Practical aspects of MX

J-L Ferrer
IBS/Synchrotron Group (Grenoble, France)

Who am I?

Initial training

- M.S. degree from the ECP (mechanics, electrotechnics,...), major: Accelerators and High Energy Physics
- M.S. degree in Chemistry - Physics (Paris VI University, France)
- Ph.D. degree in Physics (Paris XI University, France): *"Study of the non linear dynamic of the Free Electron Laser spectrum in the Compton regime"*

Experience

- construction of beamline D2AM at the ESRF
- construction of beamline FIP at the ESRF
- 2-year at the Salk Institute, USA (structure of CHS, STS, TAA1, ...)

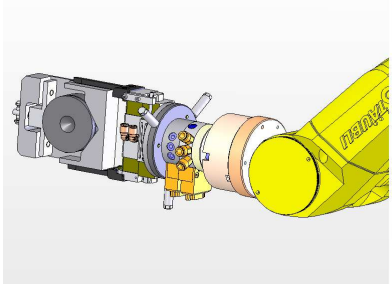
Present situation

- head of the Synchrotron Group at the “Institut de Biologie Structurale” (Grenoble)
- head of beamline FIP-BM30A at the ESRF

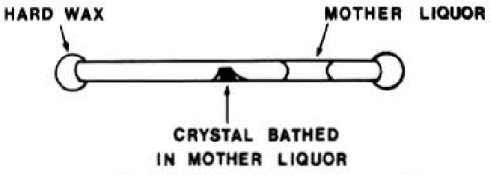
Crystal handling

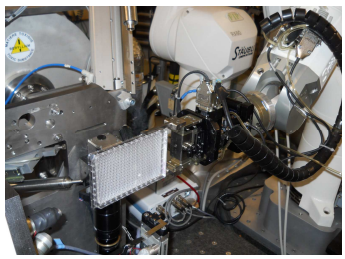
Flash-frozen crystals

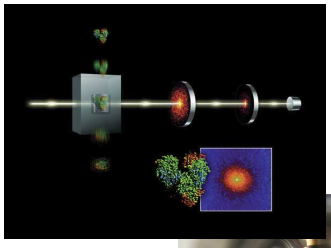
High pressure

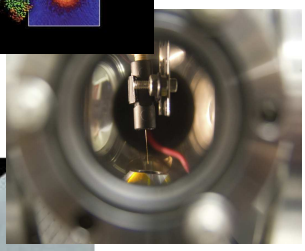


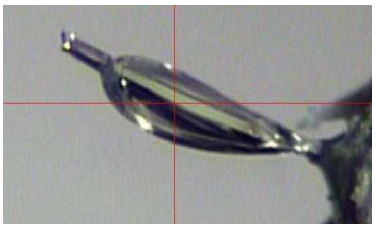
Room temperature

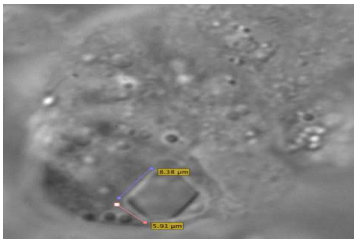












Data collection strategy

Images - What to collect?

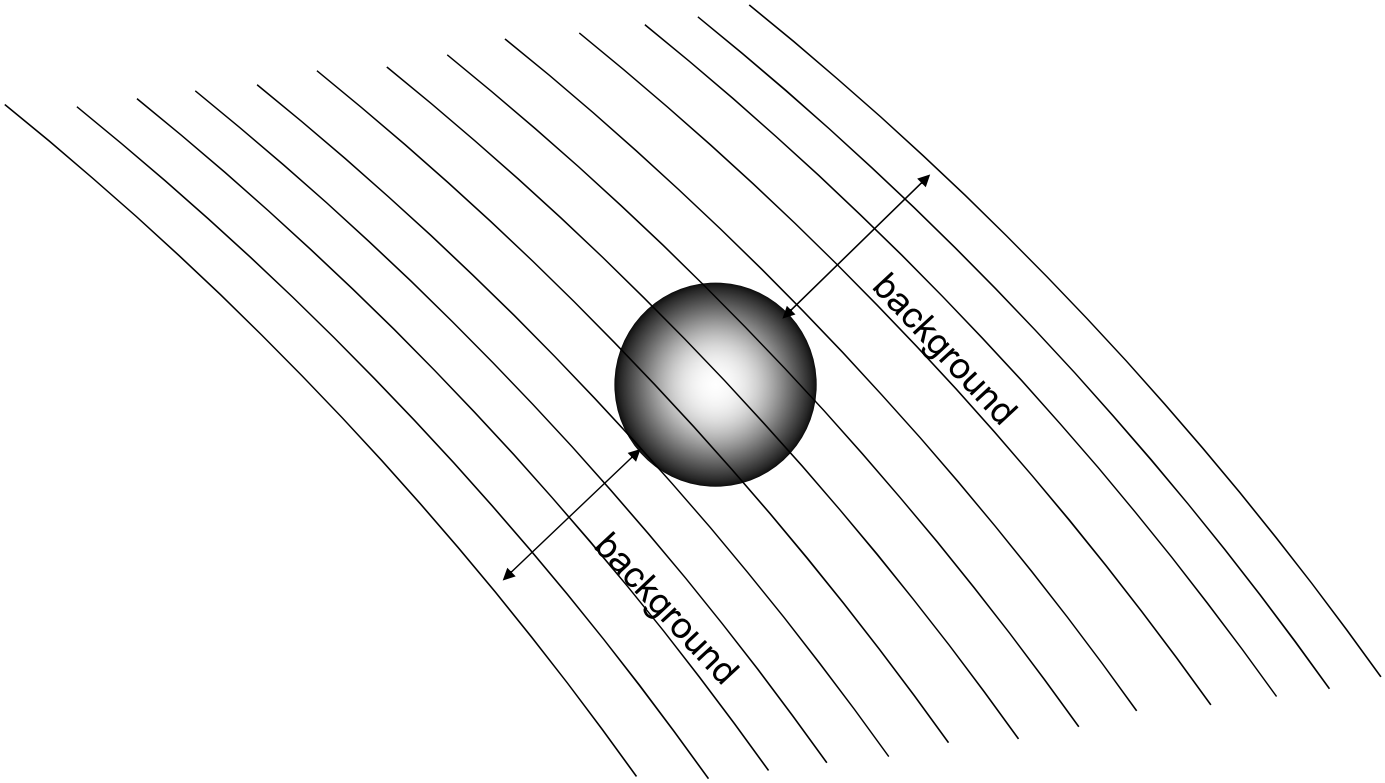
- Overall start and end
- Rotation increment
- Exposure time
- Depends on crystal, spacegroup, mosaicity
- Spatial overlaps
- Expected statistics

Classes of error in MX

Time \ Dependence on signal	none	sqrt	proportional
none	CCD Read-out	Photon counting	Detector calibration attenuation partiality Non-isomorphism Radiation damage
1/sqrt			Beam flicker
1/prop.			Shutter jitter Sample vibration

(J. Holton, ALS)

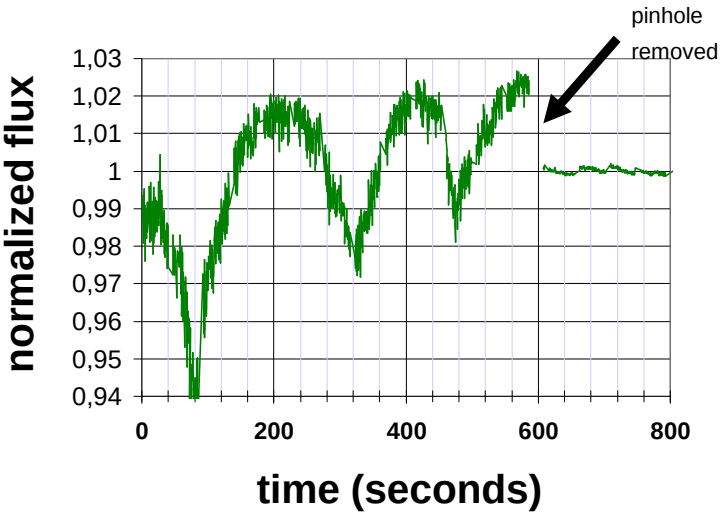
Fine Slicing

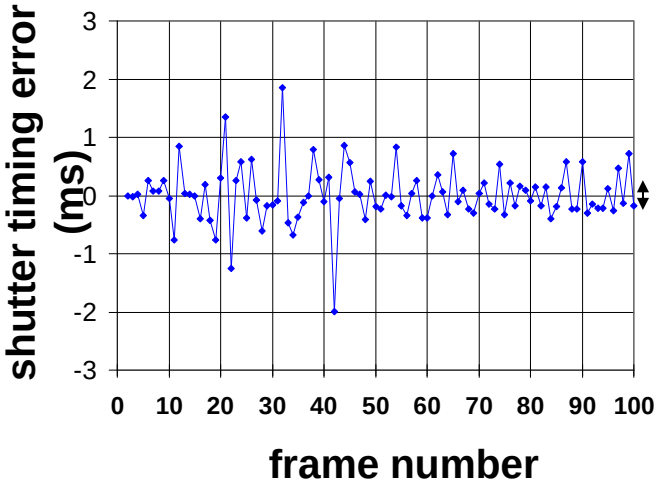


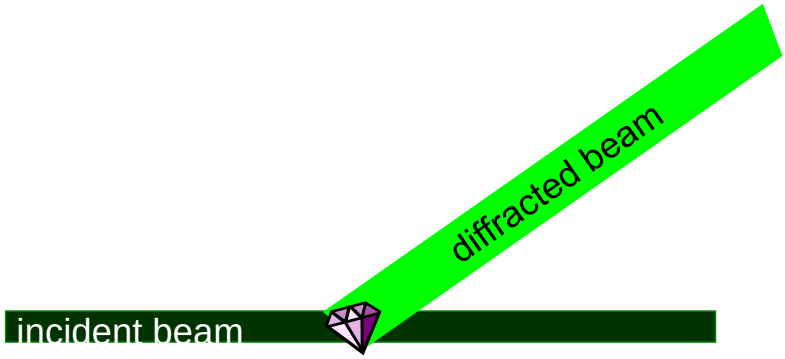
Pflugrath, J. W. (1999)."The finer things in X-ray diffraction data collection", *Acta Cryst. D* **55**, 1718-1725.

(J. Holton, ALS)

Beam flicker/Shutter jitter/Xtal vibration





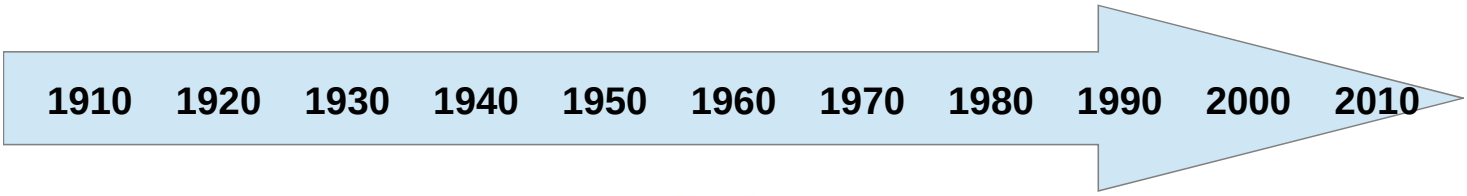


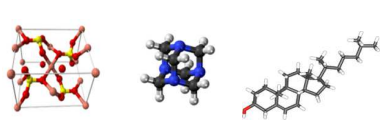
- reduce flux and increase exposure

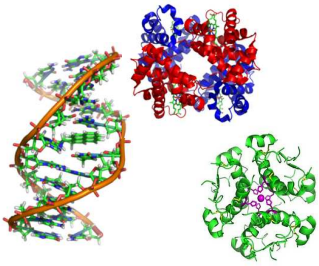
OR:

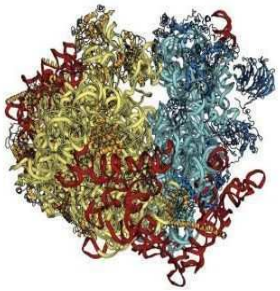
- vibrate at frequencies **different** from experiment
- average over it

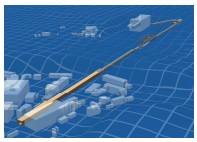
(J. Holton, ALS)









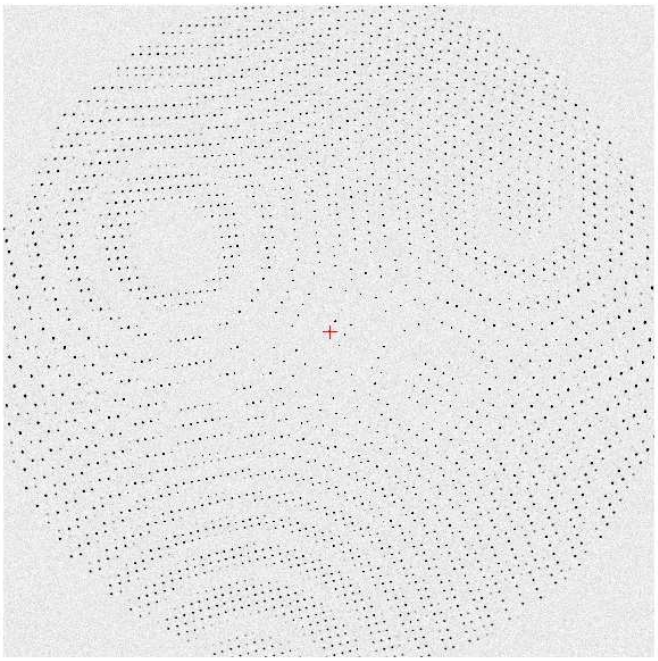


Data processing

- Indexing (finding the unit cell, orientation & space group)
- Integrating (determining the intensities of each spot)
- Merging (scaling data, averaging data & determining data quality)
- Calculating structure factor amplitudes from merged intensities

(Auto-) Indexing Methods

1. **Fourier Method** Bricogne (1986), Rossmann (1999), used in Mosflm (and most likely Denzo/ HKL2000)
   In *very* brief: Fourier Transformation along one correct lattice direction has main frequency at length of lattice constant (plus higher harmonics), for other directions peaks are smaller
   $\Rightarrow$ test many directions (mosflm: about 7300 steps)
2. **Difference Vectors** Kabsch (1993), used in XDS
   Basics:
   (a) Calculation of reciprocal lattice points from reflection data.
   (b) Fitting of parameters based on **differences** of locally close lattice points
   (c) Using small differences reduces the effect of systematic errors



(Tim Grüne, Gottingen)

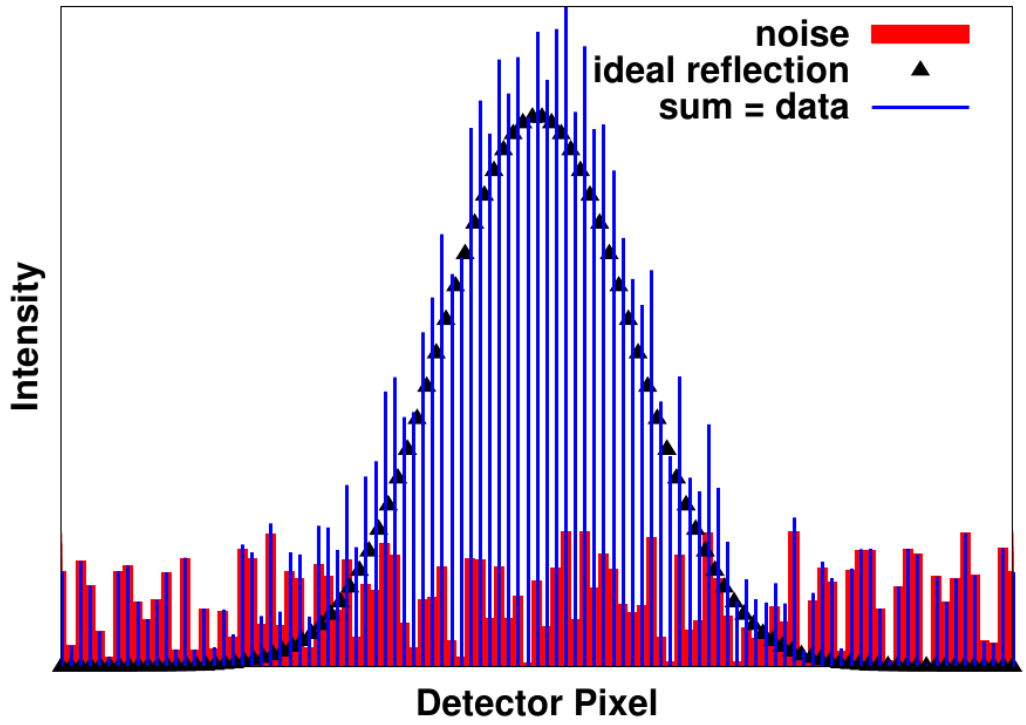
Data Integration = Data Reduction

Finding Spots per Frame:

1. Calculate expected spot positions (X_{calc}, Y_{calc}) from experimental parameters
2. Extract pixel intensities
   - Spot size/ extent (2D or 3D)
   - Background
   - Border between background and spot region
3. Calculate spot centroid (X_{obs}, Y_{obs})
4. Comparison between (X_{obs}, Y_{obs}) and (X_{calc}, Y_{calc}) allows refinement of experimental parameters.

(Tim Grüne, Gottingen)

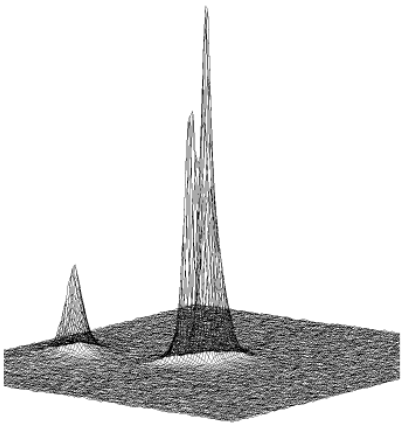
Spot & Background Noise



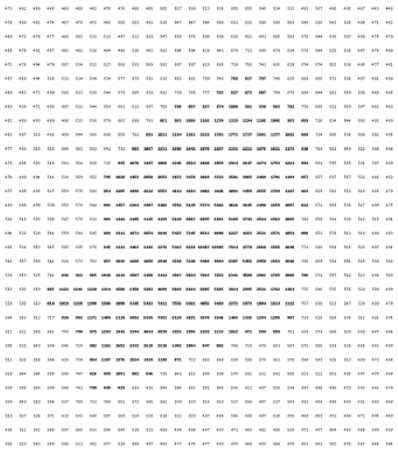
(Tim Grüne, Gottingen)

Strong Spots

3D profile



Numeric values per pixel



Dark: spot area

Light: background area

1. Sum up pixel values in **spot area**
2. Calculate average background from **flat area**
3. Substract background per counted pixel

Large spot area and sharp borders: correct intensity, even if spot area too large.

Key for overlap, weak reflections
Can lead to negative intensity (not an issue)

(Tim Grüne, Gottingen)

XDS Characteristics

Some of the special features of XDS:

- Very flexible
- Large number of supported detector formats
- 3-dimensional spot integration
- Correction for Radiation Damage
- Optimised for new Pilatus Detector
- Command-line program
- Parallelised: Fast!
- Simple to read documentation

(Tim Grüne, Gottingen)

XDS

xds, xds_par Main program for data integration

xscale scaling program for multiple datasets. Can be replaced by other scaling programs

- scala (CCP4)
- aimless (CCP4)
- scalepack (HKL Research)
- sadabs (Bruker AXS)

N.B.: Data integration makes corrections which scaling program must not repeat. Otherwise: corruption of standard deviations (σ_I). Can hamper phasing

xdsconv converts scaled data file to other formats

`XYINIT` writes files for positional corrections of the detector plane. Most modern detectors provide already corrected images so that these to files are normally flat.

`INIT` determines initial detector background

COLSPOT locates strong diffraction spots and saves their centroids

`IDXREF` indexing: unit cell dimensions and crystal orientation

`DEFPIX` set active dectector area (exclude resolution cut-off, beam stop shadow, ...)

(`XPLAN`, optional) generate "strategy" table with data completeness depending on

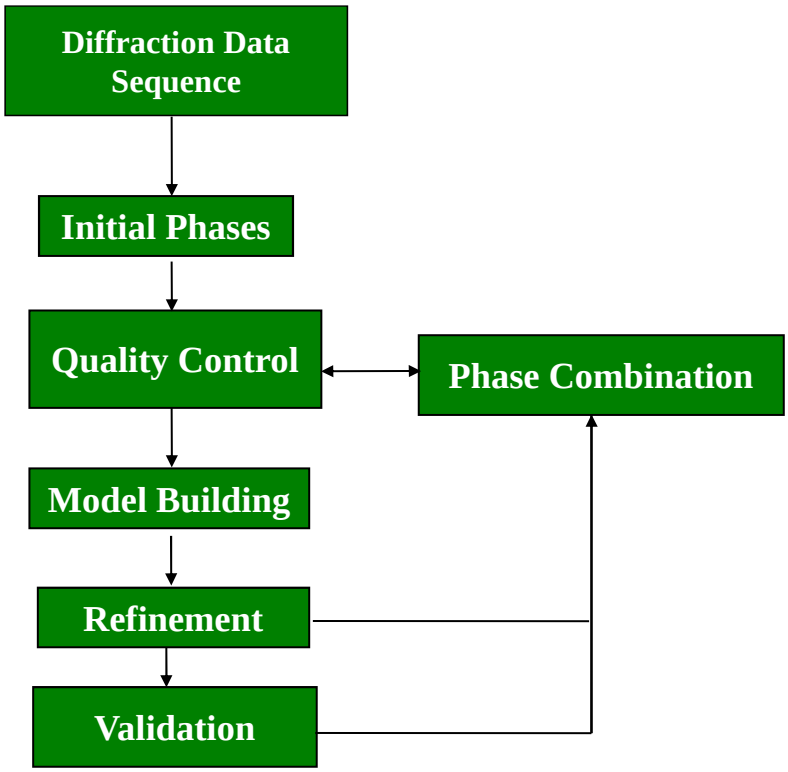
- starting angle
- total scan width

`INTEGRATE` determine reflection intensities

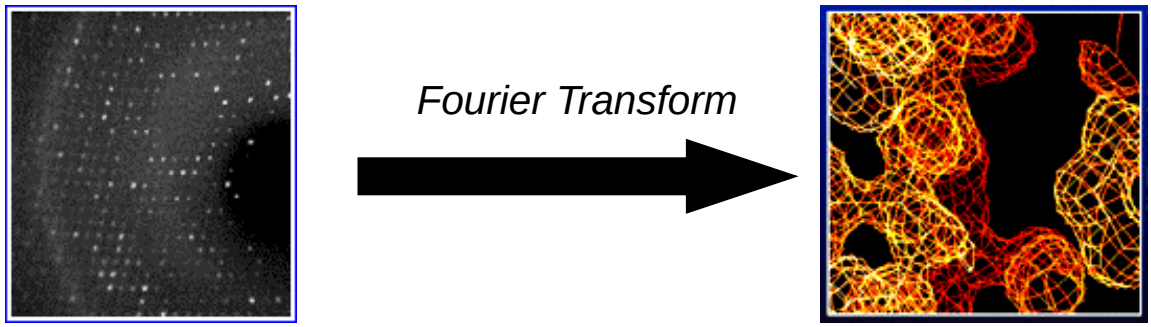
`CORRECT` applies corrections (polarisation, Lorentz-correction, ...), scales reflections, reports data statistics

(Tim Grüne, Gottingen)

The steps to solve the macromolecular crystal structure



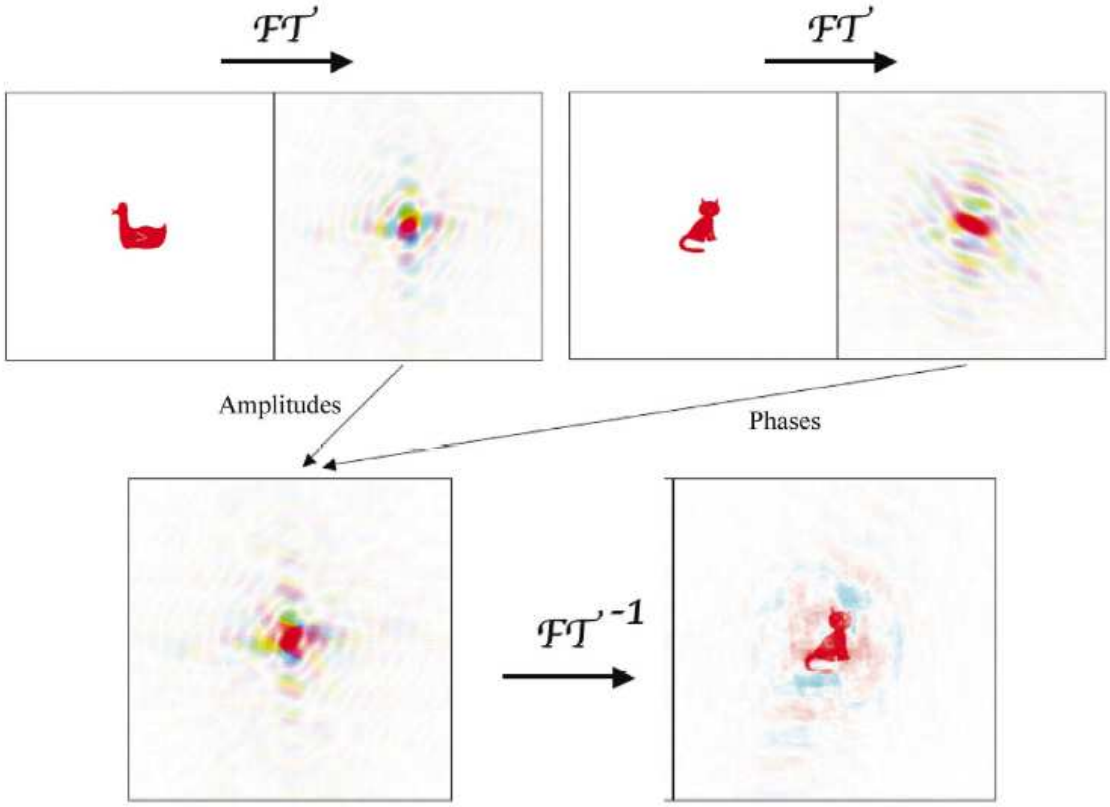
Phasing Methods in Macromolecular Crystallography



To get from the diffraction pattern to the electron density, you have to use a Fourier Transform.

- **Molecular Replacement Method (MR)**
- **Isomorphs Replacement Method (MIR, SIR)**
- **Anomalous Dispersion Method (MAD, SAD, SIRAS)**
- **Direct Method**
- **Other Methods**

Phases critically impact model quality



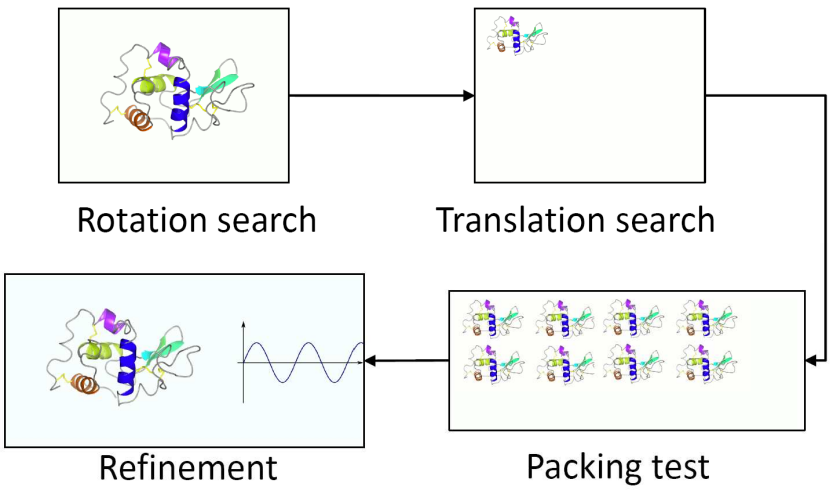
(Univ. North Carolina)

MR

The calculation involves a 6 dimensional search over all possible orientations and translations (EPMR, Genetic Algorithm).

This calculation is generally too time consuming to perform in full, so it is usually split into two parts:

- A 3 dimensional search over all possible orientations to determine the orientation of the model.
- A 3 dimensional search over all possible translations to determine the position of the orientated model.

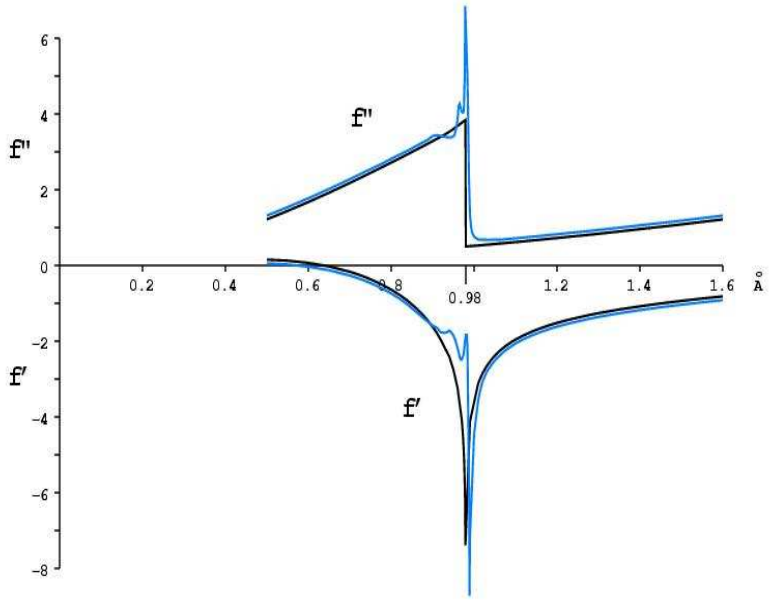


The anomalous signal

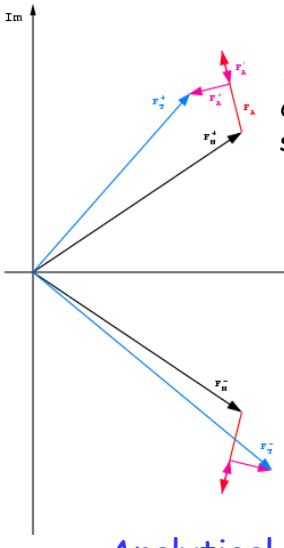
$$F(h) = \Sigma_j f_j \cdot \exp(2\pi i h \cdot r_j)$$

$$f_j = f^\circ_j(\theta) + f'_j(\lambda) + i \cdot f''_j(\lambda)$$

Anomalous correction f″ is proportional to absorption and fluorescence and f′ is its derivative

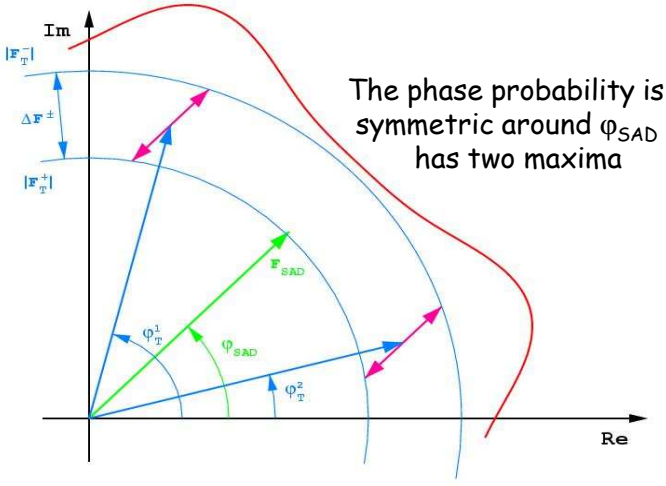


Phasing with anomalous signal



Anomalous correction f″ causes the positive shift of phase of both F(h) and F(-h)

in effect
$|F_T(h)| \neq |F_T(-h)|$
$\varphi_T(h) \neq -\varphi_T(-h)$



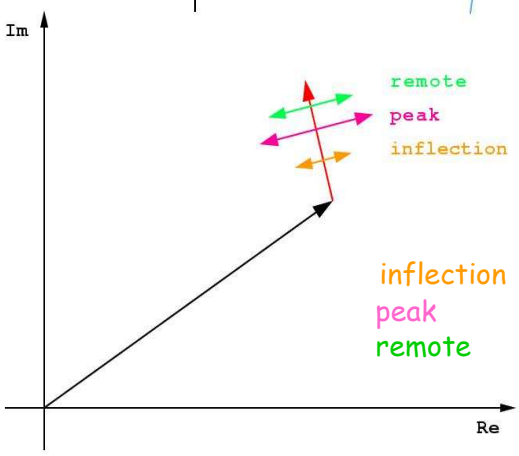
The phase probability is symmetric around φ_{SAD} has two maxima

Analytical MAD approach (Karle & Hendrickson)

$$F_T(\pm)^2 = F_T^2 + a(\lambda) \cdot F_A^2 + b(\lambda) \cdot F_T \cdot F_A \cdot \cos(\varphi_T - \varphi_A) \pm c(\lambda) \cdot F_T \cdot F_A \cdot \sin(\varphi_T - \varphi_A)$$

$$a(\lambda) = (f'^2 + f''^2)/f^{o2}$$
$$b(\lambda) = 2 \cdot f'/f^o$$
$$c(\lambda) = 2 \cdot f''/f^o$$

Three unknowns: F_T, F_A and $(\varphi_T - \varphi_A)$: system can be solved

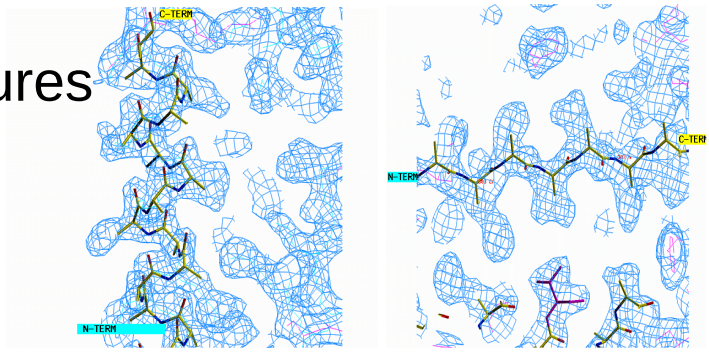


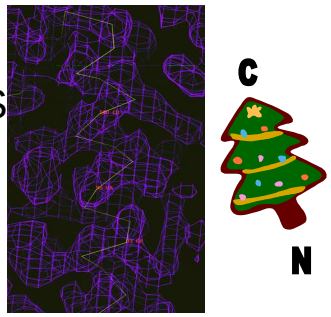
inflection	f′ = -10.5	f″ = 2.5
peak	f′ = -4.0	f″ = 8.0
remote	f′ = -0.5	f″ = 4.5

(Se K edge)

Model Building: Steps in making the first trace in electron density map

- Generating Ca chain trace
- recognize secondary structures
- Identifying chain direction
- sequence assignment
- add water molecules, etc...

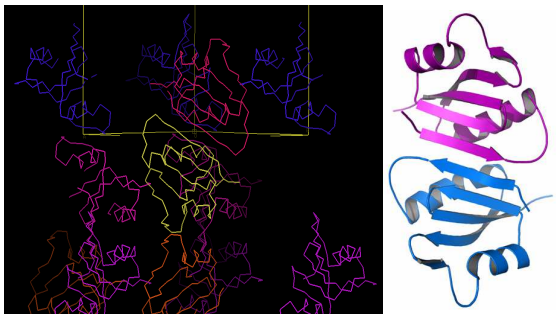




- The sulfur or Se-methionines are the perfect starting point for the sequence fitting if the map is from sulfur SAS or Se-MAD phases.
- Tryptophan is so much larger than all the other amino acids it can often be recognized.
- Hydrophilic side chains are often disordered.
- A correct fitting should be easily extended in both directions.

Structure Validation and Deposition

Generate symmetry related molecules.
(contacts >= Van der Waals packing distance)



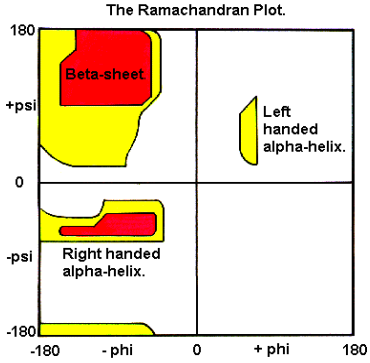
Missing density is much better than extra density

The model should make chemical sense

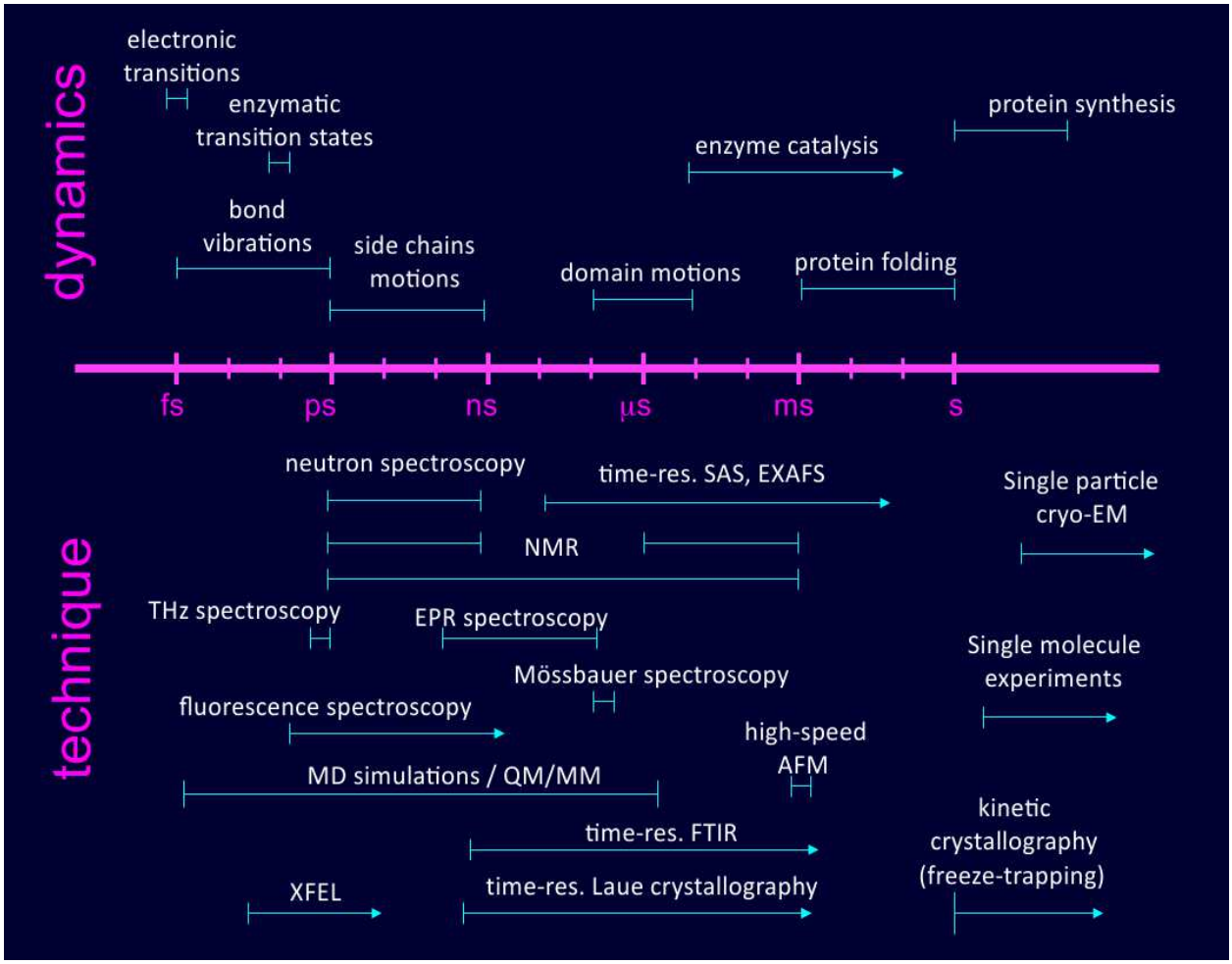
Residues identified as being in the active site:
are they close together in the model?

The stereochemical parameters such as bond length, bond angle etc, should within the standard deviation from their ideal values

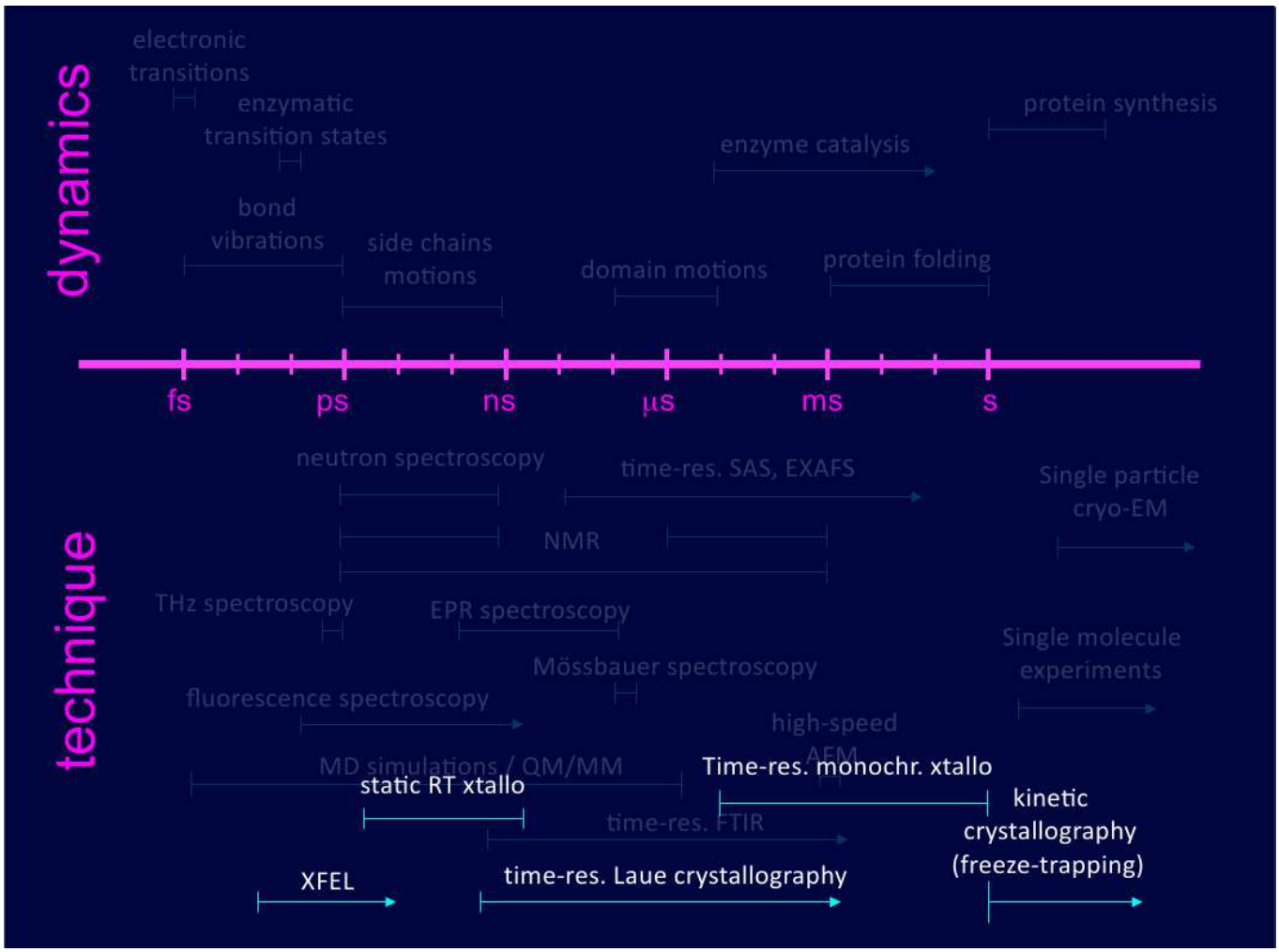
The Ramachandran Plot should be normal



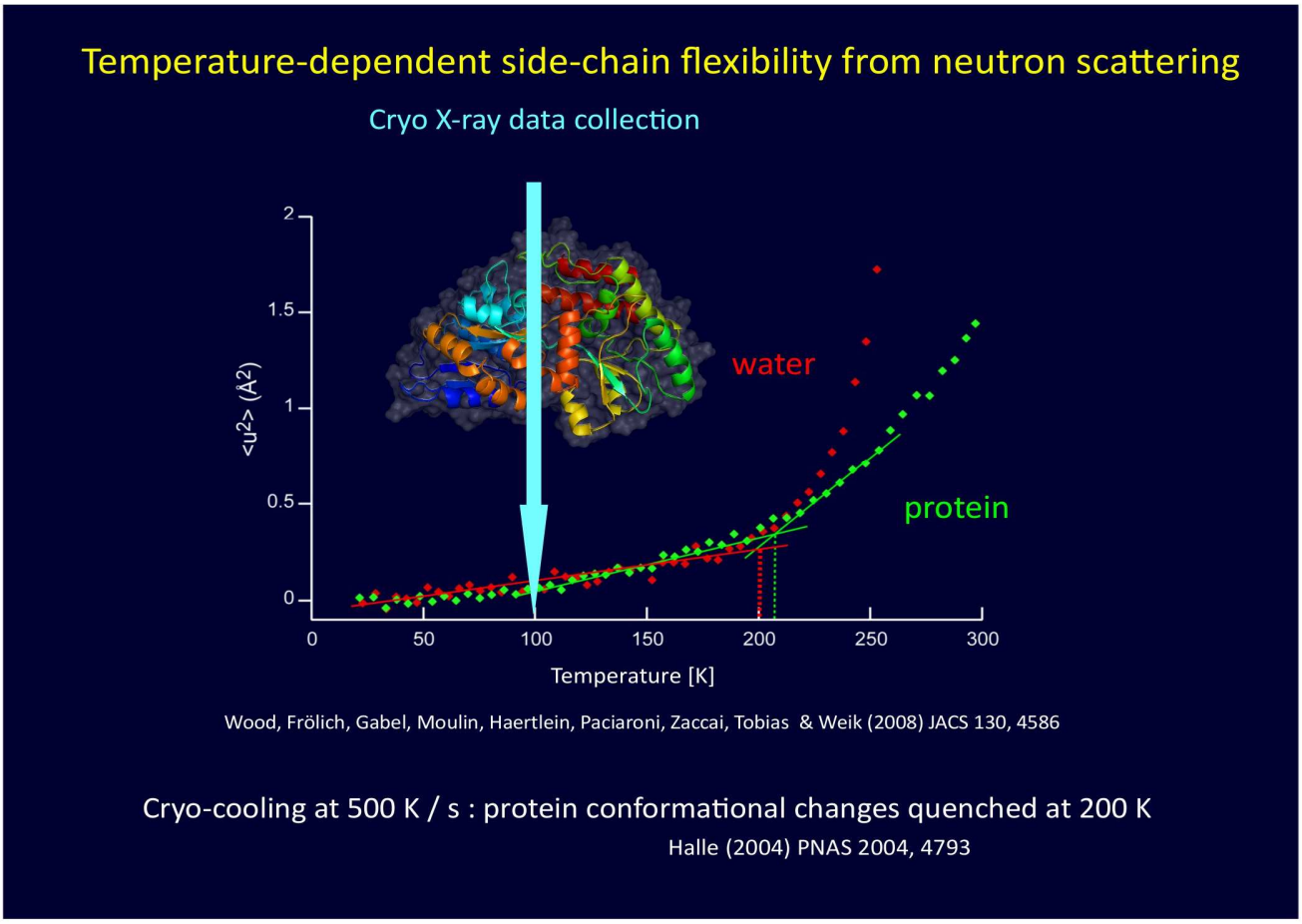
==> WHATCHECK, MOLPROBITY, ...



M. Weik, ESRF Users Meeting, 2014



M. Weik, ESRF Users Meeting, 2014



M. Weik, ESRF Users Meeting, 2014

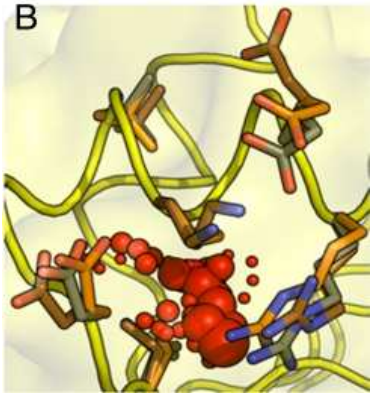
Room temperature crystallography

Flash cooling of protein crystals

- biases structural collective motions in protein crystals;
- remodels the conformation of > 35% of side chains;
- eliminates packing defects necessary for functional motions;
- induces bias toward smaller, overpacked, and unrealistically unique models.

Instead, **room-temperature** X-ray crystallography experiments, such as the *in situ* experiments, helps in revealing

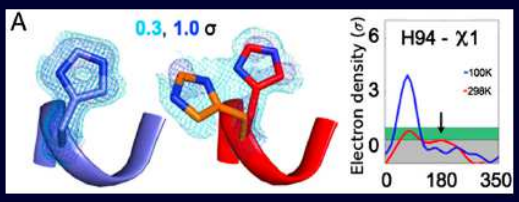
- motions crucial for catalysis,
- ligand binding,
- allosteric regulation.



In the signaling switch protein, H-Ras, an allosteric network consistent with fluctuations detected in solution by NMR was uncovered in the room-temperature, but not the cryogenic, electron-density maps (Fraser *et al.*, PNAS, 2011 (108), 16247-52).

Protein conformational heterogeneity greater in RT than in 100 K structures

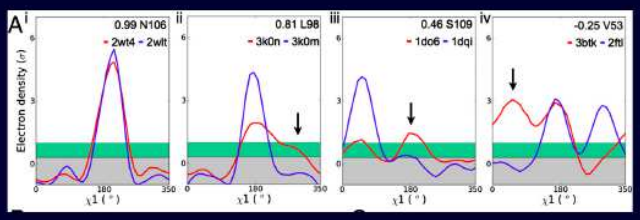
Fraser, van den Bedem, Samelson, Lang, Holton, Echols & Alber (2011) PNAS 108, 16247



100 K RT

Alternate conformation of H94
In H-Ras at RT, but not at 100 K

Cryo-cooling remodels
conformational distributions in
35% of all protein side-chains



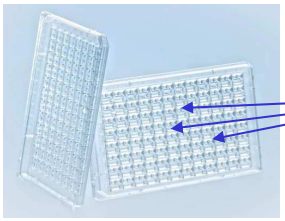
Tools to analyse conformational heterogeneity in crystal structures:

- RINGER: samples e- density around side-chain dihedrals below 1σ level (Lang *et al.* (2010) Protein Sci. 19, 1420)
- qFit: automates building of alternative polypeptide conformations (van den Bedem *et al.* (2009) Acta Cryst. D65, 1107)
- Time-averaged crystallographically restrained MD refinement of ensembles (Burnley et al. (2012) eLife 1, e00311)

- END, RAPID: place e- density maps on absolute scale and calculate noise at each position in the map (Lang *et al.* (2014) PNAS 111, 237)

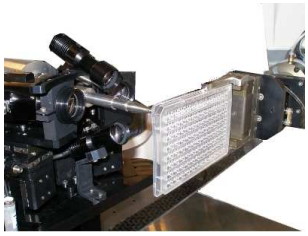
M. Weik, ESRF Users Meeting, 2014

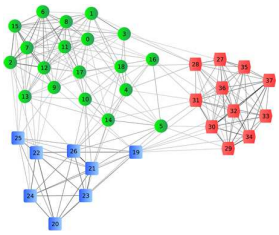
Data collection at RT



Same condition reproduced all over the plate
Or multiple crystals in capilaries

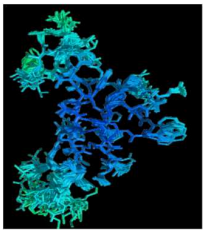
Data collection on selected crystals



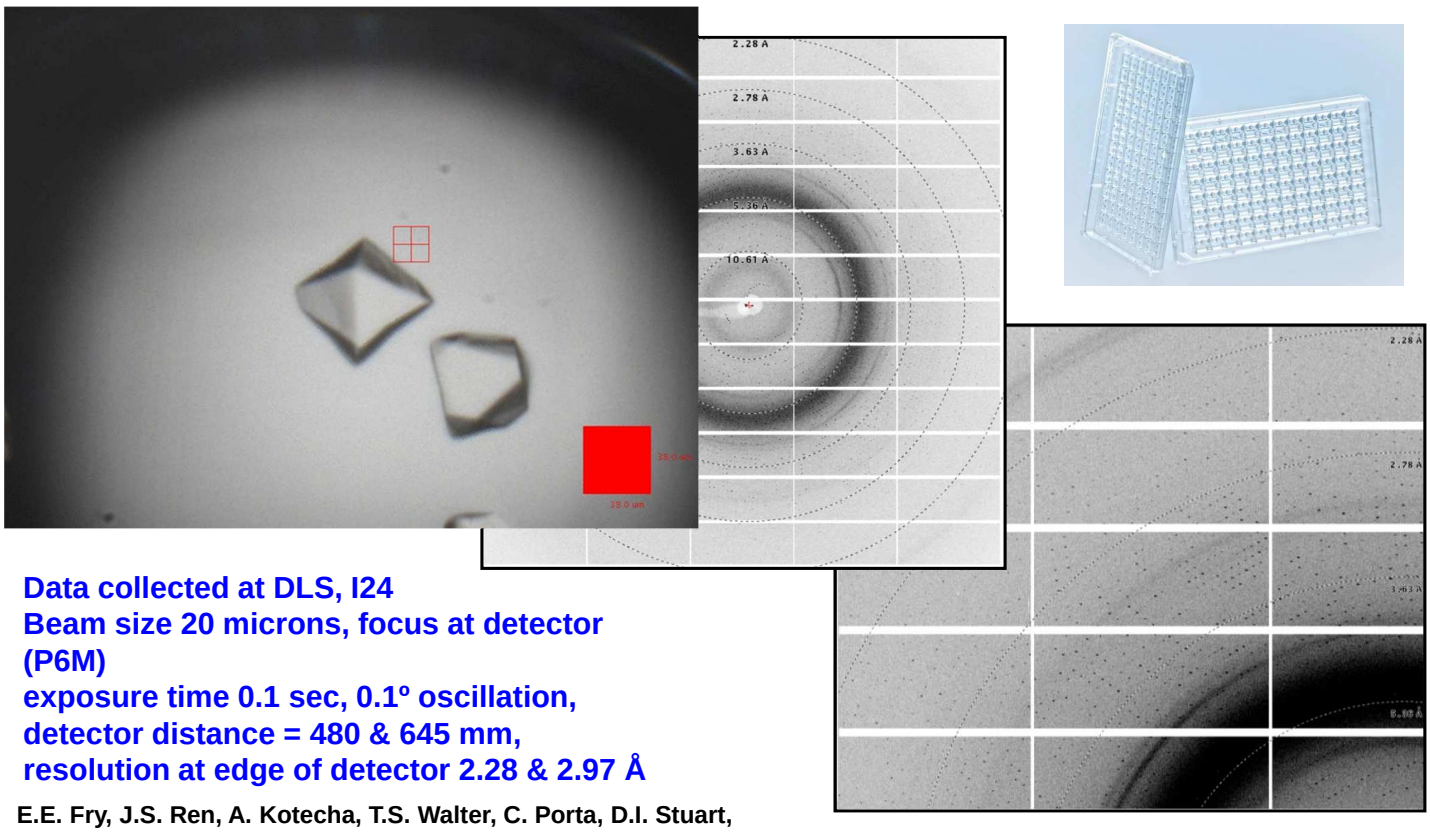


Automated clustering of data set + merging:
Similar cell param. (BLEND) or CC

« ensemble » refinement (Phenix)



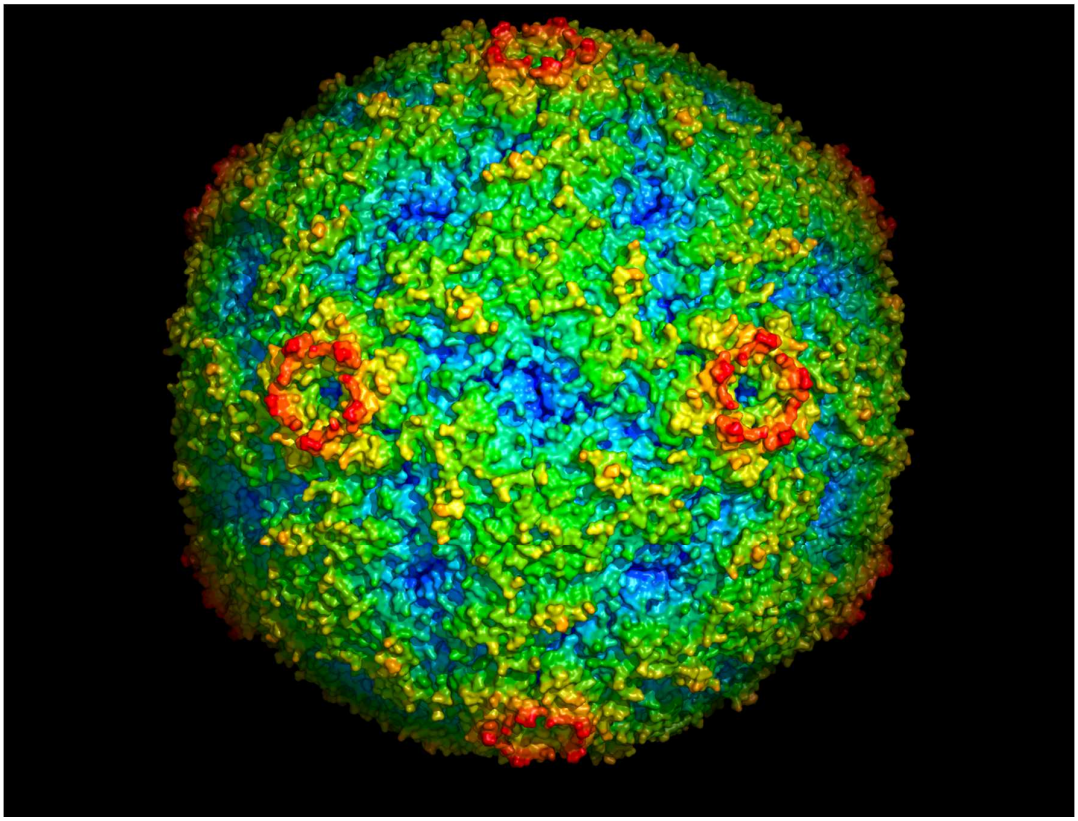
Bovine enterovirus 2
Crystallization plate screening on I24



Data collected at DLS, I24
Beam size 20 microns, focus at detector (P6M)
exposure time 0.1 sec, 0.1º oscillation,
detector distance = 480 & 645 mm,
resolution at edge of detector 2.28 & 2.97 Å

E.E. Fry, J.S. Ren, A. Kotecha, T.S. Walter, C. Porta, D.I. Stuart,
The Wellcome Trust Centre for Human Genetics, University of Oxford (UK),
D.J. Rowlands, Institute of Molecular and Cellular Biology, University of Leeds (UK) and
Gwyndaf Evans, Robin Owen, Danny Axford, Jun Ashima, I24, Diamond Light Source (UK)

A new virus structure: Bovine enterovirus 2
Crystallization plate screening on I24 (DLS)



E.E. Fry, J.S. Ren, A. Kotecha, T.S. Walter, C. Porta, D.I. Stuart,
The Wellcome Trust Centre for Human Genetics, University of Oxford (UK),
D.J. Rowlands, Institute of Molecular and Cellular Biology, University of Leeds (UK) and
Gwyndaf Evans, Robin Owen, Danny Axford, Jun Ashima, I24, Diamond Light Source (UK)