ISPyB: an information management system for synchrotron macromolecular crystallography

Solange Delagenière[1], Patrice Brenchereau[1], Ludovic Launer[2], Alun W. Ashton[3], Ricardo Leal[1], Stéphanie Veyrier[1], José Gabadinho[1], Elspeth J. Gordon[1], Samuel D. Jones[1], Karl Erik Levik[3], Seán M. McSweeney[1], Stéphanie Monaco[1], Max Nanao[4], Darren Spruce[1], Olof Svensson[1], Martin A. Walsh[2,3,*] and Gordon A. Leonard[1,*]

[1]European Synchrotron Radiation Facility, 6 Rue Jules Horowitz, 38043 Grenoble, France, [2]MRC-France (BM14), c/o ESRF, BP 220, 38043 Grenoble, France, [3]Science Division, Diamond Light Source Ltd, Harwell Science and Innovation Campus, Didcot, Oxfordshire, OX11 0DE, United Kingdom and [4]EMBL-Grenoble, 6 Rue Jules Horowitz, 38042 Grenoble, France

Associate Editor: Jonathan Wren

ABSTRACT

Motivation: Individual research groups now analyze thousands of samples per year at synchrotron macromolecular crystallography (MX) resources. The efficient management of experimental data is thus essential if the best possible experiments are to be performed and the best possible data used in downstream processes in structure determination pipelines. Information System for Protein crystallography Beamlines (ISPyB), a Laboratory Information Management System (LIMS) with an underlying data model allowing for the integration of analyses down-stream of the data collection experiment was developed to facilitate such data management.

Results: ISPyB is now a multisite, generic LIMS for synchrotron-based MX experiments. Its initial functionality has been enhanced to include improved sample tracking and reporting of experimental protocols, the direct ranking of the diffraction characteristics of individual samples and the archiving of raw data and results from ancillary experiments and post-experiment data processing protocols. This latter feature paves the way for ISPyB to play a central role in future macromolecular structure solution pipelines and validates the application of the approach used in ISPyB to other experimental techniques, such as biological solution Small Angle X-ray Scattering and spectroscopy, which have similar sample tracking and data handling requirements.

Contact: leonard@esrf.fr; martin.walsh@diamond.ac.uk

Received on April 5, 2011; revised on September 16, 2011; accepted on September 19, 2011

1 INTRODUCTION

The advances made in bioinformatics, protein production, purification and crystallization techniques when the era of structural genomics was at its height led to a greatly increased demand for the collection of X-ray diffraction data at synchrotron-based macromolecular crystallography (MX) beamlines. This, coupled with the advent of user-friendly synchrotron beamline control interfaces (Gabadinho *et al.*, 2010; McPhillips *et al.*, 2002; Skinner *et al.*, 2006), the increased use of robotic sample changers (Cipriani *et al.*, 2006; Cohen *et al.*, 2002; Cork *et al.*, 2006; Hiraki *et al.*, 2008; Jacquamet *et al.*, 2009; Snell *et al.*, 2004) and the ability to automatically screen and rank the diffraction properties of crystals using software such as DNA (Leslie *et al.*, 2002) and WEB-ICE (Gonzalez *et al.*, 2008) ushered in the age of high-throughput crystallography (HTX).

The most evident effect of HTX is that in the period 2000–2009, there was a 3-fold increase in the annual rate of deposition of macromolecular crystal structures in the Protein Data Bank (PDB; http://www.pdb.org) (Berman *et al.*, 2000). The overwhelming majority (~85%) of these depositions are based on diffraction data collected at synchrotron facilities (see http://biosync.sbkb.org/ for details). However, the increase in PDB depositions tells us only half of the story. During the same time period, the number of individual samples being analyzed per year at the MX facilities of the European Synchrotron Radiation Facility (ESRF) rose by more than an order of magnitude (currently >100 000/year at the ESRF) (Bowler *et al.*, 2010), an increase significantly boosted by the emergence of fragment-based drug design as a tool for drug discovery in the pharmaceutical industry (see Murray and Blundell, 2010 for a review). Increased automation has also meant that many more samples are being tested for diffraction quality before any full data collections are carried out (Bowler *et al.*, 2010).

Given the huge increase in the numbers of samples analyzed and the increased importance of the pre-screening of the samples, the need for a Laboratory Information Management System (LIMS) combining sample tracking and experiment reporting during synchrotron-based MX experiments was clear. In 2001, a prototype LIMS, PXWeb, was developed and deployed at the ESRF (Arzt *et al.*, 2005). The scope of PXWeb was rather limited and functionality was confined mainly to basic experiment recording and reporting. In particular, no exchange of information between PXWeb and other LIMS [i.e. for protein production (Pajon *et al.*, 2005), crystallization (Amin *et al.*, 2006; Fulton *et al.*, 2004), X-ray

*To whom correspondence should be addressed.

crystallography (Harris and Jones, 2002)] or 'electronic lab books' (Prilusky *et al.*, 2005) was possible.

Subsequent to the initial deployment of PXWeb at the ESRF, an e-science resource for MX, e-HTPX, was initiated (Allan *et al.*, 2005). The aims of e-HTPX were to provide a unified set of tools for structural biologists to plan, direct and document experiments in a site-independent manner and to afford seamless flow and storage of information, from target selection to structure deposition in the PDB, for each sample under study. As initial use of PXWeb at the ESRF illustrated the need to increase the scope of the database and highlighted restrictions imposed by the Zope (http://www.zope.org/) technology used to underpin its web interface, a joint ESRF/e-HTPX initiative was launched to reassess both the Use Cases deployed and the technology upon which PXWeb was based. This resulted in the development of an upgraded LIMS, Information System for Protein crystallography Beamlines (ISPyB), (Beteva *et al.*, 2006). Compared with PXWeb, the initial version of ISPyB provided more sophisticated sample tracking and experiment recording features and facilitated the transmission of information from/to other LIMS. In particular, enhanced communication with the ESRF beamline control Graphical User Interface (GUI) MxCuBE (Gabadinho *et al.*, 2010) and an improved harvesting and display of results (including data collection strategies) made experiment planning and execution easier.

The development of ISPyB is currently a collaboration between ESRF and Diamond Light Source (DLS). This collaboration has produced a multisite, generic LIMS for synchrotron-based MX experiments. Additional features beyond those reported by Beteva *et al.* (2006) include:

(i) extended reporting of experimental/beamline parameters,
(ii) tracking of samples to/from the home laboratory/synchrotron.
(iii) snapshots showing the disposition of a sample in the X-ray beam.
(iv) logging and display of results from the analysis of absorption edge scans and X-Ray fluorescence emission (XFE) spectra.
(v) storage and ranking of the diffraction characteristics of individual samples and (vi) archiving, display and download of the results of the automatic processing of full diffraction data sets collected (c.f. Winter, 2010).

Taken together, this functionality allows the tracking of the progress of a diffraction experiment in real-time and paves the way for ISPyB to play a central role in future macromolecular structure solution pipelines.

Access to ISPyB, its basic functionality and its role in the semi-automated data-collection pipeline developed under the auspices of the Structural Proteomics in Europe (SPINE; http://www.spineurope.org) project have been reported elsewhere (Beteva *et al.*, 2006). Here, we take the opportunity to illustrate the current status and functionality of ISPyB, to elucidate its underlying data model and data flow and to describe its means of communication with other LIMS and applications external to the database.

2 METHODS

ISPyB is written in Java J2EE and uses the Jboss Application Server (http://jboss.org/jbossas) and Struts Model-View-Controller (MVC) Web Application Framework (http://struts.apache.org/). The End-User interface of

ISPyB is a pure web-based application that is accessible using any browser. At the ESRF, the ISPyB database is MySQL (http://www.mysql.com/), whereas, at DLS, the database is Oracle (http://www.oracle.com/). Communication with other databases is performed via web services (see below).

The ISPyB data model (Fig. 1) was designed to be modular so that data management and storage pertaining to new functionality can easily be added as and when it becomes available. As experimental work at synchrotron facilities is carried out in pre-defined time slots on pre-defined beamlines, the ISPyB data model is based on the concept of an individual experimental session (an 'Experiment' at ESRF, a 'Visit' at DLS). Data concerning the scheduling of experimental sessions are retrieved from the appropriate synchrotron User Office management databases which are external to ISPyB (Fig. 2).

Once an 'experiment' (or 'visit') has been scheduled, one, or more, 'shipments' can be associated with the session. A shipment consists of one or more 'dewars' each comprising one or more 'containers' in
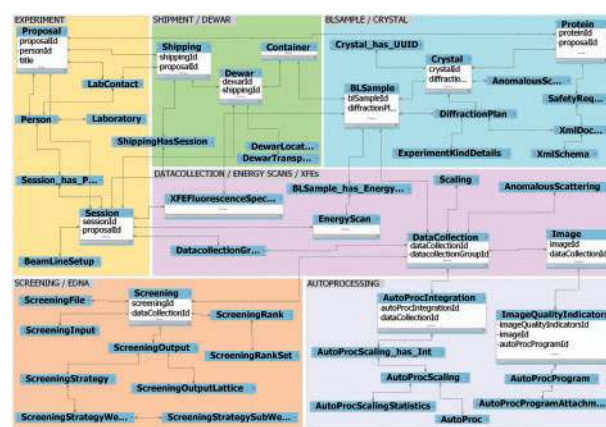

Fig. 1. A schematic representation of the ISPyB data model. Both its 'experiment/shipment/container/sample/data collection' hierarchy and its modularity are shown. The different modules are labeled and shown with different background colors. Data tables are shown as labeled boxes with links between them shown as dashed lines. For clarity, the full designation of a data table is not always given and not all entries for all data tables are shown.
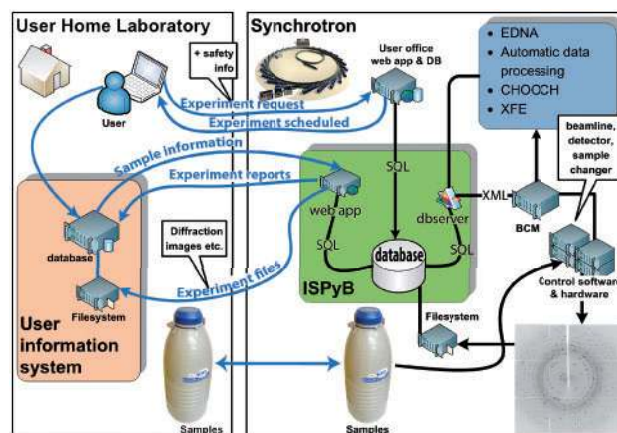

Fig. 2. The data flow and application environment of ISPyB. ISPyB (database, web application, dbserver) is highlighted in green, the User Home Laboratory information system in pink and experiment software external to ISPyB in blue. The means of communication between each component is described in the main text.

which the individual 'samples' to be studied in the experimental session are held. 'Containers' are most commonly cryocanes or sample changer baskets. Individual sample identifiers are constructed using a combination of user-defined sample names and protein acronyms. At the ESRF, protein acronyms are contained in its User Office Database (Scientific Management Information System; SMIS) and are linked to information (protein name, organism, safety class, etc.) provided during the beamtime application process.

During an 'experiment/visit' ISPyB records, and makes available immediately, all relevant experimental parameters for all data collections and stores these along with the results of automatic sample evaluation or post-operation data analysis. 'Data collections', each given an individual identifier (DataCollectionId) are referred to an individual sample. Several types of data collection are possible: initial diffraction quality screening, full diffraction data collections, absorption edge scanning and the collection of XFE spectra. Depending on the data collection operation carried out, a variety of different information and data are stored in the database.

The data flow and application environment of ISPyB are shown in Figure 2. The recording of experimental and beamline parameters and the archiving of the results produced by software external to the database is performed via the ISPyB Database Gateway Server (dbserver). Developed to provide a simple, uniform and standard interface to ISPyB for trusted client applications, this gateway is a lightweight Python HTTP server communicating in Extensible Markup Language (XML) with its clients and in Structured Query Language (SQL) with the ISPyB database. Client applications include the Beamline Control Module (BCM) and any downstream data processing software executed on the command of the BCM. Typically, the client application will send a request consisting of a command and associated XML, and the Database Gateway Server will return an XML response.

Communication between ISPyB and external LIMS or other databases is achieved using web services based on the Simple Object Access Protocol (SOAP; http://www.w3.org/TR/2007/REC-soap12-part1-20070427). Using this technology, web services are simply an extension to an existing application and do not require any particular configuration or firewall exception.

Web services facilitating information exchange between ISPyB and external systems are implemented using a 'message model'-based set of XML Schema Definitions (XSDs), developed within e-HTPX (Allan *et al.*, 2005), that define what data can be exchanged. The same software architecture forms the basis for an Application Programming Interface (API) which allows the submission of information (i.e. sample and shipment information for a particular experiment) through upload of a standard spreadsheet.

The increased demand for the screening or collection of complete diffraction data sets from crystals in their crystallization drops (Emamzadah *et al.*, 2009; Jacquamet *et al.*, 2004), has driven the tighter integration of ISPyB and external crystallization LIMS. Specific exchange mechanisms have been developed for the EMBL Grenoble Crystallization Information Management System (CRIMS, https://embl.fr/htxlab/) and the Protein Information Management System (PiMS) (Morris *et al.*, 2011) compliant xtalPiMS application (Daniel *et al.*, 2011). Since PiMS and xtalPiMS reference the same data model and can even reference the same database, these two applications can, in tandem with ISPyB, allow users to track data pertaining to a particular sample from initial cloning experiments through protein production and crystallization to diffraction data collection and beyond.

3 RESULTS

3.1 Sample management

Sample management in ISPyB involves the tracking of samples to/from and within the synchrotron and the recording of their

progress, including data logging, through the data collection experiment.

Once a shipment has been created in ISPyB, transport labels including unique barcodes are generated automatically for all the individual components (currently, these are typically transport dewars) of the shipment. These labels contain details of the experimental session to which the shipment is assigned specifically—date, beamline and experiment Local Contact (LC). The component can also be linked, by the user, with a courier company tracking number which is then also displayed in ISPyB. When the labeled component arrives at the synchrotron, the label barcode is scanned, the status of the component in ISPyB automatically updated and an e-mail confirming its reception sent to both the dispatcher and to the LC for the experiment. When the component is delivered to the beamline on which the experiment is to take place, the barcode is again scanned and the status of the component updated in ISPyB. After the experiment, the barcode of the return label is scanned and linked to a courier company tracking number, an e-mail containing a hyperlink to this sent to the user and the status of the component updated in ISPyB. External users are therefore informed of the arrival/departure of shipments, of their exact whereabouts while they are at the synchrotron and are able to track their progress between the home laboratory and the synchrotron. Given the 'experiment/shipment/container/sample' hierarchy of the ISPyB data model (Fig. 1) this feature of the database also allows the tracking of an individual sample's location at any given time. Dewar barcodes are read using Hand Held Dolphin 7600 barcode scanners (Honeywell Scanning and Mobility, USA) that run a modified instance of the ISPyB web interface and transmit information to the ISPyB database via wireless networks.

As described previously (Beteva *et al.*, 2006), various information can be associated with the individual samples in a shipment. Such information can include:

(i) sample container barcode (see Section 2 for a definition of 'containers').
(ii) individual sample holder barcode.
(iii) position in the sample container.
(iv) unit-cell dimensions and space group.
(v) rudimentary experiment plan—this usually consists of the minimum data resolution, as predicted from preliminary screening, required for a full data collection to be carried out.
(vi) experiment type (single wavelength/multiwavelength, direct/inverse beam, etc.).
(vii) additional comments deemed useful for the sample.

Uploading of sample information can be carried out by directly typing it into the ISPyB interface; by uploading information contained in an Excel spreadsheet or files in comma-separated values (CSV) format; by transferring it, via web services, from an external data base (see above); (at DLS only) by direct input into the GUI of the BCM.

At the start of an experimental session, either the 'prepare experiment' tab of the ISPyB interface or (at the ESRF only) the *in situ* scanning (Cipriani *et al.*, 2006) of the barcodes of each SPINE sample holder is used to define the position of individual samples in the sample changer. When the 'prepare experiment' tab is used, only the position of a container in the sample changer is determined at the beamline. The positions of individual samples in the container are

read from ISPyB. Sample name and sample changer positions are then passed to the sample area in the GUI of the BCM along with any relevant information (space group, unit-cell dimensions, minimum resolution, etc.). If desired, the sample names can then be used to provide the prefixes in the naming conventions for subsequent data collection operations.

The sample tracking protocol described above cannot currently be used for non-frozen crystals brought to the synchrotron in crystallization plates. However, as for frozen crystals shipped to the synchrotron in a transport dewar, information pertaining to each individual sample can be entered into ISPyB via its 'Sample' tab. ISPyB then associates all experimental details and results with a given sample in the usual way. In the case that no sample information is entered into ISPyB prior to a data collection operation, the 'image prefix' given in the GUI of the BCM is used for the association of experimental details and results.

3.2 Experiment recording and planning

3.2.1 Experiment recording. Once a user has logged in to a BCM connected to ISPyB, details of all data collection operations carried out during the current experimental session are collected, archived and made available in real-time via the ISPyB 'Data Collection' tab (Fig. 3).

For each diffraction data collection carried out, a top level summary displays diffraction image names, basic experimental information and a summary of the results (unit-cell dimensions, space group, completeness, resolution and quality) of automatic data processing. For absorption edge scans, both the raw data for the scan itself and the results of the Kramers–Kronig transformation as implemented in CHOOCH (Evans and Pettifer, 2001) are accessible (Fig. 3). For XFE spectra (Leonard *et al.*, 2009), both the spectrum itself (Fig. 3) and a full analysis of this are retrievable, the latter as a Hypertext Markup Language (HTML) page.

Lower level information is accessible by clicking on the relevant 'image prefix' in the 'Data Collection' tab. This information includes optional crystal snapshots (Fig. 4), information tables giving fuller details of experimental and beamline parameters and the results and log files of any crystal characterizations carried out (see below). When automatic data processing protocols are successful, log files and files containing the resulting scaled and merged intensity data can also be perused or downloaded (Fig. 3).

At both the ESRF and DLS, diffraction images are stored on a central file system from which they are deleted periodically. In order to restrict the space requirements of ISPyB, only Joint Photographic Experts Group (JPEG; http://www.jpeg.org) format thumbnails of diffraction images, and not the images themselves, are stored indefinitely. Diffraction images can thus only be downloaded via ISPyB while they remain in central storage.

3.2.2 Experiment planning. As well as providing the extended reporting of experiments carried out at synchrotron-based MX facilities, the information stored in ISPyB also allows for real-time (i.e. at the beamline) experiment planning.

Different crystals of the 'same' biological macromolecule often show large variations in the quality of the diffraction data that can be obtained from them (see Bowler *et al.*, 2010 for a brief discussion as to the reasons for this). In order to choose the best crystal from which to collect diffraction data and thus, to increase the likelihood

that an experiment will produce a definitive result, preliminary investigations of crystal quality are increasingly carried out in MX. Such experiments at the ESRF and DLS currently use EDNA (Incardona *et al.*, 2009). Experimental details of characterizations and the results of these (including proposed data collection strategies) are stored in ISPyB (Fig. 4). For each individual EDNA characterization, ISPyB provides snapshots showing the disposition of the sample in the X-ray beam, details of the autoindexing and integration of the test diffraction images, a series of test image quality indicators and the results of data collection strategy calculations carried out using BEST (Bourenkov and Popov, 2010). When preliminary characterization is carried out successfully, the resulting data collection plan is also provided to the BCM and can be acted upon at the experimenter's discretion.

The storage of the results of preliminary diffraction tests means that, in contrast to the semi-automated data collection pipeline described by Beteva and colleagues (2006), ISPyB can now be used to directly rank the diffraction quality of a set of samples. Ranking can be carried out for all samples (regardless of sample name root) within a single session or, via a database search, across experimental sessions for samples with the same protein acronym. Initial ranking is carried out via the 'Data Collection' tab (see 'Sample Ranking' checkboxes, Fig. 3 middle panel) and results presented either as a table or as a radar comparison chart (Fig. 5). If desired, further ranking can then be carried out interactively using different weights for different ranking criteria. After ranking, the data collection strategies for selected samples can be exported to an XML format file that can be subsequently read and used by the BCM to drive full data collections for the samples chosen.

3.3 Searching and reporting

For each data collection session, reports can be produced in formats (i.e. CSV) that allow users to update any home laboratory LIMS they might use or as a PDF or word document.

Basic database searching functionality is available through the 'Data Collection' and 'Sample' tabs of the ISPyB interface. Searches can be carried out, in particular, using the protein acronym as a parameter and will return all information stored for individual samples whose names contain this acronym, including whether or not data collections have been performed. The results of all searches can be exported as CSV files, a PDF file or a word document. Generic dashboards allow the display of statistics relating to beamline use. Currently, this functionality is rather limited and, when necessary, specific bespoke searching of ISPyB by database administrators can be carried out using SQL requests. At DLS tools supplied with the Oracle system are exploited to generate web page reports of both beamtime usage and of data quality statistics derived from automatic data processing. These can be copied into spreadsheet packages for detailed analysis and presentation.

4 DISCUSSION

ISPyB was designed to optimize information transfer between home laboratories and synchrotron-based MX beamlines before, during and after experimental sessions. The particular strength of ISPyB relates to sample management and the recording of information during an experiment. However, when properly used, the functionality of ISPyB should also allow for more successful
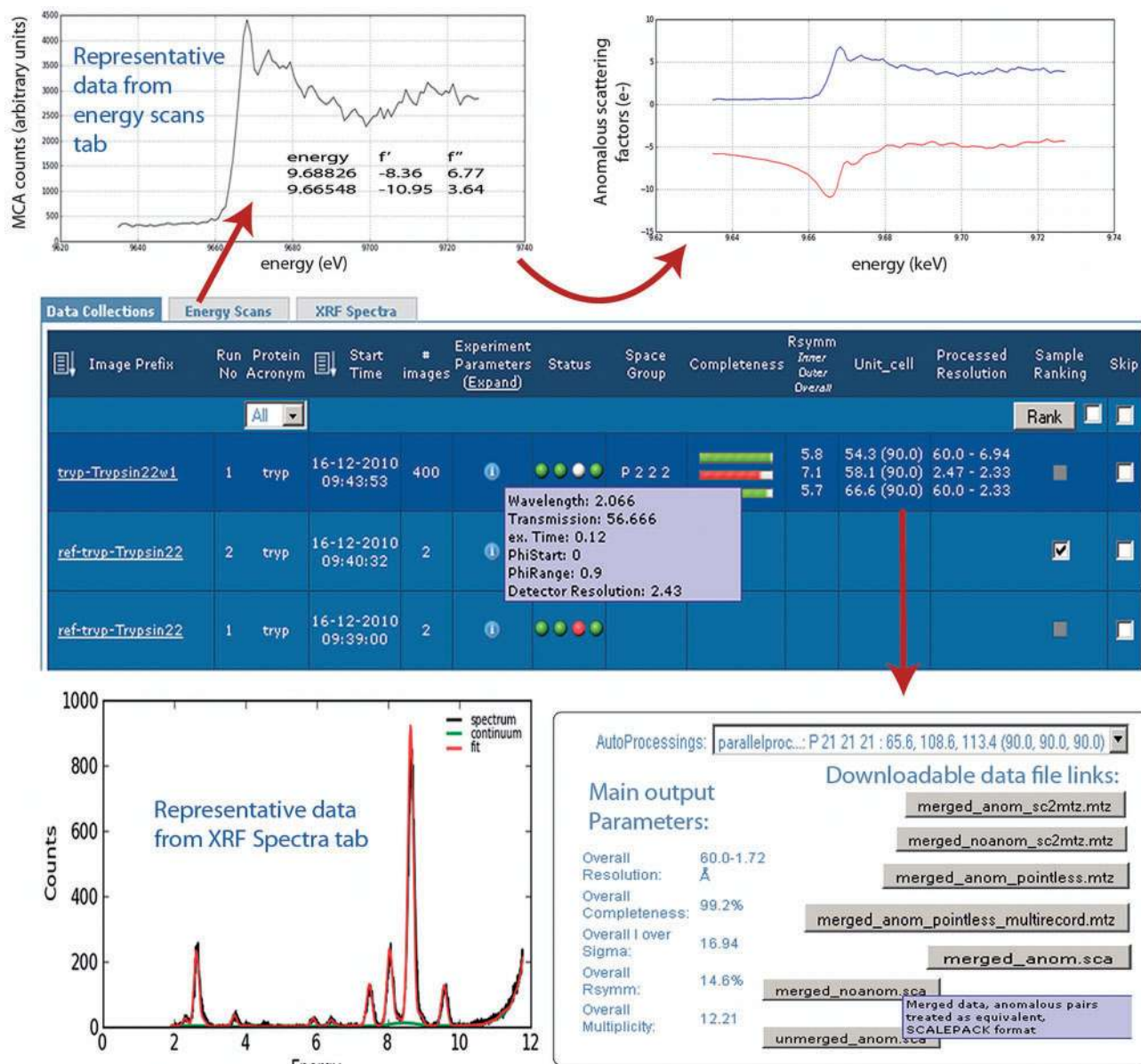
Fig. 3. Experiment recording in ISPyB. The 'Data Collection' tab (middle panel) of the ISPyB web interface provides a top level summary for each data collection, as well as access to raw data and results for absorption edge scans (top panel) and XFE spectra (bottom left). Clicking on the relevant 'image prefix' directs the user to a full reporting of the experimental protocols used, the results of sample evaluations by EDNA (Fig. 4) and the results of automatic data processing (bottom right). Mouse-over tool-tips are integrated into the HTML interface. In the examples above the tool-tip in the 'Data Collection' tab shows a summary of the data collection parameters for a selected experiment while that for the data processing results gives information concerning both the format and data content of downloadable files containing the reduced data.

and efficient experimental sessions. The possibility to rank the diffraction characteristics of samples means that, in principle, full data collections should only be carried out on the samples best suited to the experiment in hand. Moreover, the accessibility of results and output files from the automatic, on-line data processing protocols means that experimenters can also monitor the success of data collection experiments in real-time (i.e. while still at the beamline) and decide whether or not more data need to be collected for a specific project.

ISPyB is a useful tool for those who operate synchrotron-based MX facilities. It allows the generation of statistics on what type of experiments are being carried out and, thus, the elucidation of new trends that may drive improvements in beamline functionality. A specific example of this is the trend that many more samples are tested for diffraction quality before any full data collections are carried out. This observation has lead to the design of new beamlines and workflows for MX that will be implemented as part of the ESRF Upgrade
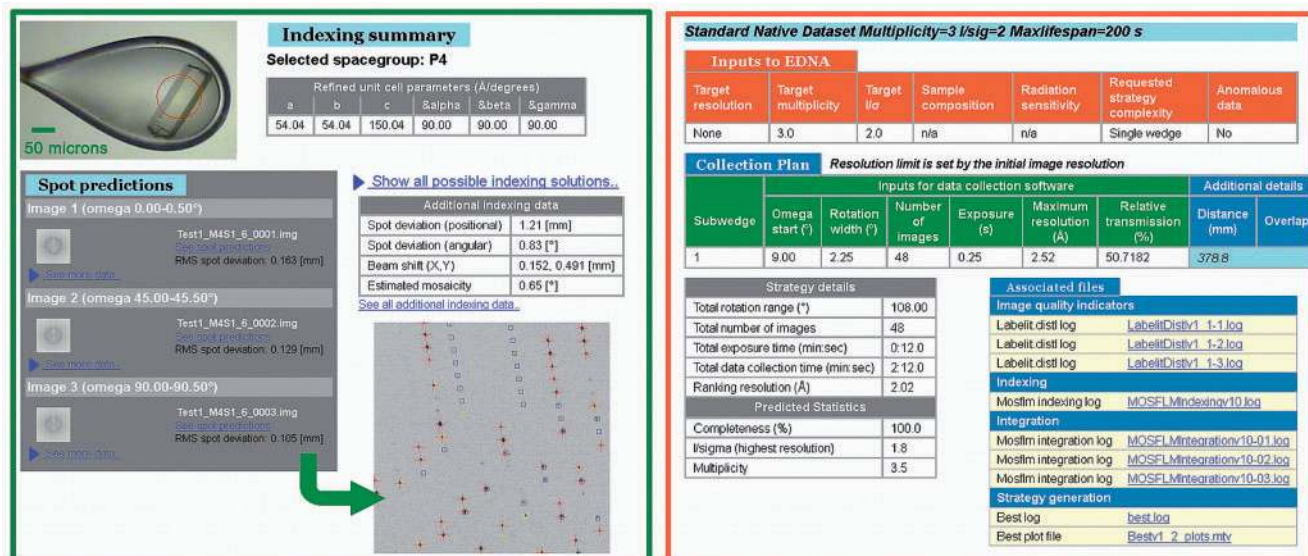
Fig. 4. MX experiment planning in ISPyB. Top left: a snapshot showing the disposition of the sample in the X-ray beam (indicated by the red circle) and details of the autoindexing of test diffraction images using EDNA. Right: the data collection plan, based on the EDNA characterization, proposed by BEST (Bourenkov and Popov, 2010) showing a summary of both the experimental protocol to be used and the results that the execution of the protocol should produce.
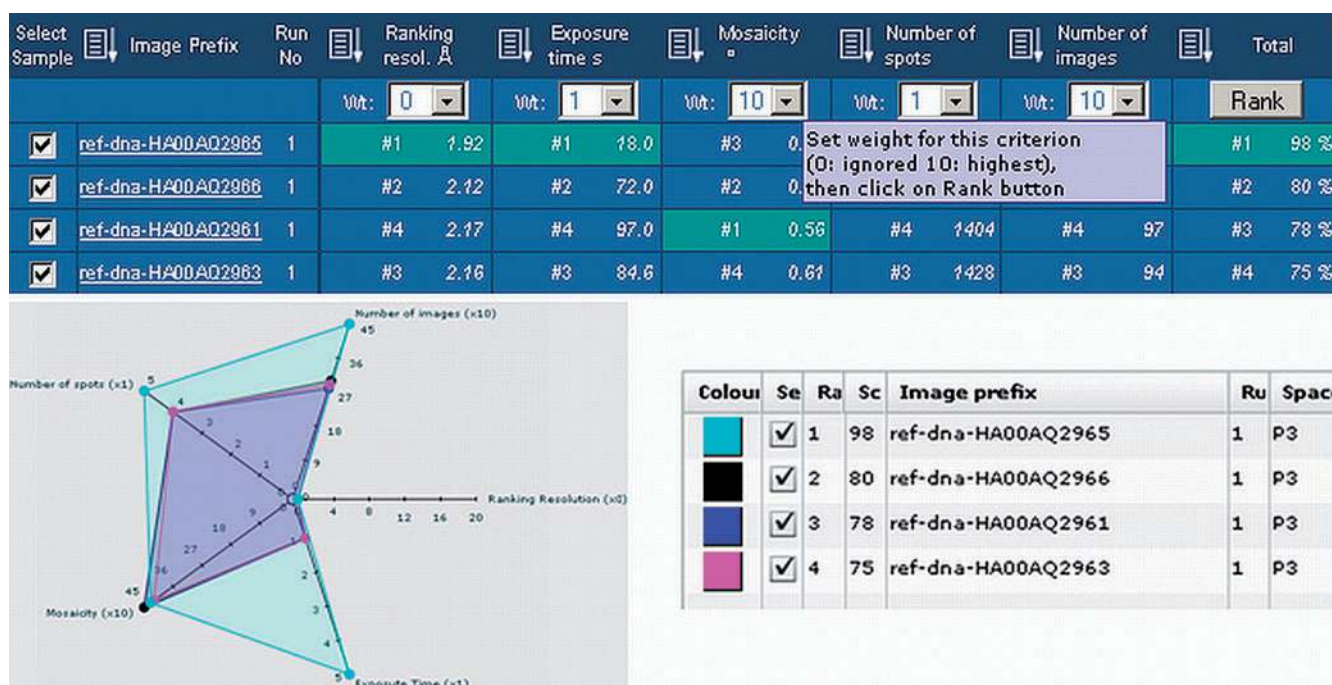

Fig. 5. Sample ranking in ISPyB. Samples can be selected and ranked either on individual criteria or using a combination of criteria that can be given different weights. Results can be visualized either as a table (top) or as a radar chart. Tool-tips provide guides for ease of use.

Program (see http://www.esrf.eu/UsersAndScience/Experiments/MX/beamline-portfolio/CDR_UPBL10_futureID30-BM29.pdf). At DLS this observation has helped drive the deployment of faster sample changers.

While ISPyB is currently exclusively exploited during MX experiments, its data model hierarchy can also be applied to other techniques with analogous workflows. Examples of such techniques include Small Angle X-ray Scattering (SAXS), particularly when applied to solutions of biological macromolecules (BioSAXS), small molecule X-ray crystallography and powder diffraction. In each of these cases, the development of ISPyB-like resources for use at synchrotron facilities is under serious consideration.

Although the growing needs of experimenters has resulted in ISPyB becoming an integral part of MX experiments on the ESRF and DLS Structural Biology beamlines, it is clear that the database will need to evolve further if it is to meet the ever more challenging requirements of Structural Biologists and synchrotron beamlines dedicated to the MX technique. It is likely that many future projects in Structural Biology will involve the screening of the diffraction properties of several thousand rather than several hundred samples in order to find those most suitable for data collection. Such projects are also likely to involve the combination of partial data sets collected either from different crystals or from different parts of the same crystal. Moreover, sample pre-screening will increasingly take place at different beamlines—if not different synchrotron facilities—prior to subsequent full data collections. Future developments to ISPyB will certainly require workflows modified to take into account previously collected data and, thus, a revised data model based on projects rather than experimental sessions.

ISPyB has evolved from a database confined to simple sample tracking, experiment recording and reporting to a sophisticated and interactive system for experiment planning, management and execution. MX facilities at European synchrotron sources other than the ESRF and DLS have expressed an interest in making use of the database. This would enable ISPyB as a pan-European LIMS at the hub of synchrotron-based macromolecular structure solution pipelines.

ACKNOWLEDGEMENTS

Conflict of Interest: none declared.

REFERENCES

Allan,R.J. *et al.* (2005) Portal design, Synchrotron and HPC services in e-HTPX – A resource for High Throughput Protein Crystallography. In Cox,S.J and Walker,D.W. (eds) *Proceedings of the UK e-Science All Hands Meeting 2005*, ISBN 1-904425-53-4.

Amin,A.A. *et al.* (2006) Managing and mining protein crystallization data. *Proteins*, **62**, 4–7.

Arzt,S. *et al.* (2005) Automation of macromolecular crystallography beamlines. *Prog. Biophys. Mol. Biol.*, **89**, 124–152.

Berman,H.M. *et al.* (2000) The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242.

Beteva,A. *et al.* (2006) High-throughput sample handling and data collection at synchrotrons: embedding the ESRF into the high-throughput gene-to-structure pipeline. *Acta Crystallogr. D Biol. Crystallogr.*, **D62**, 1162–1169.

Bourenkov,G.P. and Popov,A.N. (2010) Optimization of data collection taking radiation damage into account. *Acta Crystallogr. D Biol. Crystallogr.*, **D66**, 409–419.

Bowler,M.W. *et al.* (2010) Diffraction cartography: applying microbeams to macromolecular crystallography sample evaluation and data collection. *Acta Crystallogr. D Biol. Crystallogr.*, **D66**, 855–864.

Cipriani,F. *et al.* (2006). Automation of sample mounting for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, D62, 1251–1259.

Cohen,A.E. *et al.* (2002). An automated system to mount cryo-cooled protein crystals on a synchrotron beamline, using compact sample cassettes and a small-scale robot. *J. Appl. Cryst.,* **35**, 720–726.

Cork,C. *et al.* (2006). Advanced beamline automation for biological crystallography experiments. *Acta Crystallogr. D Biol. Crystallogr.,* **D62**, 852–858.

Daniel,E. *et al.* (2011) xtalPiMS: A PiMS-based web application for the management and monitoring of crystallization trials. *J. Struct. Biol.*, **175**, 230–235.

Emamzadah,S. *et al.* (2009) Cyclic olefin homopolymer-based microfluidics for protein crystallization and in situ X-ray diffraction. *Acta Crystallogr. D Biol. Crystallogr.*, **D65**, 913–920.

Evans,G. and Pettifer,R.F. (2001) CHOOCH: a program for deriving anomalous-scattering factors from X-ray fluorescence spectra. *J. Appl. Cryst.*, **34**, 82–86.

Fulton,K. *et al.* (2004) CLIMS: Crystallography Laboratory Information Management System. *Acta Crystallogr. D Biol. Crystallogr.*, **D60**, 1691–1693.

Gabadinho,J. *et al.* (2010) MxCuBE: a synchrotron beamline control environment customized for macromolecular crystallography experiments. *J. Synchrotron Radiat.* **17**, 700–707.

Gonzalez,A. *et al.* (2008) Web-Ice: integrated data collection and analysis for macromolecular crystallography. *J. Appl. Cryst.* **41**, 176–184.

Harris,M. and Jones,T.A. (2002) Xtrack - a web-based crystallographic notebook. *Acta Crystallogr. D Biol. Crystallogr.*, **D58**, 1889–1891.

Hiraki,M. *et al.* (2008) High-throughput operation of sample-exchange robots with double tongs at the Photon Factory beamlines. *J. Synchrotron Radiat.* **15**, 300–303.

Incardona,M.-F. *et al.* (2009) EDNA: a framework for plugin-based applications applied to X-ray experiment online data analysis. *J. Synchrotron. Radiat.* **16**, 872–879.

Jacquamet,L. *et al.* (2004) Automated Analysis of Vapor Diffusion Crystallization Drops with an X-Ray Beam. *Structure*, **12**, 1219–1225

Jacquamet,L. *et al.* (2009) Upgrade of the CATS sample changer on FIP-BM30A at the ESRF: towards a commercialized standard. *J. Synchrotron Radiat.* **16**, 14–21.

Leonard,G.A. *et al.* (2009) Online collection and analysis of X-ray fluorescence spectra on the macromolecular crystallography beamlines of the ESRF. *J. Appl. Cryst.* **42**, 333–335.

Leslie,A.G.W. *et al.* (2002) Automation of the collection and processing of X-ray diffraction data - a generic approach. *Acta Crystallogr. D Biol. Crystallogr.*, **D58**, 1924–1928.

McPhillips,T.M. *et al.* (2002) Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. *J. Synchrotron. Radiat.* **9**, 401–406.

Morris,C. *et al.* (2011) The Protein Information Management System (PiMS): a generic tool for any structural biology research laboratory. *Acta Crystallogr. D Biol. Crystallogr.*, **D67**, 249–260.

Murray,C.W. and Blundell,T.L (2010) Structural biology in fragment-based drug design. *Curr. Opin. Struct. Biol.*, **20**, 497–507.

Pajon,A. *et al.* (2005) Design of a data model for developing laboratory information management and analysis systems for protein production. *Proteins*, **58**, 278–284.

Prilusky,J. *et al.* (2005) HalX: an open-source LIMS (Laboratory Information Management System) for small- to large-scale laboratories. *Acta Crystallogr. D Biol. Crystallogr.*, **D61**, 671–678.

Skinner,J.M. *et al.* (2006) Integrated software for macromolecular crystallography synchrotron beamlines II: revision, robots and a database. *Acta Crystallogr. D Biol. Crystallogr.*, **D62**, 1340–1347.

Snell,G. *et al.* (2004) Automated sample mounting and alignment system for biological crystallography at a synchrotron source. *Structure*, **12**, 537–545.

Winter,G. (2010) xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Cryst.* **43**, 186–190.