

**FIP -BM30A:  
an automated beamline  
for protein crystallography**

**Automated data collection (Xnemo)  
Cryogenic Automated Transfer System (CATS)  
Automated Data Processing (ADP)**



*J.-L. Ferrer  
FIP-BM30A / ESRF  
Institut de Biologie Structurale  
Groupe Synchrontron*





## *The FIP beamline*

- installed on a Bending Magnet section (BM30A) at ESRF
- devoted only for protein crystallography
- tunable energy for anomalous scattering use (MAD/SAD)
- financed by:

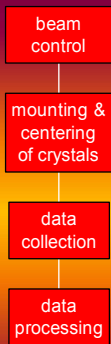


### Beam time

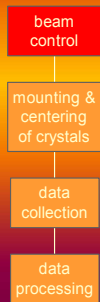
- 33% ESRF users
- 7 % Belgium users
- 60 % French users



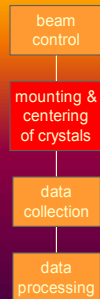
# Automation of the crystallography experiment



Automation of the beam control (1998)



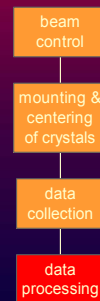
Automation of mounting and centering (2003)



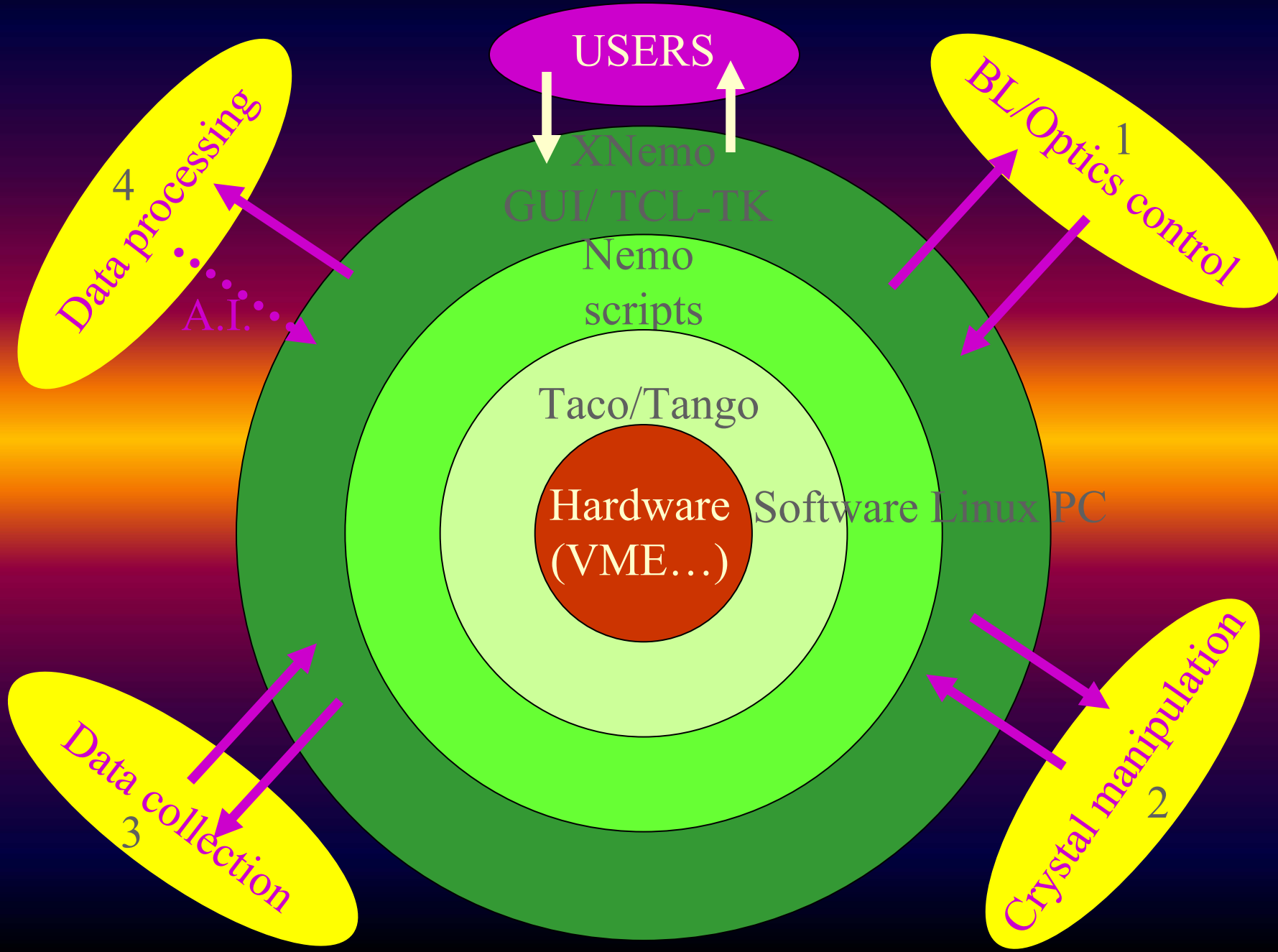
Automation of data collection (1999)



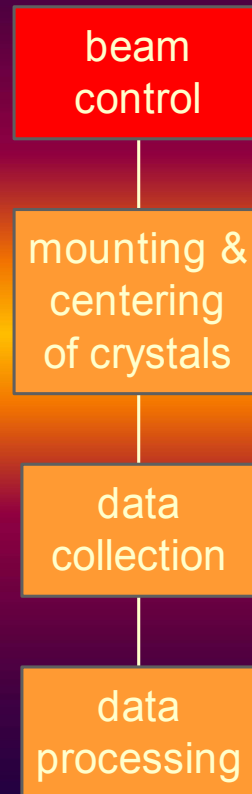
Automation of data processing (2001)



# Architecture for experiments management



# 1- Automated experiment: beam control



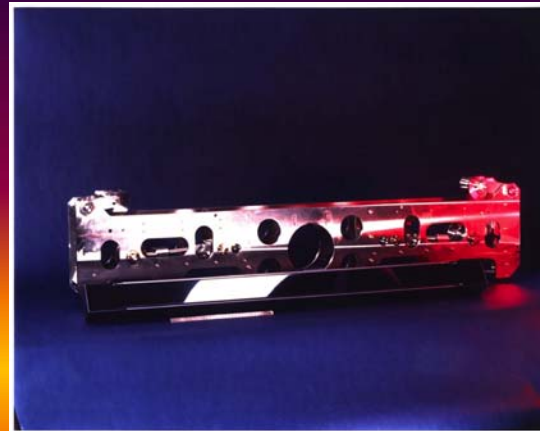
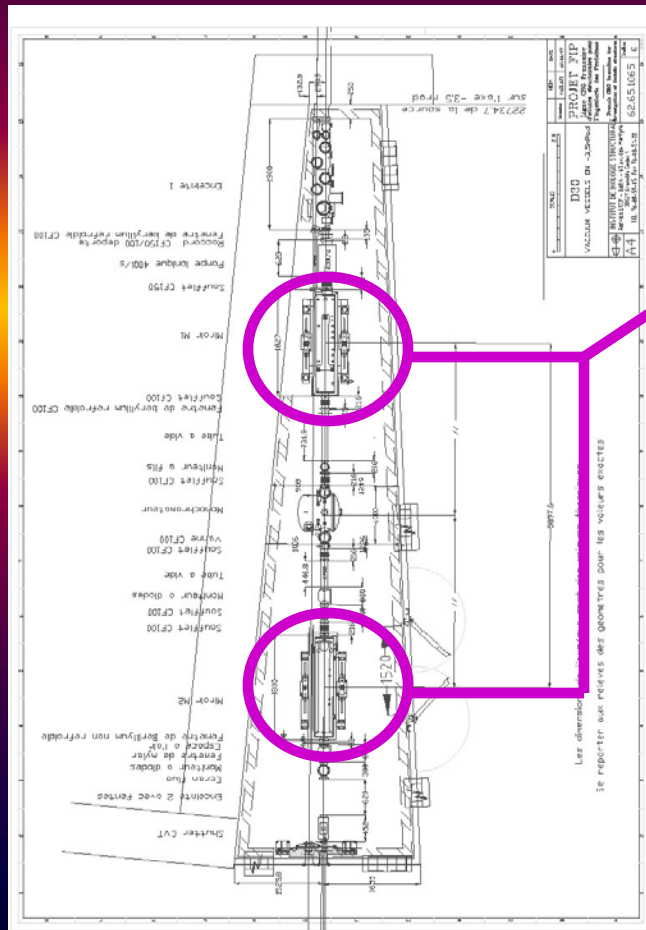
beam control

mounting & centering of crystals

data collection

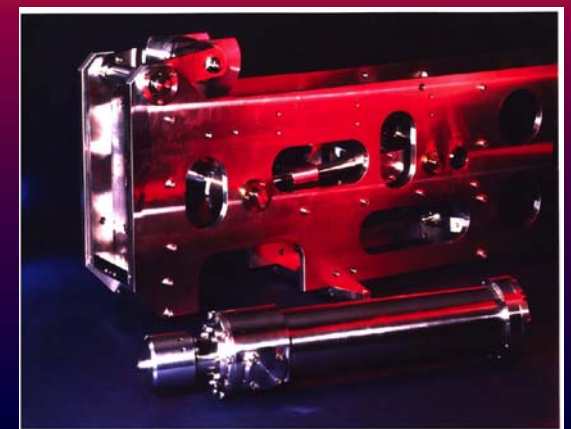
data processing

# Two grazing angle mirrors



1.3 m long, 6 cm large  
Si / Zerodur™  
water cooling (Ga bath)

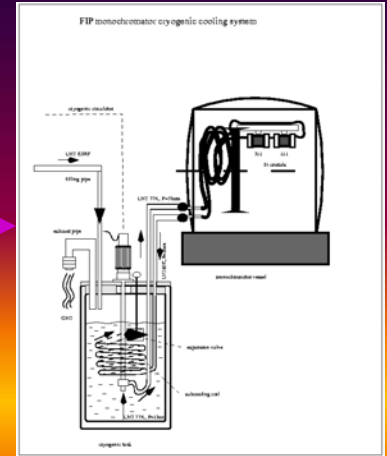
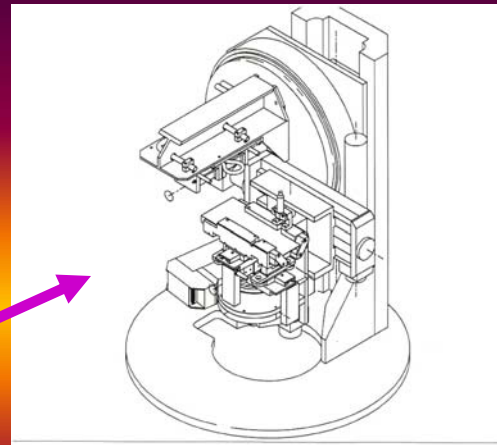
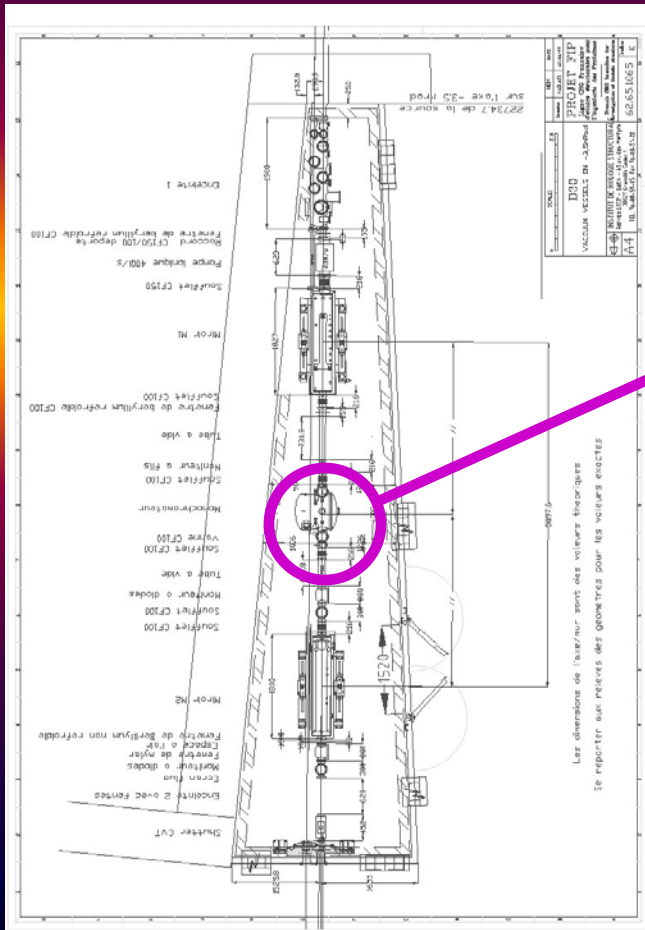
vertical focusing  
harmonics rejection  
heat load reduction



- beam control
- mounting & centering of crystals
- data collection
- data processing

# Two crystals monochromator (13 motors!)

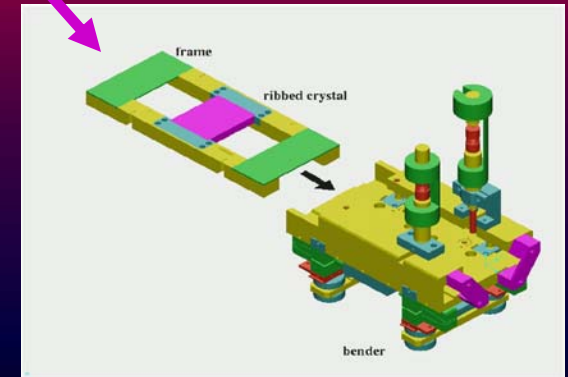
2Si(111)



heat load reduction

Monochromatic:  
 $7 \text{ keV} < E < 17 \text{ keV}$   
 $\Delta E/E \sim 10^{-4}$

Focal spot:  
 $300 \times 300 \mu\text{m}^2$



horizontal focusing

beam control

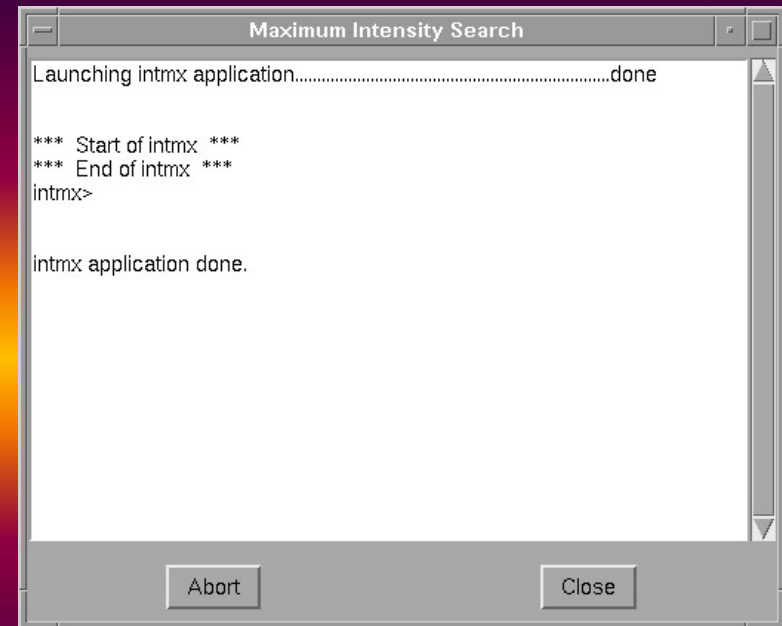
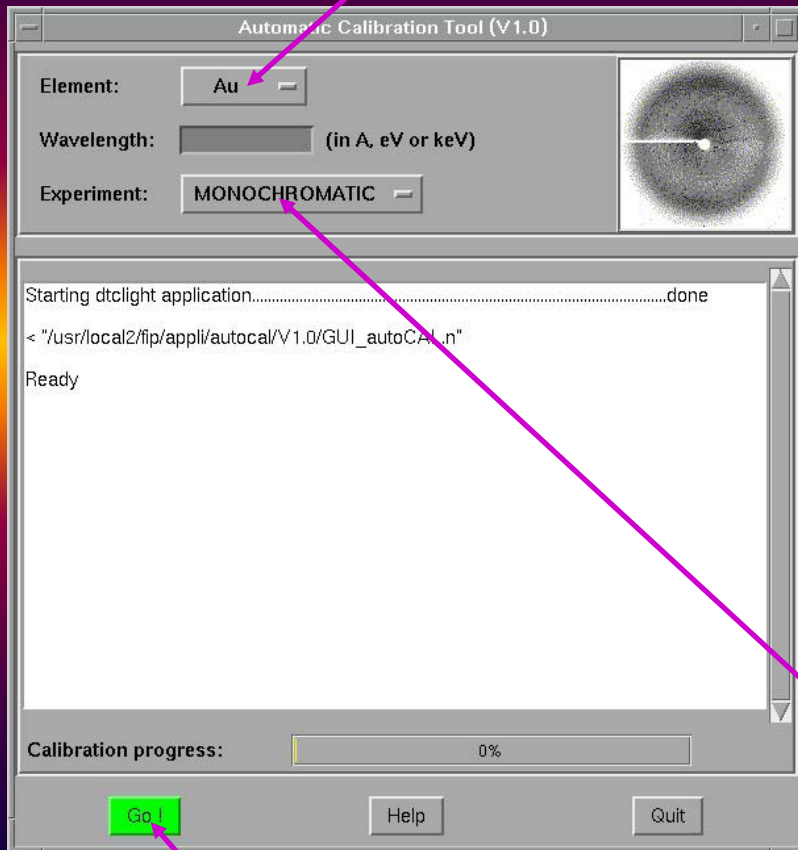
mounting & centering of crystals

data collection

data processing

# Setting up the X-ray beam

1. select edge or wavelength



2. theoretical position, or automated calibration with metal foil

3. and then: Go...



# Setting up the X-ray beam

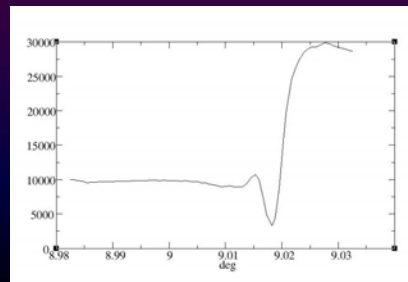
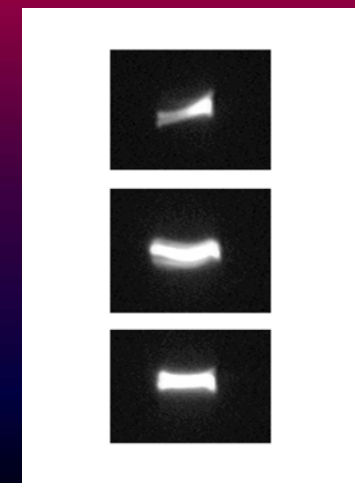
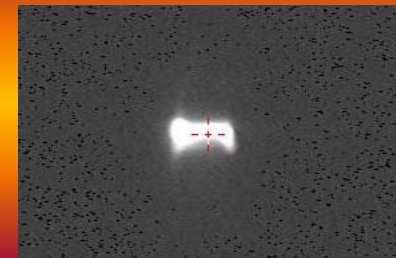
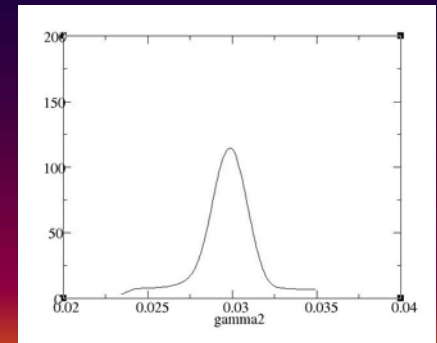
## - optics geometry (theory)

- \* based on absolute position of origin
- \* tabulated corrections

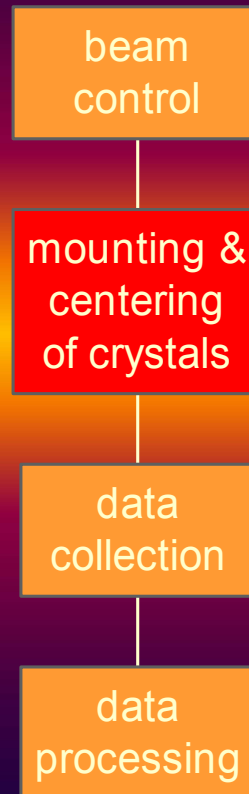
## - beam optimization

- \* scans (before collimator)
- \* beam position on fluo.
- \* scans (after collimator)
- \* beam shape on fluo.

## - energy calibration (optional)



## 2- Automated experiment: sample mounting and centering



beam control

mounting & centering of crystals

data collection

data processing

# The experiment

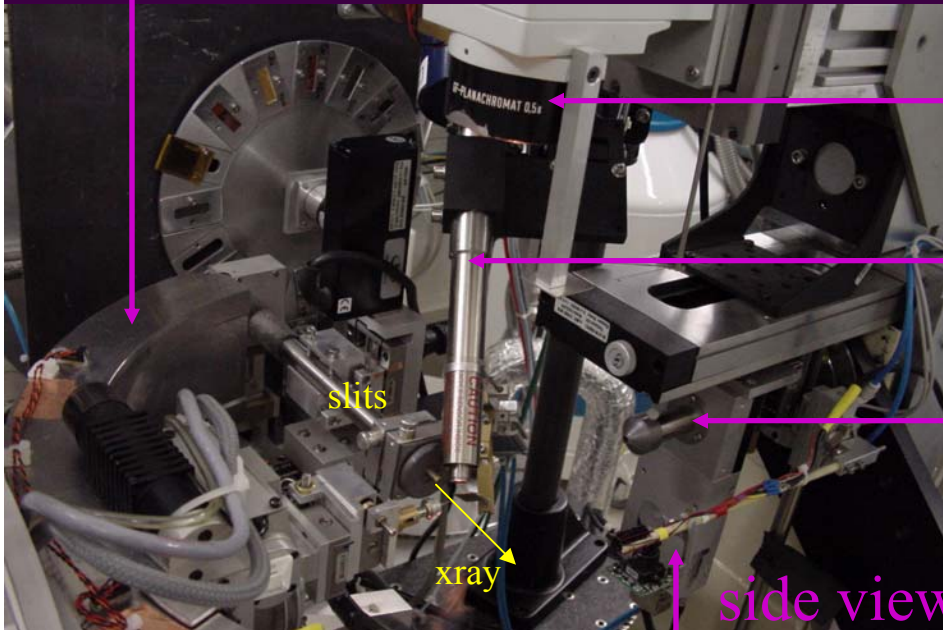
## A 5-circle diffractometer including:

- sample orientation (MAD)

- microscope + video

- Oxford cryo. cooling

- fluorescence detector

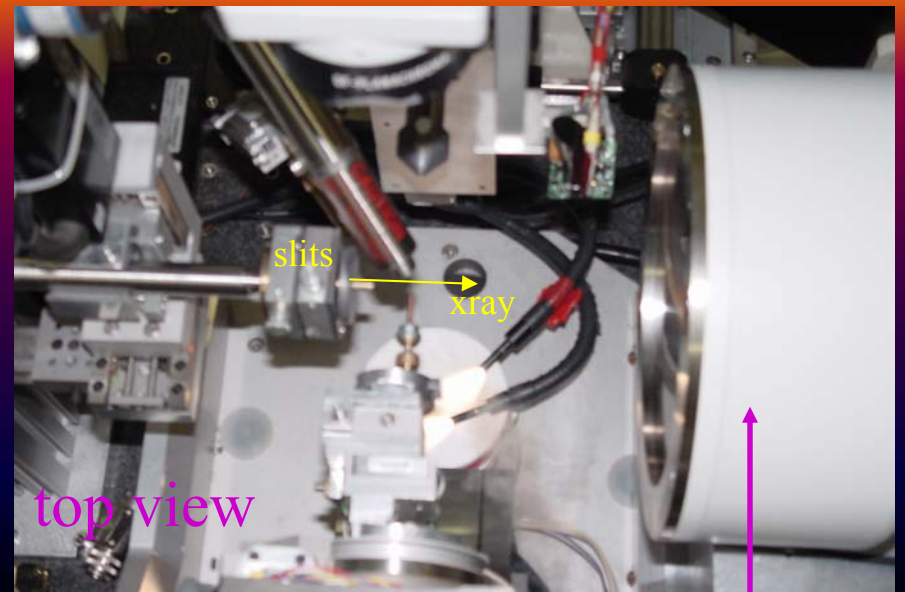


side view

- beam monitors

- 2-theta for high resolution

- Mar CCD detector



top view

beam control

mounting & centering of crystals

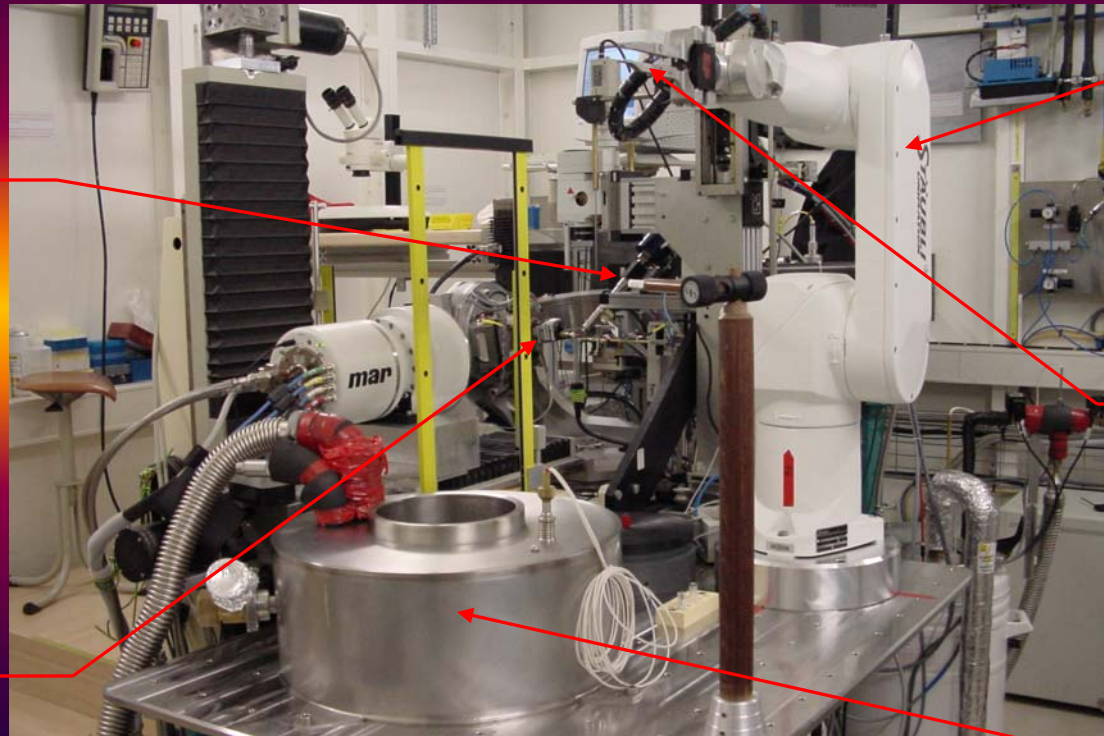
data collection

data processing

# Integrated system

Flipping tong

Goniometer head



Robot

Actuator

Storage Dewar

beam  
control

mounting &  
centering  
of crystals

data  
collection

data  
processing

# Sample unmounting (magnetic cap)



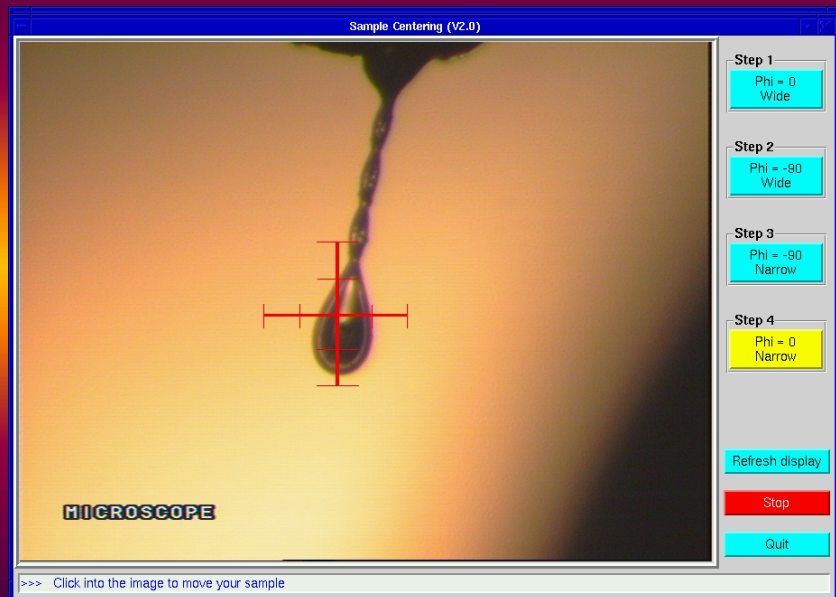
beam control

mounting & centering of crystals

data collection

data processing

# Semi-automated crystal centering



... steps 1, 2, 3 and 4  
in chronological order



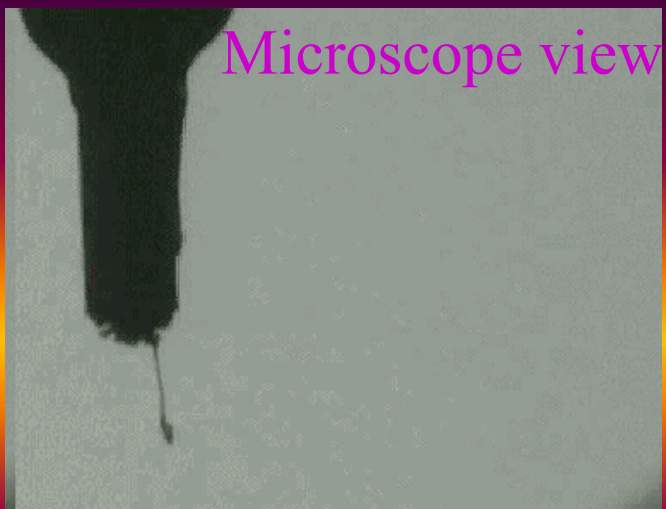
beam control

mounting & centering of crystals

data collection

data processing

# Automated crystal centering



Microscope view



Running action

## Crystal recognition

standard

UV



lysozyme

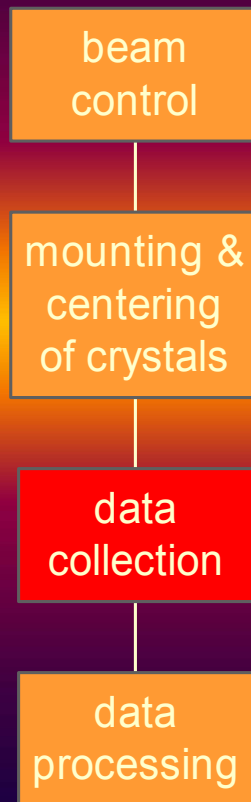


unknown



CCoAOMT

## 3-Automated experiment: data collection





beam control

mounting & centering of crystals

data collection

data processing

# Preparing data collection

Changing detector position

Distance:	563.000000
Angle:	0.000000
Resolution at top:	4.889952
Resolution at bottom:	4.889952

Move detector      Cancel

Collections list

/bioHP\_dsk1/local/control/nemo

- collect.params.mad-invers
- collect.params.mono
- collect.params.old\_style
- collect.params.sav
- collect.params.single
- det\_marip.params
- param1
- param2

Add file >>      IntMax >>

Delete      Remove

Done      Cancel

Xnemo V3.1 - Beamline: d30 ehl - User: control/ferrer

File View Commands Shutter Tools Beam-Stop Help

Set Wavelength      Refresh positions

Instruments Control

Motor Position Control:	Current	Destination	Relative
phi	0.000000	0.0	0.0
alpha	0.000000	0.0	0.0
rtm00	0.000000	0.0	0.0
detfwd	563.000000	0.0	0.0
pmY	0.000000	0.0	0.0
stopY	0.000000	0.0	0.0
stopZ	0.000000	0.0	0.0
colY	0.000000	0.0	0.0
colZ	0.000000	0.0	0.0
mono	6.345150	0.0	0.0

CCD Detector

CCD file format: Save corrected image

Resolution: 2048x2048

Detector status messages

Take background data at start of collection

Distance: 563.000000

Angle: 0.000000

Resolution at top: 4.889952

Resolution at bottom: 4.889952

Change ?

Data Collection Input Parameters

Exposure Mode: MAD      Beam checked: Yes

Collections list:      Automatic processing: Enable

Current collection: unnamed      Intensity optimization: On level and WaveChange

Run number	1	First image number	1
Rotation	phi	Sector width (MAD)	0
Wavelength	12650	Disk name	/bioDEC_dsk3
Sweep Start (deg)	0.0	Directory	test
Sweep End (deg)	120.	Prefix	img_
Sweep step (deg)	1.	Template	img_###.mccd
Total exposure time (sec)	15.	Disk space	7375392 kb (878 images)
Oscillations	1	Total execution time	00:03:00

Start data collection      Restart previous collection      Start program      Suspend data collection      Abort Now

```
1
2
3
4
5
6
7
8
Offsets:
offset alpha1: -0.0015
offset alpha2: 0
offset detfwd: 563.000 (MarCCD detector)
*** fin de chargement de dtclight ***
```

Enter Wavelength

12650  
12655

Delete (selected)

Enter values in A, eV or keV:

12655      Append

Done      Cancel

beam control

mounting & centering of crystals

data collection

data processing

# Xnemo

Xnemo V3.1 - Beamline: d30 ah1 - User: "control/ferrer"

File View Commands Shutter Tools Beam-Stop Help

Set Wavelength Refresh positions

Instruments Control

Motor Position Control:	Current	Destination	Relative
phi	0.000000	0.0	0.0
alpha	0.000000	0.0	0.0
rtm80	0.000000	0.0	0.0
defrwd	563.000000	0.0	0.0
pmY	0.000000	0.0	0.0
stopY	0.000000	0.0	0.0
stopZ	0.000000	0.0	0.0
colY	0.000000	0.0	0.0
colZ	0.000000	0.0	0.0
mono	6.345150	0.0	0.0

CCD Detector

CCD file format: Save corrected image

Resolution: 2048x2048

Detector status messages

Take background data at start of collection

Distance: 563.000000

Angle: 0.000000

Resolution at top: 4.889952

Resolution at bottom: 4.889952

Change ?

Data Collection Input Parameters

Exposure Mode: MAD Beam checked: Yes

Collections list: Automatic processing: Enable

Current collection: unnamed Intensity optimization: On level and WaveChange

Run number: 1 First image number: 1

Rotation: phi Sector width (MAD): 0

Wavelength: 12650 Disk name: /bioDEC\_dsk3

Sweep Start (deg): 0.0 Directory: test

Sweep End (deg): 120.0 Prefix: img\_

Sweep step (deg): 1.0 Template: img\_###.mccd

Total exposure time (sec): 15.0 Disk space: 7375392 kb (878 images)

Oscillations: 1 Total execution time: 00:03:00

Start data collection Restart previous collection Start program Suspend data collection Abort Now

Offsets:  
offset alpha1: -0.0015  
offset alpha2: 0  
offset defrwd: 563.000 (MarCCD detector)  
\*\*\* fin de chargement de dtclight \*\*\*

NetScape: Dataset collected on FIP (BM30A)

File Edit View Go Communicator

Back Forward Reload Home Search NetScape Print Security Str

Bookmarks Location File /home/ferrer/T088/ibccapp\_sel1/wm/ What's Related

Dataset collected on FIP (BM30A)

This page provides a direct access to the last operations performed on this data collection. Some of these facilities are under development.

- Generalities
  - crystal [mapplot](#)
- Data collection
  - data collection [logfile](#)
  - [frames](#)
- Data collection strategy
  - [pre-orientation](#) of the crystal
  - best [xyzxyz](#) for data collection
- Data reduction
  - energy [x01](#), [RENZO](#) and [SCALEPACK](#) log files
- Data processing
  - [transmission](#) of reduced data



NetScape

File Edit View Go Communicator

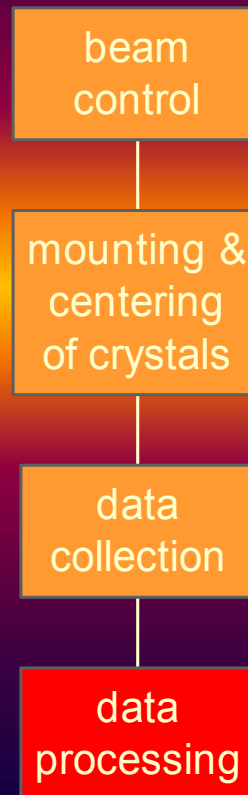
Back Forward Reload Home Search NetScape Print Security Str

Bookmarks Location File /home/ferrer/T088/ibccapp\_sel1/obj/ What's Related

```
# 27 Apr 01 01:08:43 collection mode: mad - sweep_end: phi
# Wavelength: 0.97541 Ang
# chi: 0.00 deg
# alpha: 0.00 deg
# defrwd: 563.00 mm
# angle: 0.00 deg
# Output directory: /home/ferrer/T088/ibccapp_sel1/obj/

#####
# Image name: /logged-time: /gnidra-start: /gnidra-end: /gnidra: /intensity: /m: /temp
# 27 Apr 01 01:09:15: /wavelength: 0.97570 /chi: 0.719168
#####
#1_img_001.mccd 114 0.000 1.000 3480271 644.76 3070705 -131.5
#1_img_002.mccd 148 0.000 2.000 3480284 636.10 3018035 -131.5
#1_img_003.mccd 218 0.000 3.000 3480288 649.10 2818277 -131.5
#1_img_004.mccd 270 0.000 4.001 3480288 642.80 3330264 -131.2
#1_img_005.mccd 326 0.001 5.001 3480288 641.60 3050565 -131.2
#1_img_006.mccd 374 0.001 6.001 3480288 641.10 3346512 -131.9
#1_img_007.mccd 426 0.001 7.001 3480288 639.10 3050565 -131.6
#1_img_008.mccd 460 7.001 8.002 3480285 672.40 3573415 -131.1
#1_img_009.mccd 530 0.002 9.002 3480279 630.10 3005121 -131.0
#1_img_010.mccd 580 9.002 10.002 3480279 635.10 3370122 -131.0
#1_img_011.mccd 630 10.002 11.002 3480279 636.40 3370121 -131.3
#1_img_012.mccd 680 11.002 12.002 3480286 630.10 3087239 -131.6
#1_img_013.mccd 742 12.002 13.002 3480276 637.10 3025101 -131.7
#1_img_014.mccd 781 13.001 14.001 3480276 637.20 3087239 -131.0
#1_img_015.mccd 845 14.001 15.001 3480282 670.10 3025101 -131.6
#1_img_016.mccd 899 15.001 16.002 3480279 635.90 3013341 -131.5
#1_img_017.mccd 953 16.002 17.002 3480280 637.50 3025101 -131.0
#1_img_018.mccd 1004 17.002 18.002 3480280 630.10 3007013 -131.6
#1_img_019.mccd 1061 18.002 19.002 3480277 710.40 3042350 -132.0
#1_img_020.mccd 1117 19.002 20.002 3480280 630.10 3042350 -132.7
#1_img_021.mccd 1180 20.002 21.002 3480280 711.10 3051093 -131.7
#1_img_022.mccd 1219 21.002 22.002 3480280 711.10 3051093 -131.9
#1_img_023.mccd 1270 22.002 23.002 3480280 729.80 3070424 -131.6
#1_img_024.mccd 1323 23.002 24.002 3480280 727.80 3051093 -131.4
#1_img_025.mccd 1373 24.002 25.002 3480281 719.40 3070424 -131.0
#1_img_026.mccd 1425 25.002 26.002 3480281 742.10 3109327 -131.0
#1_img_027.mccd 1476 26.002 27.002 3480284 719.10 3120495 -131.6
#1_img_028.mccd 1527 27.002 28.002 3480281 718.10 3112491 -131.1
#1_img_029.mccd 1579 28.002 29.002 3480286 705.70 3120495 -131.2
#1_img_030.mccd 1630 29.002 30.002 3480280 718.40 3120495 -131.1
#1_img_031.mccd 1681 30.002 31.002 3480282 703.40 3120495 -131.0
#1_img_032.mccd 1734 31.002 32.001 3480281 716.40 3114164 -131.4
#1_img_033.mccd 1787 32.001 33.001 3480281 740.10 3144057 -131.0
#1_img_034.mccd 1838 33.002 34.001 3480282 657.40 3144056 -131.4
#1_img_035.mccd 1889 34.001 35.001 3480280 716.50 3144056 -131.7
#1_img_036.mccd 1940 35.001 36.001 3480280 646.40 3144056 -131.7
#1_img_037.mccd 1991 36.001 37.001 3480280 646.40 3144056 -131.7
#1_img_038.mccd 2042 37.001 38.001 3480285 560.30 4077792 -131.7
#1_img_039.mccd 2093 38.001 39.001 3480285 560.30 4077792 -131.7
#1_img_040.mccd 2144 39.001 40.001 3480285 560.30 4077792 -131.7
#1_img_041.mccd 2195 40.001 41.001 3480285 560.30 4077792 -131.7
#1_img_042.mccd 2246 41.001 42.001 3480285 560.30 4077792 -131.7
#1_img_043.mccd 2297 42.001 43.001 3480285 560.30 4077792 -131.7
#1_img_044.mccd 2348 43.001 44.001 3480285 560.30 4077792 -131.7
#1_img_045.mccd 2399 44.001 45.001 3480285 560.30 4077792 -131.7
#1_img_046.mccd 2450 45.001 46.001 3480285 560.30 4077792 -131.7
#1_img_047.mccd 2501 46.001 47.001 3480285 560.30 4077792 -131.7
#1_img_048.mccd 2552 47.001 48.001 3480285 560.30 4077792 -131.7
#1_img_049.mccd 2603 48.001 49.001 3480285 560.30 4077792 -131.7
#1_img_050.mccd 2654 49.001 50.001 3480285 560.30 4077792 -131.7
```

## 4- Automated experiment: data processing





# adp: steps

peak search (MarSearch)

indexation (denzo)

*mosaicity*

crystal reorientation (oXo)

strategy (strategy/best)

integration (denzo)

*partials*

scaling (scalepack)

extinctions

Patterson map, ... (CCP4)



site search (solve)

molecular replacement (molrep)

*heavy atoms*

SAD phasing (solve)

solvent fl. (resolve)

model building (resolve)

⇒ peak size    ⇒ box size  
 ⇒ laue group, nbr. residues  
 ⇐ *denzo log file (iterations)*

⇒ expected completeness

⇐ *interruptions*

⇒ anomalous signal

⇒ final space group

⇒ max resol. for anomalous

beam control

mounting & centering of crystals

data collection

data processing

# adp: log file

```
NetScape: Automated Data Processing on FIP (EMRBA)
File Edit View Go Communicator Help

Automated Data Processing

Initialization
shp running on ds (help: input file)
graphic mode is on
space group: p41212 (provided by user)
image in: img (provided by user)
reference infocenter in: info01 (provided by user)
resolution: 2.585 A

peaks search
cutoff at 8 sigma
232 peaks (log file)
shp peaks per spot: 92.61
radius of spot = 42915 nm

indexation
indexation done (input file, log file)

evaluate space group
space group according to shp: P4
space group from reference: p41212
let's try with spacegroup p41212 anyway
cell: 50.569 59.569 152.267 90.000 90.000 90.000

evaluate mosaicity
mosaicity increased to 50
mosaicity increased to 80
mosaicity increased to 70
70 deg mosaicity OK

strategy
strategy not available

data integration
initial processing of image 1 (phi_start = -90.000)
processing of image 1 (phi_start = -90.000)
processing of image 2 (phi_start = -88.700)
processing of image 2 (phi_start = -87.400)
processing of image 4 (phi_start = -86.100)
processing of image 5 (phi_start = -84.800)
processing of image 6 (phi_start = -83.500)
processing of image 7 (phi_start = -82.200)
processing of image 8 (phi_start = -80.900)
processing of image 9 (phi_start = -79.600)
processing of image 10 (phi_start = -78.300)
processing of image 11 (phi_start = -77.000)
processing of image 12 (phi_start = -75.700)
processing of image 13 (phi_start = -74.400)
processing of image 14 (phi_start = -73.100)
processing of image 15 (phi_start = -71.800)
processing of image 16 (phi_start = -70.500)
processing of image 17 (phi_start = -69.200)
processing of image 18 (phi_start = -67.900)
processing of image 19 (phi_start = -66.600)
processing of image 20 (phi_start = -65.300)
processing of image 21 (phi_start = -64.000)
processing of image 22 (phi_start = -62.700)

processing of image 56 (phi_start = -5.500)
processing of image 67 (phi_start = -4.200)
processing of image 68 (phi_start = -2.900)
processing of image 69 (phi_start = -1.600)
integration done (data statistics, log file)

scaling
analysis of data collection (log file)
run scaling (previous cycle: 0 rejections)
run scaling (previous cycle: 1532 rejections)
run scaling (previous cycle: 1418 rejections)
scaling coverage (log file)
images = 1 to 217 A
run final scaling
scaling done (log file)

test for extinctions
no systematic extinctions

translation to mtz
211 residues per asym. unit (according to shp)
translation to mtz format done (log file)
translation done (log file)

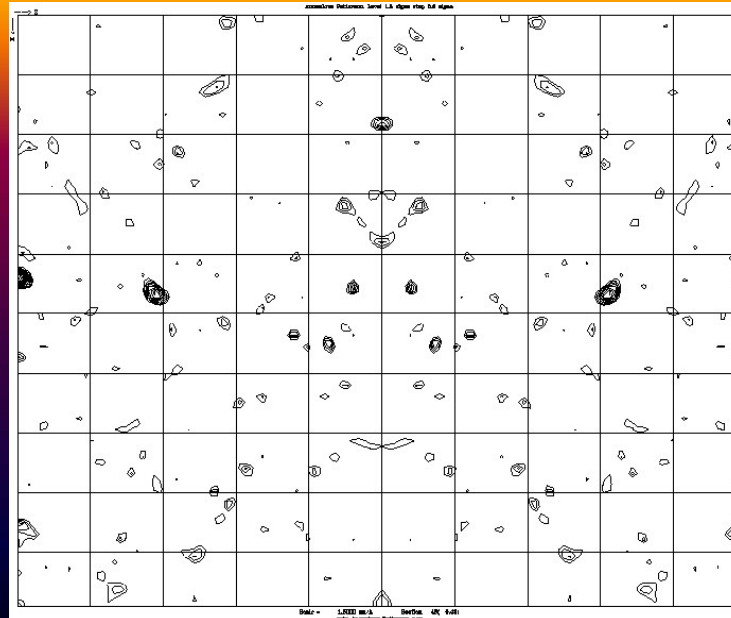
anomalous map
run scaling with merged structure (log file)
run scaling with merging (log file)
comparisons of anomalous signal
I sigma = 2 to 2.06 A
sigma cut done (log file)
max diff out to 37.84
fit done (log file)
app done (log file)
pattern map

processing started: Fri Aug 3 16:01:15 CEST 2001
processing ended: Fri Aug 3 16:16:20 CEST 2001
```

NetScape: Communicator

Volume of the primitive cell 956563:

Lattice	Metric tensor	Best cell (symmetrized)
	distortion index	Best cell (without symmetry restrictions)
primitive cubic	19.944 115.66 109.04 109.82 115.26 121.04 59.51	111.50 111.50 111.50 90.00 90.00 90.00
I centered cubic	6.444 117.17 115.66 147.14 88.66 89.83 80.64	116.65 116.65 116.65 90.00 90.00 90.00
F centered cubic	3.154 165.55 163.71 147.14 88.84 80.86 89.25	159.89 159.90 159.89 90.00 90.00 90.00
primitive rhombohedral	3.096 115.66 109.04 109.82 64.74 53.96 53.51	111.50 111.50 111.50 81.07 81.07 82.07
primitive hexagonal	15.074 109.82 111.07 109.04 87.69 64.74 115.96	110.48 110.48 109.04 90.00 90.00 120.00
primitive tetragonal	19.944 109.82 115.66 109.04 120.49 64.74 121.04	112.74 112.74 109.04 90.00 90.00 90.00
I centered tetragonal	0.726 117.17 115.66 147.14 88.66 89.83 80.64	116.41 116.41 147.14 90.00 90.00 90.00
primitive orthorhombic	15.944 109.04 109.82 115.66 58.36 53.51 64.74	109.04 109.82 115.66 90.00 90.00 90.00
C centered orthorhombic	15.104 109.82 116.41 109.04 110.81 64.74 82.75	109.82 116.41 109.04 90.00 90.00 90.00
I centered orthorhombic	0.654 115.66 117.17 147.14 80.07 88.66 89.26	115.66 117.17 147.14 90.00 90.00 90.00
F centered orthorhombic	0.674 147.14 163.71 165.55 89.75 89.04 88.84	147.14 163.71 165.55 90.00 90.00 90.00
primitive monoclinic	15.234 109.82 109.04 111.07 96.11 116.96 115.26	109.92 109.04 111.07 90.00 116.96 90.00
C centered monoclinic	0.274 184.85 117.17 115.66 90.64 127.28 89.55	184.85 117.17 115.66 90.00 127.28 90.00
primitive triclinic	0.004 109.04 109.82 111.07 116.86 96.11 115.26	
orthorhombic unit cell	117.16 117.16 147.09 90.00 90.00 90.60	
crystal cuts, cutp, cuts	72.933 -141.923 -43.530	
antiferro Idem, Idem	89.49 81.72	





# 3.8 Å dataset processed by user vs adp

I	sigl	Average stat.	Norm. Chi**2	Linear R-fac	Square R-fac	
3356.8	285.7	267.0	0.451	<u>0.125</u>	0.083	← user
3281.8	259.9	238.4	0.552	<u>0.128</u>	0.090	← adp

% of reflections with I / Sigma less than								
0	1	2	3	5	10	20	>20	total
6.9	23.1	34.4	42.5	53.6	69.3	84.1	15.4	<u>99.5</u> ← user
6.1	21.4	32.6	40.5	51.5	68.0	83.4	16.2	<u>99.6</u> ← adp

Comparison of a typical dataset, processed by a crystallographer using *HKL* and with *adp*. Each data processing is illustrated with statistics calculated by *scalepack* (upper part: I,  $K^2$  and R-factors; lower part:  $I/\sigma(I)$  and completeness).





# First structure solved in automated mode

Structure solved in  
3h data collection  
2h phasing/building

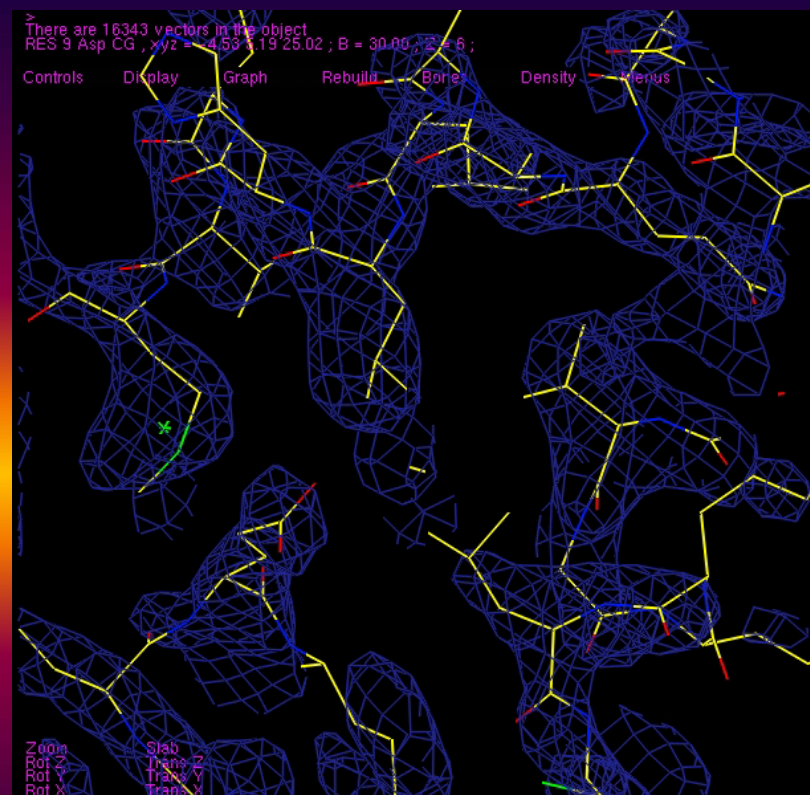
## Data collection statistics:

p212121, 1 mol./asym. unit  
completeness at 2.1 Å : 98.7%  
<Rsym> : 4.2%  
Rsym (last shell) : 12.5%

## Phasing statistics (solve/resolve):

fig. of merit: 0.25  
180/233 residues built autom.

80% of the model in 5 hours



Farnesoid X Receptor (FXR),  
M. Downes et al.,  
Molecular Cell 11 (2003), 1079

crystallization  
tray

crystallization  
analysis

# Analysis of crystallization drops

## Aim:

automated analysis of Greiner™ box  
discrimination salt/protein  
precipitate analysis ?

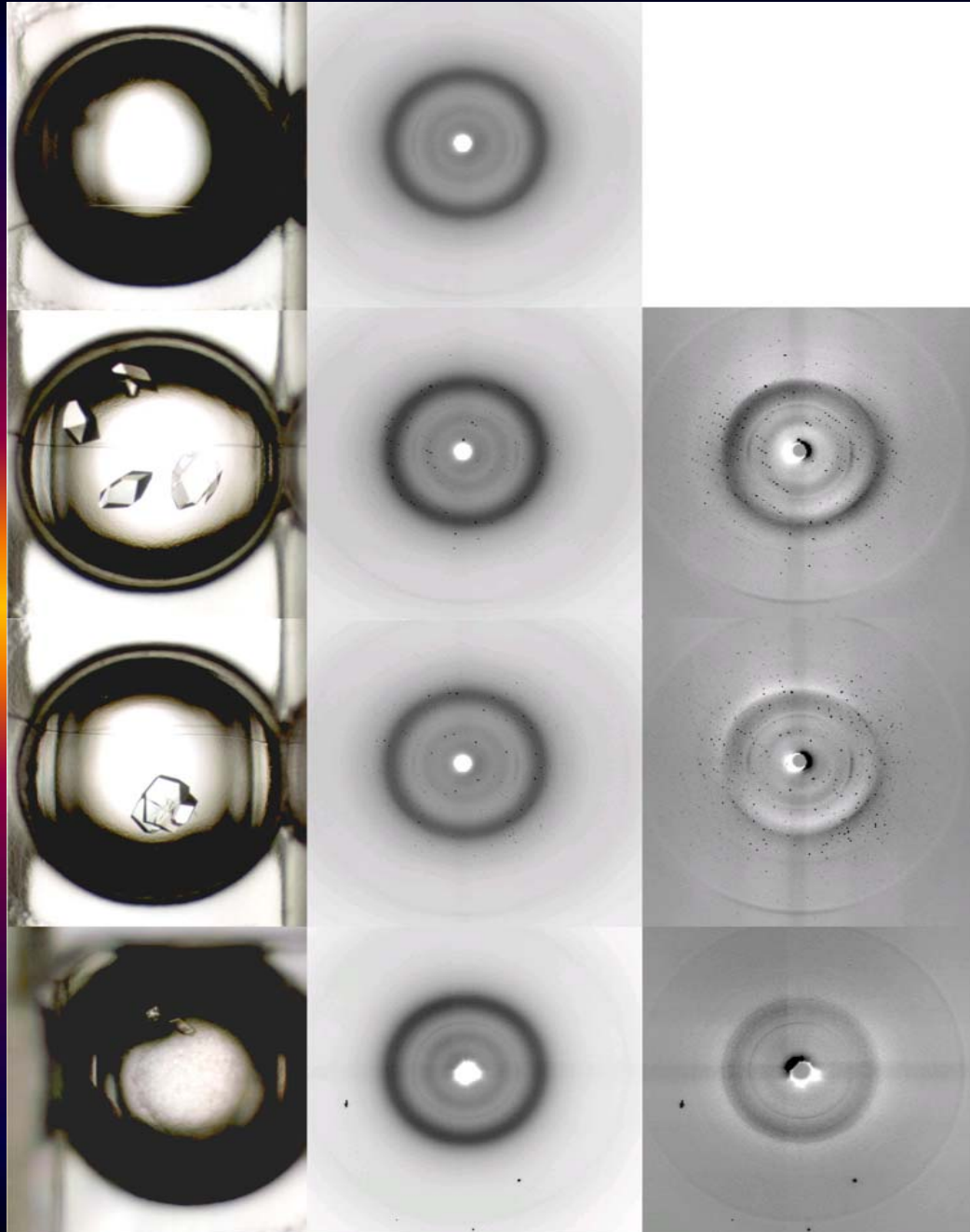
## Means:

beam: 2x2 mm, 0.8 Å  
oscillations: 1 deg  
10 to 30 sec / drop (1 to 2 h per box)  
image processing





# Protein vs. Salt



Protein crystal

Salt crystal

Precipitate analysis !

# Data collection *in situ*

<i>Plate-drop</i>	L-B3b	T-C12a	T-B4b	C-B1c	Y-A11a
<i>sample</i>	lysozyme	lysozyme	thaumatin	chalcone s.	kinase
<i>Space group</i>	p43212	p43212	p41212	p3121	c2
<i>Resolution (Å)</i>	1,8	1,8	2,2	3,0	2,2
<i>Completeness (%)</i>	89,8	85,1	95,1	70,0	48,2
<i>Redundancy</i>	2,6	2,8	2,4	2,6	1,3
<i>I/σ</i>	11,0	9,1	6,1	4,2	6,1
<i>Rsym (%)</i>	6,2	8,4	11,5	23,9	7,1
<i>Rfree (%)</i>	26,9	26,7	27,2	24,2	28,9

L. Jacquamet et al., *accepted in Structure*

J.-L. Ferrer	(beamline responsible)
L. Jacquamet	(crystallographer)
J. Joly	(software/hardware)
P. Charrault	(electronic)
M. Pirocchi	(vacuum )
J. Ohana	(robotics)
R. Kahn	(adviser)
F. Borel	(local contact)
S. Fieulaine	(local contact)
L. Serre	(local contact)
J. Dupuis	(local contact)
P. Israel-Gouy	(administration)
A. Bertoni	(machining)
M. Roth (retired)	(adviser)
P. Carpentier (IBS/BM16)	(instrumentation)
E. Fanchon (IBS/LCM)	(software)