

Optimising the Diamond experience from a user's perspective

Dave Lawson

Department of Biological Chemistry

John Innes Centre, Norwich, UK



East of England Macromolecular X-ray Crystallography BAG



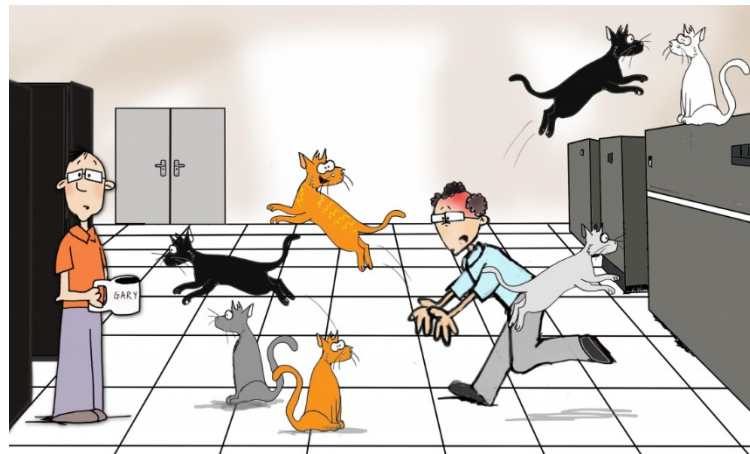
Norwich



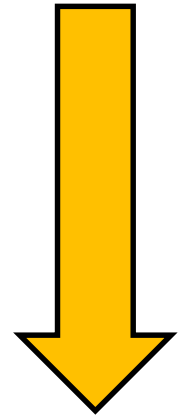
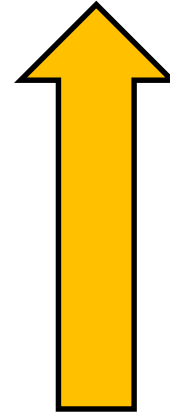
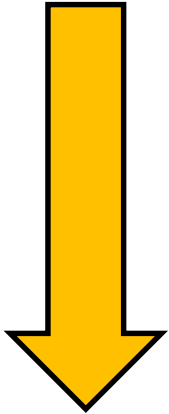
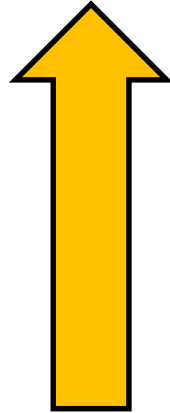
Colchester



8 groups: ~30 potential users



Data collection philosophy...



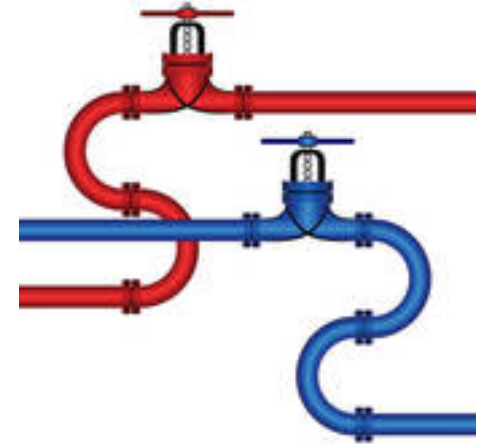
Strategy



Tools



Pipelines



Remote



- Maximizing efficiency
- Minimizing time commitment
- Managing your data
- How to ~~collect~~ data

“Routine” data collection:

- on non-hazardous samples
- at cryogenic temperatures
- on pre-cooled crystals in pucks
- at conventional wavelengths

Disclaimer: the thoughts, views, and opinions expressed here are entirely my own and do not reflect the opinions of my employer, my wife, my kids, my...

Primary goals (data quantity **AND** quality...)

- To collect as much data as possible...
- To collect the best possible data...
- To collect the data that will enable me to:
 - Solve my structure
 - Extend the resolution of my structure
 - Show that my ligand is bound

Secondary goal (get a life!)

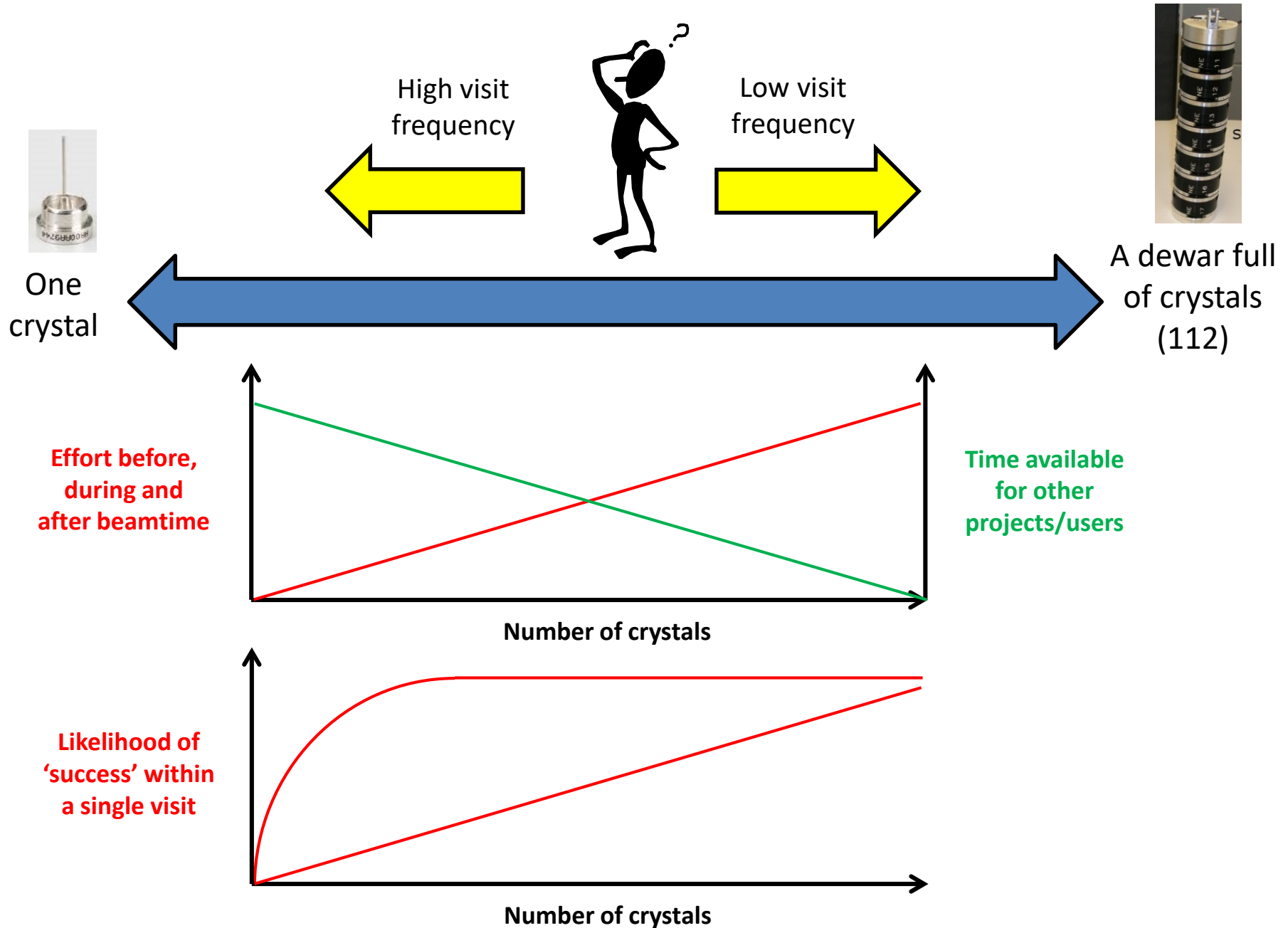
Primary goals (data quantity AND quality...)

- To collect as much data as possible...
- To collect the best possible data...
- To collect the data that will enable me to:
 - Solve my structure
 - Extend the resolution of my structure
 - Show that my ligand is bound

Secondary goal (get a life!)

- Try to answer these questions during the visit:
 - Can I solve my structure?
 - Can I extend the resolution of my structure?
 - Is my ligand bound?
- Make best use of time and minimise the amount of follow-up work...

How many samples to take/send per project...



Data collection is always a compromise...

*How much
data do I
need?*

*What
resolution
do I need?*

*What are
the data
for?*

*How much
time do I
have?*

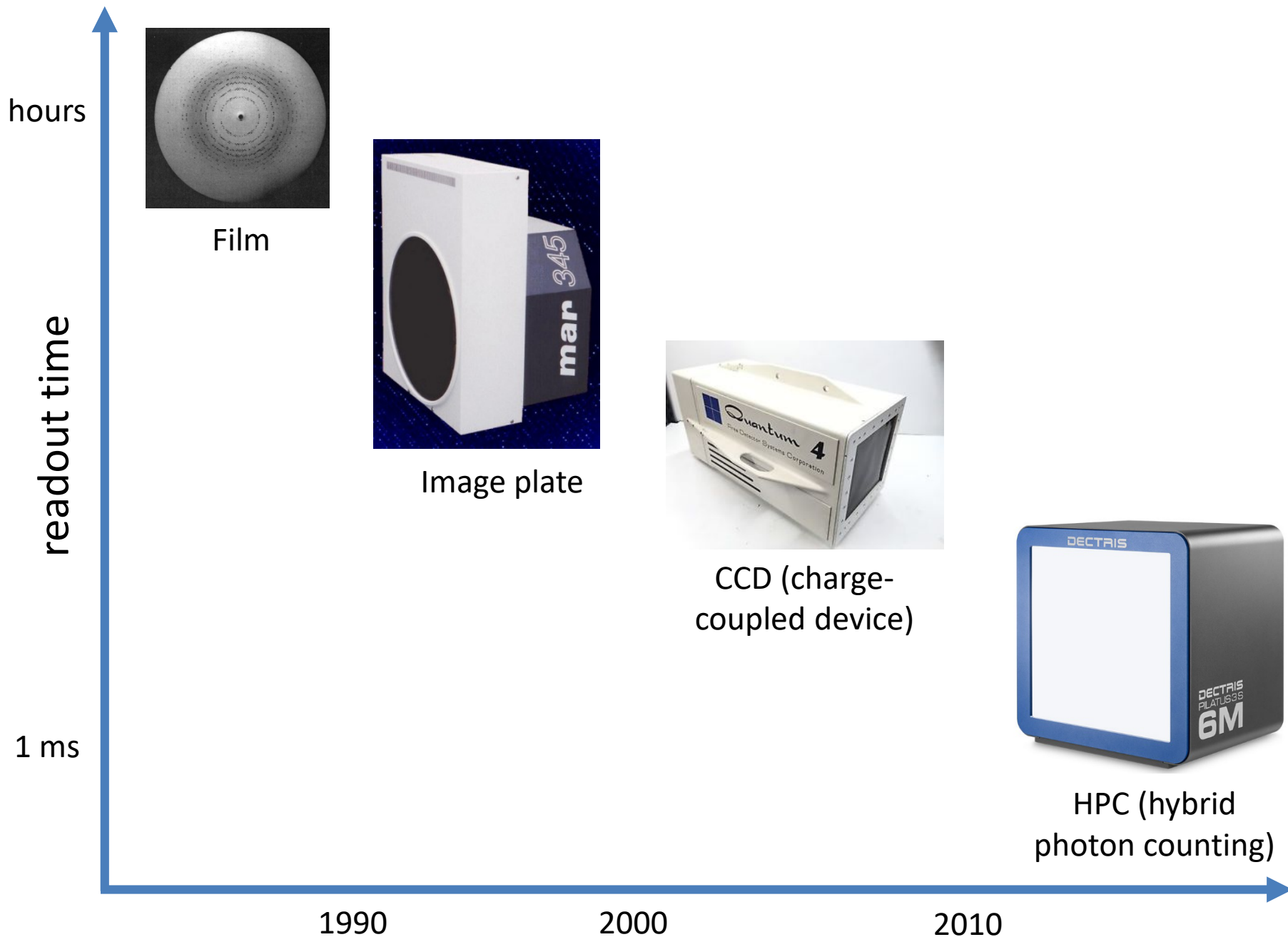


*How many
crystals do
I have?*



Be prepared!

Do as much of this thinking **before** your beamtime as possible!

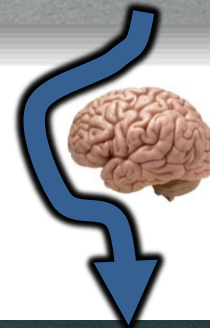


MX data collection has become faster...

20th century data collection



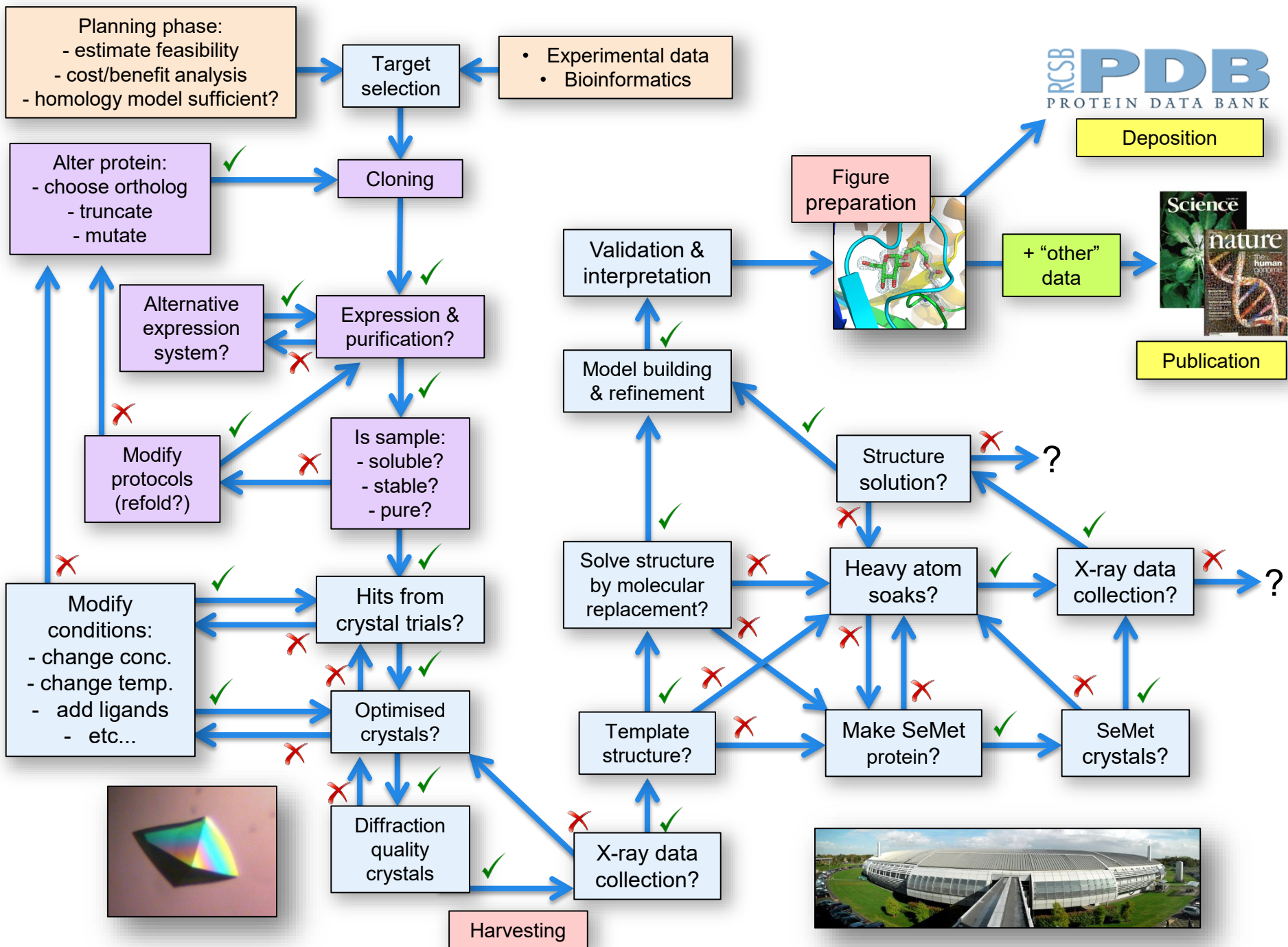
21st century data collection

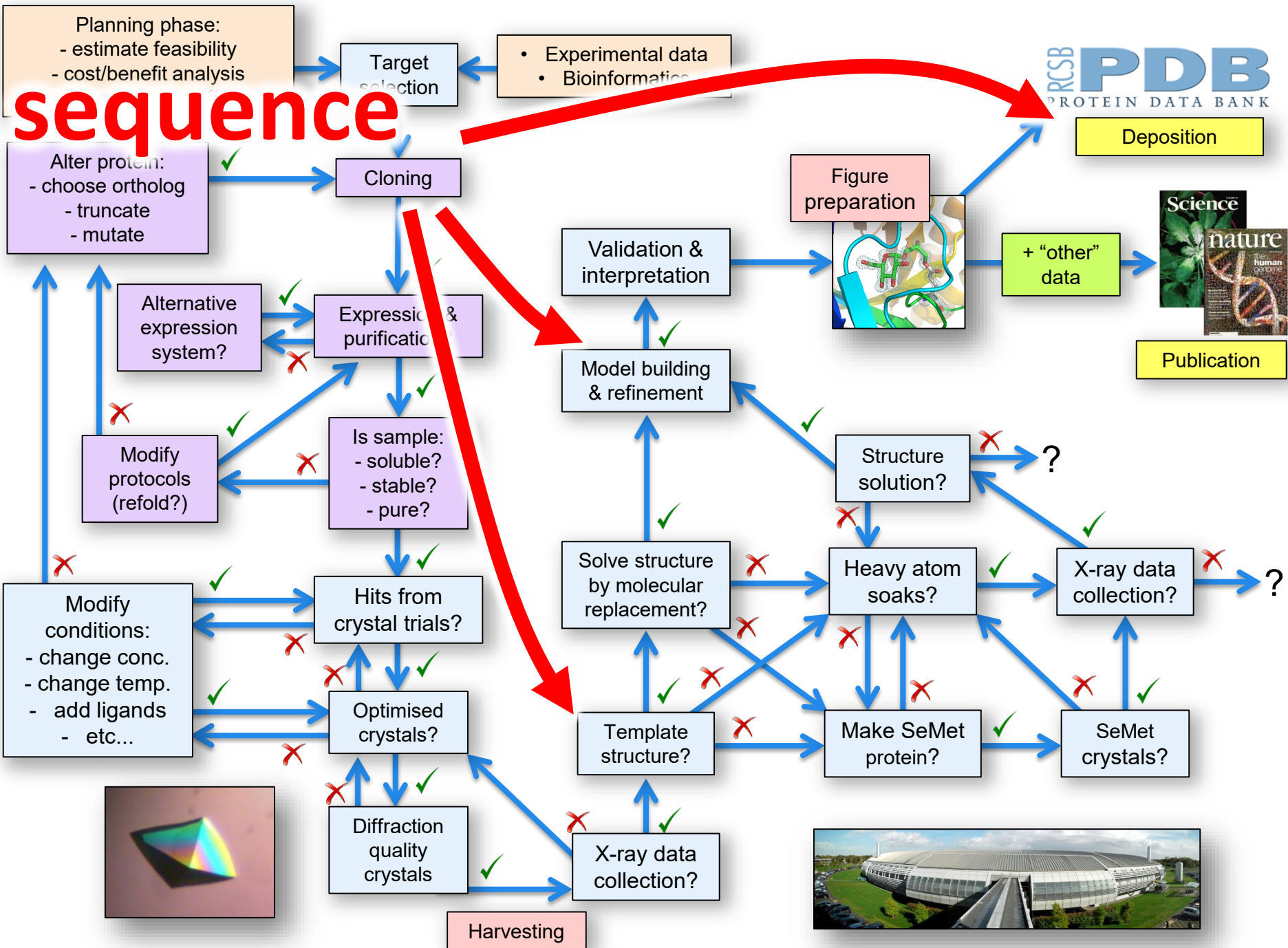


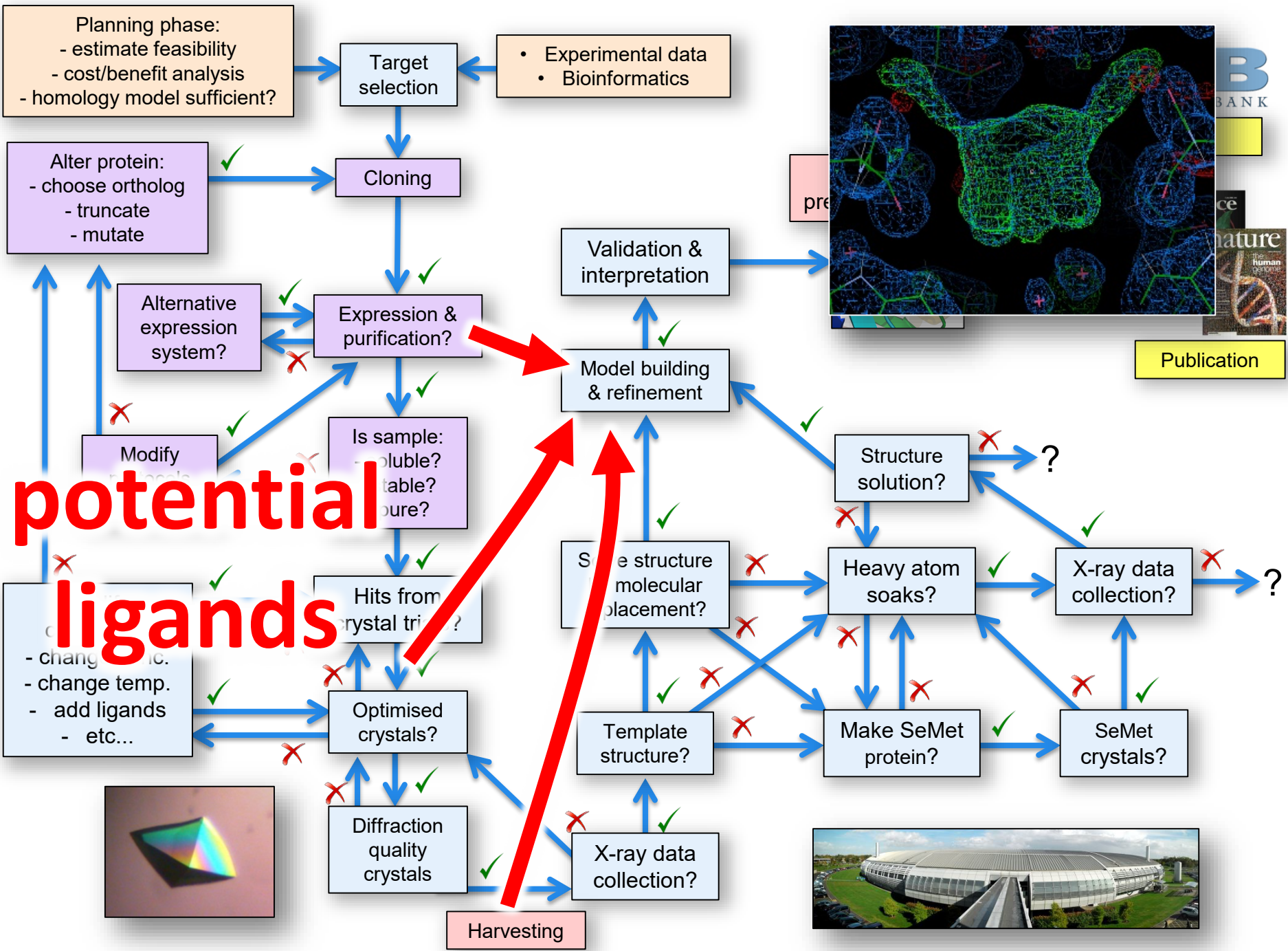
MX data collection has become faster...

- Typical dataset takes <3 min to collect
 - 10 - 20 GB raw data per data set!
 - >1.0 TB data possible in 24 hrs!
- In practice, 4 “useful” data sets per hour is good
 - includes screening for “best” crystal
 - ...and thinking!
- But - beamtime is still in high demand
 - users are generating crystals more rapidly
 - can get away with “marginal” samples
- Each session is often split between several groups
 - therefore need to be efficient and organised...

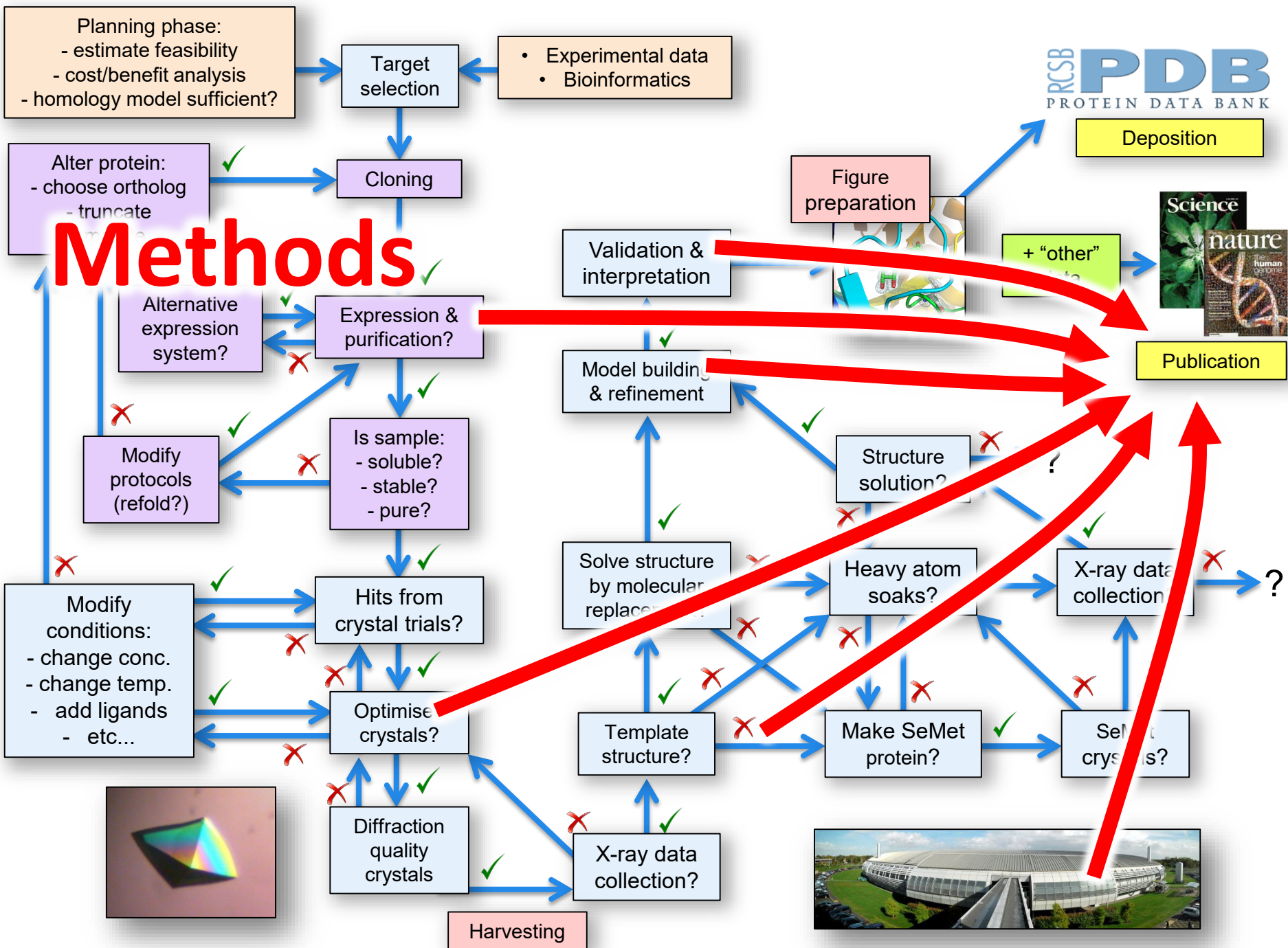








Methods



Taking back control...



ISPyB

+



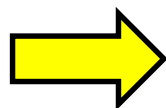
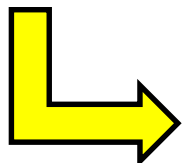
CCP4i2

F•R•I•E•N•D•S



ISPyB database

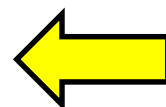
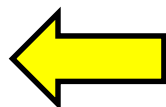
Sample info.



(test images)

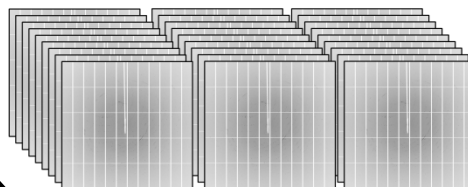


Strategy:
EDNA
Xia2

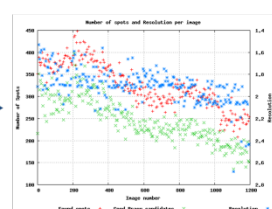


Time / s
0
60
120
180
240

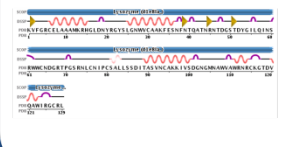
Data Collection



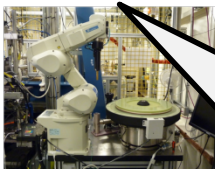
Per-image Analysis



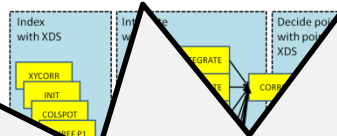
Sequence Search
(MR Bump)



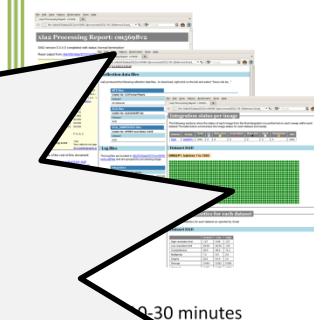
Change Sample



Fast Data Processing: Fast_X



Data Processing
xia2 – 2d and 3d

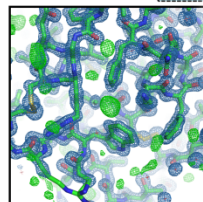
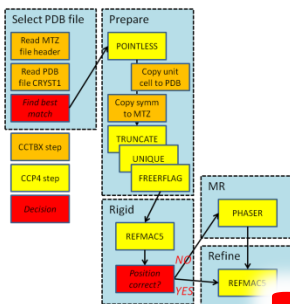


~30 minutes

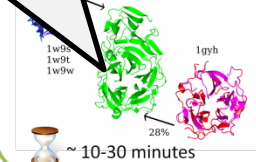
Archiving
icat.diamond



Diff Map: DIMPLE

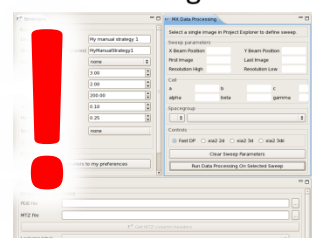


Molecular Replacement
MR Bump



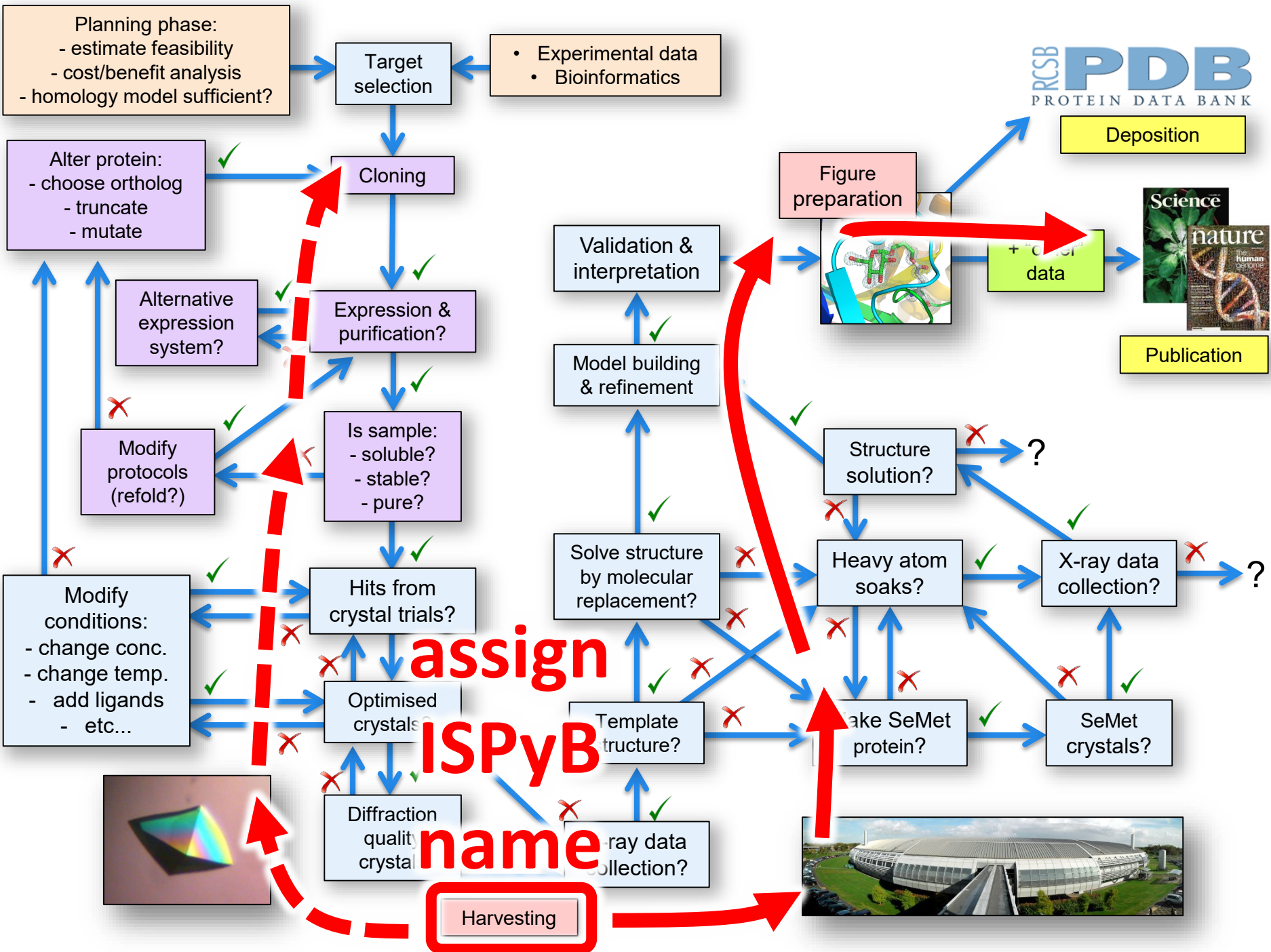
~10-30 minutes

Rerunning Jobs



Each crystal
has a unique
"ISPyB name"!

Use it!



Date: 29-JUL-2017

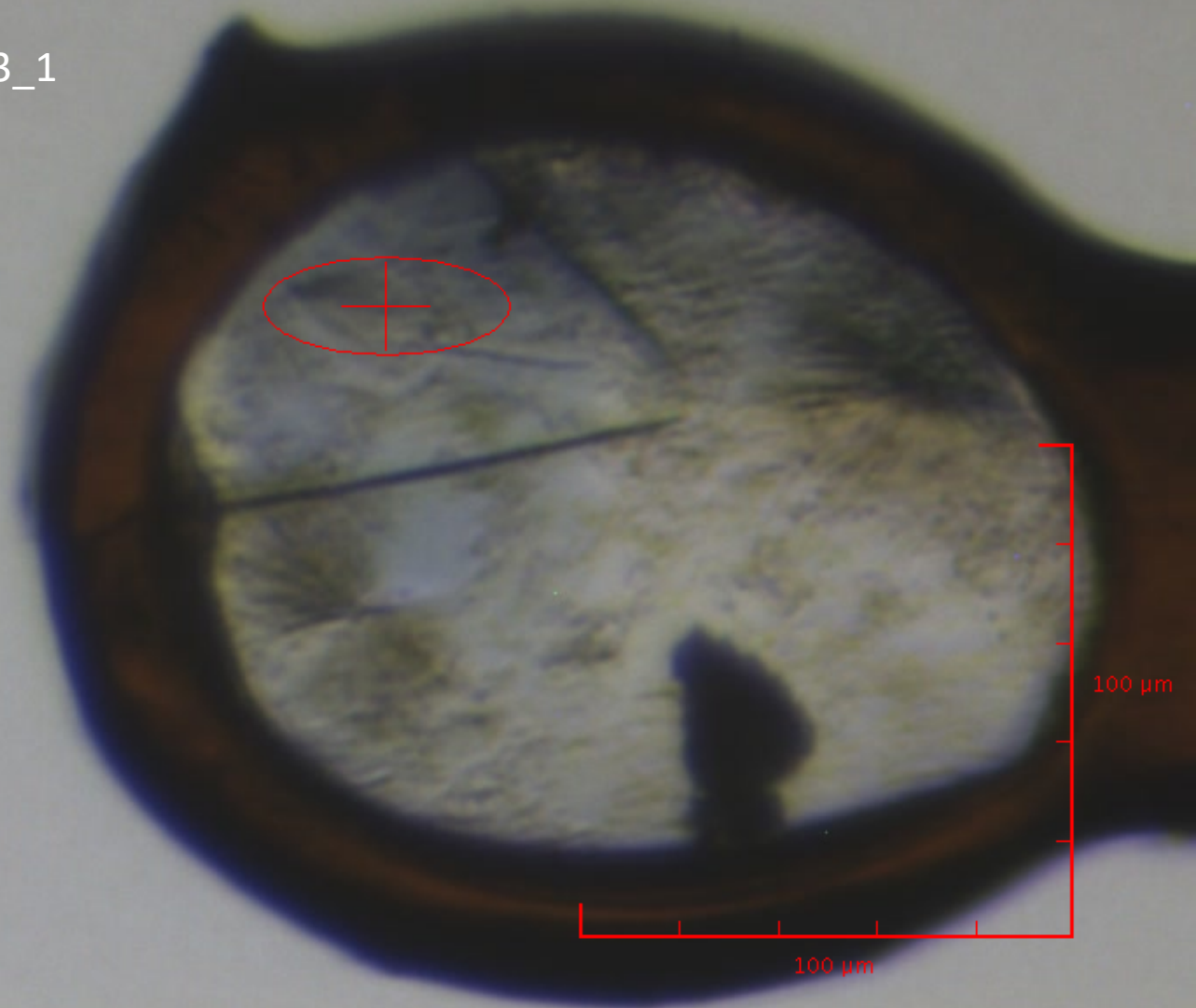
Beamline: i03 Diamond

Visit ID: mx13467-41

Protein acronym: NmADH9

ISPyB name: NmADH9_23

Dataset name: NmADH9_23_1



Beam size: 50.0 × 20.0 μm

Harvest your crystals and enter sample info into ISPyB

Container: DLS-442

This page shows the contents of the selected container. Samples can be added and edited by clicking the pencil icon, and removed by clicking the x

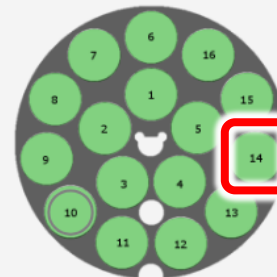
This container is currently assigned and in use on a beamline sample changer. Unassign it to make it editable

Shipment [JIC 260717 i03](#)
Dewar DLS-MX-0002
Container Type Puck
Registered Container DLS-442 [\[View\]](#)
Barcode [Click to edit](#)
Automated Collection [+ Queue](#) this container for Auto Collect
Comments [Click to edit](#)

Location History

Date	Status	Location	Beamline
08-09-2017 10:49	at facility		
04-08-2017 15:58	at DLS		
29-07-2017 11:26	processing	7	i03
21-07-2017 12:04	at DLS		

10 Page << < 1 > >>



this is **UNIQUE**

Location	Protein Acronym	Abundance	Components	Name	Spacegroup	Barcode	Comment	Status
9	NmADH9			NmADH9_18				Q
10	NmADH9			NmADH9_19				Q
11	NmADH9			NmADH9_20				Q
12	NmADH9			NmADH9_21				Q
13	NmADH9			NmADH9_22				Q
14	NmADH9			NmADH9_23				Q
15	NmADH9			NmADH9_24				Q
16	NmADH9			NmADH9_25				Q

Puck number – DLS442

Position	Person	Protein name	Isyb name	Plate	Well	Conditions	Soak/cryo	Space group	Test	Resolution	Comments
9	Benjy	NmADH9	NmADH9_18	MCBL0004	A11.1	"	Cis cis nepetalactol+ 20%EG				10mM soak for approx 1 h
10	Benjy	NmADH9	NmADH9_19	MCBL0004	A11.1	"	8 oxogeralial+ 20%EG				10mM soak for approx 1 h
11	Benjy	NmADH9	NmADH9_20	MCBL0004	A11.1	"	8 oxogeralial+ 20%EG				10mM soak for approx 1 h
12	Benjy	NmADH9	NmADH9_21	HD01	B6	PEG 4k 29%, 0.1M Mes pH 6.5, 1 mM NAD	8-oxocitronellal+ 20%EG				5mM soak for approx 1.5 h
13	Benjy	NmADH9	NmADH9_22	HD01	B4	"	8-oxocitronellal+ 20%EG				5mM soak for approx 1.5 h
14	Benjy	NmADH9	NmADH9_23	HD01	B3	"	Cis cis nepetalactone+ 20%EG				5mM soak for approx 1.5 h
15	Benjy	NmADH9	NmADH9_24	HD01	B2	"	Cis cis nepetalactol+ 20%EG				5mM soak for approx 1.5 h
							8-oxogeralial+ 20%EG				5mM soak for approx 1 h

MX13467-41 7 hrs I03 from 12:00 Sat 29 Jul 2017 to 19:00 Sat 29 Jul 2017

First Local Contact: Dr Katherine McAuley (01235 778405)

Beamline I03: 01235 778707

Experimental Hall Coordinator: 01235 778787

Institution	JIC	Group(s)	Lawson
Contact name(s) and number(s)	Dave Lawson: 01603 123456 / 07778 123456		
Approx. timings	12:00 – 15:00	On-site/remote?	Remote
Dewar ID	Puck ID (# samples)	Users and comments	
DLS-MX-0002	DLS0086 (16 samples)	Sue	
	DLS0087 (12 samples)	JC	
	DLS0090 (12 samples)	Sue	
	DLS442 (16 samples)	Benjy	
Anomalous elements:	Zn		
Other requirements:	MCA scans		
Comments/ questions:			

annotate during data
collection – helps
decision making

Decide on a running order for
users and samples – share your
plans with LC and EHC

Institution	UEA	Group(s)	Hemmings
Contact name(s) and number(s)	Andrew Hemmings: 01603 123456/07799 123456		
Approx. timings	15:00 – 19:00	On-site/remote?	Remote
Dewar ID	Puck ID (# samples)	Users and comments	

Data collection setup in GDA

Select required sample from drop-down menu:

- no need to enter sample information or specify sample location
- less likely to get the wrong sample

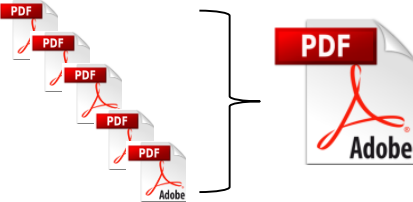
Put your data into your own directory....

...especially important if there are multiple users from several institutions – simplifies backing up too....

use the “screening” tab for test images – keeps them separate from datasets

Gives: `/dls/{beamline}/data/{year}/{visit-ID}/{your-name}/{proteinacronym}/{samplename}/{samplename}_{run_no.}_{image_no.}.cbf`

e.g.: `/dls/i03/data/2017/mx13467-41/JIC/NmADH9/NmADH9_23/NmADH9_23_1_3600.cbf`



simple to search this for an ISPyB name

Visit List

This page lists the

Start

19:00 15-09-2017

02:00 11-09-2017

10:00 05-08-2017

12:00 29-07-2017

Sample	Images	Res	λ	Q Osc	Spacegroup	Unit Cell	Processed Resolution	Rmerge	Completeness	Comments
NmADH9_23	3600	1.5	0.9763	0.10	P 1 2 1	63.84, 107.75, 90.00, 104.25, 90.00	29.77 - 1.5, 29.77 - 6.7, 1.54 - 1.5	0.085, 98.8, 0.036, 98.9, 0.024, 96.0		(-262, -192, 1150) EDNAStrategy1: subWedge:1Aperture: Medium
NmADH9_24	3	1.5	0.9763	0.50						(-107, -190, 1012) Aperture: Medium
NmADH9_24	3	1.5	0.9763	0.50						(-156, -228, 1396) Aperture: Medium
NmADH9_25	3	1.5	0.9763	0.50						(-213, -228, 1406) Aperture: Medium
NmADH9_15	3	1.5	0.9763	0.50						(-247, -388, 861) Aperture: Medium
NmADH9_15	42	3.0	0.9763	0.00						Diffraction grid scan of 7 by 6 images, Top left [350,177], Bottom right [746,517]
NmADH9_15	12	3.0	0.9763	0.00						Diffraction grid scan of 3 by 4 images, Top left [472,245], Bottom right [642,472]
NmADH9_15	3	1.5	0.9763	0.50						(276, -373, 844) Aperture: Medium
NmADH9_15	3	2.0	0.9763	0.50						(276, -373, 844) Aperture: Medium
NmADH9_15	3600	2.0	0.9763	0.10	P 1 2 1	62.17, 107.70, 71.76, 90.00, 105.05, 90.00	30.02 - 3.04, 29.02 - 13.6, 3.12 - 3.04	0.167, 99.4, 0.040, 99.8, 0.033, 93.3		(276, -373, 844) EDNAStrategy1: subWedge:1Aperture: Medium
NmADH9_11	3	2.0	0.9763	0.50						(-477, -294, 1143) Aperture: Medium
NmADH9_12	3	2.0	0.9763	0.50						(-696, -216, 1130) Aperture: Medium
NmADH9_12	3	1.6	0.9763	0.50						(-696, -216, 1130) Aperture: Medium
NmADH9_12	3600	1.6	0.9763	0.10	P 1 2 1	63.84, 108.11, 69.29, 107.16, 90.00, 104.25, 120.00	29.74 - 1.69, 29.74 - 7.55, 1.97 - 1.92	0.085, 99.8, 0.036, 96.6, 0.053, 97.1		(-696, -216, 1130) EDNAStrategy1: subWedge:1Aperture: Medium
TAPHY_27_1	4	3.0	0.9763	0.00						Diffraction grid scan of 1 by 4 images, Top left [540,265], Bottom right [592,435]
TAPHY_27_1	3	1.8	0.9763	0.50						(-109, -296, 1303) EDNAStrategy1: subWedge:1Aperture: Medium
TAPHY_27_1	1200	1.5	0.9763	0.10	H 3					(-109, -296, 1303) EDNAStrategy1: subWedge:1Aperture: Medium
TAPHY_27_2	5	3.0	0.9763	0.00						Diffraction grid scan of 1 by 5 images, Top left [534,251], Bottom right [576,463]
TAPHY_27_2	3	1.5	0.9763	0.50						(136, 104, 1462) EDNAStrategy1: subWedge:1Aperture: Medium

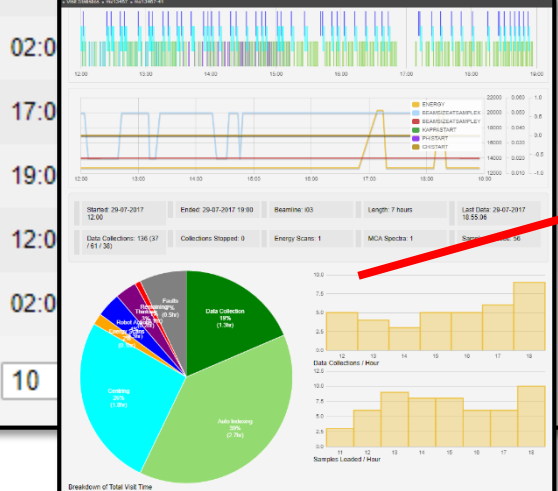
Page 5 of 12

29.07.2017_03_rml13647.41 - Excel

File Home Insert Page Layout Formulas Data Review View ACROBAT Tell me what you want to do																														
Cut		Copy		Format Painter		Clipboard		Font		Alignment		Number		Conditional Formatting		Table		Check Cell		Explanatory...		Input		Neutral Cell		Ca				
Calibri		11		A		Wrap Text		General																						
B		F		I		Merge & Center																								
NmADH9_23																														
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U										
40	NmADH9_103	1	1	0.97625	NmADH9_NAD-depr	3	0.97625	289.0478	0.04	0	0.5	211.6297	206.4712	1.6	(272,-419,809)	Aperture: Medium														
41	NmADH9_103	1	1	0.97625	NmADH9_NAD-depr	3	0.97625	289.0478	0.04	0	0.5	211.6297	206.4712	1.6	(-231,-681,1252)	Aperture: Medium														
42	NmADH9_103	1	1	0.97625	NmADH9_NAD-depr	3	0.97625	289.0478	0.04	0	0.5	211.6297	206.4712	1.6	(-249,-175,1442)	Aperture: Medium														
43	NmADH9_103	1	1	0.97625	NmADH9_NAD-depr	3600	0.97625	264.5335	0.01	122	0.1	211.6215	206.6305	1.5	(-249,-175,1442)	EDNAStrategy1: subWedge:1Aperture: Medium														
44	NmADH9_103	1	1	0.97625	NmADH9_NAD-depr	3	0.97625	289.0478	0.04	0	0.5	211.6297	206.4712	1.6	(-533,7,1513)	Aperture: Medium														
45	sample_103	8	1	0.97625	NmADH9_NAD-depr	120	0.97625	602.2545	0.1	-0.00016	0	211.7341	204.4354	2.9966	Diffraction grid scan of 12 by 10 images, Top left [358,106], Bottom right [746,517]															
46	sample_103	9	1	0.97625	NmADH9_NAD-depr	12	0.97625	602.2545	0.1	89.99975	0	211.7341	204.4354	2.9967	Diffraction grid scan of 3 by 4 images, Top left [475,268], Bottom right [642,472]															
47	NmADH9_103	2	1	0.97625	NmADH9_NAD-depr	3	0.97625	427.9818	0.04	0	0.5	211.676	205.5681	2.2	(-488,25,1516)	Aperture: Medium														
48	NmADH9_103	1	1	0.97625	NmADH9_NAD-depr	3	0.97625	427.9818	0.04	0	0.5	211.676	205.5681	2.2	(-496,457,1456)	Aperture: Small														
49	NmADH9_103	1	1	0.97625	NmADH9_NAD-depr	3	0.97625	427.9818	0.04	0	0.5	211.676	205.5681	2.2	(-510,1069)	Aperture: Small														
50	sample_103	10	1	0.97625	NmADH9_NAD-depr	35	0.97625	602.2545	0.1	89.99975	0	211.7341	204.4354	2.9967	Diffraction grid scan of 7 by 5 images, Top left [327,239], Bottom right [642,472]															
51	sample_103	11	1	0.97625	NmADH9_NAD-depr	12	0.97625	602.2545	0.1	179.9997	0	211.7341	204.4354	2.9966	Diffraction grid scan of 3 by 4 images, Top left [469,262], Bottom right [642,472]															
52	NmADH9_103	2	1	0.97625	NmADH9_NAD-depr	3	0.97625	427.9818	0.04	0	0.5	211.676	205.5681	2.2	(-33,96,1067)	Aperture: Small														
53	NmADH9_103	1	1	0.97625	NmADH9_NAD-depr	3600	0.97625	359.8787	0.01	106	0.1	211.6533	206.0108	1.9	(-33,96,1067)	EDNAStrategy1: subWedge:1Aperture: Small														
54	NmADH9_103	1	1	0.97625	NmADH9_NAD-depr	3	0.97625	427.9818	0.04	0	0.5	211.676	205.5681	2.2	(-241,-208,1371)	Aperture: Medium														
55	NmADH9_103	2	1	0.97625	NmADH9_NAD-depr	3	0.97625	427.9818	0.04	0	0.5	211.676	205.5681	2.2	(-262,-192,1150)	Aperture: Medium														
56	NmADH9_103	3	1	0.97625	NmADH9_NAD-depr	3	0.97625	264.5335	0.04	0	0.5	211.6215	206.6305	1.5	(-262,-192,1150)	Aperture: Medium														
57	NmADH9_103	1	1	0.97625	NmADH9_NAD-depr	3600	0.97625	213.4382	0.01	45	0.1	211.6045	206.9627	1.3	(-262,-192,1150)	EDNAStrategy1: subWedge:1Aperture: Medium														
58	NmADH9_103	1	1	0.97625	NmADH9_NAD-depr	3	0.97625	264.5335	0.04	0	0.5	211.6215	206.6305	1.5	(-107,-190,1012)	Aperture: Medium														
59	NmADH9_103	2	1	0.97625	NmADH9_NAD-depr	3	0.97625	264.5335	0.04	0	0.5	211.6215	206.6305	1.5	(-107,-190,1012)	Aperture: Small														
60	NmADH9_103	1	1	0.97625	NmADH9_NAD-depr	3	0.97625	264.5335	0.04	0	0.5	211.6215	206.6305	1.5	(-156,-228,1396)	Aperture: Medium														
61	NmADH9_103	2	1	0.97625	NmADH9_NAD-depr	3	0.97625	264.5335	0.04	0	0.5	211.6215	206.6305	1.5	(-215,-228,1406)	Aperture: Medium														
62	NmADH9_103	2	1	0.97625	NmADH9_NAD-depr	3	0.97625	264.5335	0.04	0	0.5	211.6215	206.6305	1.5	(-247,-388,861)	Aperture: Medium														
63	sample_103	12	1	0.97625	NmADH9_NAD-depr	42	0.97625	602.2595	0.1	-0.00016	0	211.7341	204.4353	2.9966	Diffraction grid scan of 7 by 6 images, Top left [350,177], Bottom right [746,517]															
64	sample_103	13	1	0.97625	NmADH9_NAD-depr	12	0.97625	602.2595	0.1	89.99975	0	211.7341	204.4353	2.9966	Diffraction grid scan of 3 by 4 images, Top left [472,245], Bottom right [642,472]															
65	NmADH9_103	3	1	0.97625	NmADH9_NAD-depr	3	0.97625	264.5335	0.04	0	0.5	211.6215	206.6305	1.5	(276,-373,844)	Aperture: Medium														
66	NmADH9_103	4	1	0.97625	NmADH9_NAD-depr	3	0.97625	382.8205	0.04	0	0.5	211.6609	205.8617	2	(276,-373,844)	Aperture: Medium														
67	NmADH9_103	1	1	0.97625	NmADH9_NAD-depr	3600	0.97625	385.1022	0.01	103	0.1	211.6617	205.8468	2.01	(276,-373,844)	EDNAStrategy1: subWedge:1Aperture: Medium														

19:00 26-07-2017

02:00 27-07-2017



10

12:00 29-07-2017	19:00 29-07-2017	41	i03	Dr Katherine McAuley	136	Compulsorily Remote
19:00 26-07-2017	02:00 27-07-2017	40	i03	Mr Mark Williams	71	Compulsorily Remote
02:00			i03	Dr Neil Paterson	38	Compulsorily Remote
17:00			i03	Dr Neil Paterson	76	Compulsorily Remote
19:00			i04	Dr Melanie Vollmar	69	Compulsorily Remote
12:00			i04	Dr Dave Hall	106	Compulsorily Remote
02:00			i03	Dr Neil Paterson	82	Compulsorily Remote

What you did...

(1) ISPyB interface



(2) Visit PDF (combine with others...)

mx13467-41 on i03 at 29-07-2017 12:00										
Sample	Images	Res	λ	Ω Osc	Spacegroup	Unit Cell	Processed Resolution	Rmeas	Completeness	Comments
NmADH9_23	3600	1.3	0.9763	0.10	P 1 2 1	63.92, 107.75, 69.36 90.00, 104.27, 90.00	29.77 - 1.5 29.77 - 6.7 1.54 - 1.5	0.085 0.036 0.824	98.8 98.7 96.6	(-262,-192,1150) EDNAStrategy1: subWedge:1Aperture: Medium

(3) Visit Excel sheet (annotate...)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	
1		mx13467-41_i03_29-JUL-2017 - 7 hr - Pilatus3 6M detector (100 Hz)																			
2																					
3		data:Image prefix	Run	Sta	Sam	Prot	# imag	Wavele	Dist	Exp.	Phi s	Phi ra	Xbe	Ybe	Det	auto/m	resoln	space gr	cell	twinn	in comments
54		NmADH9_23	2	##	NmADH9_23	3	0.9763	428	0.04	0	0.5	212	206	2.2							
55		NmADH9_23	3	##	NmADH9_23	3	0.9763	265	0.04	0	0.5	212	207	1.5							
56	12	NmADH9_23	1	##	NmADH9_23	3600	0.9763	213	0.01	45	0.1	212	207	1.3	a/3dii	1.37	P21	64 108 69 / 90 104 90	2.09	Binary complex with NAD - best data so far	
57		NmADH9_24	1	##	NmADH9_24	3	0.9763	265	0.04	0	0.5	212	207	1.5							
58		NmADH9_24	2	##	NmADH9_24	3	0.9763	265	0.04	0	0.5	212	207	1.5							



Retain ISPyB name in CCP4i2...

CCP4i2

Create a separate
“project” for
each dataset

If you have more than one dataset for a given protein – create a separate “dummy project” and group the “dataset projects” under this:

Create a New Project

Name of project/folder: NmADH9_23_1

By default all projects go in the 'CCP4I2_PROJECTS' directory in your home area - click 'Select directory' to choose an alternative.
Hint to organise your projects: in the 'Manage projects' window you can use a project as a folder and drag other projects into it

Description of project: High res data from binary complex with NAD

Choose tag.. Choose tag.. Choose tag..

New tag

Save

Create project Select directory Cancel Help

Manage projects

Name	Directory	Created	Last active	Tags
NmADH9	C:\Users\la...	18 Oct 18	24 Nov 18	
NmADH9_23_1	C:\Users\la...	24 Nov 18	24 Nov 18	
NmADH9_26_1	C:\Users\la...	24 Nov 18	27 Nov 18	
	C:\Users\la...	24 Nov 18	24 Nov 18	

Open

Add project or folder

Rename project

Edit description

Can rename “projects” later... (when you know more...)*

Manage project

Name	Directory	Created	Last active	Tags
NmADH9	C:\Users\la...	18 Oct 18	24 Nov 18	
NmADH9_23_1_NAD	C:\Users\la...	24 Nov 18	24 Nov 18	
NmADH9_26_1_apo	C:\Users\la...	24 Nov 18	27 Nov 18	
	C:\Users\la...	24 Nov 18	24 Nov 18	

Open

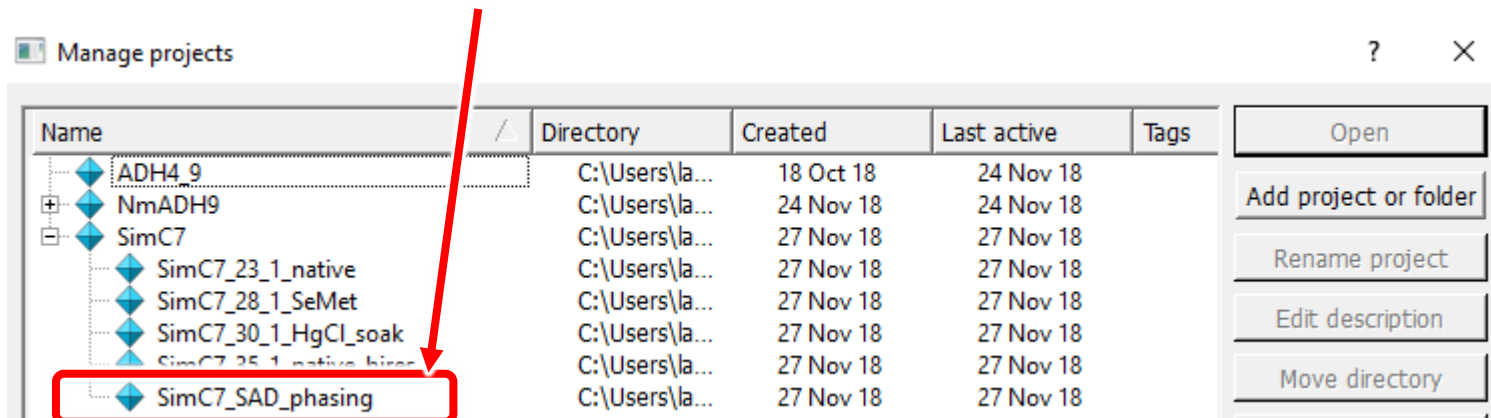
Add project or folder

Rename project

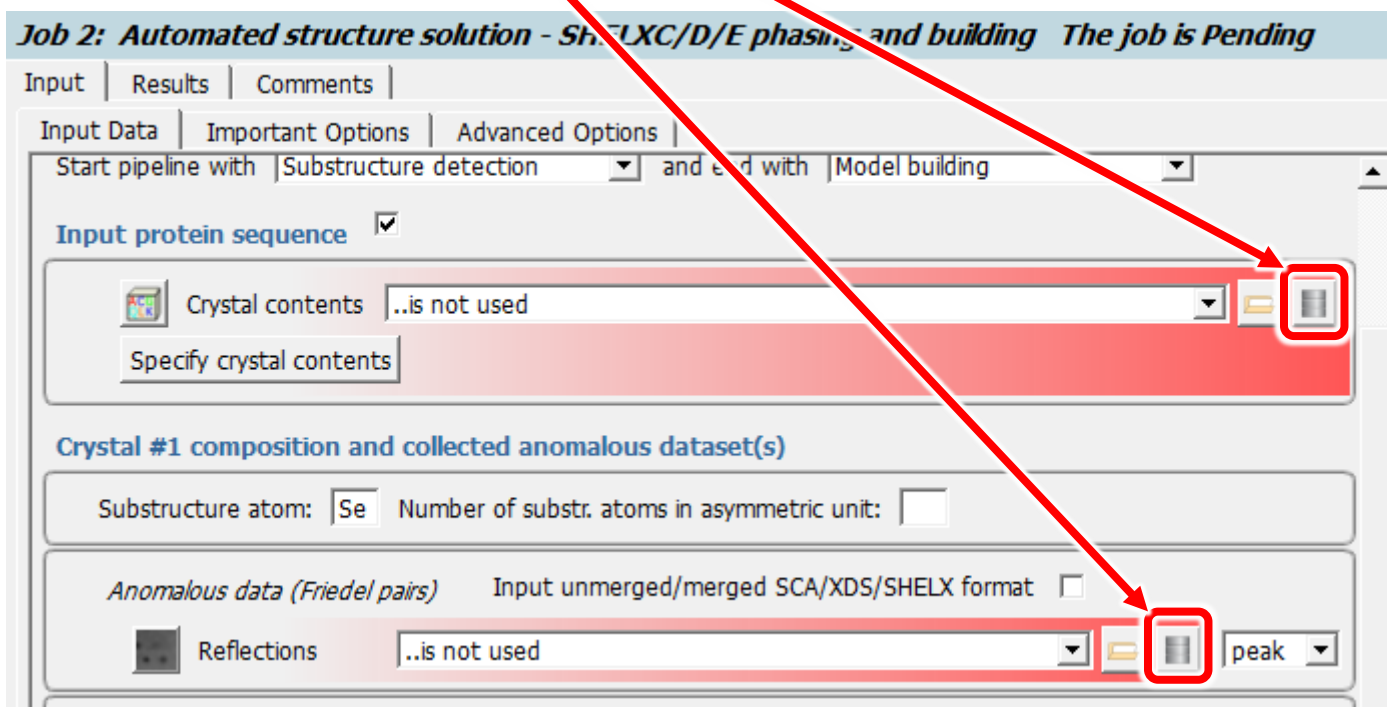
Edit description

*but directory names remain the same

Could create “non-dataset projects” for specific activities such as phasing...



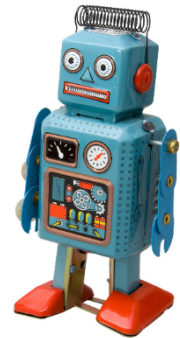
...then borrow files from other projects...

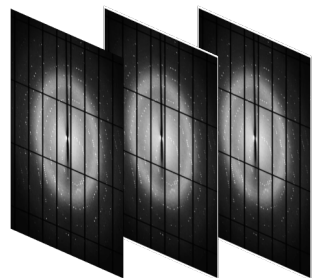


Ideal scenario during session:

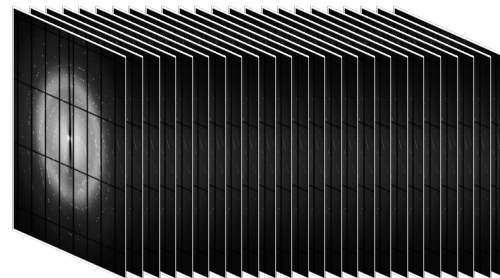
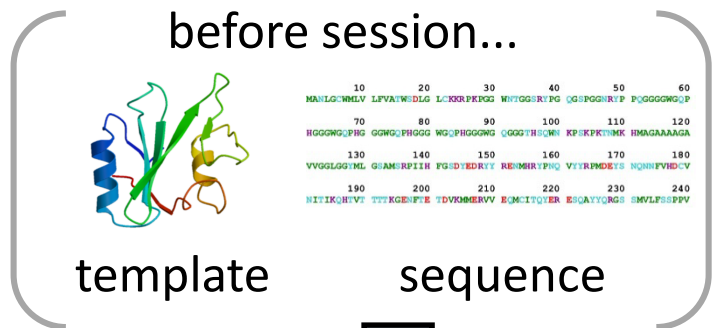
- Load first sample
- Collect test images
- Based on these, decide:
 - to collect...
 - not to collect...
 - to revisit later...
- For a “suitable” sample:
 - devise a data collection strategy
 - collect data set
- Analyse data as they are collected
- Based on this analysis, revise plans if appropriate
- Move onto next sample...

– not practical
without automation





test images



dataset



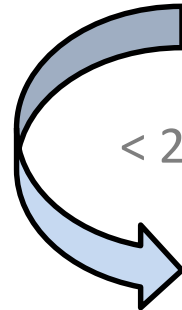
ISPyB

< 1 min



Strategies:
Mosflm
EDNA
Xia2

< 2 min



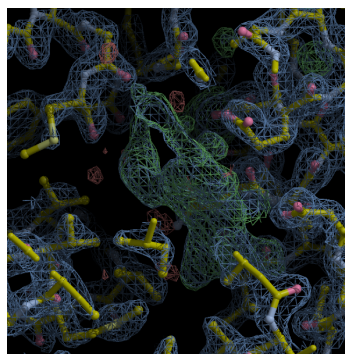
Fast DP

h	k	l	plus	SIGplus	lminus	SIGlminus
-45	0	4	-1.00	1.20	-1.00	1.20
-45	0	5	-0.03	1.82	-0.03	1.82
-45	0	6	2.17	2.01	2.17	2.01
-45	0	7	-0.22	1.24	-0.22	1.24
-45	0	8	0.63	1.33	0.63	1.33
-45	0	9	1.46	1.40	1.46	1.40
-45	0	10	0.11	1.34	0.11	1.34
-45	0	11	2.02	1.41	2.02	1.41
-45	0	12	0.63	1.33	0.63	1.33

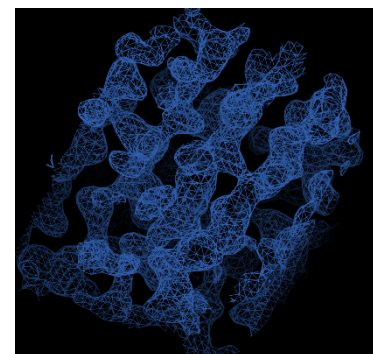
< 5 min



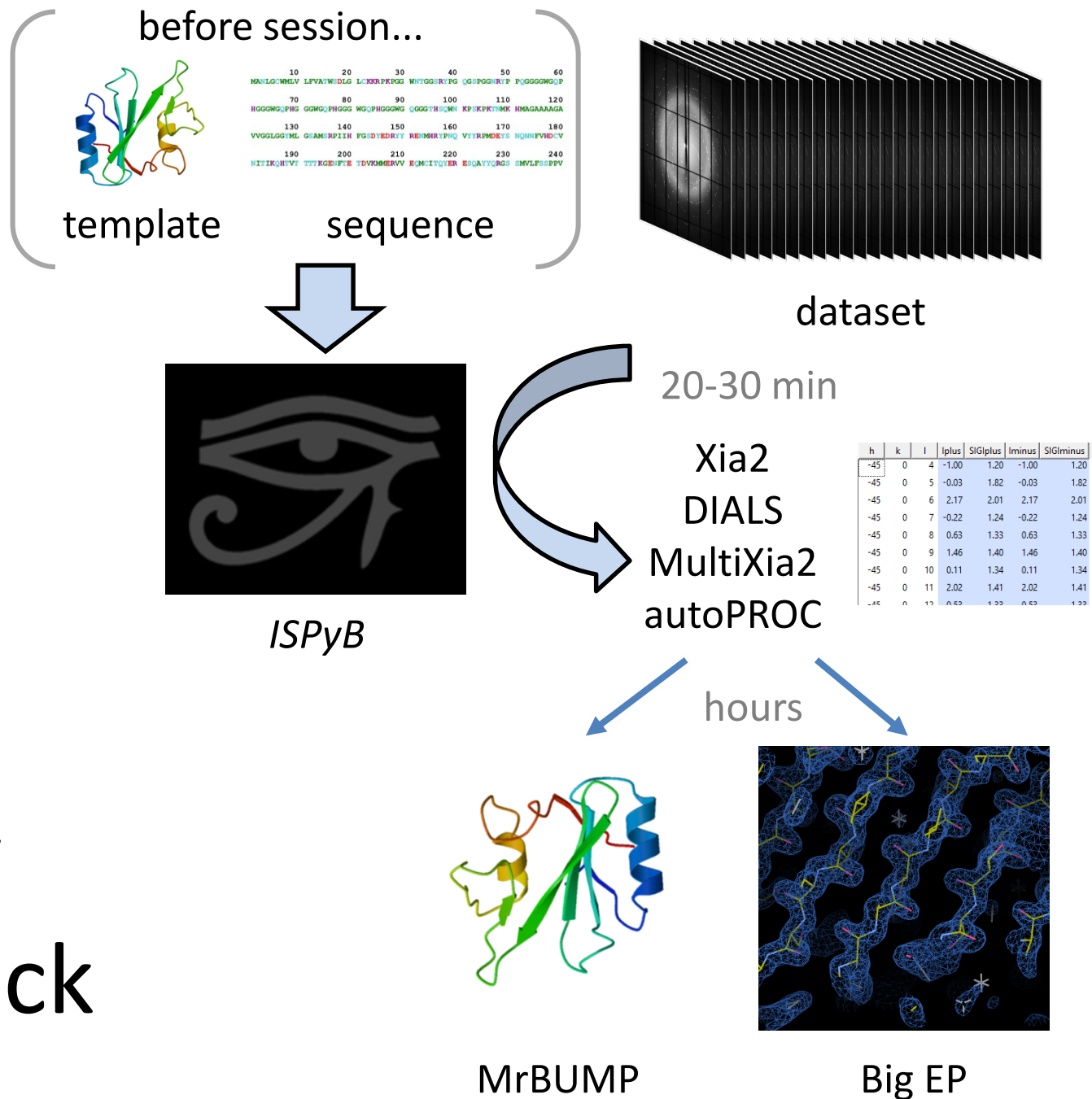
Quick
feedback



Dimple



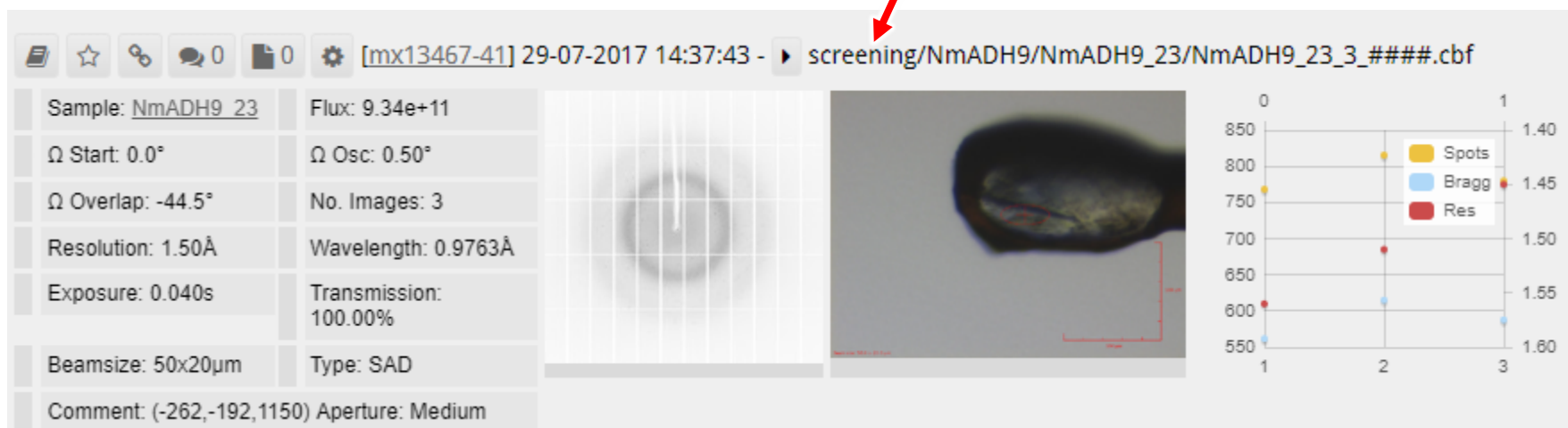
Fast EP



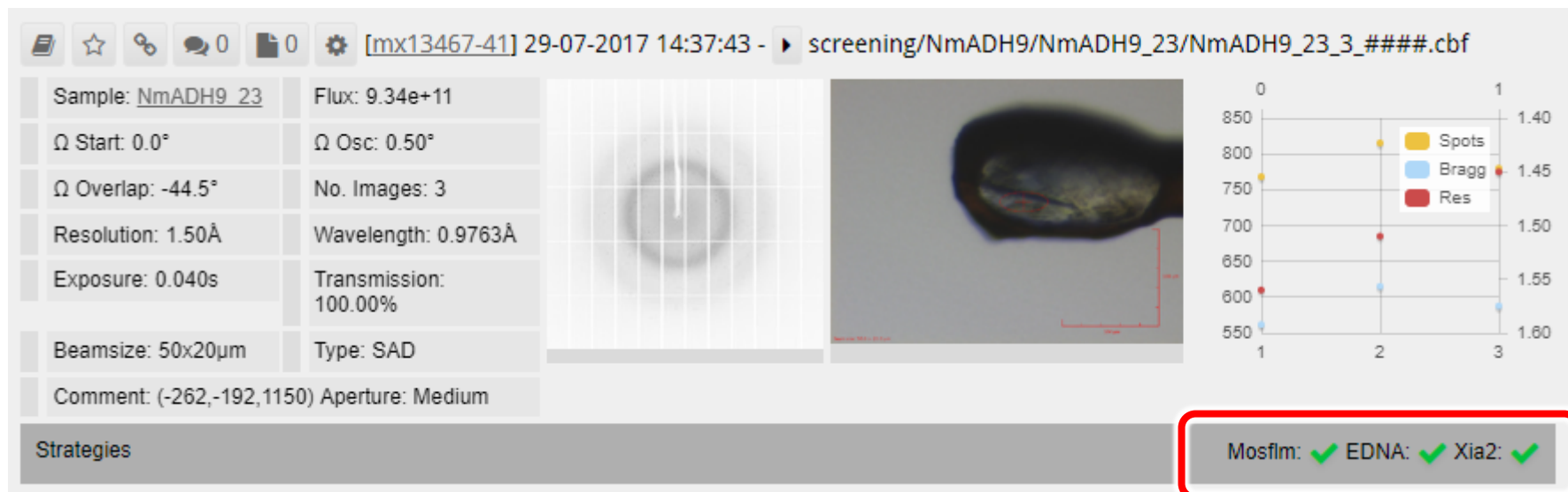
Slow
feedback

Screening...

(using the “screening” tab...)



Screening...



Screening...

Strategies

Mosfilm: ✔ EDNA: ✔ Xia2: ✔

xia2.strategy

Space Group	A	B	C	α	β	γ
P 1 2 1	63.87	107.69	69.31	90.00	104.25	90.00

🔍 Lookup Cell

Strategy	Description	Ω Start	Ω Osc	Res (Å)	Ranking Res (Å)	Rel Trn (%)	Abs Trn (%)	Exposure (s)	No. Images
anomalous Wedge1	Standard Anomalous Dataset Multiplicity=3 I/sig=2 202s	0	0.15	1.51	1.17	4.0	4	0.010	2214
gentle Wedge1	Gentle: Target Multiplicity=2 I/sig=2 Maxlifespan=20s	6	0.15	1.51	1.29	4.0	4	0.010	2000
high multiplicity Wedge1	Strategy with target multiplicity=16 I/sig=2 202s	0	0.15	1.51	1.17	4.1	4.1	0.010	2400
native Wedge1	Standard Native Dataset Multiplicity=3 I/sig=2 Maxlifespan=202s	162	0.15	1.51	1.14	2.2	2.2	0.010	1947

EDNA MXv1

Space Group	A	B	C	α	β	γ
P2	63.99	107.49	69.28	90.00	104.37	90.00

🔍 Lookup Cell

Strategy	Description	Ω Start	Ω Osc	Res (Å)	Ranking Res (Å)	Rel Trn (%)	Abs Trn (%)	Exposure (s)	No. Images
Strategy1 Wedge1	Standard Native Dataset Multiplicity=3 I/sig=2 Maxlifespan=202 s	45	0.10	1.51	1.21	11.0	11	0.010	1630
Strategy2 Wedge1	Standard Anomalous Dataset Multiplicity=3 I/sig=2 Maxlifespan=202 s	78	0.10	1.51	1.24	12.4	12.4	0.010	2740
Strategy3 Wedge1	strategy with target multiplicity=16, target I/sig=2 Maxlifespan=202 s						5.1	0.010	3600
Strategy4 Wedge1	Gentle: Target Multiplicity=2 and target I/Sig 2 and Maxlifespan=202 s						14.8	0.010	1110
Strategy5 Wedge1	UnderDEV Anomalous Dataset, RadDamage of standard						12.4	0.010	2740

mosfilm

Space Group	A	B	C	α	β	γ
P2	63.80	107.65	69.44	90.00	104.13	90.00

🔍 Lookup Cell

Strategy	Description	Ω Start	Ω Osc	Res (Å)	Ranking Res (Å)	Rel Trn (%)	Abs Trn (%)	Exposure (s)	No. Images
anomalous Wedge1		196	0.20	1.46	0.00	0.0	0	0.000	525
native Wedge1		211	0.20	1.46	0.00	0.0	0	0.000	525

The screenshot shows the 'Data Collection Settings' dialog box. The 'Sample' field is set to 'NmADH9_23'. The 'Strategies...' button is highlighted with a red box. A red arrow points from the 'Strategy1 Wedge1' row in the table above to the 'Sample' field.

Collecting a dataset...

Space Group	A	B	C	α	β	γ
P2	63.99	107.49	69.28	90.00	104.37	90.00

Q Lookup Cell

Strategy	Description	Ω Start	Ω Osc	Res (Å)	Ranking Res (Å)	Rel Trn (%)	Abs Trn (%)	Exposure (s)	No. Images
Strategy1 Wedge1	Standard Native Dataset Multiplicity=3 I/sig=2 Maxlifespan=202 s	45	0.10	1.51	1.21	11.0	11	0.010	1830

☐ Data Collection Settings ☒ Plate View ☐ Screening ☐ Command Queue

Data Collection Settings

Run Scan

Sample

NmADH9_23 ✖

Strategies...

Barcode NR

Holder 2

Position 14

Files

Visit directory
/dls/i03/data/2017/mx13467-41

Folder [Configure Defaults](#)
JIC/\${proteinacronym}/\${samplena
JIC/NmADH9/NmADH9_23

Prefix [Configure Defaults](#)
\${samplename}
NmADH9_23

☒ Automatic run number

Run number 0

Comment
EDNAStrategy1: subWedge:1

Omega

Start 45.00 °

Oscillation 0.100 °

Total oscillation 360.0 °

Delta 0.00 °

Image

Number of images 3600

Exposure time 0.010 s

Total exposure time 36.0 s

First image number 1

Beam and Detector

Maximum resolution 1.3000 Å

Detector distance 213.5 mm

Wavelength 0.97623 Å

Energy 12700.3 eV

☐ Use current energy

Transmission 50.156283 %

make sure you collect enough data – still could be P1...

push the resolution a little...

increased transmission...

Checking the results...



Checking the results...

Import directly into CCP4i2



Auto Processing Fast DP: Xia2: ✓✓✓ MultiXia2: ? ? autoPROC: ✓

Type Resolution Spacegroup Mn<I/sig(I)> Rmeas I/sig(I) Completeness Cell Status

fast_dp

xia2 3d

xia2 3dii

xia2 dials

autoPROC 1.0.5 (see: <http://www.globalphasing.com/autoproc/>)

fast_dp xia2 3d xia2 3dii xia2 dials autoPROC

Beam Centre X Y

Start 211.60 206.96

Refined 211.50 206.92

Δ 0.10 0.04

Space Group A B C α β γ

P 1 21 1 63.92 107.75 69.36 90.00 104.27 90.00

Shell Observations Unique Resolution Rmeas I/sig(I) CC Half Completeness Multiplicity Anom Completeness Anom Multiplicity CC Anom

outerShell 64205 9220 1.37 - 1.39 1.685 1.1 0.5 96.7 7.0 96.0 3.5 0.0

innerShell 64845 9660 3.72 - 42.06 0.038 44.2 1.0 99.9 6.7 98.6 3.5 0.1

overall 1285570 187412 1.37 - 42.04 0.099 10.8 1.0 98.5 6.9 98.0 3.5 0.1

Attachments

File Type

mx13467v41_xNmADH9231_free.mtz Result Download

xia2.html Log Download View

15 Page < 1 >

Close

Archive Logs & Files Lookup Cell

But there are other ways... (see later)

N.B. space group is only a hypothesis at this stage! – see talk about data processing

Re-running jobs:

- Most pipelines will run from the command line (Terminal window)
- Also through ISPyB interface...

The screenshot displays the ISPyB web interface for managing X-ray diffraction jobs. At the top, a toolbar includes icons for file operations and a gear icon for settings, which is highlighted with a red circle and a red arrow pointing to a text box. The main panel shows job details for 'NmADH9_23' and a 'Reprocess Data' button. A text box titled 'Change:' lists options for reprocessing: space group, cell parameters, maximum resolution, and reject images. Below this, a 'Process Individually' checkbox and a 'Pipeline' dropdown menu are visible. The bottom section shows a detailed view of the job 'NmADH9_23_1' with its parameters and a data plot. The plot shows 'Spots' (yellow), 'Bragg' (blue), and 'Res' (red) data points across a range of resolution values. The 'Integrate' and 'Close' buttons are at the bottom right.

Change:

- space group
- cell parameters
- maximum resolution
- reject images

Sample: NmADH9_23 Flux: 1.60e+11
 Ω Start: 45.0° Ω Osc: 0.10°
 Ω Overlap: 0° No. Images: 3600
Resolution: 1.30Å Wavelength: 0.9763Å
Exposure: 0.010s Transmission: 50.16%
Beamsize: 50x20µm Type: SAD
Comment: (262, 102, 1150) EDNA Strategy 1
sub

Reprocess Data

Auto

Multi Crystal

☐ Process Individually | Pipeline: Xia2 3dii High Res: Å | Space Group / Cell Options

NmADH9_23_1 - JIC/NmADH9_23/NmADH9/

Sample: NmADH9_23 Ω Start: 45.0°, Osc: 0.10°
Resolution: 1.30Å Wavelength: 0.9763Å
Start End +

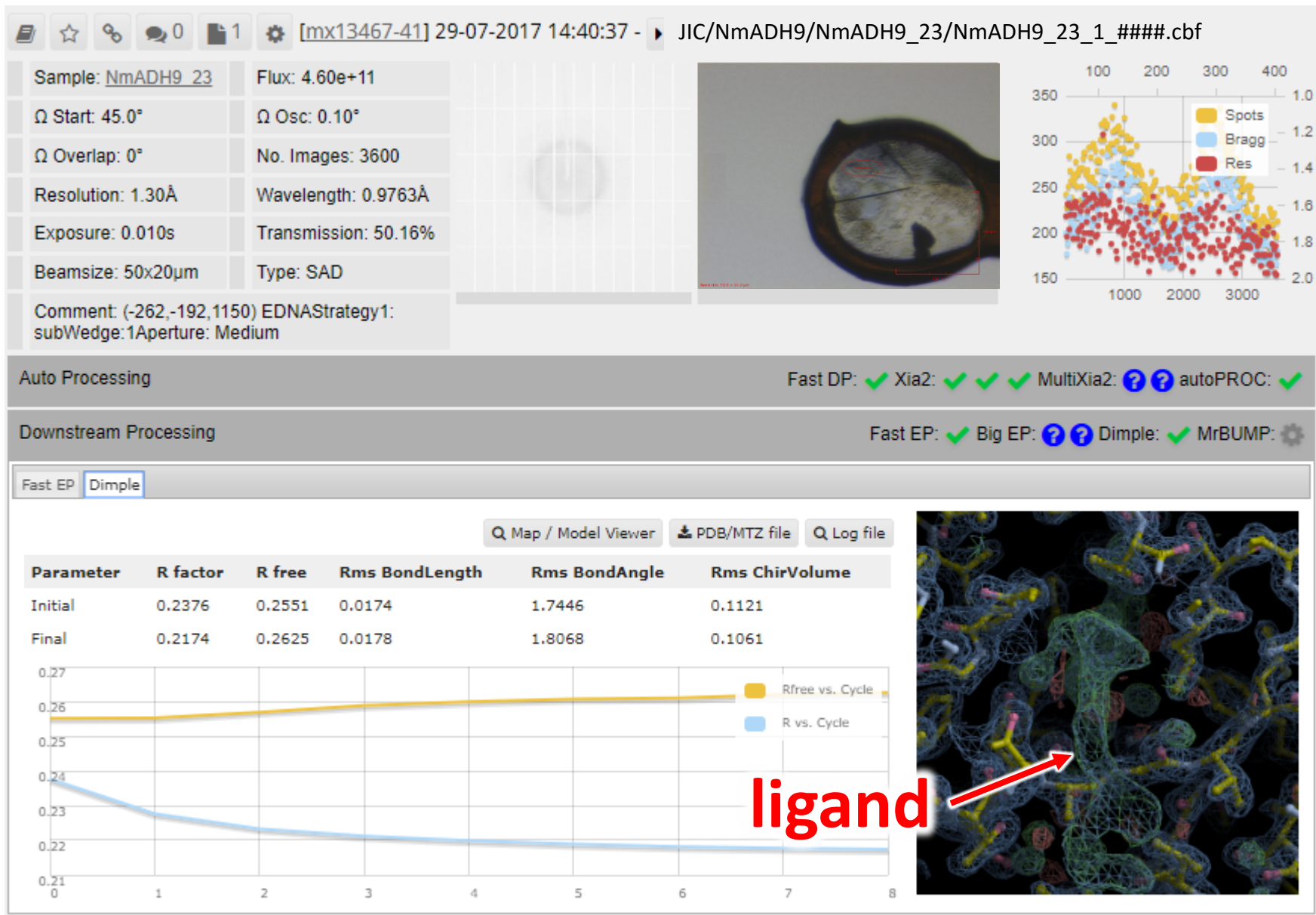
50 100 150 200 250 300 350 400
350 300 250 200 150
500 1000 1500 2000 2500 3000 3500
Spots Bragg Res

Integrate Close

Checking the results...



Checking the results...

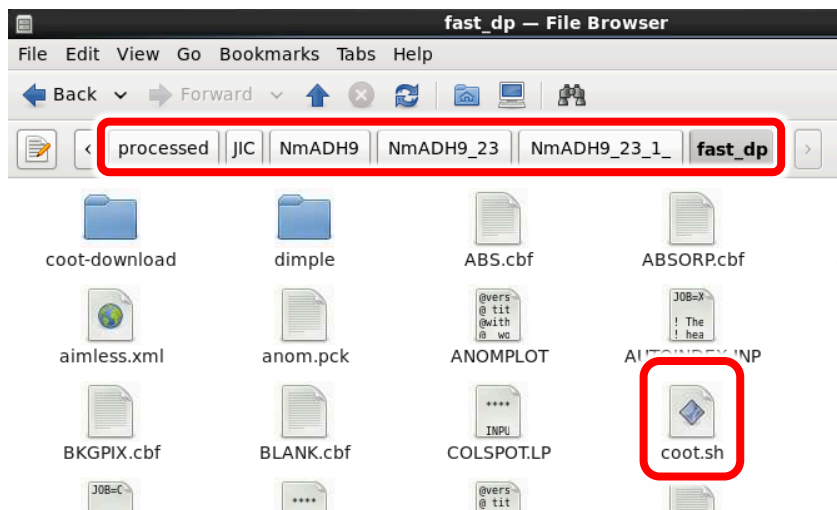


Looking at maps in Coot (on DLS computer):

...go to the appropriate fast_dp/fast_ep directory in “processed”

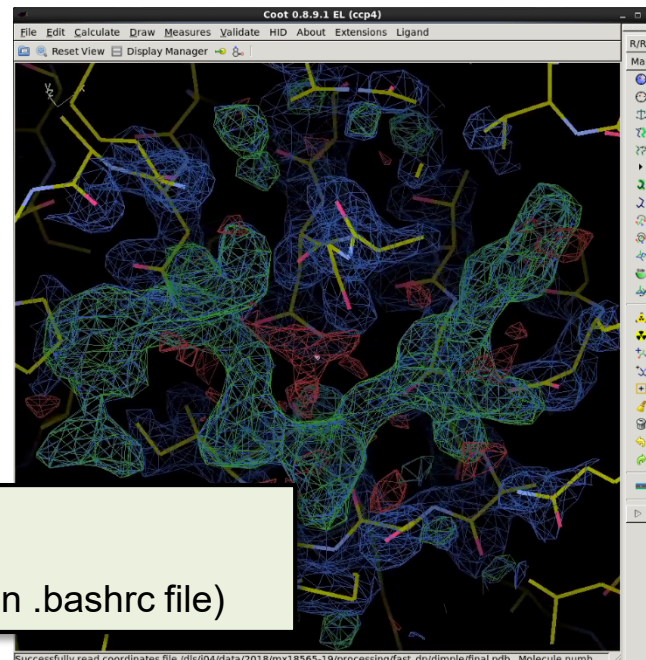
Using File Browser...

double click on: “coot.sh” icon



Using Terminal...

type: `./coot.sh`



Top tip: setup directory shortcuts/aliases to avoid too much typing:

- e.g. alias m41="cd /dls/i03/data/2017/mx13467-41" (put in .bashrc file)

On the basis of a “positive” result, may decide to move onto a different project...

Ideal scenario after session:

- Each data set characterised as:
 - useful/may be useful/not useful
- Some datasets processed to your satisfaction
- You may have interpretable experimental maps
- You may have preliminary structures
- You **will** have less follow-up work to do!

What to do with all the data...

Raw data (images)

- removed from disk after 40 days – still available through TopCAT (tape archive)



compress images (gzip) if space is a problem (use a script)



“Meta” data – all the other “stuff” – removed from disk after 40 days – not backed up

>90% of useful datasets derived from **Meta data** rather than going back to **Raw data**

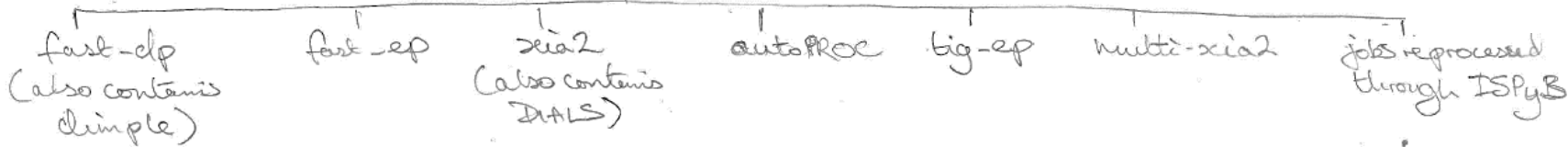
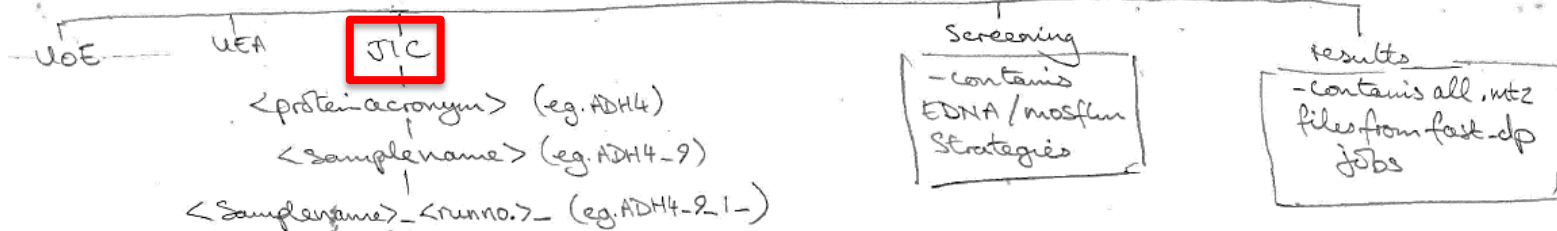
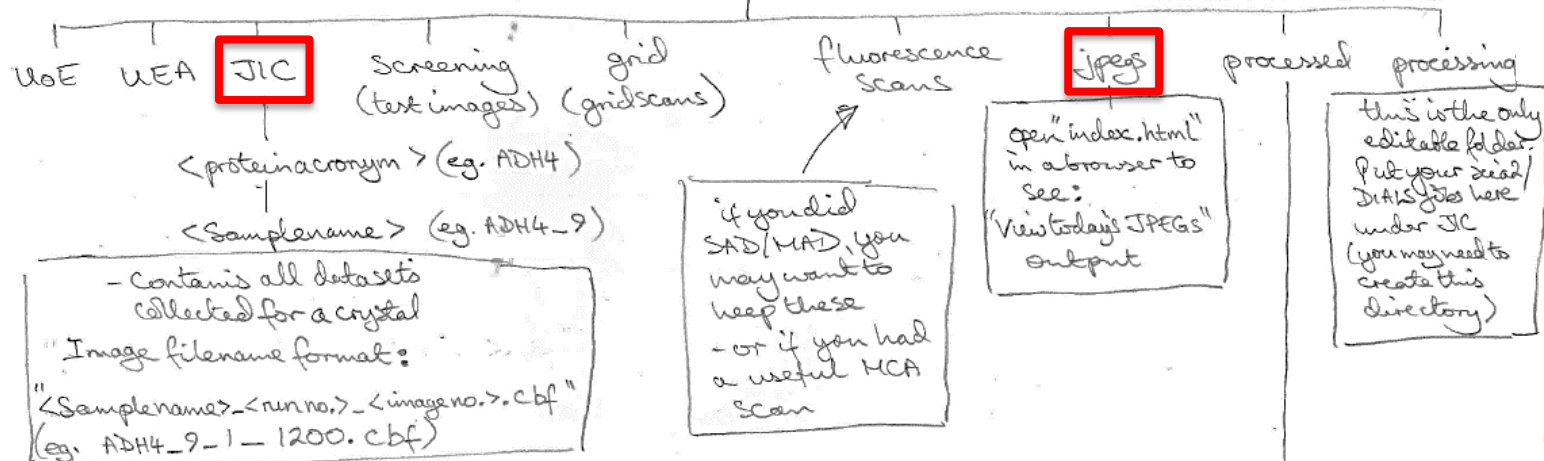
...just be thankful it's not cryo-EM!

Directory Structure on Diamond Computers

dls
 <beamline> (eg. i03)
 |
 data
 |
 <year> (eg. 2017)
 |
 <visit-ID> (eg. mx13467-20)

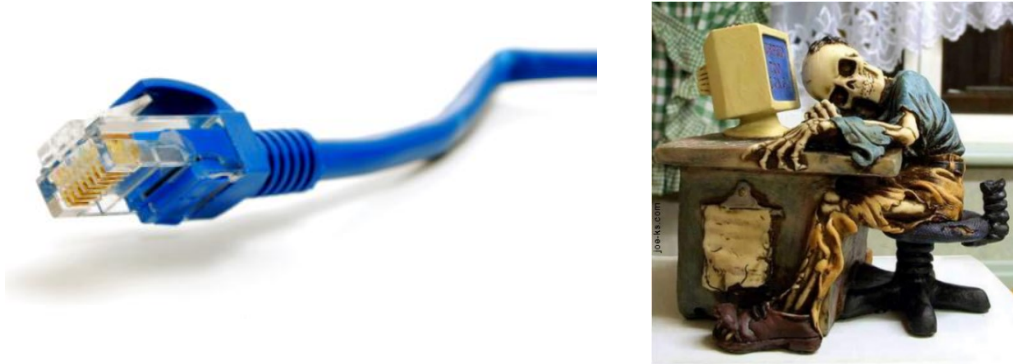
 Copied "home" by default

dataset
 ← directories →



Getting your data home...

- FTP data home (use a script)



Very slow!



Faster?

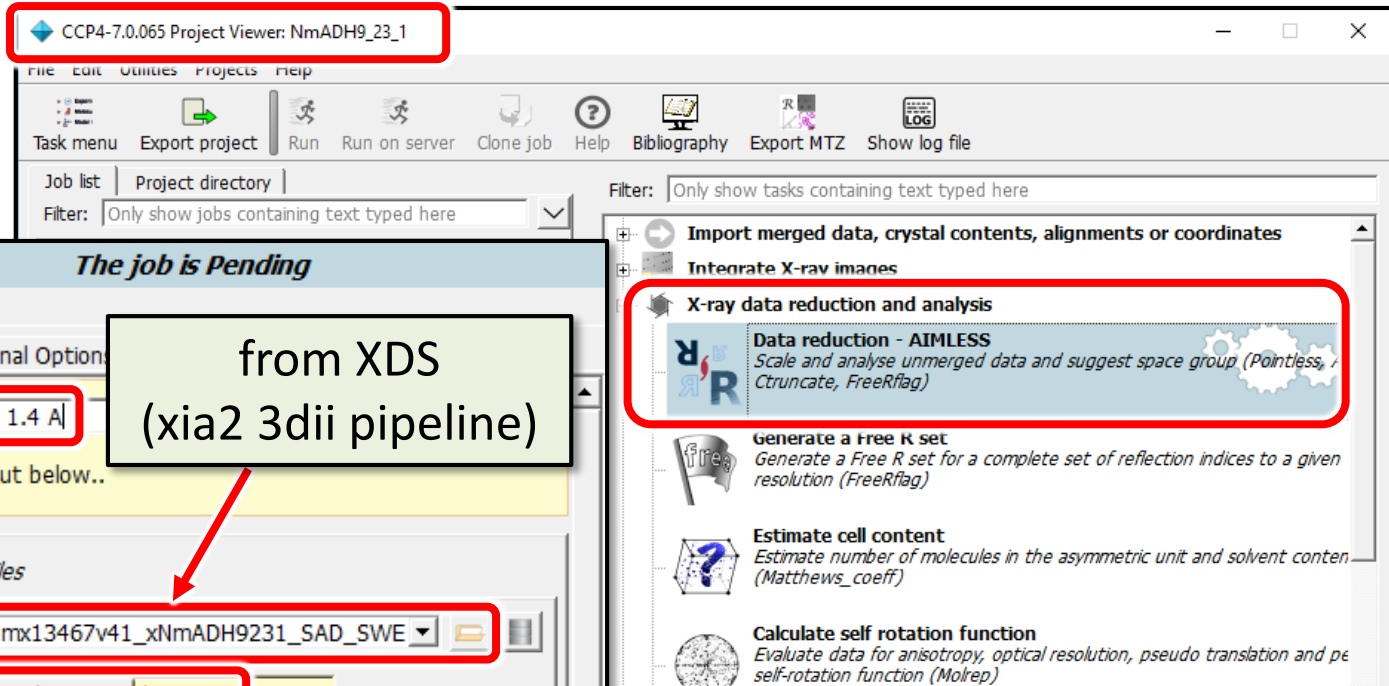
In the meantime:

- Use autoprocessing output or...
- (re)process data remotely on DLS computers and FTP output only
- archive (compressed) raw data later

My preferred option for processing...

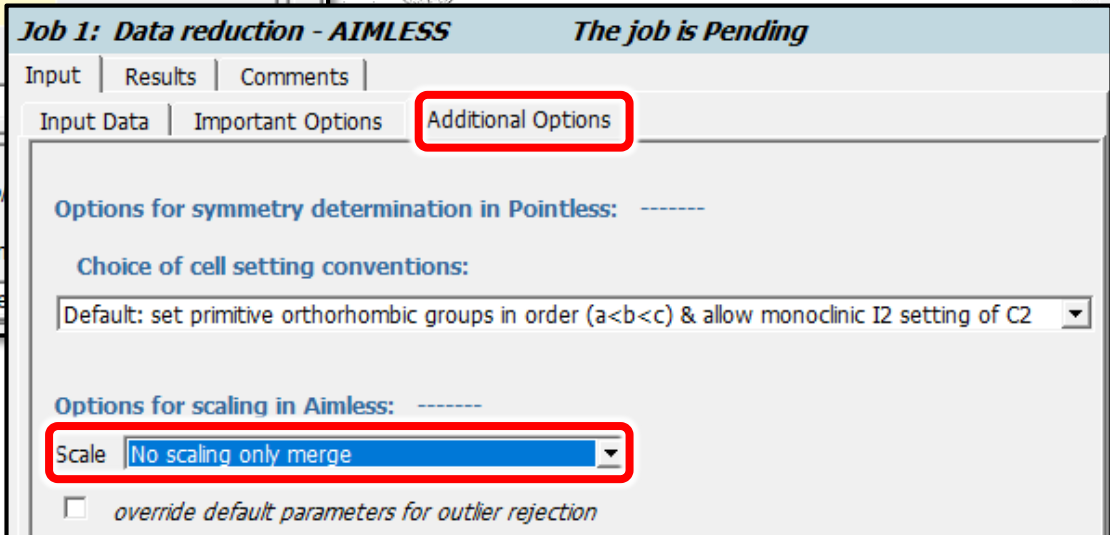
- where the pipelines have done a good job – take part-processed dataset (from the “Meta” data) and re-run the merging step in CCP4i2

use informative
job titles



from XDS
(xia2 3dii pipeline)

may want to change
resolution, exclude
images etc...



Gives lots of output – useful for troubleshooting...

for your
paper/thesis

Job 1: Data reduction - AIMLESS *The job is Finished*

Input Results Comments

Headline Summary SG details MergingGraphs SDanalysis MergingDetails Istats Biblio

Data reduction - full dataset to 1.4 Å

▼ Key summary

Selecting space group P 1 21 1
as there is a single space group with the highest score

Solution probability: 0.872, Confidence 0.866 (high resolution limit for symmetry testing 1.495 Å)

Key statistics for Dataset: NmADH9_23_1/NmADH9_23/1

Unit cell: 63.921 107.750 69.361 90.000 104.265 90.000, wavelength 0.976250 Å
Resolution of input data: 1.40 Å, resolution estimate: beyond 1.40 Å
Anisotropic limits: - Along 0.99 a° - 0.16 c° 1.48 Å CC(1/2), 1.58 Å I/σ - Along k axis 1.40 Å° CC(1/2), 1.42 Å I/σ - Along -0.09 1.40 Å I/σ
Rmeas: overall 0.096, inner bin 0.037
In outer bin: Mean(I/sd) 1.2 CC(1/2) 0.559
Overall filtered Mean(chi^2): 1.03
Anomalous CC(1/2) in inner bin 0.093
No significant anomalous signal detected
NOTE: no scaling was done, just merging

SD correction information:
SD correction parameters were not refined

✓ No evidence of twinning

✓ No evidence of possible translational non-crystallographic symmetry

● Warning: Some anisotropy detected. This may affect the quality of the data.

● Warning: Completeness test shows some issues.

✗ Warning: Severe deviation from Wilson plot.

✗ Warning: Possible ice rings found.

A free-R set has been created, fraction of the data = 0.05

Show Pointless logfile Show Aimless logfile Show Ctruncate logfile

▼ Overall summary

Job 1: Data reduction - AIMLESS *The job is Finished*

Input Results Comments

Headline Summary SG details MergingGraphs SDanalysis MergingDetails Istats Biblio Run

▼ Overall summary

Space group determination
Selecting space group P 1 21 1
as there is a single space group with the highest score

Solution type: space group

Group name	P 1 21 1
Reindex	[h,k,l]
Space group confidence	0.866
Laue group confidence	0.821
Laue group probability	0.882
Systematic absence probability	0.988

Scores for each symmetry element
Lattice group name P 1 21 1

Likelihood	CC	R	Symmetry
0.880	0.87	0.087	identity
0.882	0.87	0.087	2-fold k (0 1 0) [-h,k,-l]

Data internal consistency statistics

Summary of merging statistics for dataset
NmADH9_23_1/NmADH9_23/1

	Overall	Inner	Outer
Low resolution limit	47.16	47.16	1.42
High resolution limit	1.40	7.67	1.40
Rmerge(within I+ /I-)*	0.081	0.031	1.402
Rmerge(all I+ and I-)*	0.090	0.036	1.563
Rmeas (within I+ /I-)*	0.096	0.037	1.657
Rmeas (all I+ & I-)*	0.097	0.039	1.690
Rpim (within I+ /I-)	0.051	0.020	0.876
Rpim (all I+ & I-)	0.037	0.015	0.636
Rmerge in top intensity bin*	0.048		
Number of observations	1205637	7127	59269
Number unique	175909	1122	8610
Mean(I)/sd(I)	11.2	46.7	1.2
Half-set correlation CC(1/2)	0.999	0.998	0.559
Completeness %	98.6	99.2	97.0
Multiplicity	6.9	6.4	6.9
Filtered Mean(chi^2)	1.03	1.12	1.03
Anomalous completeness %	98.1	97.0	96.4
Anomalous multiplicity	3.4	3.4	3.5
DelAnom CC(1/2)	0.055	0.093	0.034
Mid-Slope of Anom Probability	1.038		

Download

all the important data processing statistics
are now in your CCP4i2 project database

...or reprocess from scratch using the Dials User Interface
(N.B. can read compressed images!)



Integrate images with Dials
Launch DUI and capture output

...somebody else may talk about that...

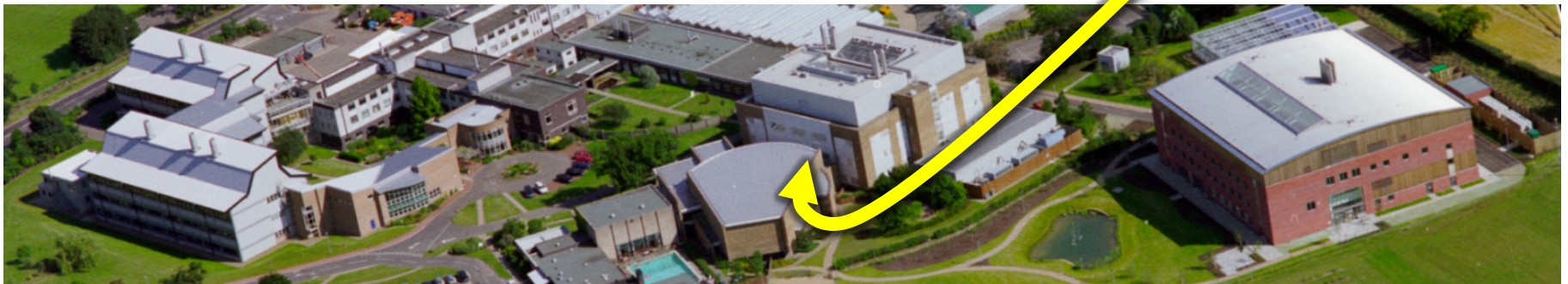
Summary - why use ISPyB & pipelines?

- Faster sample changing (select by ISPyB name)
 - essential for remote...
- Informs the decision making process
 - make decisions sooner
 - revise strategy on the fly
 - (e.g. recollect dataset x... no more data required for project y...)
 - make better overall use of beamtime
- Easier to cope with a small team
- Reduces amount of post-beamtime follow-up work
- Simple to keep track of your samples and data

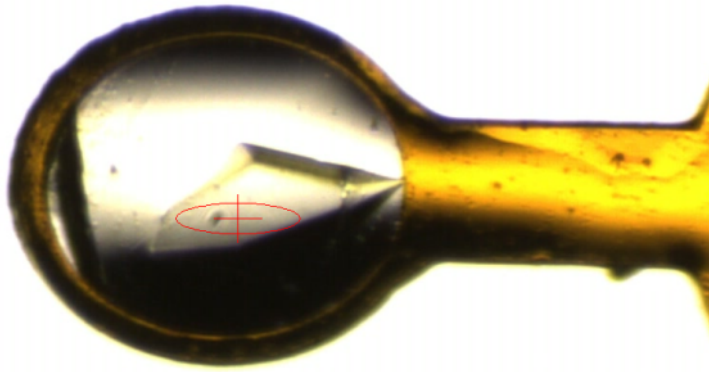
data Remote/collection



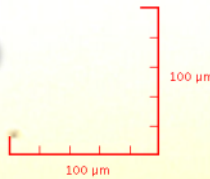
...from here



Some real data...

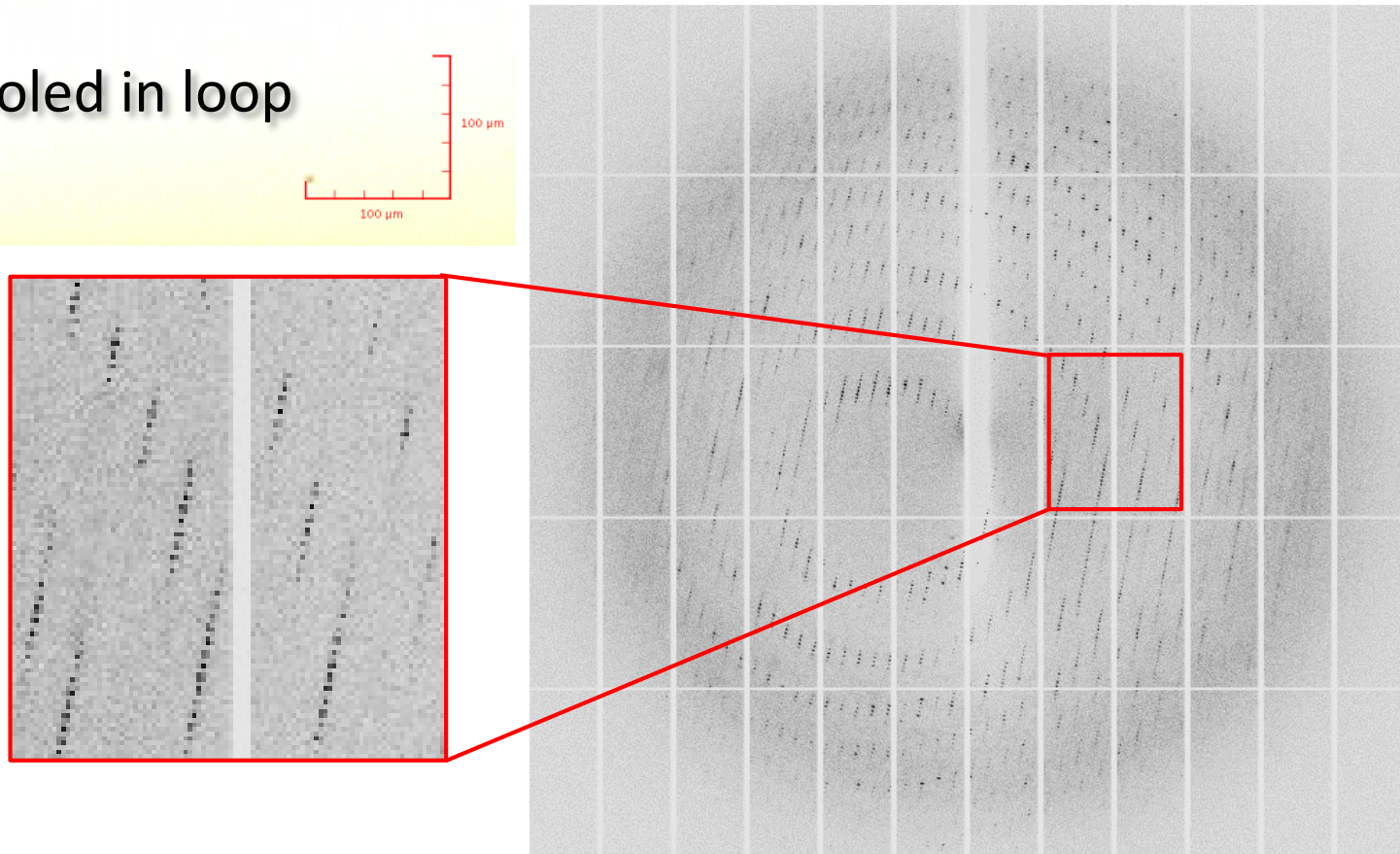


Crystal cooled in loop



Beam size: $82.0 \times 21.0 \mu\text{m}$

Diffraction pattern recorded
from crystal at Diamond



Who should collect the data?

- Someone who knows the project
 - what data are required?
 - what space group/cell parameters/resolution to expect
- Someone who has a stake in the project
- Person who grew and mounted the crystals
- Someone who is not tired!

Ideal scenario:

- everyone collects data from their own crystals
- work in short shifts



Solutions:

1. everyone travels to Diamond (not practical ... and expensive)
2. small team travels to Diamond – everyone else collects data remotely (mixed access)
3. everyone collects data remotely (fully remote access)

Who travels to Diamond?

Without remote users:

- 3 people minimum – at least 2 are experienced users

With remote users (mixed access):

- 2-3 people – only 1 needs to be an experienced user
 - good way to introduce new users to DLS



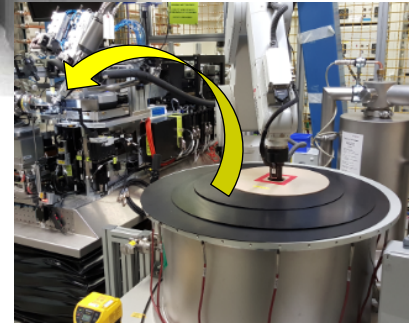
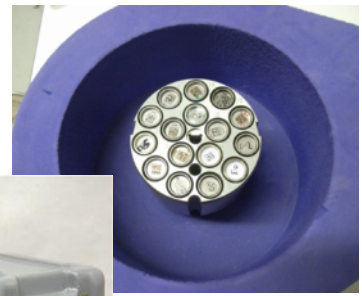
Who collects data remotely?

- Users at multiple sites; at each of these:
 - 1 experienced user can cope, but better with 2 (second person analyses data)
 - any number of new users (+observers)



For routine data collection at 100 K:

- Samples prepared and cryo-cooled in home lab
- Transported to Diamond in dry shipping dewars
- Samples mounted robotically
- Data collection controlled through GDA
- Only manual operation at DLS: loading/unloading pucks
 - Up to 592 samples can be loaded at once
 - may only need to enter hutch at beginning and end of visit
- Everything else computer driven...



Therefore you don't need to be there!

iMac with extra monitor:

- (i) data collection
- (ii) data evaluation

View today's jpegs/ISPyB

2nd iMac:

- (i) data evaluation
- (ii) data (re)processing
- (iii) Skype

webcams

Useful phone numbers

GDA interface

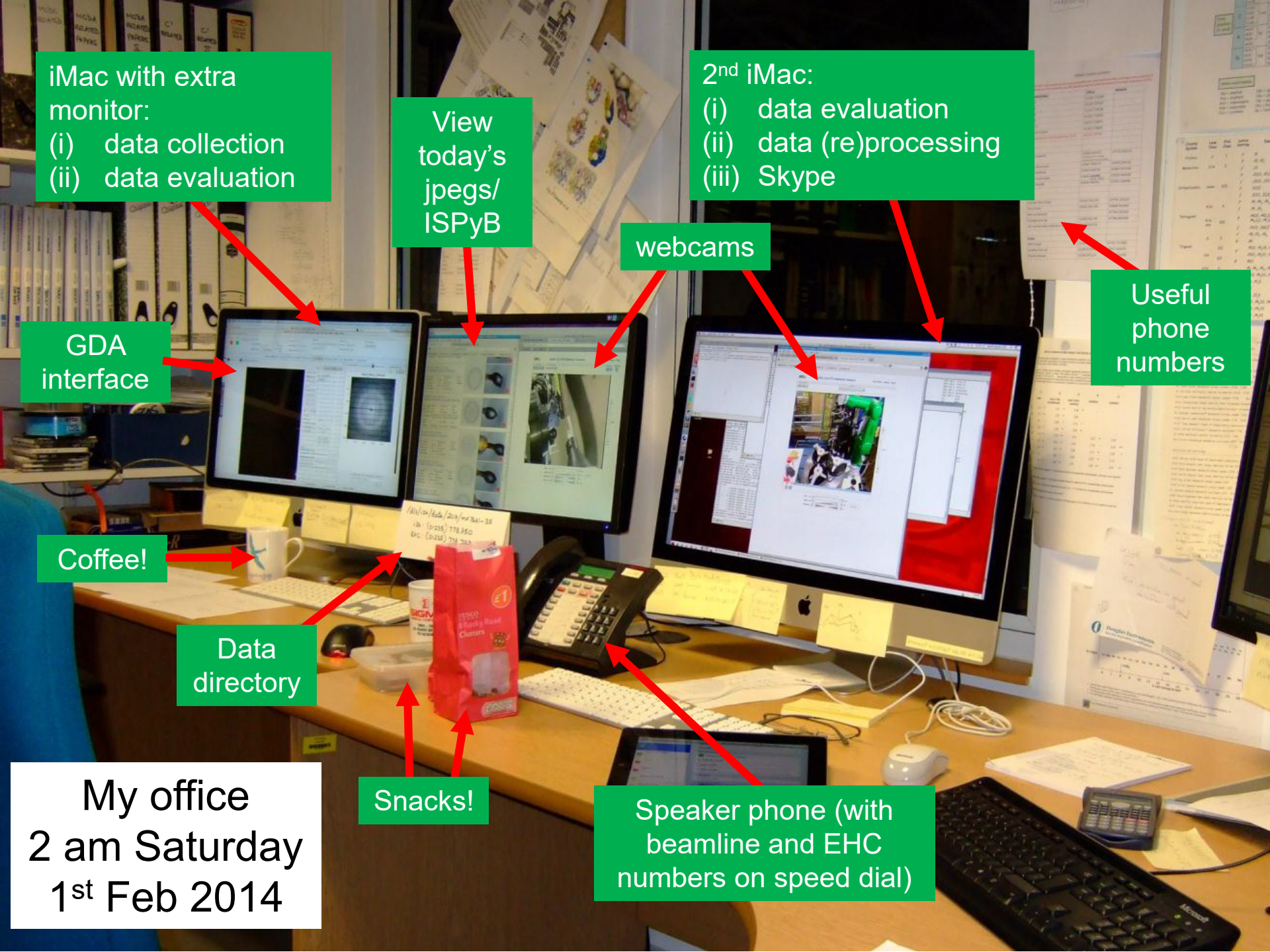
Coffee!

Data directory

My office
2 am Saturday
1st Feb 2014

Snacks!

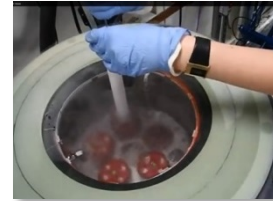
Speaker phone (with beamline and EHC numbers on speed dial)



Timeline for fully remote data collection

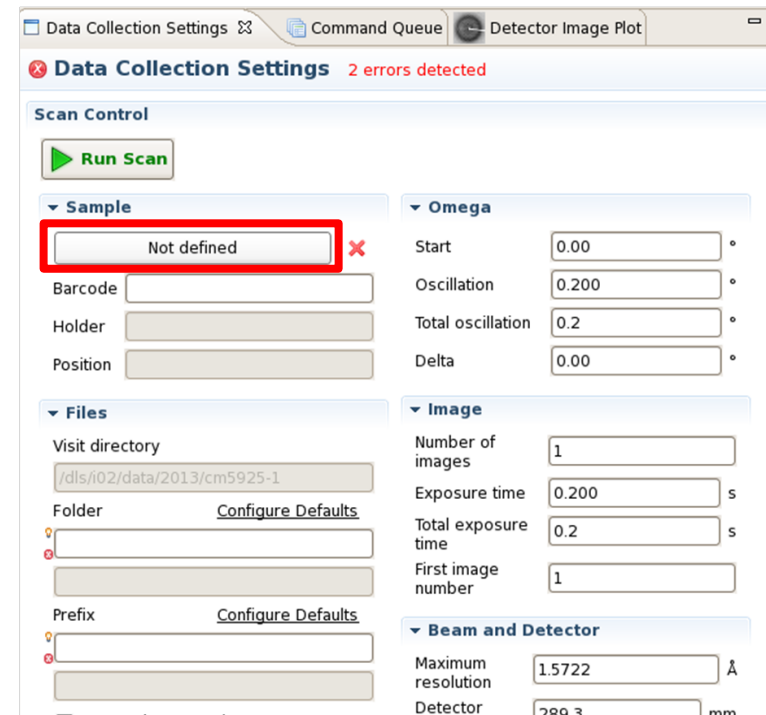
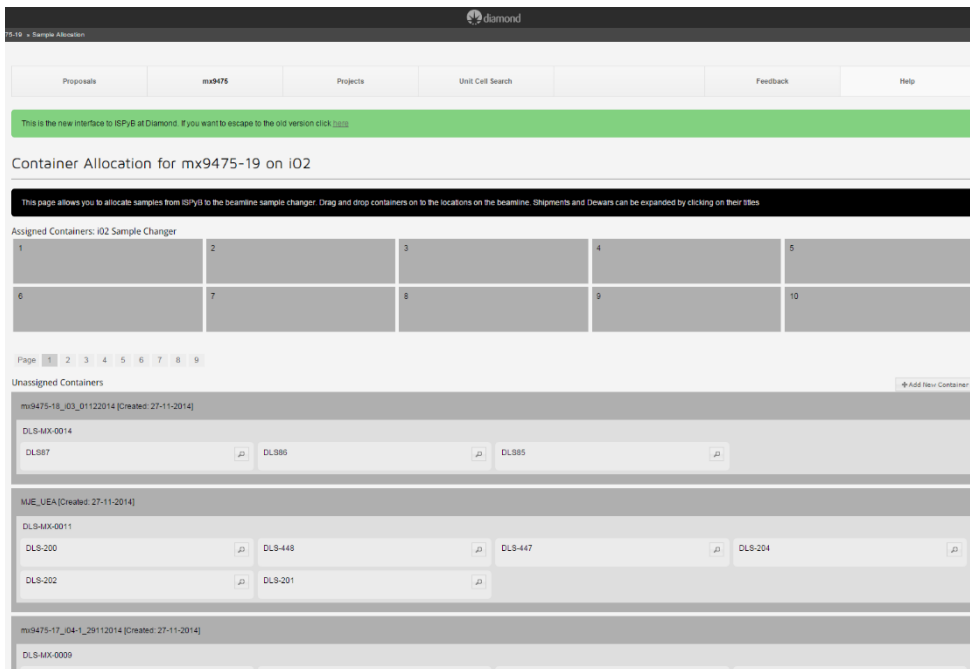
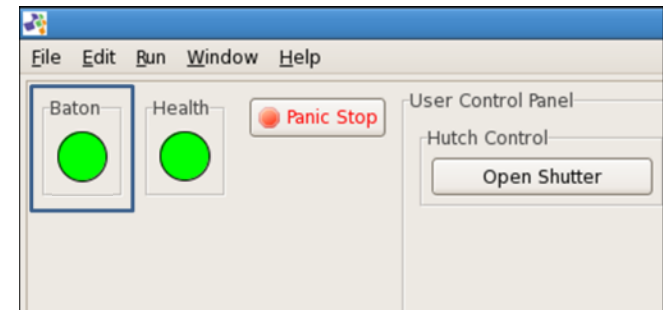
Sometime prior to shift:

- LC/EHC loads pucks into beamline dewar
- ~1 hr before start of shift (or slot in running order):
 - connect to beamline computer using NX client
 - start GDA
 - enter puck positions into ISPyB database (LC/EHC)



Zero hour:

- take the baton, select sample from menu, collect data!



When things go wrong...

Can you fix things?

- probably not!

On-site user:

- During normal working hours – call local contact
- Out of hours – call EHC

Remote user:

- During normal working hours – call local contact
- Out of hours – call EHC

Therefore you don't need to be there!

Check the webcams!

i04 Webcams & Beamline Status

Ring Current
299.991

Refill
255.758

Hutch
Locked

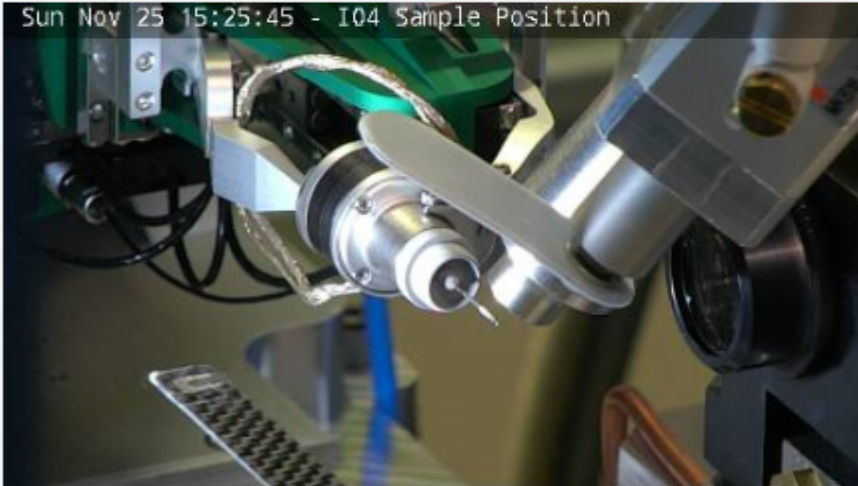
Port Shutter
Open

Expt Shutter
Open

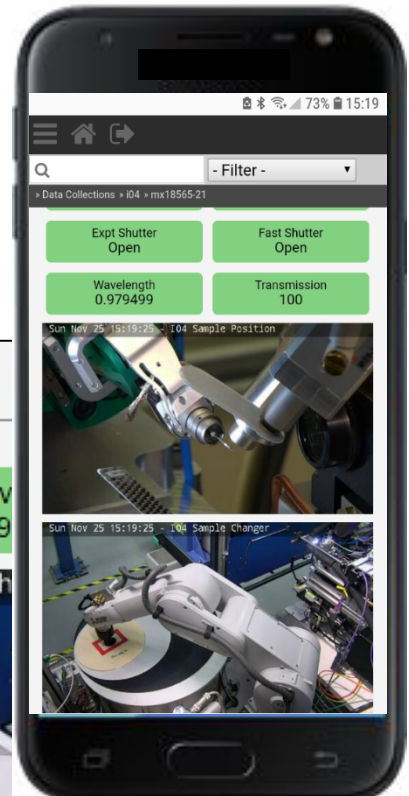
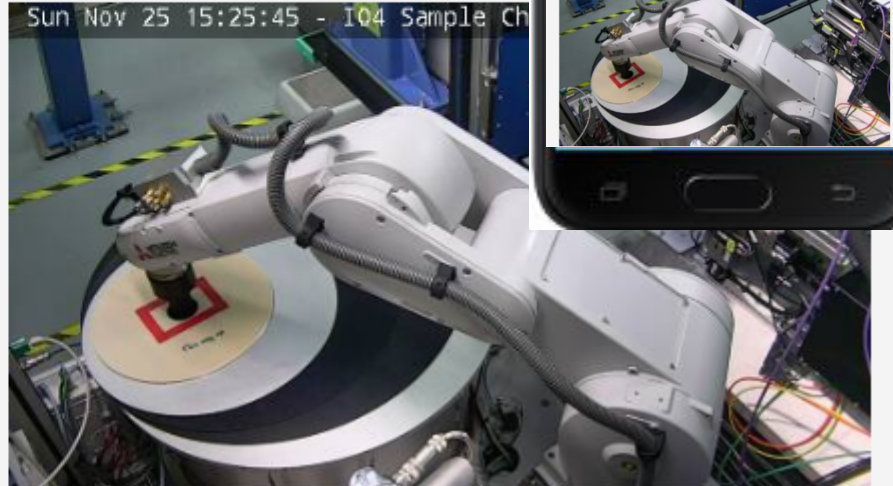
Fast Shutter
Open

Wav
0.9

Sun Nov 25 15:25:45 - I04 Sample Position



Sun Nov 25 15:25:45 - I04 Sample Ch

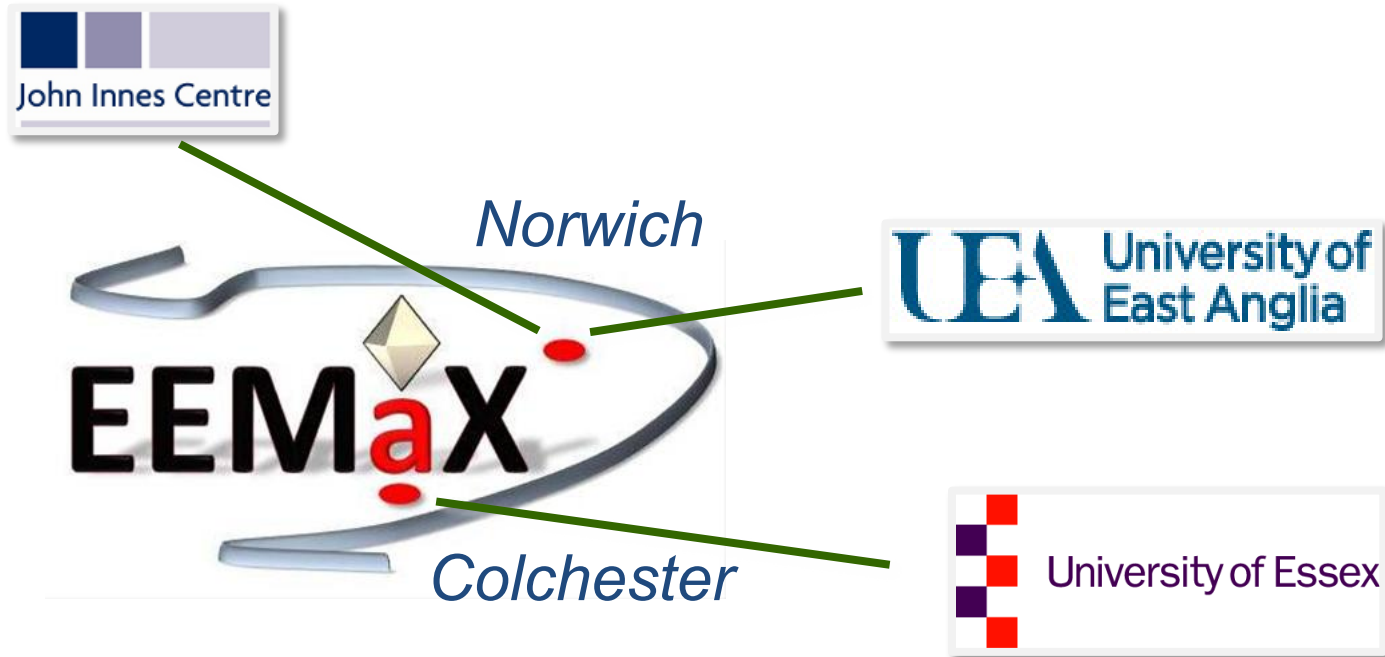


If you suspect a problem – call the EHC!

Advantages of remote data collection:

- Beamtime is fully used
- Users collect data on their own crystals
- Most/all of the users stay at home labs
 - time commitment is low for remote users
 - stakeholders can observe data collection
 - useful for training non-experts
- Time can be used flexibly
- Less stressful
- Everyone gets some sleep!
- Difficult to “break” the beamline

East of England Macromolecular X-ray Crystallography BAG



8 groups: ~30 potential users

123 sessions since Jan 2014

- 8 mixed mode
- 115 fully remote!

Take home messages

- If you are not using remote data collection – why not?
- If you are not using ISPyB – why not?
- If you are not using the software tools/pipelines – why not?
- If you don't have a life – you know why....



Data collection is your last
experimental step
– don't mess it up!

Acknowledgements

- Access to MX beamlines at Diamond
- Excellent support of:
 - Beamline staff
 - EHC/Control Room staff
 - User Office
- Software developers
- BBSRC