Automated Pair-Distribution Function Data Processing

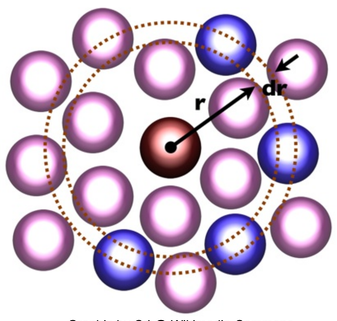
Timothy Spain[1], Dean Keeble[1], Philip Chater[1], Michael Wharmby[1], Jacob Filik[1], Alun Ashton[1], Heribert Wilhelm[1]

[1]Diamond Light Source, Harwell Science and Innovation Campus, Didcot, Oxon OX11 0DE

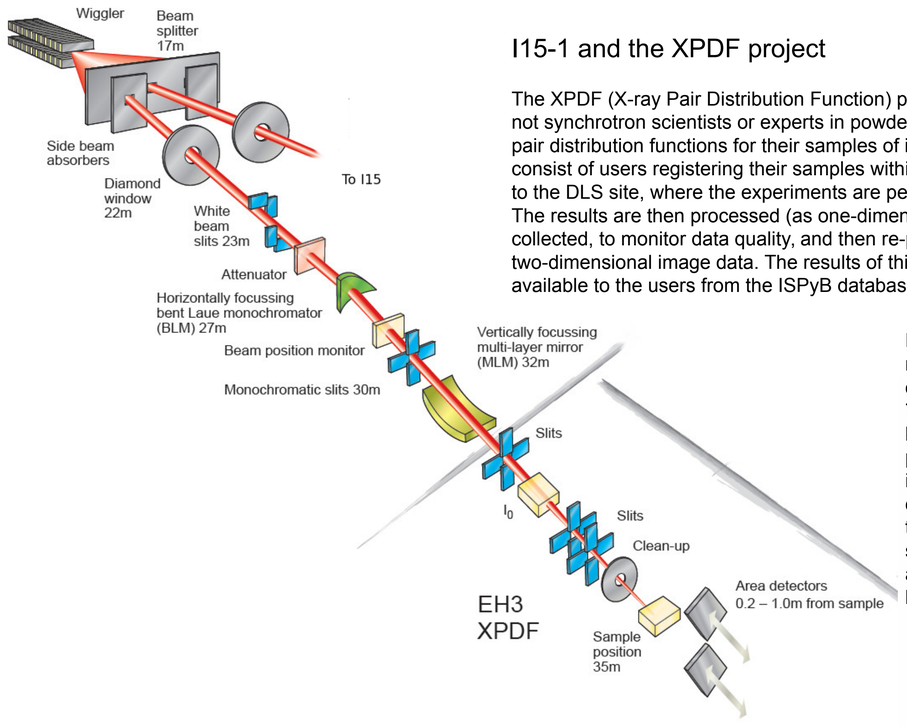
The XPDF project aims to provide a hands-off data service for users to produce X-ray Pair Distribution Function data. The software will be accessible to non-expert users, allowing scientists from a variety of disciplines to get PDF data without knowledge of the details of X-ray powder diffraction and with minimal intervention from beamline staff. The software for the project uses several of the technologies already developed by Diamond Light Source and other facilities to allow the automatic processing of X-ray powder diffraction data, including DAWN (Data Analysis Workbench), ISPyB and SynchWeb.

Pair distribution function (PDF)

A (radial) pair distribution function characterizes the structure of a material in a one-dimensional function. It records, as a function of separation, the probability of finding two atoms with that separation. The pair distribution function can be recovered from the total scattering data, and so can be obtained from a wide variety of materials, not only polycrystalline powders, but also liquids, amorphous solids, nano-particles, *&c*. Measuring the pair distribution function can characterize the structure of a material of known composition, but unknown structure, or to investigate the changes in the structure as the conditions of the material are changed, such as the temperature of the sample, or whilst undergoing an electrochemical reaction.



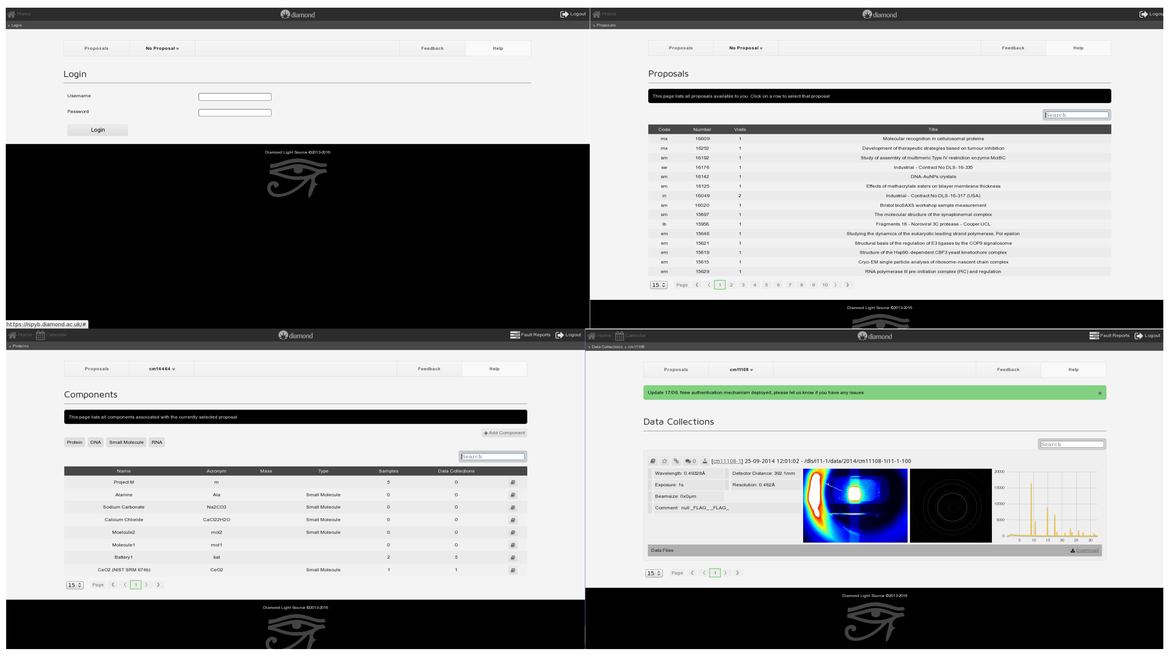
Graphic by Cri @ Wikimedia Commons



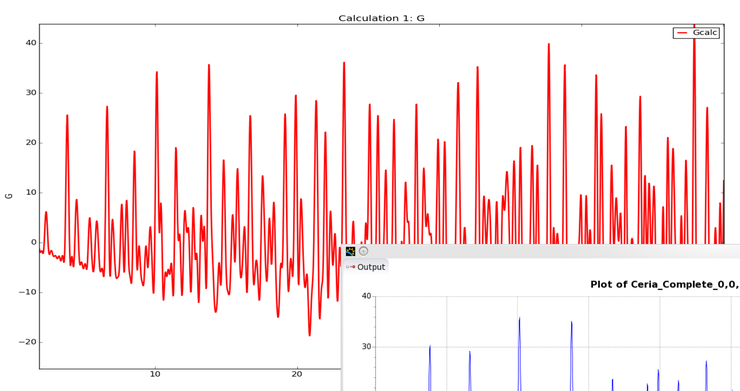
I15-1 and the XPDF project

The XPDF (X-ray Pair Distribution Function) project aims to let users who are not synchrotron scientists or experts in powder diffraction obtain high quality pair distribution functions for their samples of interest. The process will consist of users registering their samples within the database, shipping them to the DLS site, where the experiments are performed on the I15-1 beamline. The results are then processed (as one-dimensional data) as they are collected, to monitor data quality, and then re-processed as full two-dimensional image data. The results of this processing are then made available to the users from the ISPyB database.

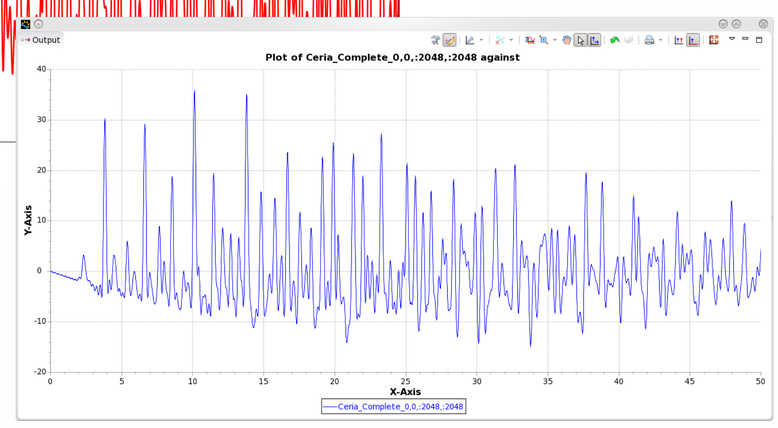
For the XPDF experiments, the radiation is monochromated to one of three energies (40, 65, 76 keV). At present, the beamline has one 16 million pixel Perkin-Elmer detector, but in future will use two identical detectors at different distances to capture high resolution small-angle data, whilst also allowing data at large angles to be simultaneously recorded.



SynchWeb interface to ISPyB



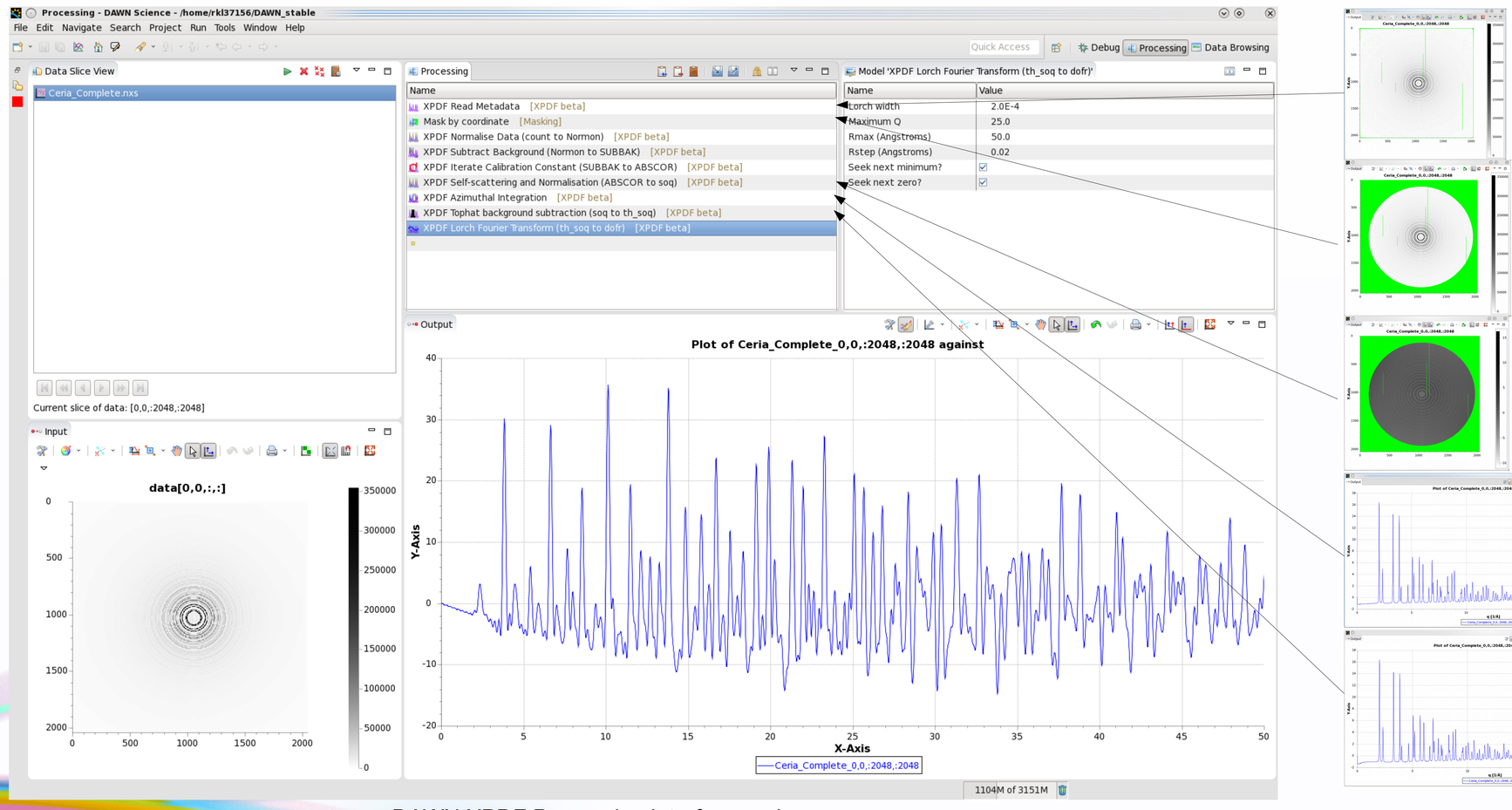
Simulated ceria SRM PDF (PDFgui)



Processed ceria SRM PDF (DAWN XPDF)

DAWN (Data Analysis Workbench)

The processing of the data is performed using Diamond Light Source's in-house data analysis system, DAWN (www.dawnsci.org). Dawn provides the framework for reading and writing the acquired and processed data, stepping through the acquired data frames and initiating the data processing on the user's computer, or on the DLS data analysis compute cluster. The data is currently analysed in the DAWN Processing Perspective. The steps of the processing pipeline can be exported and shared with users to allow them to process their data without learning the intricacies of the DAWN Processing system. In the future, data processing will be done automatically, with data obtained from the ISPyB LIMS database. This will make the processed data available through the SynchWeb interface to ISPyB.



DAWN XPDF Processing interface and intermediate results

XPDF Processing

The goal of the processing of the XPDF data is to take the recorded data in detector counts, and eliminate the scattering from the sample containers (capillary, furnace, &c.), the fluorescence of the sample and the scattering from the air and components of the beamline. The effects of the transmission and geometry of the detector are also corrected for.

The processing of the XPDF data requires several background datasets, in addition to the data taken of the sample. The diffraction image of the identically set up beamline must be taken with the sample containers, with the innermost container or sample removed in turn. Using these data, the recorded data can be corrected for the scattering of the containers, as well as the attenuation of the sample scattering by the containers.

Some processes are not yet handled. Multiple scattering is neglected, with x-rays being only scattered, attenuated and then absorbed by the detector. Any fluorescence of the sample containers is also ignored, although this should have been minimized by careful selection of container materials.

Once the data has been cleaned of these unwanted effects, the remaining values should be the scattering by the sample alone, unaffected by any containers or fluorescence. The data is then normalized to the total scattering of all the atoms in the substance, calculated from first principles. The normalized data is then integrated into a radial profile as a function of momentum transfer, *Q*, filtered to remove an remaining background. This clean signal is then transformed from momentum (*Q*) space to radial (*r*) space using an apodized form of the Fourier transform. The result of this is the pair-distribution function of the sample material.

References

Dawn: Basham *et al.*, J. Synchrotron Rad., **22**, pp. 853—858 (2015)
ISPyB: Delagenière *et al.*, Bioinformatics, **27** (22), pp. 3186—3192 (2011)
SynchWeb: Fisher *et al.*, J. Appl. Cryst., **48**, pp. 927—932 (2015)
PDFgui: Farrow *et al.*, J. Phys.: Condens. Mat., **19**, 335219 (2007)

For more information please visit www.diamond.ac.uk or contact Timothy Spain at timothy.spain@diamond.ac.uk