Automation on Beamlines for Structural Biology

Seán McSweeney

European Synchrotron Radiation Facility
Grenoble, France

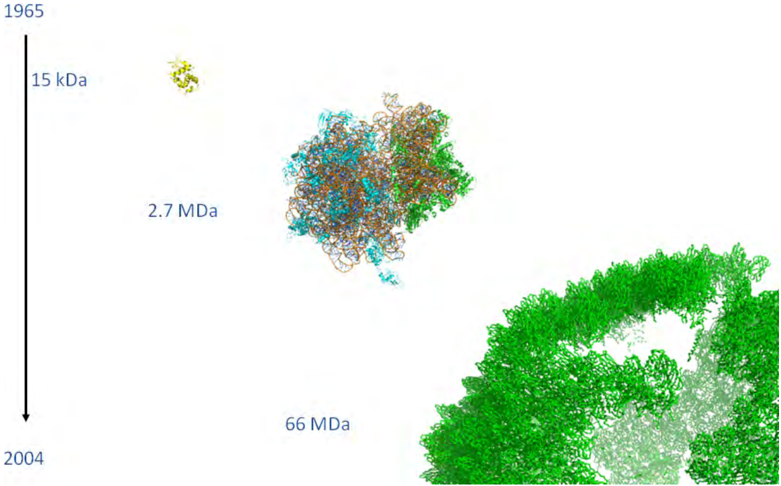
May 18, 2011

Outline

without whom

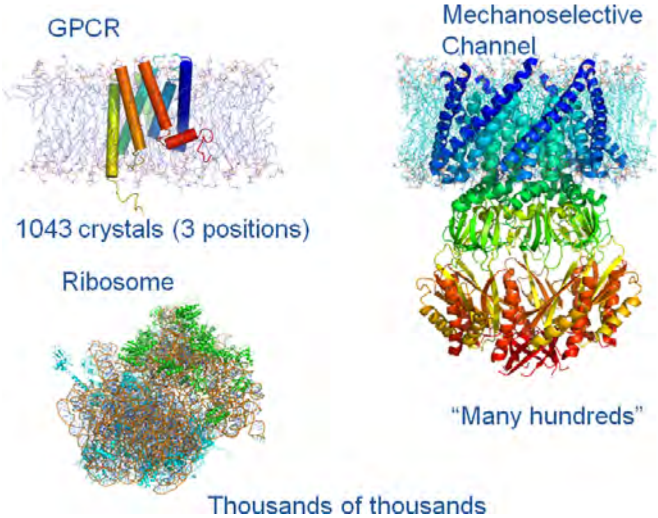
- Gordon Leonard
- ID14-1 - Philippe Carpentier
- ID14-2 - Matthew Bowler
- ID14-3 - Petra Pernot / Adam Round (EMBL)
- ID14-4 - Andrew McCarthy (EMBL) / Sandor Brockhauser (EMBL)
- ID23-1 - Sasha Popov
- ID23-2 - Max Nanao (EMBL)
- ID29 - Daniele de Sanctis / GAL
- Christoph Müller-Dieckmann
- Didier Nurizzo
- David Flot
- Joanna Timmins
- Debby Davison
- Claudine Romero
- Laurence Serre
- Trevor Mairs
- Pascal Theveneau
- Werner Schmid
- John Surr
- Thierry Giraud
- Fabien Dobias
- Mario Lentini
- Hugo Cassarotto
- Antonia Beteva
- Ricardo Nogueira Fernandes
- Darren Spruce
- Matias Guijarro
- Vicente Rey
- Ulrike Kapp
- Samira Acajjaoui
- Elspeth Gordon
- Stephanie Malbet-Monaco
- Gianluca Cioci
- Ed Mitchell
- Mats Ökvist (NORSTRUCT)
- Antoine Royant (IBS Cryobench)
- Martin Weik (IBS Cryobench)
- Carlo Petosa (IBS)
- Cyril Dian (IBS)
- Jens Radzimanowski
- Meike Stelter
- Tobias Klar
- Ricardo Leal
- Dario Piano
- Sofia Caria
- Amandine Lallemand
- Tommaso Tosi
- Simone Pellegrino
- Rita Giordano
- Joana Rocha
- Erika Pellegrini
- Chloe Zubieta

little of what you will see would have been possible without the close cooperation between EMBL and ESRF.

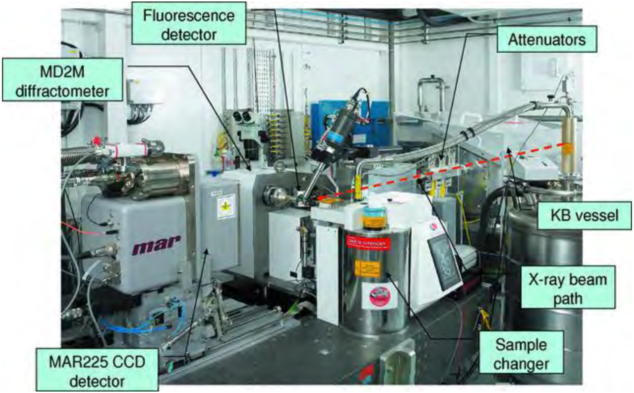


The increase in the complexity of the subjects of study in Structural Biology

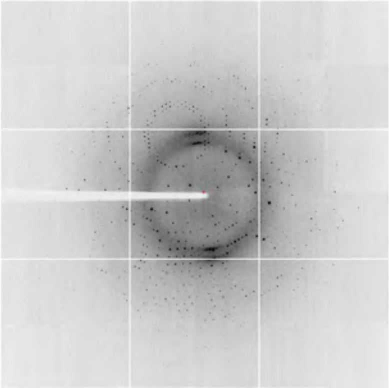
real samples



Exptl Hutch.


Fluorescence detector
Attenuators
MD2M diffractometer
mar
KB vessel
X-ray beam path
Sample changer
MAR225 CCD detector

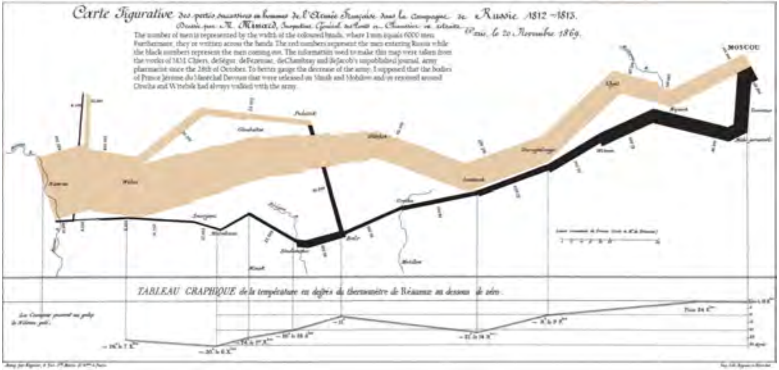
Crystals



Typical Crystal and an atypical diffraction pattern

Data ≠ knowledge

European Synchrotron Radiation Facility
A Light For Science



We have no problem generating data, or data sets. However the generation of Scientific knowledge is an issue.

Crystallographic Data Collection.

Crystallographic data collection is the measurement of **Intensities**.

$$I \propto F_{hkl}^2$$

where the structure factor is

$$\begin{aligned} F_{hkl} &= \sum_{hkl} f_n exp 2\pi i(hx + ky + lz) \\ &= |F_{\mathbf{H}}| exp(i\alpha_{\mathbf{H}}) \end{aligned}$$

but in real cases

$$F_{hkl} = \sum_n O_n f_n exp(\frac{-Bsin^2\theta}{\lambda^2}) exp(2\pi i(hx + ky + lz))$$

Electron Density Equation.

Relationship to the structure factor

$$\rho_{(x,y,z)} = \frac{1}{V} \sum_{hkl} F_{hkl} exp(-2\pi i(hx + ky + lz) + i\alpha(hkl))$$

with an integrating detector the phase ($\alpha(hkl)$) is lost in the measurement of $I(hkl)$.

To understand the experiments we are performing, we need to think about the form of the structure factor and electron density equations.

Measurement of Structure Factors

The absolute magnitude of the integrated intensity is proportional to δV the volume of the sample illuminated.

$$I_{\mathbf{H}} = Q\delta V$$

Where the constant Q is

$$Q = (e^2/mc^2)N_C^2\lambda^3 L_p|F_{hkl}|^2$$

where N_c is the number of unit cells and λ the wavelength of the X-rays. F_{hkl} is the structure factor.
In general we measure on the magnitude $|F_{hkl}|$ the observed intensity, but not the phase component. **Estimating the phase is the core of our business**.

Considerations for a crystallographic dataset.

Data collection requires the optimisation and consideration of (at least) the following properties of the dataset.

- Resolution
- Completeness
- Accuracy
- Measurement Redundancy

Each type of phasing experiment will require compromises with respect to these properties and will be conditioned by the behaviour of the crystal in the X-ray beam. The compromises that will be acceptable will depend upon the type of phasing experiment to be undertaken.

We have largely automated all steps of planning reasonable experiments.

The effect of resolution.

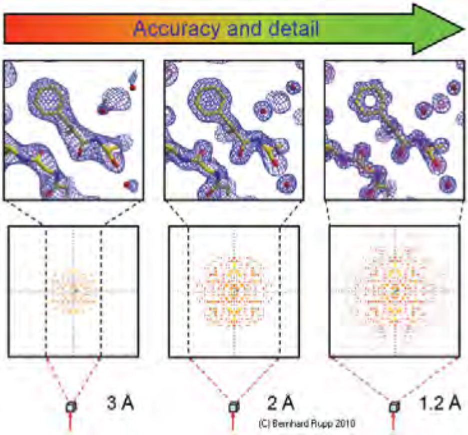


Figure 1-6 Data quality determines structural detail and accuracy. The qualitative relation between the extent of X-ray diffraction, the resulting amount of available diffraction data, and the quality and detail of the electron density reconstruction and protein structure model are evident from this figure: The crystals are labeled with the nominal resolution d_{min} given in Å (Ångström) and determined by the highest diffraction angle (corresponding to the closest sampling distance in the crystal, thus termed d_{min}) at which X-ray reflections are observed. Above each crystal is a sketch of the corresponding diffraction pattern, which contains significantly more data at higher resolution, corresponding to a smaller distance between discernible objects of approximately d_{min}. As a consequence, both the reconstruction of the electron density (blue grid) and the resulting structure model (stick model) are much more detailed and accurate. The non-SI unit Å (10^{-8} cm or 0.1 nm = 10^{-10} m) is frequently used in the crystallographic literature, simply because it is of the same order of magnitude as atomic radii (~0.77 Å for carbon) or bond lengths (~1.54 Å for the C–C single bond).

Anomalous Scattering

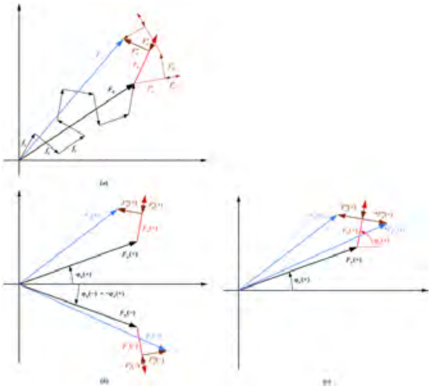
Resonance effect that occurs most strongly at absorption edges.

- $f(\theta,\lambda) = f^0(\theta) + f'(\lambda) + if''(\lambda)$
- the structure factor becomes a function of both resolution and wavelength.
- the effect is element specific.
- in the presence of anomalous scattering $F(h,k,l) \neq F(-h,-k,-l)$

Thus with anomalous scattering presence we have a way of measuring differences corresponding to a sub-set of the atoms in the structure. This gives a method for determining their postion. This knowledge allows us to estimate the phase ($\alpha(hkl)$).

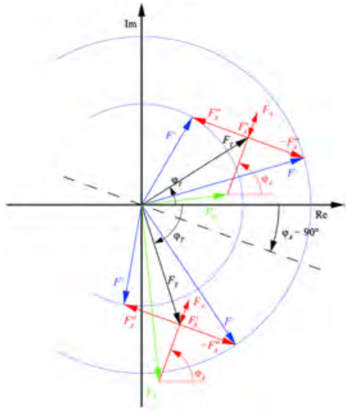
European Synchrotron Radiation Facility
A Light For Science

The effect of Anomalous scattering on the Structure Factor



in the presence of anomalous scattering $|F_{hkl}| \rightarrow |F_{hkl}(\lambda)|$

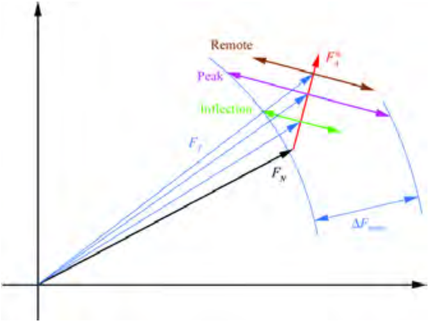
Phase Ambiguity



With one λ there are two phases that satisfy the constraints : SAD

More than one wavelength removes the Ambiguity

With MAD experiments the ambiguity is resolved by combining information from several data sets. SAD needs other information.

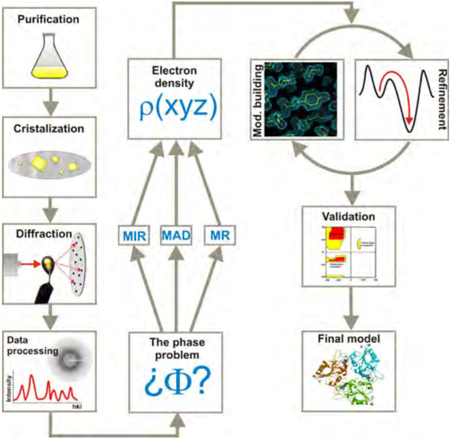


SAD and MAD data

Phases are developed following an examination of $F(h,k,l), F(-h,-k,-l)$

- **Resolution:** ideally no problem of non-isomorphism, therefore high resolution reflections contribute a lot of signal.
- **Completeness:** Need high completeness
- **Accuracy:** difference between $F(h,k,l)$ and $F(-h,-k,-l)$ is generally of order 1 to 3 %. i.e. signal is of the same level as the measurement noise !
  - Accuracy is very important.
  - Outliers need to be removed.
  - Anomalous scatterers found via Patterson methods (dominated by the strong reflections)
  - Redundancy of measurements helps outlier detection and precision.
  - **Beware** radiation damage.

Crystallographic Experiment



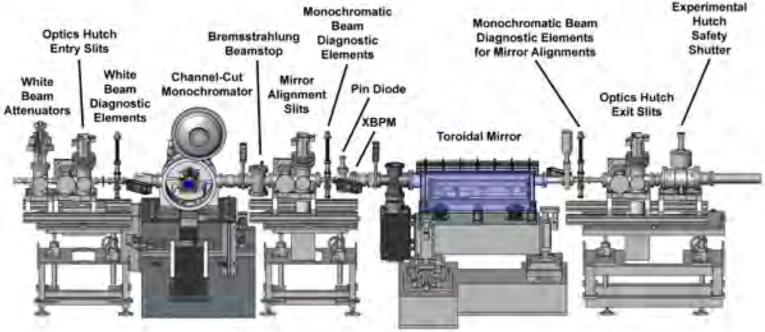
Each and every step of this process has benefited from some degree of automation in the last 10 – 15 years.

Beam Delivery

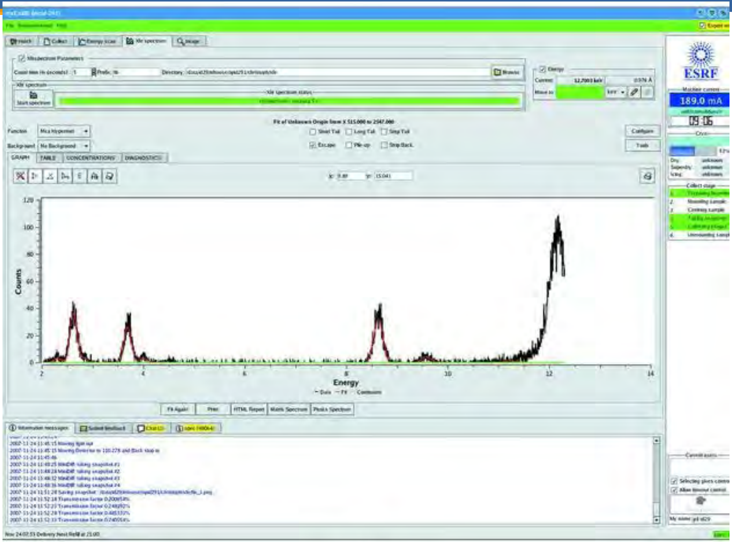
X–rays have to get to the right place with the right properties.

- Any experiment needs the X-rays to be in the right place with the right parameters.
- Provide diagnostics. . .
- Each optical element is self-contained.
- Optimisation of one flows into the next.
- reconfiguration is triggered by " significant " modifications.

Beamline Optics

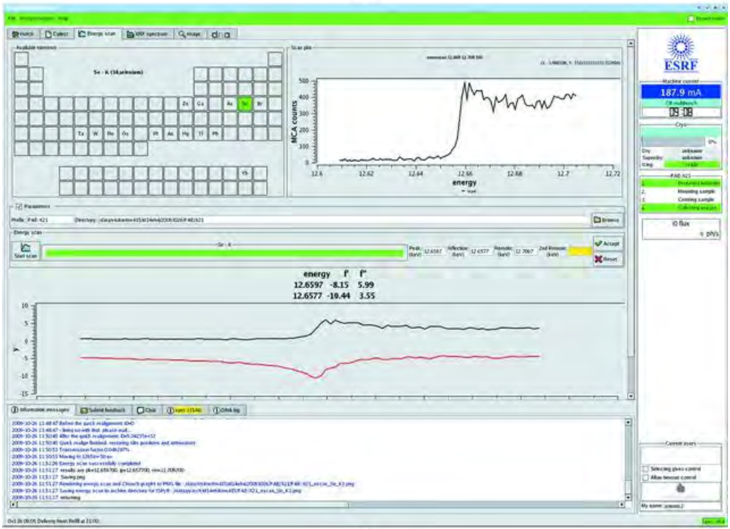


XFE


ESRF
189.0 mA
Energy
Counts
GRAPH
TABLE
CONCENTRATIONS
DIAGNOSTICS
Fit Again
Print
HTML Report
Match Spectrum
Peaks Spectrum

(b)

Finding the best $\lambda's$

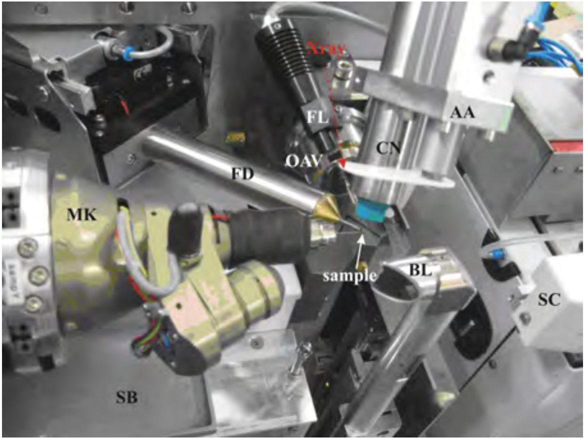


(b)

What we did . . .

- Modify source, prepare primary slits.
- tune monochromator.
- re–align mirror.
- re–align experimental table.
- check beam is at sample position.
- find lowest attenuation consistent with useful spectrum.
- perform and analyse XANES spectrum.
- pass results to data collection software. . . .

real sample environment.



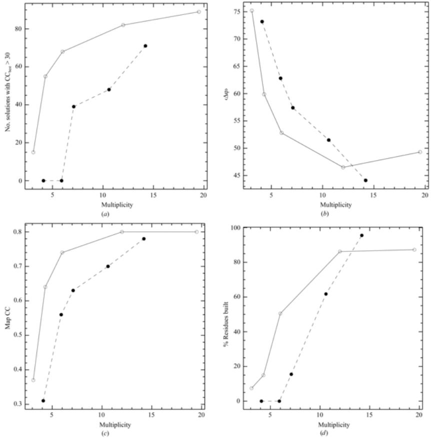
Experiment Strategy.

With a crystal in place.

- collect two or more images
- determine the crystal lattice.
- measure the Bragg peaks and the background.
- feed in beam parameters (size, λ, flux)
- calculate crystal life time.
- calculate strategy and data collection time.
- make appropriate compromises.

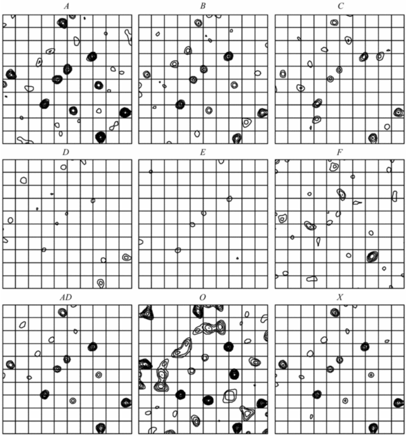
These elements are tied together with EDNA and intelligence provided with BEST.

Multiplicity of Experiment.



The small signal means we need precise measurements – multiplicity.

Killing off the signal.



measurment reduces the signal in the data set – radiation damage

Aim of Structural Biology Plans

Crurrent Status
why we need new tools.

- Currently measuring more than 10,000 crystals/month
- **in 2009** 113912 crystal tests . . .
- leading to 19735 data sets (17.3%). . .
- in PDB 931 depositions so far (or 4.7%). (852 in 2008 or 4.3 %)

Question.

Can we get better at deciding which crystal is **the right one**?

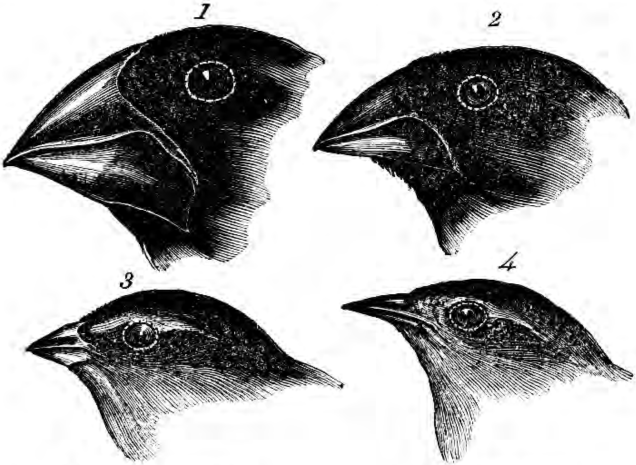
Latest from 2010

- Total evaluations: 123,309
- Total data sets: 21,540
- Top data collector: ID29 with 5897 data sets.
- Most evaluations: ID23-2 with 33798 evaluations
- Top screener: ID23-2 with only 10% of samples collected
- lowest screener: ID29 with 27.5% of sample collected
- Top(?) MAD screener: ID23-1 with 18.4% collections

Summary.

This represents an 8% increase in evaluations and a 9% increase in data sets over 2009. ID23 (1 and 2) account for 50% of datasets collected. . .

Evolution



1. Geospiza magnirostris.
3. Geospiza parvula.
2. Geospiza fortis.
4. Certhidea olivacea.

Aims and concepts

A plan has been generated along with our User community. **This plan links all Structural Biology Group beamlines in co–operative upgrades**. The expectation is to achieve more in this way than we would treating each individually. Multi–disciplinary research should be enabled.

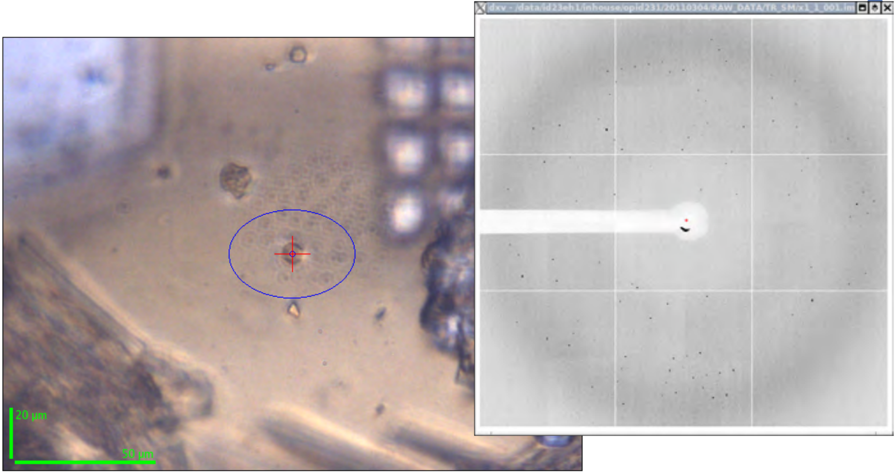
The Thinking.

- Find the **"BEST"** crystal.
- Understand the diffraction properties of this crystal.
- Perform the best experiment.

The Process.

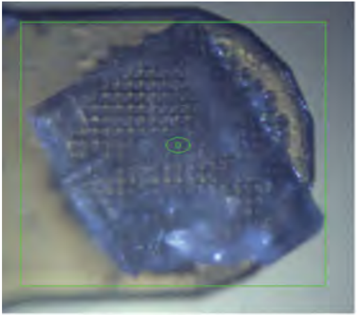
However, the Diffraction experiment is a means to the end not the object in itself. Implementation must enable access to the Biological question. *Thus each plan must be adapted to the sample and experimental aims*

Crystals

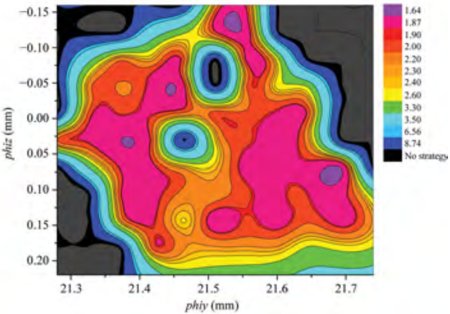


Data Collection from micro crystals.

Micro-Beam Big Xtal -1

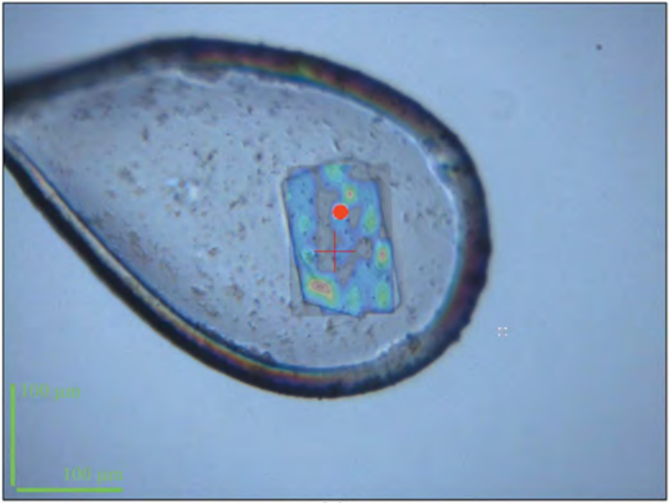


(a)



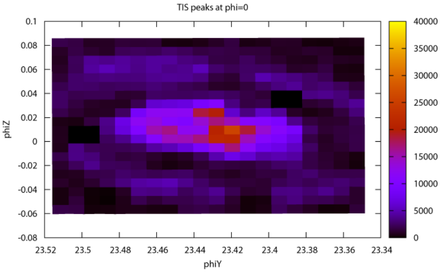
(b)

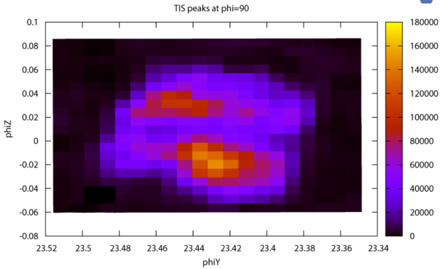
Micro-Beam Big Xtal -2



(c)

micro-beam again

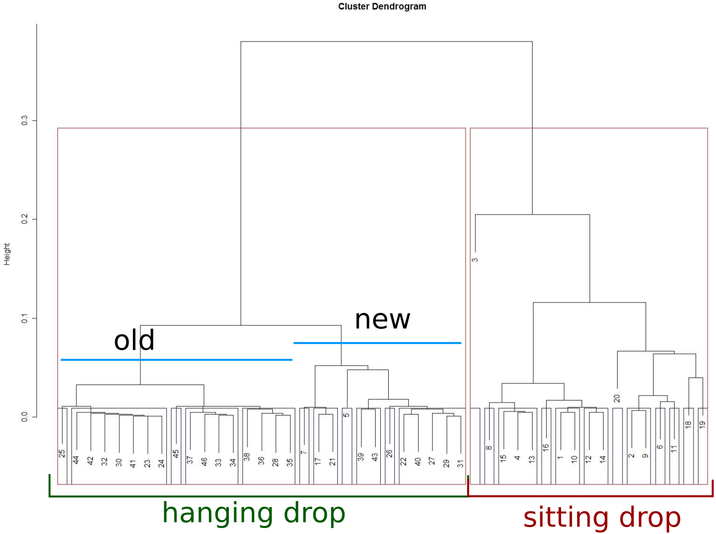




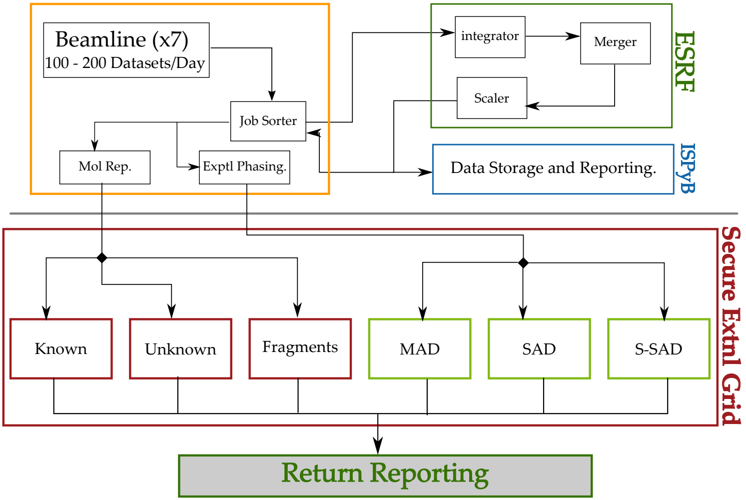
Mechanism.

Each point corresponds to the analysis of a diffraction image (`labelit`), false colour coded according to a Z-score. Map generated in quasi real time.

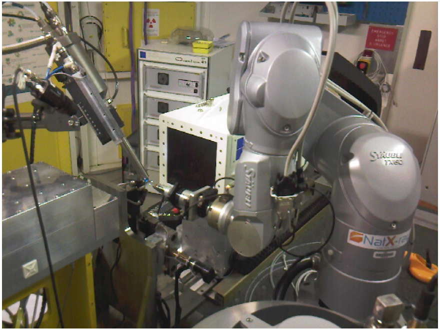
Sorting Samples



Outsourcing dealing with data



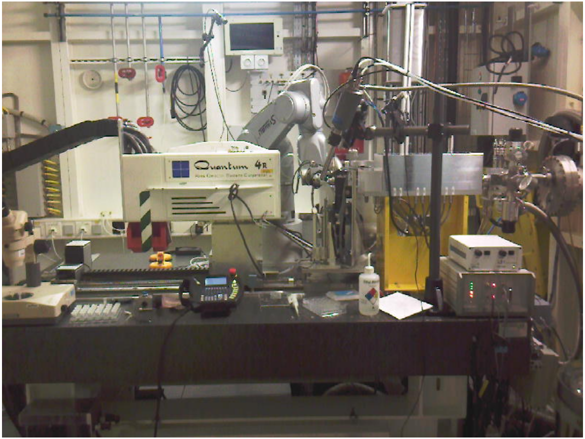
Instrumentation.



Direct data collection

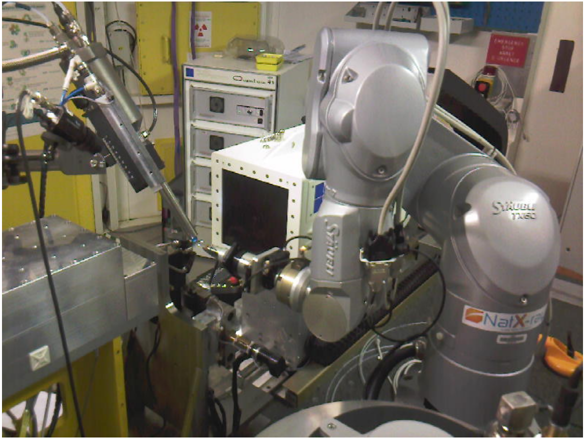
The possibility to collect data using the sample changer robot has been examined. Following extensive work with the supplier, we are now able to collect data using the CATS robot, with control from mxCuBE. Further work underway to improve utility with very small samples and X–ray beams. . . Work is under way to provide a new, standard, high-density sample holder.

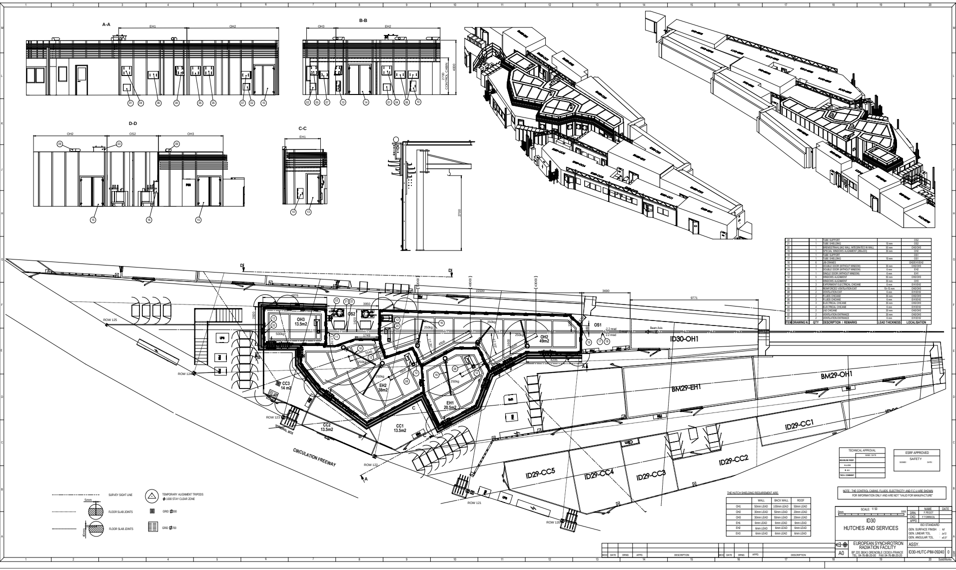
Sorting Samples Collection Data



Sorting Samples Collection Data

European Synchrotron Radiation Facility
A Light For Science





ITEM	DRAWING N.	QTY	DESCRIPTION / REMARKS	LEAD THICKNESS	LOCALISATION
22		1	TUBE SUPPORT		OS2
21		1	TUBE SHIELDING	10 mm	OS2
20		1	BREMSSTRAHLUNG WALL INTEGRATED IN WALL	50 mm	OH2/OH3
19		1	SPECIAL WINDOWS ALIGNMENT (280x220)	30 mm	OH2
18		1	TUBE SUPPORT		OS1
17		1	TUBE SHIELDING	10 mm	OS1
16		6	JIB-CRANES		OH2/EH1/EH2
15		3	DOUBLE DOOR (WITHOUT WINDOW)	30 mm	OH2/OH3
14		1	DOUBLE DOOR (WITHOUT WINDOW)	6 mm	EH2
13		1	SINGLE DOOR (WITHOUT WINDOW)	6 mm	EH1
12		5	WINDOWS ALIGNMENT	50 mm	OH2/OH3
11		1	WINDOWS ALIGNMENT	30 mm	OH3
10		3	EXPERIMENT ELECTRICAL CHICANE	6 mm	EH1/EH2
09		2	REINFORCED VENTILATION EXIT	10+10 mm	OH2/OH3
08		2	VENTILATION EXIT	6 mm	EH1/EH2
07		2	FLUIDS CHICANE	30 mm	OH2/OH3
06		3	FLUIDS CHICANE	6 mm	EH1/EH2
05		3	ELECTRICAL CHICANE	30 mm	OH2/OH3
04		4	ELECTRICAL CHICANE	6 mm	EH1/EH2
03		2	LN2 CHICANE	30 mm	OH2/OH3
02		2	VENTILATION ENTRANCE	30 mm	OH2/OH3
01		2	VENTILATION ENTRANCE	6 mm	EH1/EH2

TECHNICAL APPROVAL	NAME / DATE
BEAMLINE RESP.	
A.L.G.E.	
B. & I.	
TECH. COMMENT	

ESRF APPROVED

SAFETY

SIGNED : DATE :

NOTE : THE CONTROL CABINS, FLUIDS, ELECTRICITY AND F.C.U ARE SHOWN FOR INFORMATION ONLY AND ARE NOT "VALID FOR MANUFACTURE"

THE HUTCH SHIELDING REQUIREMENT ARE:

	WALL	BACK WALL	ROOF
OH1	50mm LEAD	100mm LEAD	50mm LEAD
OH2	30mm LEAD	50mm LEAD	20mm LEAD
OH3	30mm LEAD	50mm LEAD	20mm LEAD
EH1	6mm LEAD	6mm LEAD	6mm LEAD
EH2	6mm LEAD	6mm LEAD	6mm LEAD
EH3	6mm LEAD	6mm LEAD	6mm LEAD

MOD	DATE	DRND.	APPD.	DESCRIPTION	MOD	DATE	DRND.	APPD.	DESCRIPTION

SCALE: 1/50 (0mm – 5000)

ID30 HUTCHES AND SERVICES

	NAME	DATE
DRN.	F.PEGOT	
CKD.	F.TORRECIL	
APPD.		

ISO STANDARD
GEN. SURFACE FINISH : N7
GEN. LINEAR TOL. : Js12
GEN. ANGULAR TOL. : ±0,5°

A0

EUROPEAN SYNCHROTRON RADIATION FACILITY
BP 220 38043 GRENOBLE CEDEX-FRANCE
TEL. 04-76-88-20-00 FAX 04-76-88-20-20

ASSY.

ID30-HUTC-PIM-09240 0

ESRF

SolidWorks