


PiMS, xtalPiMS and beyond: proteins, crystals and data

Chris Morris

STFC Daresbury Laboratory…

…and the PIMS development team

CCP4 Study Weekend, Nottingham, 7/1/2010

Outline of talk

- What is information management?

- PiMS concepts
- Use of PiMS for real projects

- xtalPiMS for crystallization

- Getting information to the synchrotron

What is information management?




- **The process of storing experimental information for later retrieval**
  - Storage medium can be human memory, paper, electronic files or relational databases
  - Quality assurance (QA) experiments are typically repetitive, research experiments are constantly changing
  - Purpose of retrieval could be supporting the next experiment, providing reports, publishing papers or depositing public data
  - Automated systems may require electronic information management

Research laboratory processes




- **Full projects (solving a protein structure?)**
  - Long workflow with many decision points
  - No two projects are identical
  - Comprise different experimental steps…
- **Routine experiments (purification, QA?)**
  - Performed the same way for several projects
- **High-throughput experiments (cloning?)**
  - Routine experiments using robots
  - Tracking of samples is paramount
- **One-off experiments (functional assays?)**
  - Unique to a particular project

Systems for managing information




- **Paper-based records**
  - Well suited to independent research
  - Long-term archive for independent researcher
- **Electronic Laboratory Notebooks (ELNs)**
  - Electronic version of paper systems
  - Can be shared remotely
- **Laboratory Information Management Systems (LIMS)**
  - Based on a relational database
  - Requires a model for laboratory processes
  - Snapshot of current state of laboratory

Benefits of LIMS




- **Standardization**
  - Use of controlled vocabulary and standards
  - Comparing different experiments
  - Depositing data in standardized form
- **Distributed and collaborative processes**
  - Information can be accessed anywhere
  - Different people record into same store
- **Miniaturized processes**
  - Labelling of samples becomes impossible
- **Automated and high-throughput processes**
  - Handling layouts in plates *etc.*




Potential problems with LIMS

- **IT overhead**
  - Shared data should be on backed-up server – use a web-based service
- **Data loss**
  - Hardware failure or data corruption – hardware failures manageable
- **Data integrity**
  - Data need to be entered properly or LIMS can default to being ELN
  - Recording data takes time for no immediate benefit – easy input essential
  - Unrecorded data are lost and incomplete data may break data “chain”




What is PiMS?

- **Protein Information Management System**
  - Collaborative academic effort to produce a general free-to-use fully featured LIMS
- **PiMS needs to support anyone & everyone**
  - Data for complex ever-changing workflows
  - Data need to be recorded in standardized ways
  - Many labs have similar processes
  - Most have some unique processes
- **PiMS needs to be very easy to use**
  - Schema reflects workflow: specific to one site, specific to one workflow, hard to modify
  - Schema generalizes samples and processes: flexible but potentially hard to use




Technologies used in PiMS

- **PiMS is a web application**
  - Client is Mozilla Firefox or Internet Explorer
  - No client software to install (perhaps plugins)
  - Windows, Macintosh and Linux clients
- **PiMS requires a server – or two**
  - Web server uses Apache Tomcat
  - Database server uses PostgreSQL or Oracle
  - Windows and Linux servers
  - Managed public PiMS server hosted at RAL
- **Technologies used by developers**
  - Java1.5, Hibernate, JUnit, BioJava, dot, batik, AJAX, ...

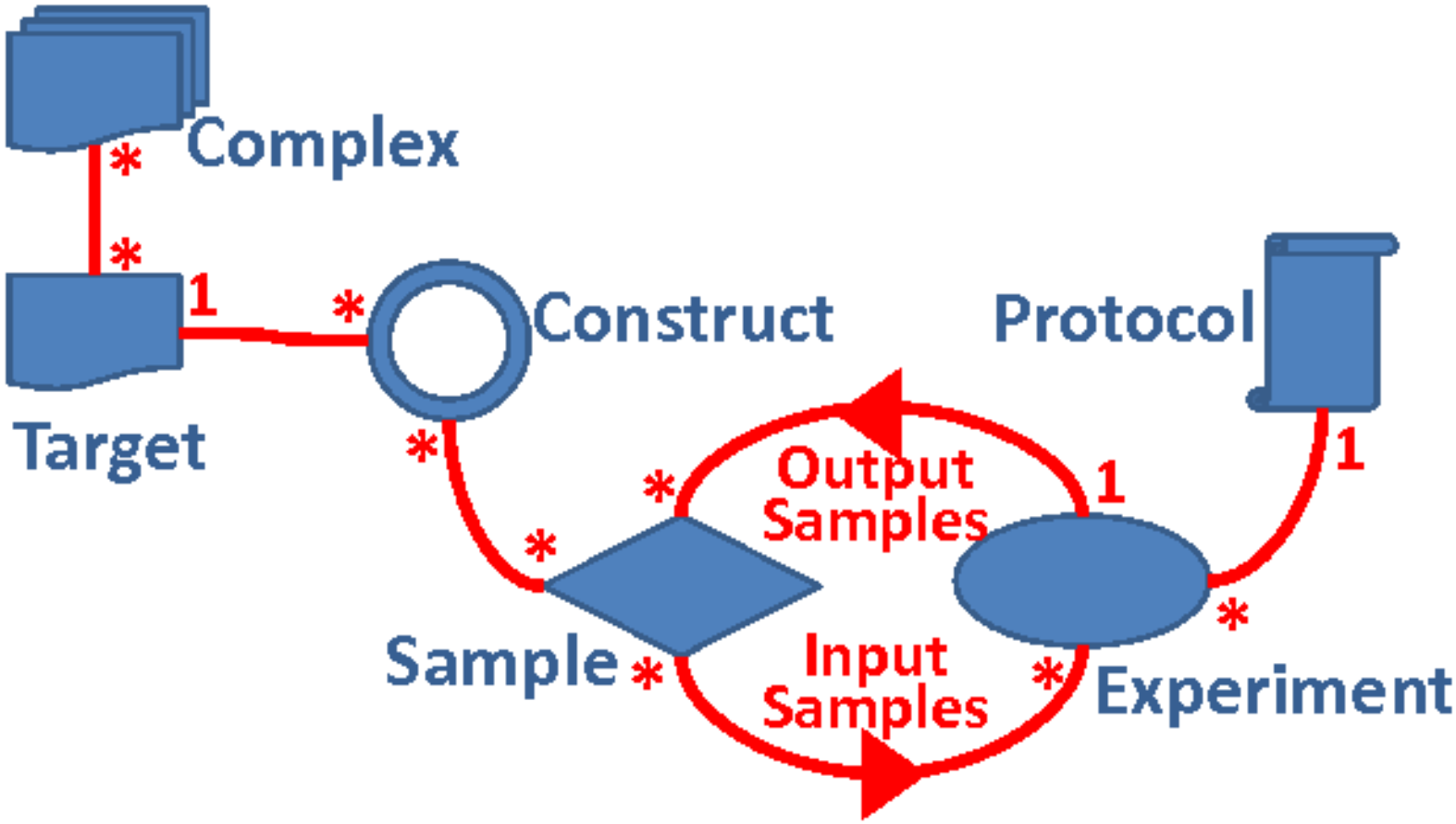
PiMS uses simple key concepts




- **Complexes & Targets**
  - Descriptions of proteins and complexes
  - Can contain bioinformatics annotations
- **Constructs**
  - Starting point for real experiments
  - Describes actual sequences
- **(Typed) Samples**
  - Tracked samples made & used by experiments
  - Samples have types, owners, locations *etc.*
- **Experiments & Protocols**
  - Consume (input) and produce (output) samples
  - Protocols are user-defined reusable templates

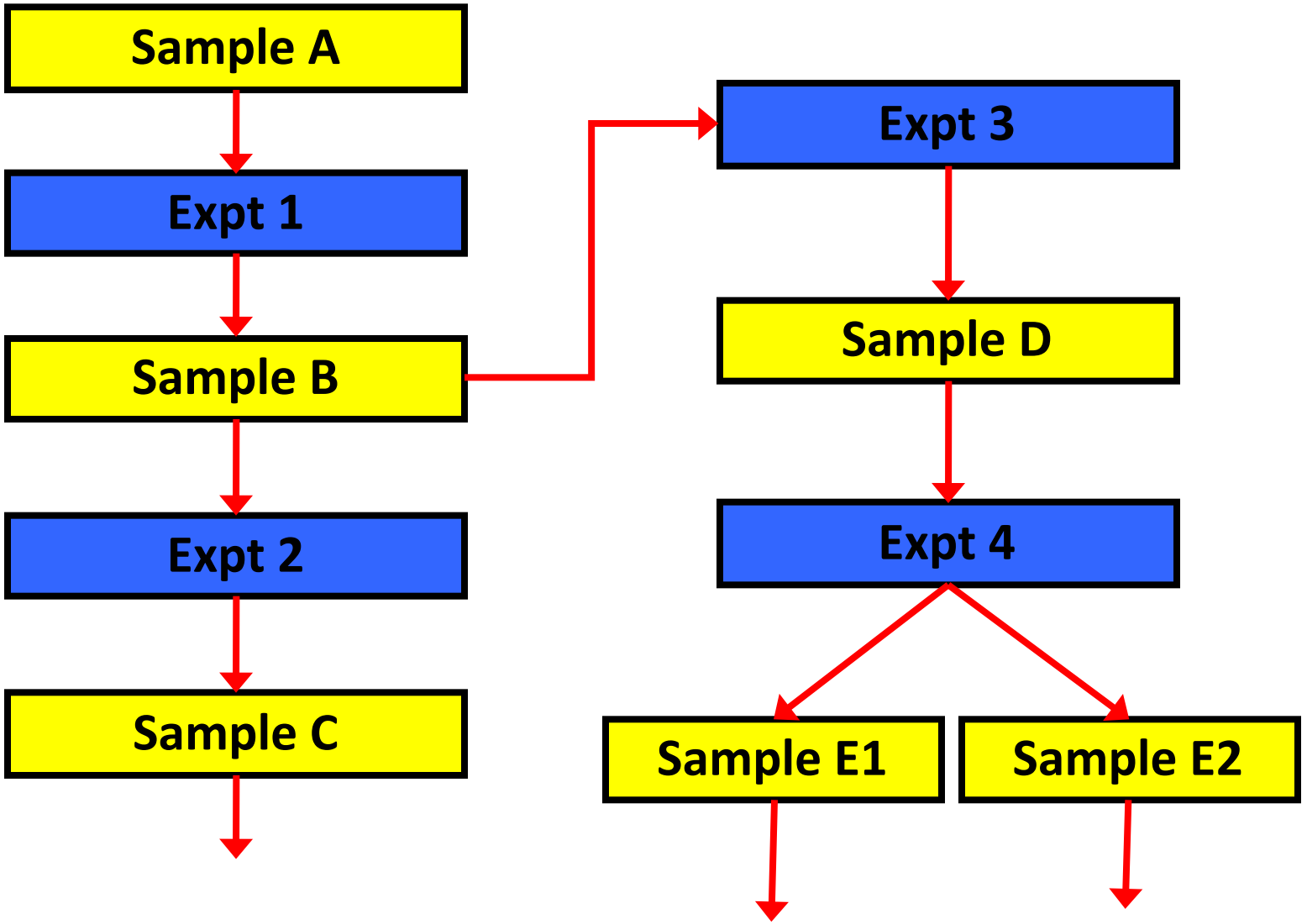
How it all links together



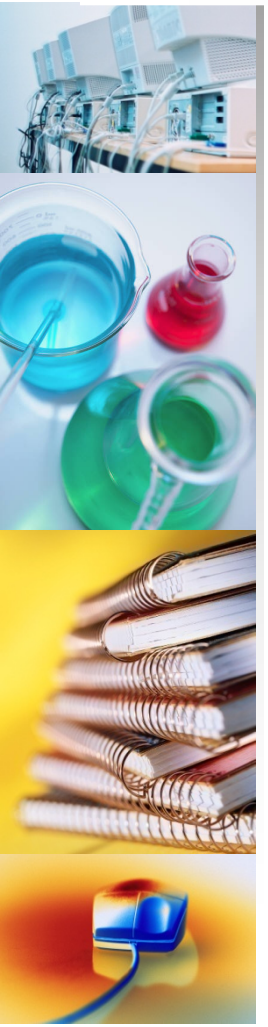



Experiments + Samples = Workflows



Basic protocols used at OPPF



PCR

PCR Clean Up

InFusion

Transformation

Plasmid Prep

Trial Expression

Verification

Sequencing

Scale Up

Lysis

Purification

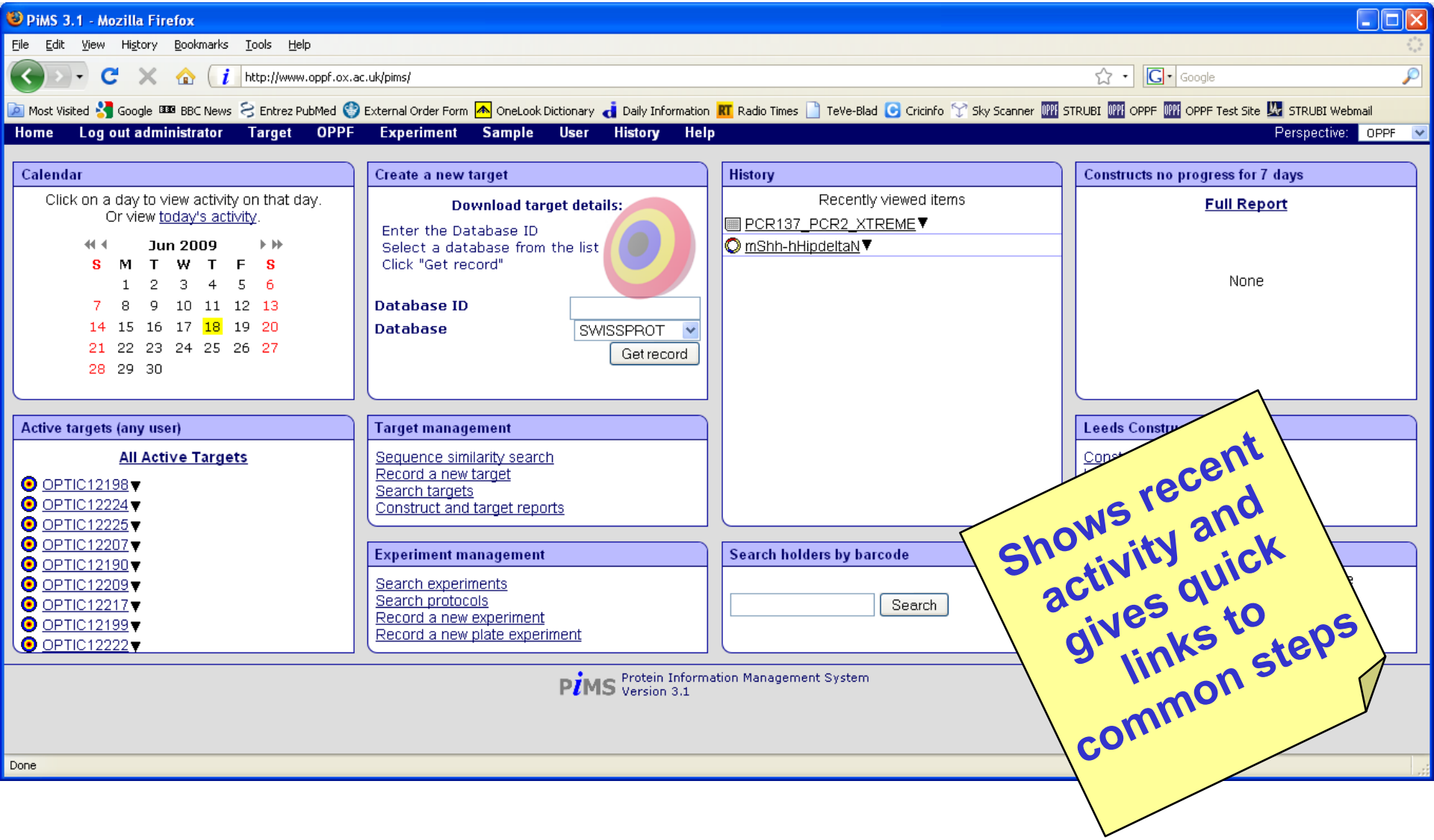
Concentration

Key:

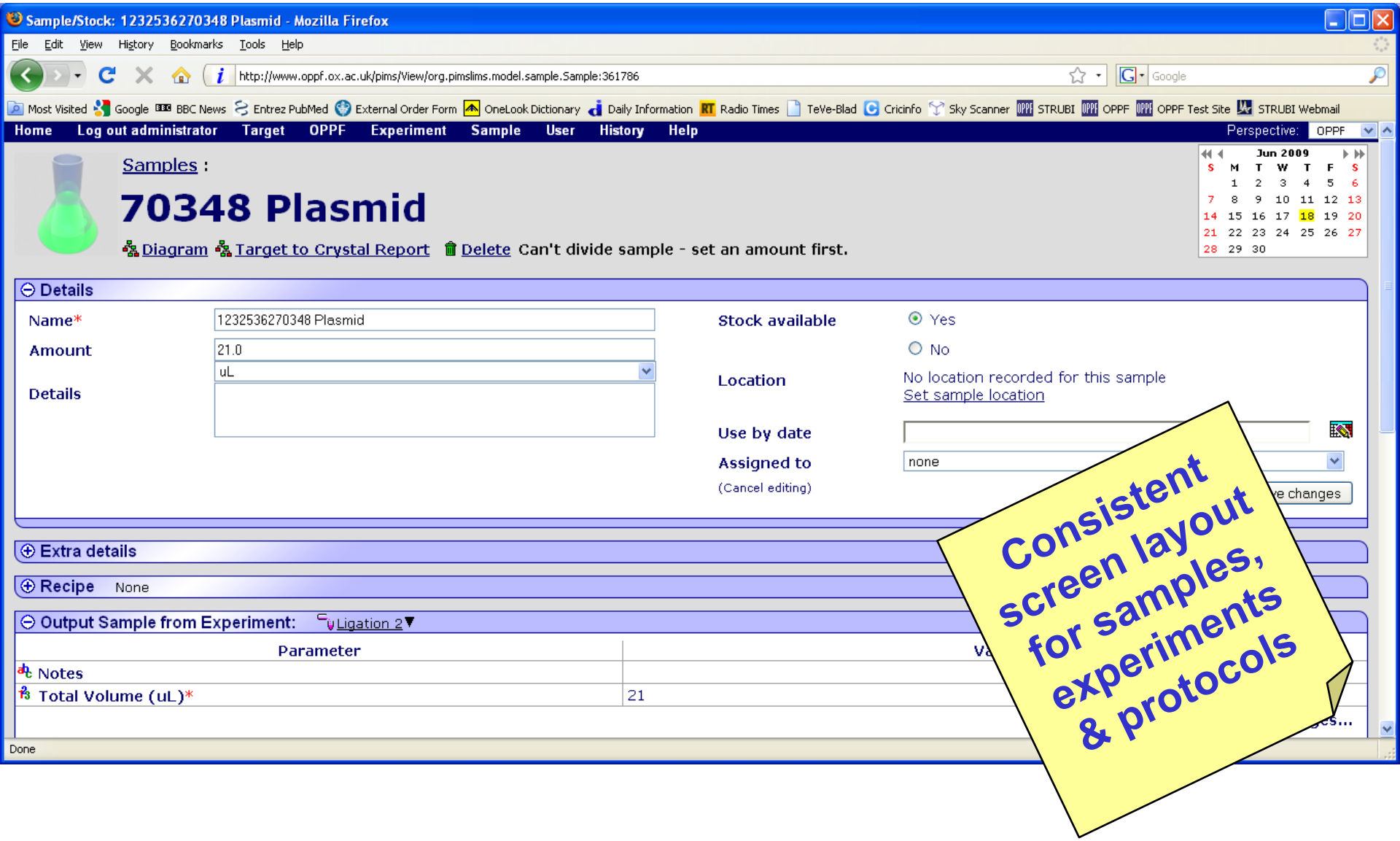
Plate Expts.

Single Expts.

The PiMS (v3.1) home page

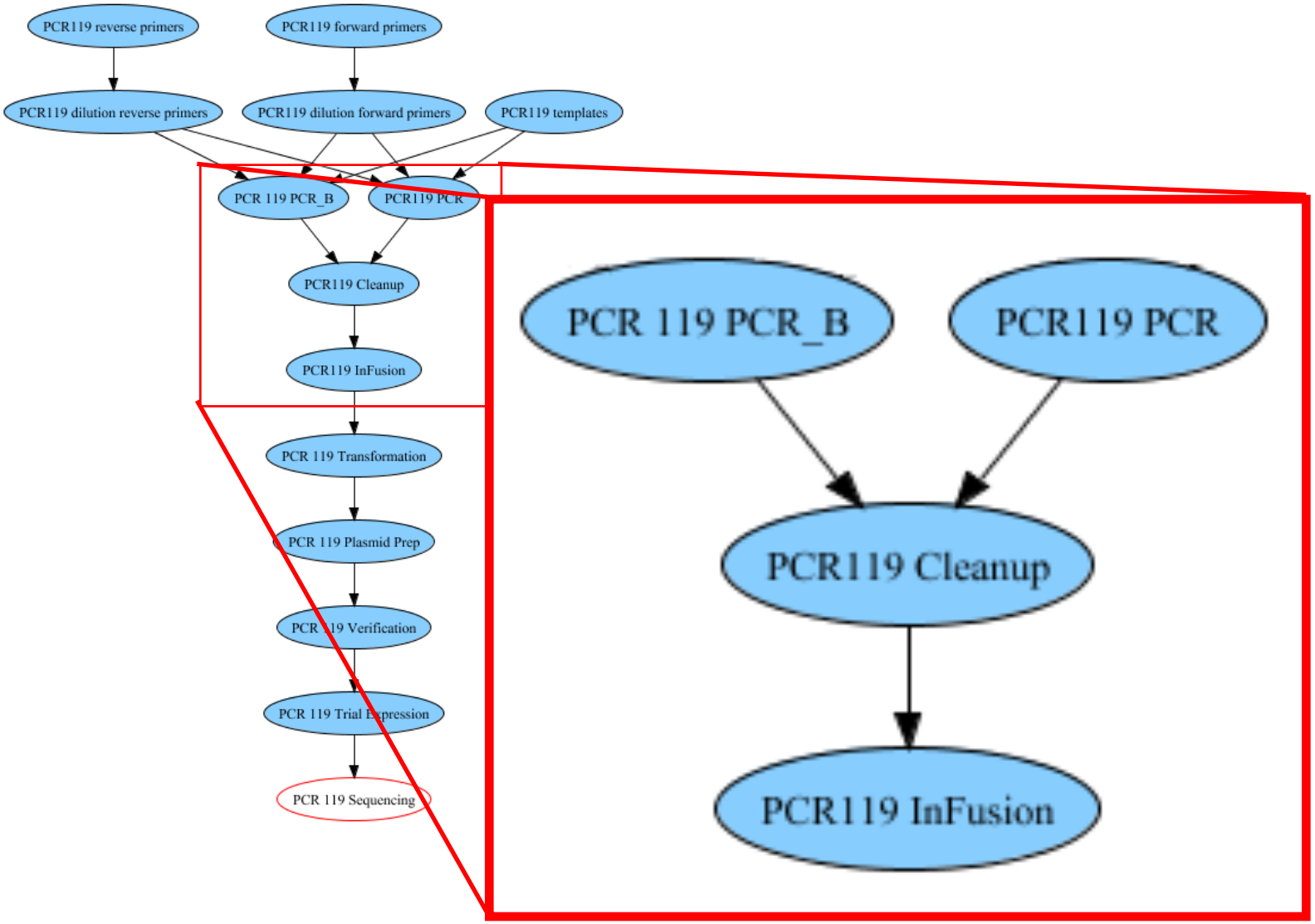


Standardized PiMS pages

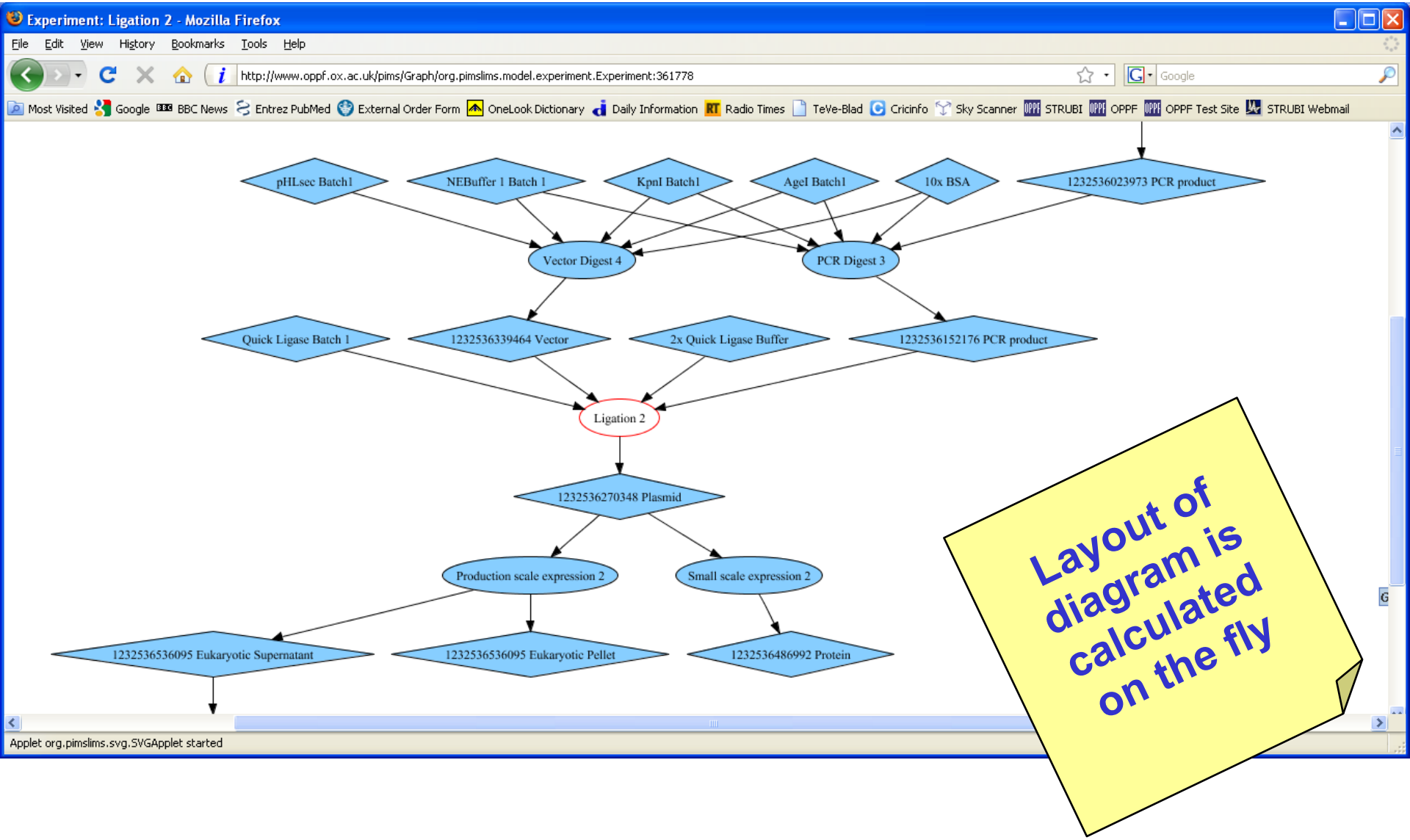


Navigating through your PiMS data: interactive diagrams



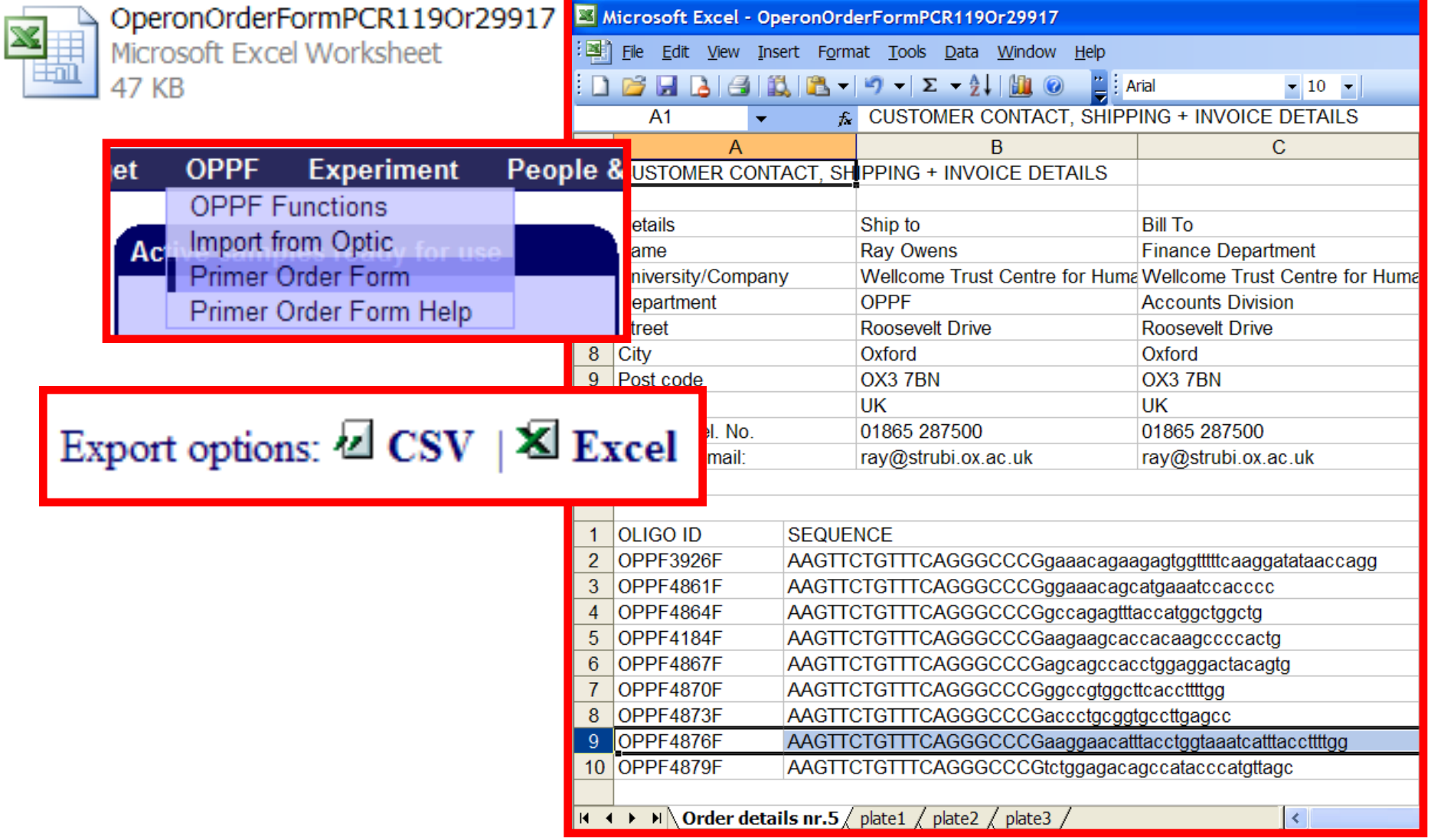



Example of a complex workflow



Layout of diagram is calculated on the fly

Experiments can read/write data

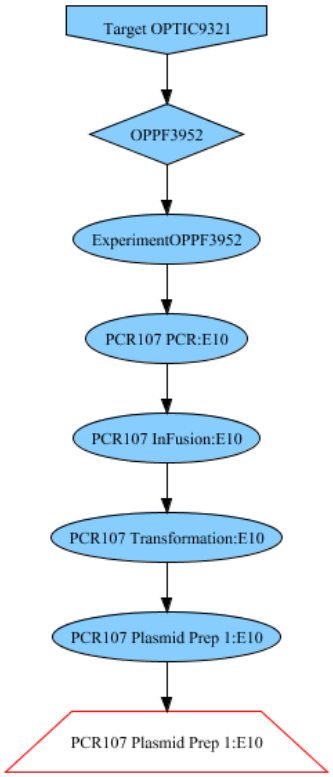


Simple reports of how samples are made

- Construct: OPPF3952
- Experiment: [illegible]erimentOPF
- Experiment: PCR107 [illegible]CR:E
- Experiment: PCR107 InFus
- Experiment: PCR107 Transfo
- Experiment: PCR107 Plasmi

Export to PDF

Export to Spine2



Experiment: ExperimentOPPF3952

Details

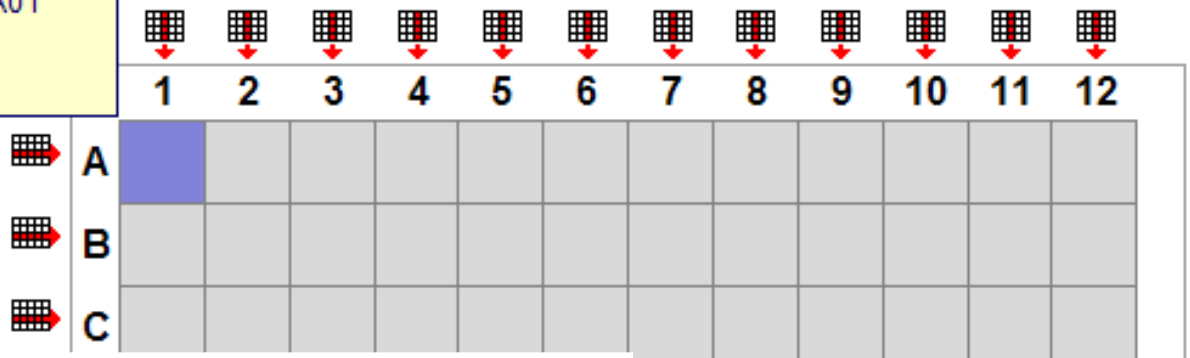
Input Samples

Parameters

Forward Primer Tm : -1.0
Reverse Overlap : 22
Forward Overlap : 31
Reverse Primer : ATGGTCTAGAAAGCTT
Forward Primer : AAGTTCTGTTTCAGGG(
Reverse Primer Tm : -1.0
Forward Tag : Opti3CInffwd
Reverse Tag : Infusion 3' site

Graphical interface for plate experiments

Plate PCR119 Cleanup, well A1
Target OPPF3926
Kit: 90.0uL, Unspecified
PCR product: 50.0uL PCR119 PCR:A01
Checked on gel?: Yes
Status: To be run

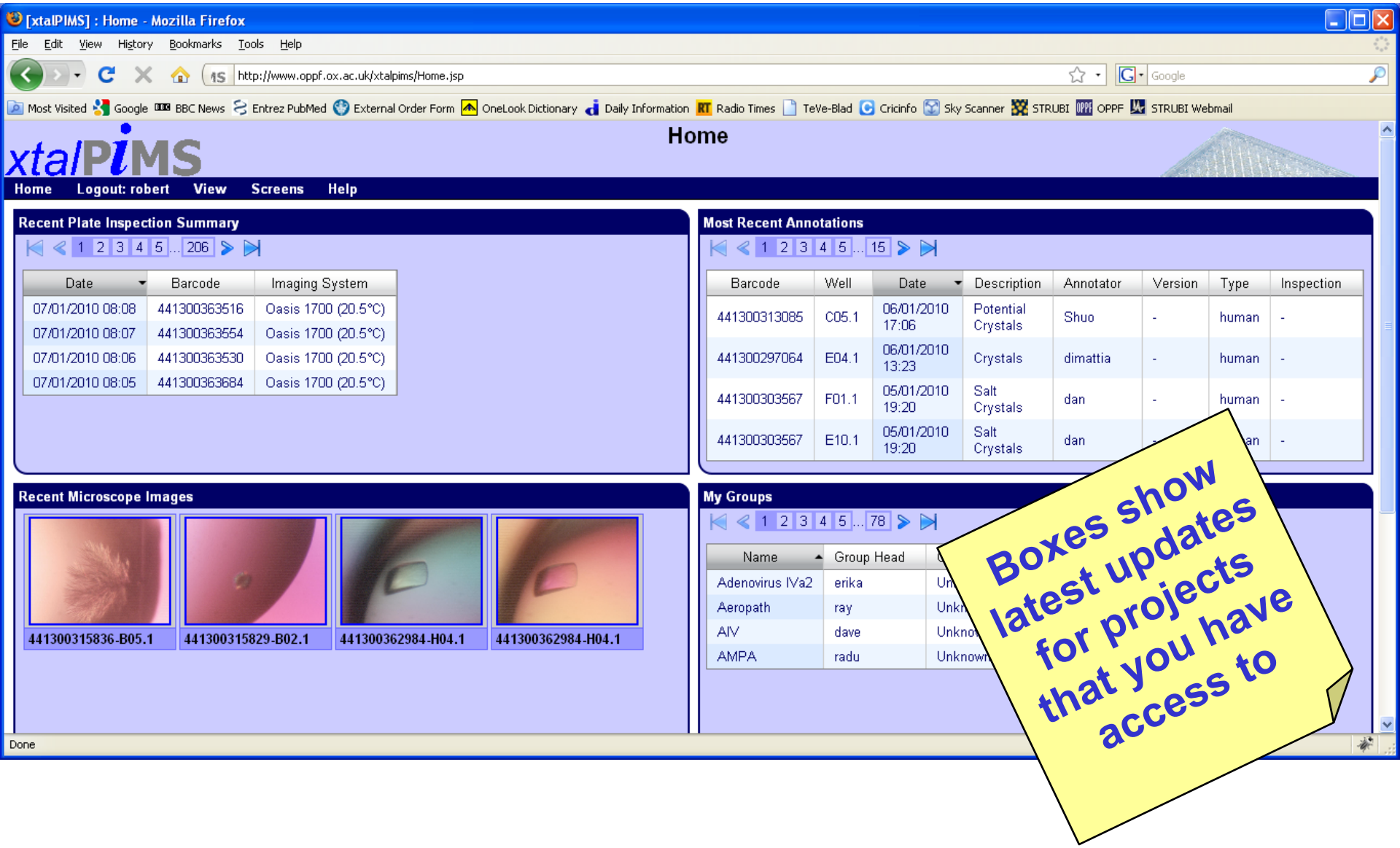


Basic details | Quick setup | Plate view | **Spreadsheet view** | Files

Tray	Row	Col	Target	Status	Output	Kit	
						Sample	Vol uL
PCR119 Cleanup	A	1	OPPF3926	To be run	PCR119 Cleanup:A01	(None)	90.0
PCR119 Cleanup	B	1	OPPF4861	To be run	PCR119 Cleanup:B01	(None)	90.0
PCR119 Cleanup	C	1	OPPF4864	To be run	PCR119 Cleanup:C01	(None)	90.0
PCR119 Cleanup	D	1	OPPF4184	To be run	PCR119 Cleanup:D01	(None)	90.0
PCR119 Cleanup	E	1	OPPF4867	To be run	PCR119 Cleanup:E01	(None)	90.0
PCR119 Cleanup	F	1	OPPF4870	To be run	PCR119 Cleanup:F01	(None)	90.0
PCR119 Cleanup	G	1	OPPF4873	To be run	PCR119 Cleanup:G01	(None)	90.0

Interface being reworked for next release of PiMS

The xtalPiMS home page



[xtalPIMS] : View Trial Drops - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.oppf.ox.ac.uk/xtalpims/ViewTrialDrops.jsp?barcode=441300356020&name=441300356020-20080814-19

Most Visited Google BBC News Entrez PubMed External Order Form OneLook Dictionary Daily Information Radio Times TeVe-Blad Cricinfo Sky Scanner

Home Logout: robert View Screens Help

Navigation - plate 441300356020

Delay: - 0.2 +

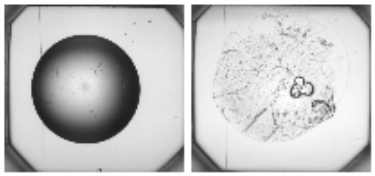
	1	2	3	4	5	6	7	8	9	10	11	12
A												
B												
C												
D												
E												
F												
G												
H												

Information - well C01.1

16/5/2009 walter Crystals

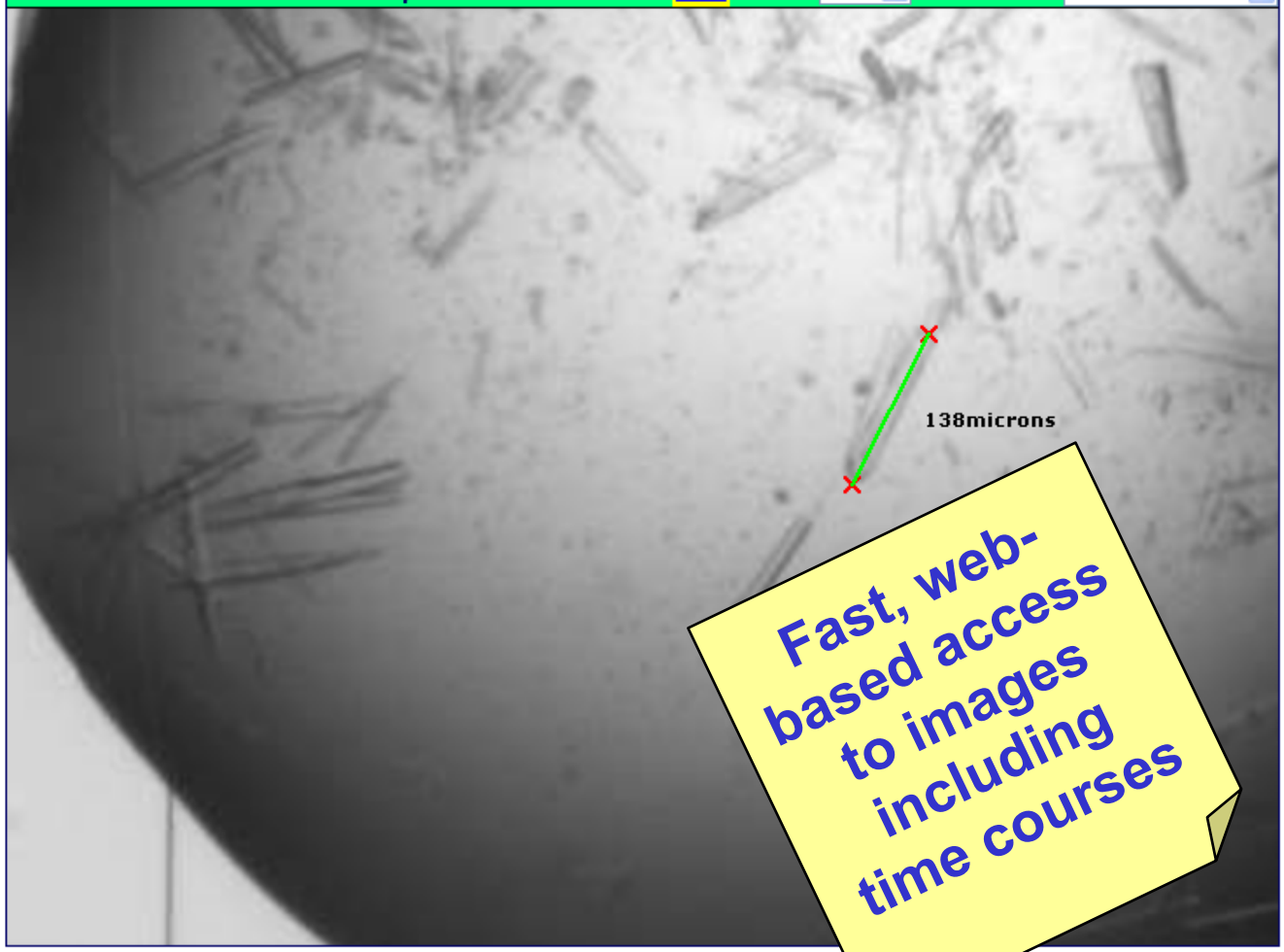
ATAD2 BROMODOMAIN#3 C-terminal HisTagged - SeMet, 9 mg/ml + DTT - Additive screen 96-well Hampton
Taken at: Thu, 14 Aug 2008 18:46:17 GMT
Taken by: Oasis 1700 (20.5°C)
Screen:
Condition: Unspecified
View Microscope Images

Earliest Latest



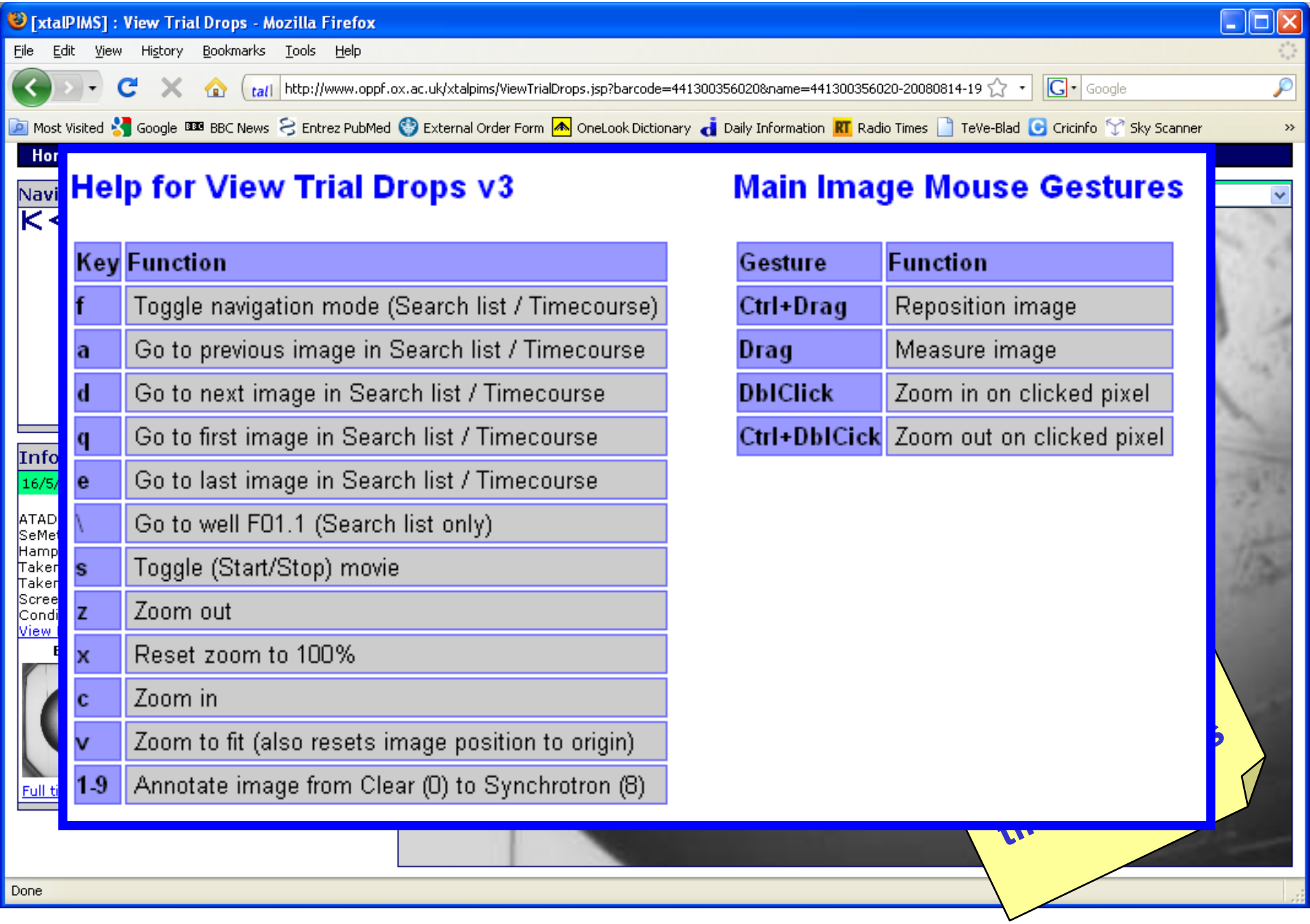
Full time course...

441300356020-C01.1 +8 days 01:07:16 PIX Zoom: 200% Annotate:



Fast, web-based access to images including time courses

Done



Help for View Trial Drops v3

Key	Function
f	Toggle navigation mode (Search list / Timecourse)
a	Go to previous image in Search list / Timecourse
d	Go to next image in Search list / Timecourse
q	Go to first image in Search list / Timecourse
e	Go to last image in Search list / Timecourse
\	Go to well F01.1 (Search list only)
s	Toggle (Start/Stop) movie
z	Zoom out
x	Reset zoom to 100%
c	Zoom in
v	Zoom to fit (also resets image position to origin)
1-9	Annotate image from Clear (0) to Synchrotron (8)

Main Image Mouse Gestures

Gesture	Function
Ctrl+Drag	Reposition image
Drag	Measure image
DblClick	Zoom in on clicked pixel
Ctrl+DblCick	Zoom out on clicked pixel

Moving along the pipeline: linking to X-ray data collection




- **Changing ways of using synchrotrons**
  - Rapid data collection means more samples
  - Expert users collect data for many projects
  - Increasing use of remote access
  - Ever more automation which needs data
- **Can PiMS and xtalPiMS help?**
  - Easy report of sample details from PiMS
  - xtalPiMS can mark crystal for data collection
  - Need simple project planning tool in xtalPiMS
  - *In situ* diffraction screening plan for xtalPiMS
  - eHTPX-style message model to get data to ISPyB *etc.*




Acknowledgments

- **The PiMS project people**
  - Kim Henrick, Dave Stuart, Keith Wilson, Colin Nave, Neil Isaacs, Jim Naismith, Richard Blake
  - Chris Morris
  - Ed Daniel, Alain Da Souza, Jon Diprose, Susy Griffiths, Bill Lin, Anne Pajon, Katya Pilicheva, Marc Savitsky, Petr Troshin, Jo van Niekerk
- **xtalPiMS (and Vault) people**
  - Ian Berry, Chris Mayo
- **PiMS and xtalPiMS users in Oxford**
  - Ben Bishop and Christian Siebold