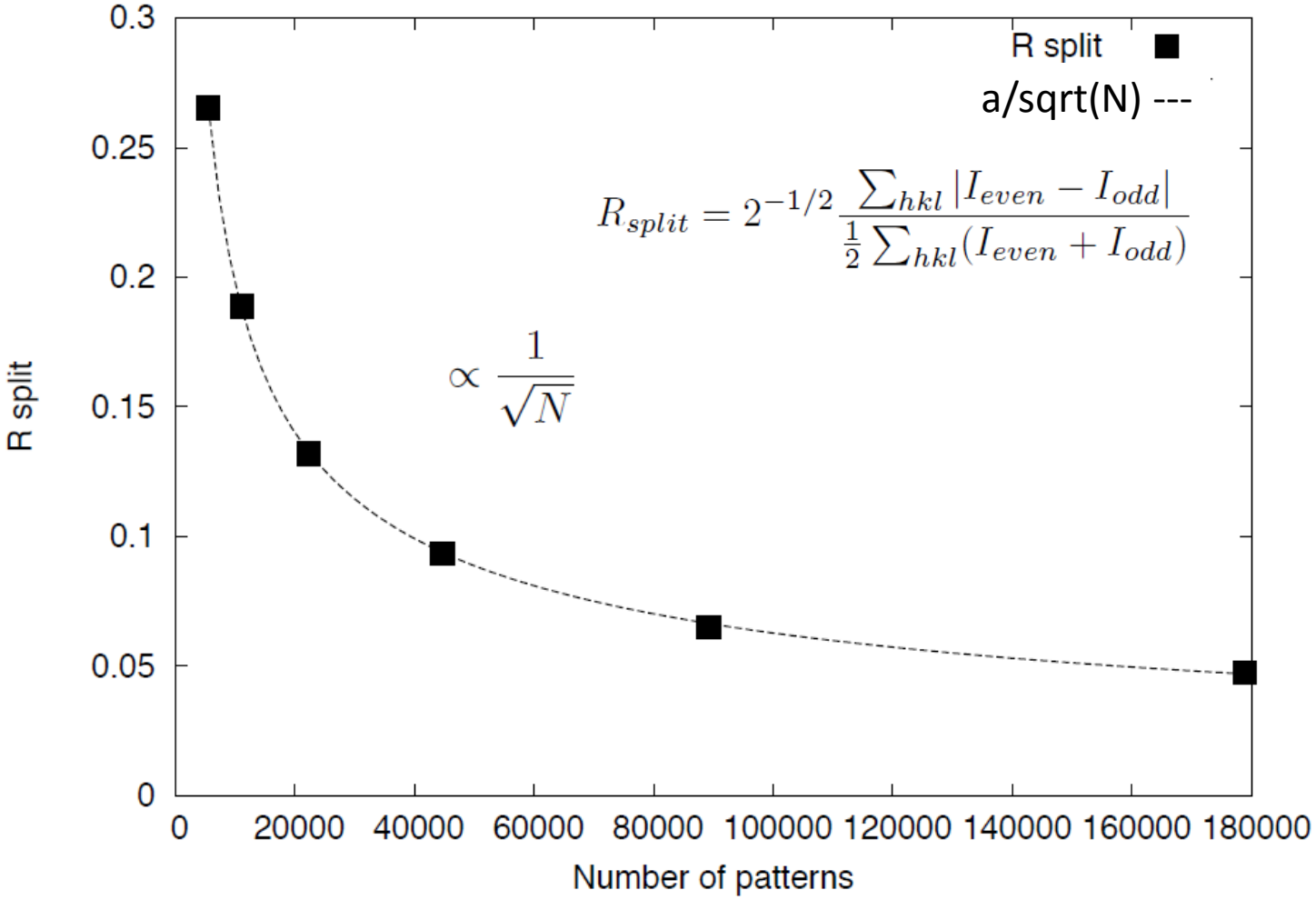




$$\overline{I_{hkl}} = \frac{\sum I_n}{N}$$

$$\sigma(I_{hkl}) = [\sum (I_n - \overline{I_{hkl}})^2]^{1/2}/N$$

Monte Carlo convergence



CASS – CFEL – ASG Software Suite

- **CASS was developed by:**

Lutz Foucar *(MPI for Medical Research, Heidelberg)*

Joachim Ullrich *(MPI für Kernphysik, Heidelberg)*

Jochen Küpper, Uwe Hoppe[@] *(CFEL, [@]Fritz-Haber-Institut der MPG, Berlin)*

Stephan Kassemeyer *(MPI für med. Forschung, Heidelberg)*

Robert Hartmann, Nils Kimmel, Peter Holl, U. Hoppe, Lothar Strüder *(MPI Halbleiterlabor, München)*

Nicola Coppola[#], Thomas White, Anton Barty *(CFEL, XFEL[#] Hamburg)*

Mirko Scholz *(MPI für biologische Chemie, Göttingen)*

Lutz Foucar et al. “CASS—CFEL-ASG software suite”
Computer Physics Communications, (2012), **183**, 2207 – 2213.

Features of CASS

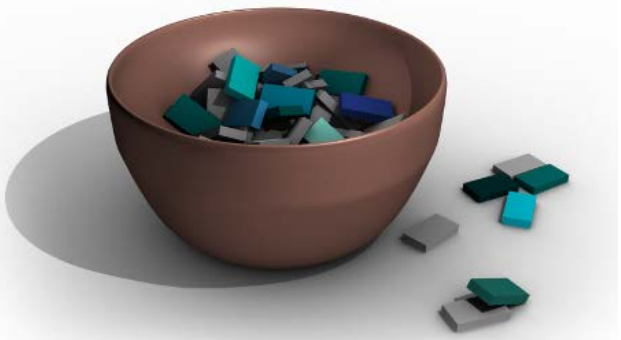
- Correct, combine and analyze correlated data from each shot
- Multiple threads allow parallel processing
- Analysis based on modular ini file architecture
- Pre-defined functions build the analysis pipeline
- Documentation:

http://www.mpi-hd.mpg.de/personalhomes/gitasg/cass/

```
; retrive the raw pnCCD image
[PostProcessor]
Image\ID = 100
Image\Device = 0
Image\Detector = 0
Image\Write = false ; do not write to file


; the pnCCD spectrum
[PostProcessor]
pnCCDSpectrum\ID = 143
pnCCDSpectrum\Detector = FrontPnCCD
pnCCDSpectrum\SplitLevelLowerLimit = 0
pnCCDSpectrum\SplitLevelUpperLimit = 20000
pnCCDSpectrum\XNbrBins = 1000
pnCCDSpectrum\Xlow = 0
pnCCDSpectrum\XUp = 10000

; summed pnCCD spectrum
summedPnCCDSpectrum\ID = 62
summedPnCCD\HistName = pnCCDSpectrum
summedPnCCD\Write = false ; do not write to file
```



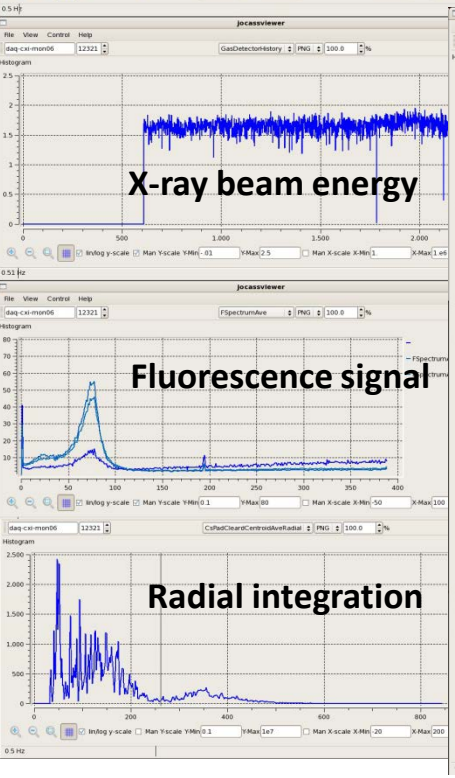
Using CASS on SFX data

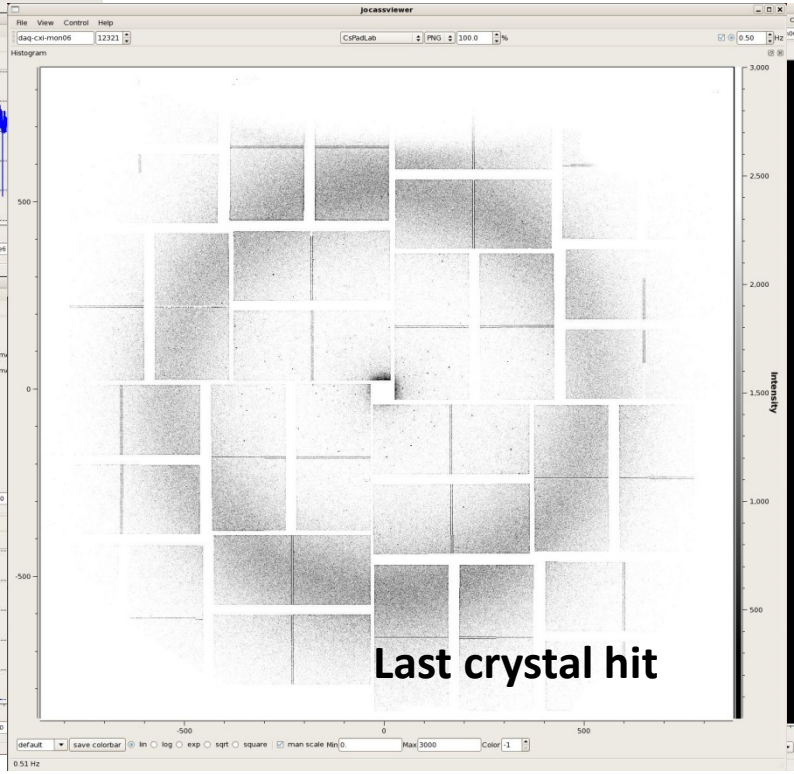
- **On-line (on the live data stream):**
  - Find crystal hits (hit-rate “on the fly” for jet alignment)
  - Accumulate virtual powder pattern (data completeness)
  - Check pixel overload
  - Meet my criteria (was this pulse seeded or not?)
  - Obtain the time delay of pump laser using raw TimeTool data, etc…

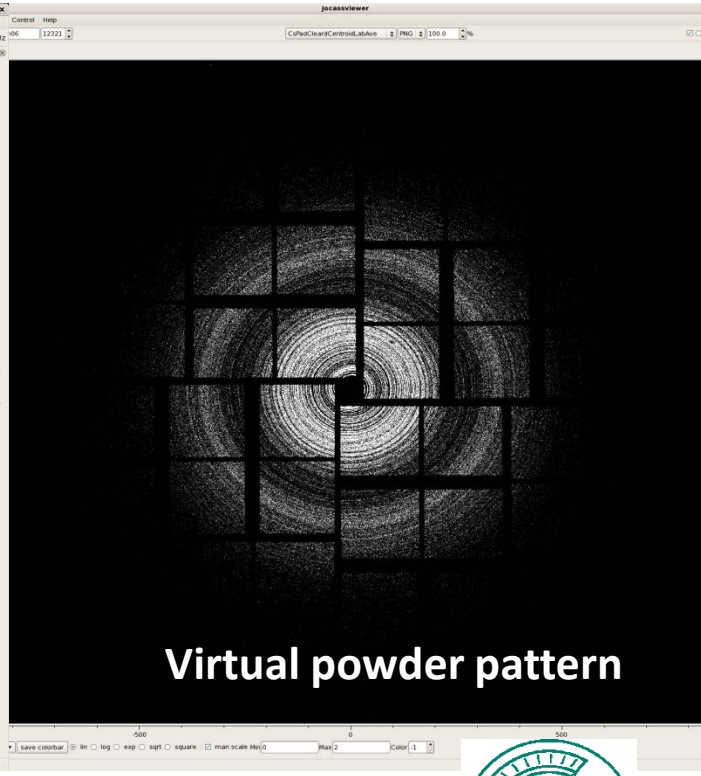
- **Off-line (on XTC files):**
  - The same and:
    - Write data to HDF5 or CBF files.

On-line data analysis using CASS:

- monitoring and feedback from CASS data during experiment





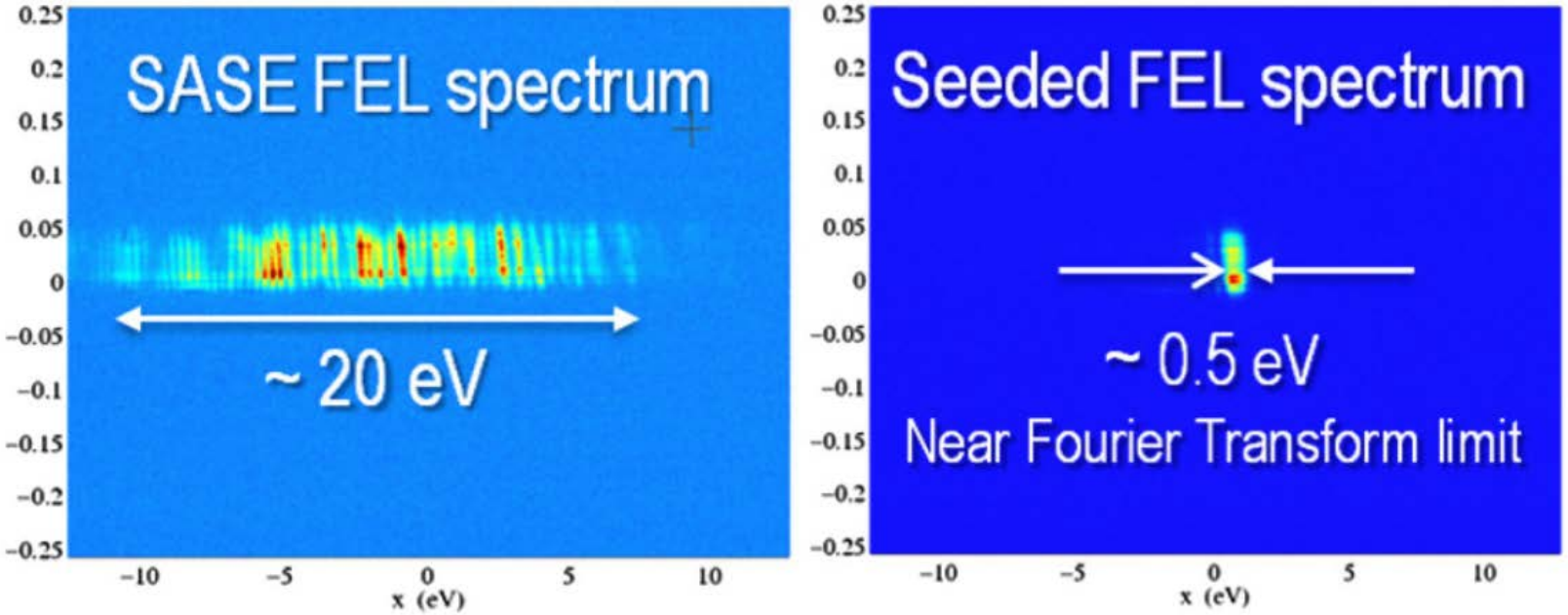


Foucar et al. *Comp. Phys. Comm.* **183**, 2207-2213 (2012).

Used CASS to analyze other types of data

- SAXS, WAXS avg. radial profiles,
- Fixed target diffraction images
- Single particles (viruses in aerosol injector)
- Fluorescence signal
- X-ray pulse spectrum (seeded pulse identification)
- pump laser delay, etc.

Self-seeded beam at LCLS – possible improvement for SFX ?



J. Phys. B: At. Mol. Opt. Phys. **46** (2013) 164003

How "seeded" is seeded?

Seeded: $\Delta\lambda/\lambda \sim 0.8 \times 10^{-4}$

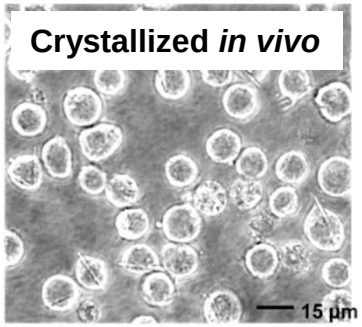
SASE: $\Delta\lambda/\lambda \sim 2.0 \times 10^{-3}$ plus jitter

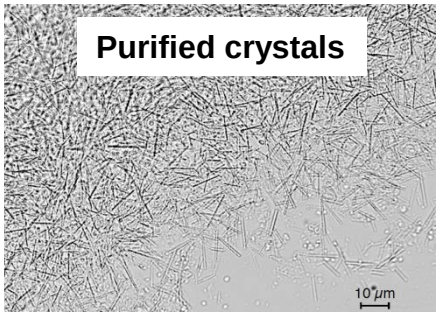
Does it really improve SFX results?

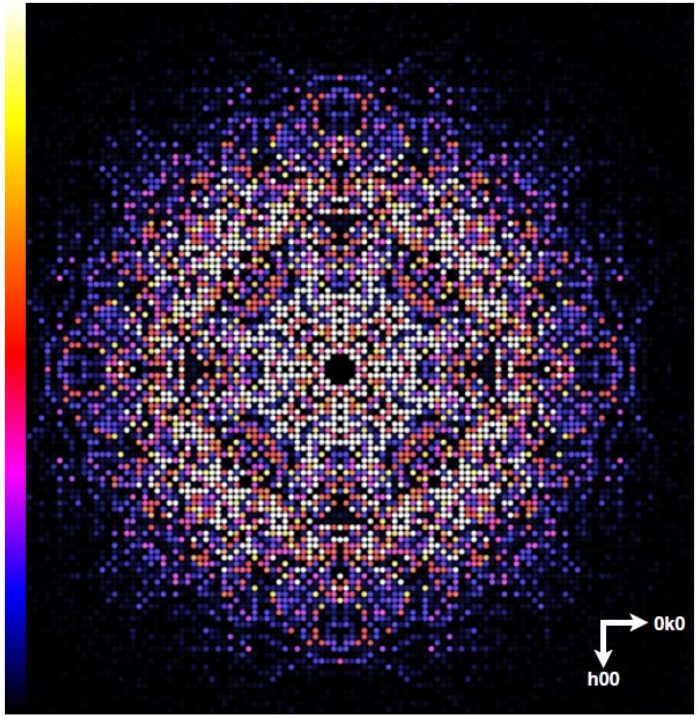
First new structural information from LCLS

Redecke', Nass' et al. Science 2013

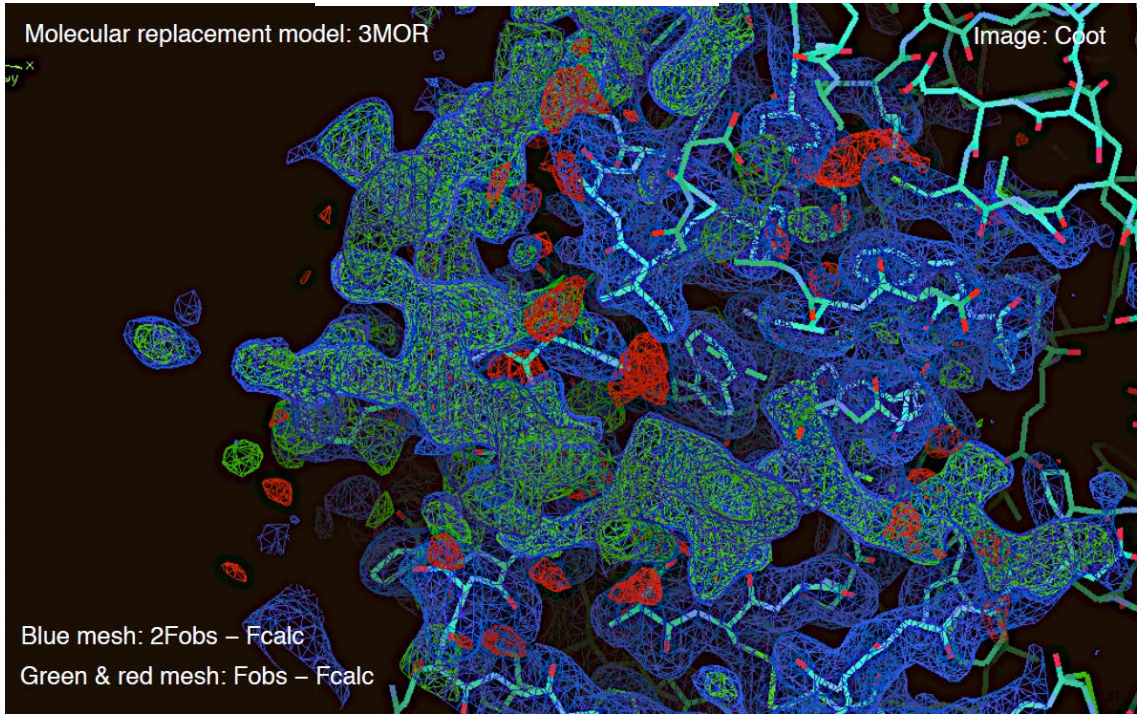
New structural information obtained from SFX data
PHASED BY MOLECULAR REPLACEMENT







Experimental integrated intensities



Tb-Cathepsin B with pro-peptide

Test case for *de novo* phasing of SFX data

- We want a really, really strong anomalous signal and something that scatters really, really well: Lysozyme:Gd

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Gd-HPDO3A, a complex to obtain high-phasing-power heavy-atom derivatives for SAD and MAD experiments: results with tetragonal hen egg-white lysozyme

Éric Girard, Laurent Chantalat, Jean Vicat and Richard Kahn*

Laboratoire de Cristallographie Macromoléculaire, Institut de Biologie Structurale J.-P. Ebel CEA–CNRS–UJF, 41 Rue Jules Horowitz, 38027 Grenoble CEDEX 01, France

Correspondence e-mail: kahn@ibs.fr

Received 18 July 2001
Accepted 3 October 2001

PDB Reference: Gd derivative of lysozyme, 1h87.

A neutral gadolinium complex, Gd-HPDO3A, is shown to be a good candidate to use to obtain heavy-atom derivatives and solve macromolecular structures using anomalous dispersion. Tetragonal crystals of a gadolinium derivative of hen egg-white lysozyme were obtained by co-crystallization using different concentrations of the complex. Diffraction data from three derivative crystals (100, 50 and 10 m*M*) were collected to a resolution of 1.7 Å using Cu $K\alpha$ radiation from a rotating anode. Two strong binding sites of the gadolinium complex to the protein were located from the gadolinium anomalous signal in both the 100 and 50 m*M* derivatives. A single site is occupied in the 10 m*M* derivative. Phasing using the anomalous signal at a single wavelength (SAD method) leads to an electron-density map of high quality. The structure of the 100 m*M* derivative has been refined. Two molecules of the gadolinium complex are close together. Both molecules are located close to tryptophan residues. Four chloride ions were found. The exceptional quality of the SAD electron-density map, only enhanced by solvent flattening, suggests that single-wavelength anomalous scattering with the Gd-HPDO3A complex may be sufficient to solve protein structures of high

Girard *et al.*, *Acta Cryst.* **D**58, 1-9

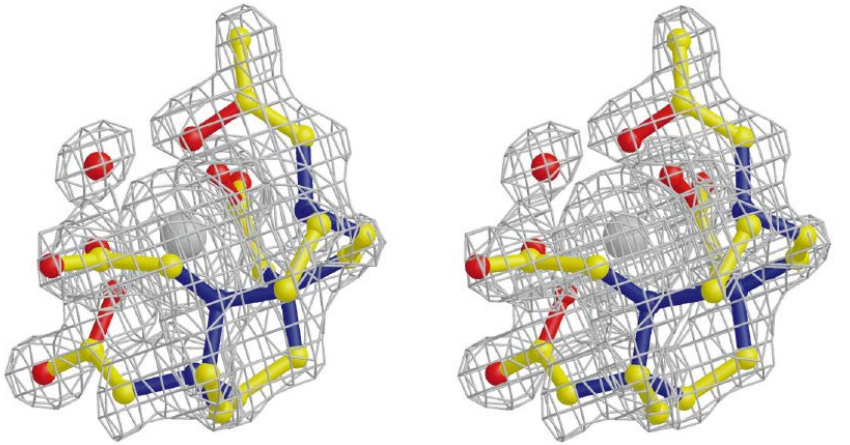
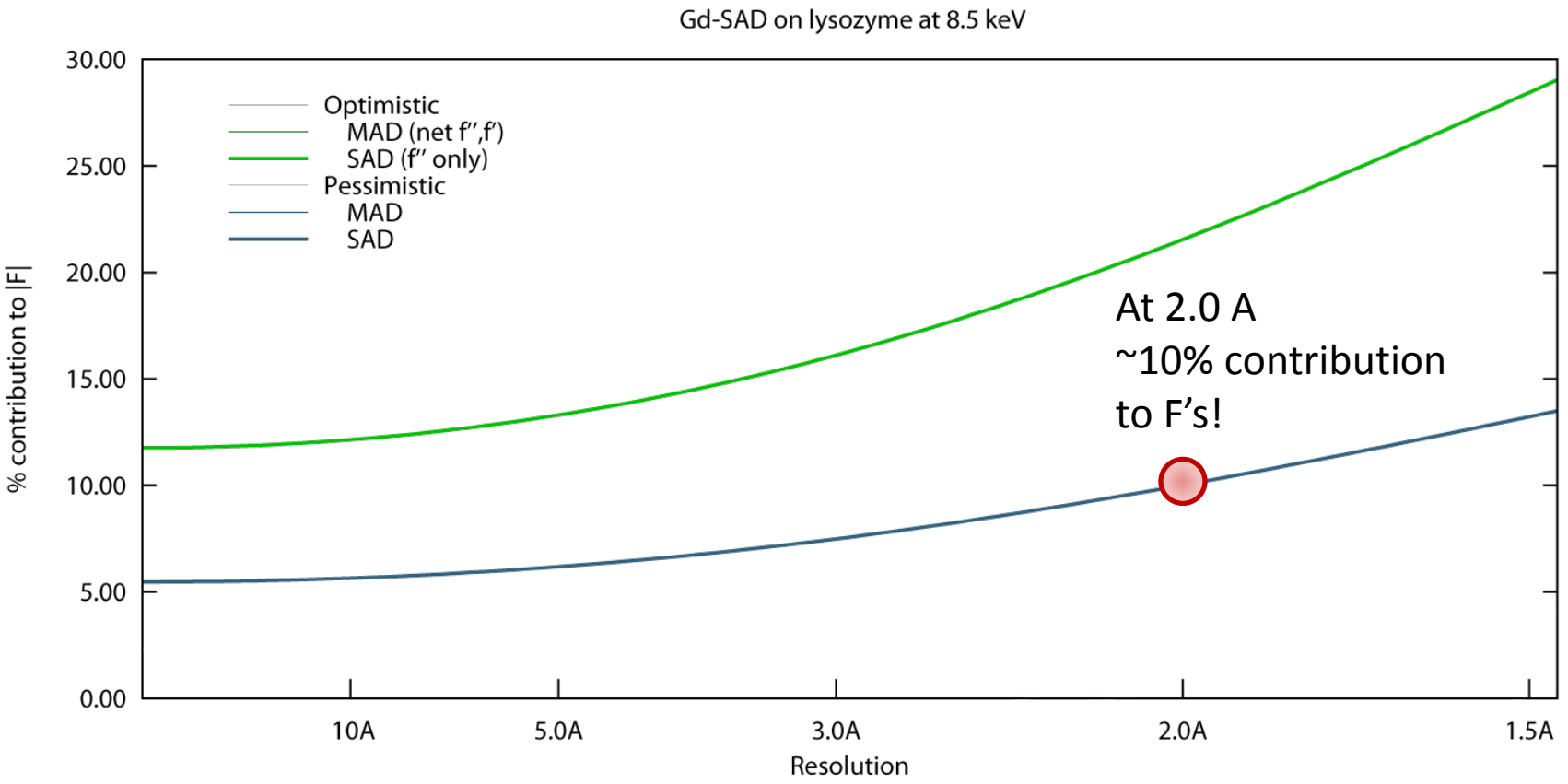
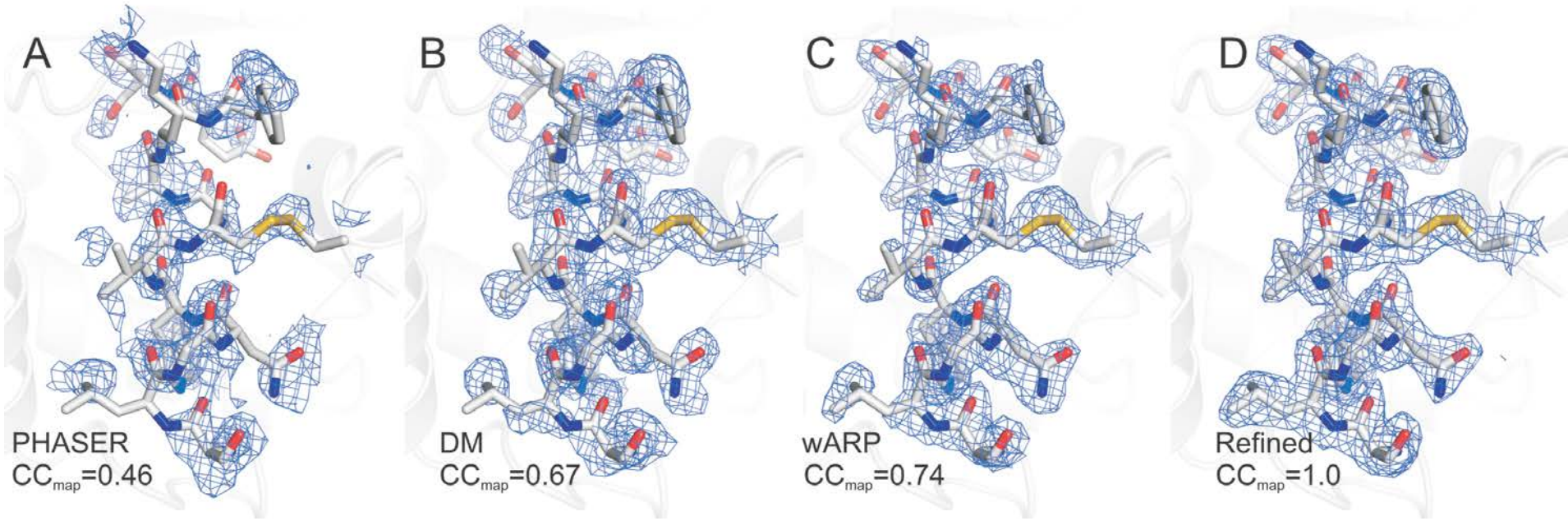


Figure 5
Stereoview of the final $2|F_o| - |F_c|$ electron-density map around Gd-HPDO3A (site Gd 1) and its ninth ligand, water molecule Wat152. The map is contoured at the 1σ level.

Anomalous differences from Gd



Successful *de novo* phasing of Lysozyme-Gd structure

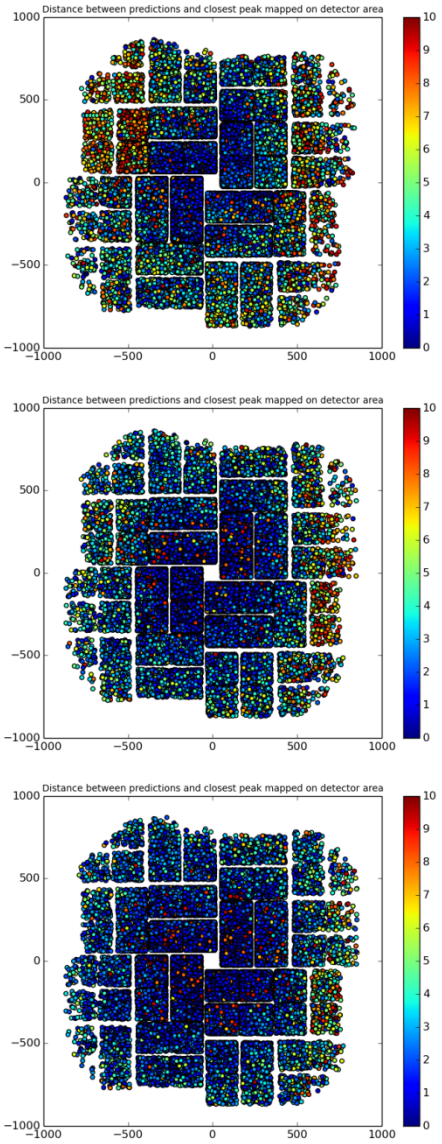


→SFX intensities can be suffiently accurate for phasing
→Despite the very strong anomalous signal, needed 60,000 images!

T. Barends et al. Nature 505, 244–247 (2014)

John R Helliwell. *Crystallography Reviews*, Vol. 20, No. 3, 207–209 (2014)

Development of detectors/software etc. will make this much easier



Geometry opt.

Reprocessed Lysozyme - Gd data :

- Find crystal hits using CASS
-Index them using a newer version of CrystFEL
-Use a program written to refine detector geometry: (as done in cctbx.xfel)

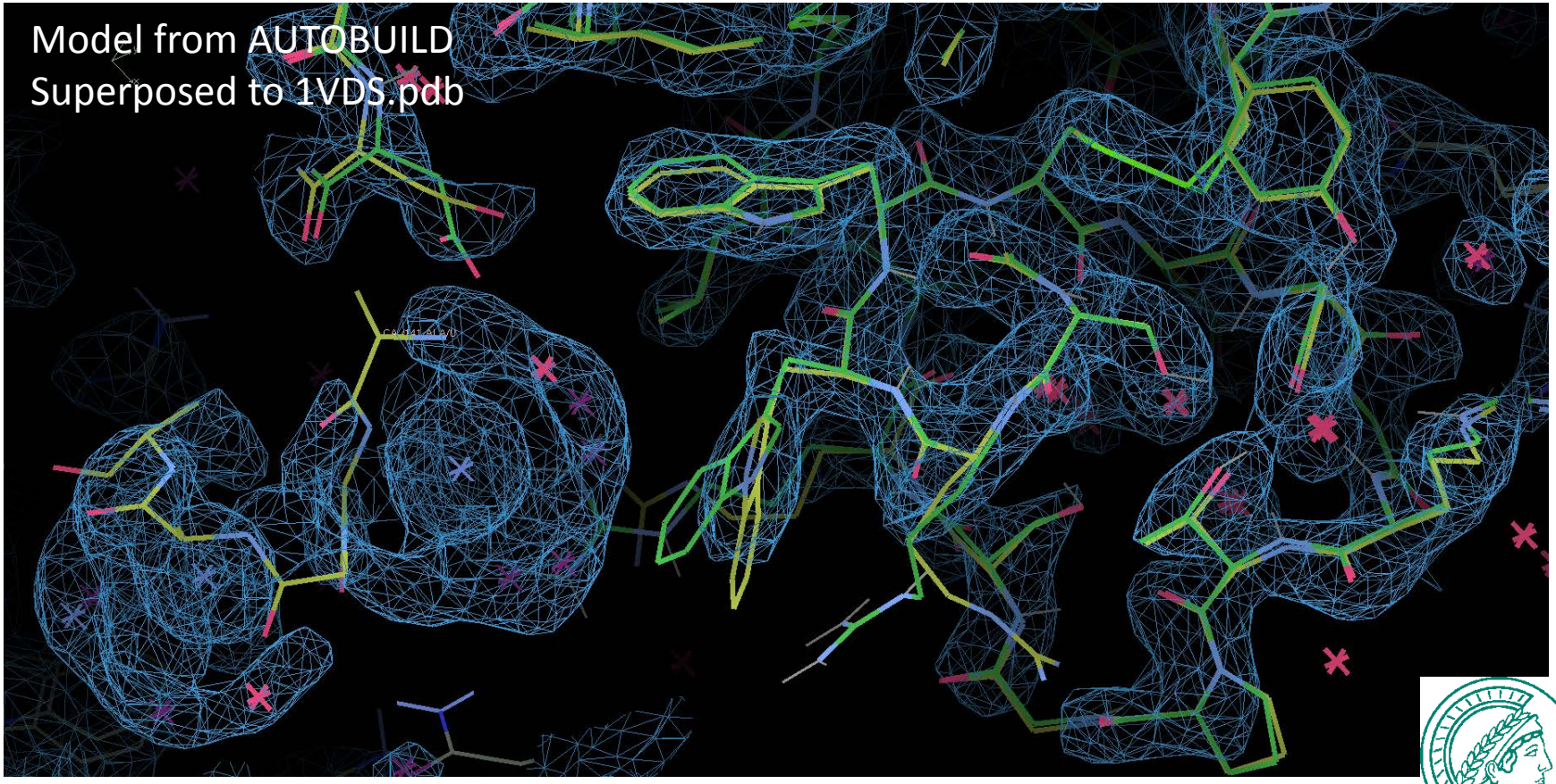
* **Indexing rate** **(was 31% is 44%)**
* **CC_{ano}** **(was 0.48 is 0.69)**
* **R_{ano}/R_{split}** **(was 1.8 is 2.6)**

Automated structure solution in PHENIX using re-processed Lysozyme-Gd SFX data

Wizard status

FINISHED

Top solution:	1	Sites:	4	Space group:	P 43 21 2	FOM:	0.615
BAYES-CC:	56.05	Residues:	126	Side-chains:	126	Chains: :	1
Model CC:	0.88	R-work:	0.2187	R-free:	0.2478		



during the tutorial tomorrow

MPI FÜR MEDIZINISCHE FORSCHUNG

SLAC

Lutz Foucar
Thomas Barends
Elisabeth Hartmann
Sabine Botha
Robert Shoeman
R. Bruce Doak
Wolfgang Kabsch
Lukas Lomb
Stephan Kassemeyer
Ilme Schlichting

Sébastien Boutet
Garth Williams
Marc Messerschmidt
Jason Koglin
Despina Milathianaki
Henrik Lemke

DESY

ASU ARIZONA STATE UNIVERSITY

UH Universität Hamburg

Collaborating groups

Henry Chapman and group, DESY
John Spence, Uwe Weierstall and group, ASU
Petra Fromme and group, ASU
…..