

A public database of macromolecular diffraction experiments

Marek Piotr Grabowski¹, Marcin Cymborowski¹, Piotr Sroka¹, Przemek Porebski¹, David Cooper¹, Wladek Minor¹

¹*Molecular Physiology And Biological Physics, University Of Virginia, Charlottesville, United States*

E-mail: marek@iwonka.med.virginia.edu

The Integrated Resource for Reproducibility in Macromolecular Crystallography (IRRC) has grown into a substantial public repository of primary data from protein diffraction experiments. Currently, the resource contains data from over 3200 macromolecular diffraction experiments (~6200 datasets), accounting for around 3% of all deposits in the Protein Data Bank (PDB). Each experiment is associated by partially curated metadata. The resource utilizes a distributed storage architecture, which accommodates a confederation of individual storage servers and provides both scalability and sustainability. The IRRC is available through its web portal at <http://www.proteindiffraction.org> and can be searched using various criteria. The resource is open to data submission from the worldwide crystallographic community, including datasets that have failed to yield X-ray structures. The availability of unsolved data will facilitate collaborative efforts to improve structure determination methods and software. Additionally, we expect the IRRC unsolved data will yield new protein structures as more molecular replacement models become available. An automated diffraction image reprocessing pipeline, as well as other tools for "wrangling" data and metadata from diffraction images have been developed. These tools have already identified situations where a modification of the data collection protocol could have significantly increased the quality of data, as well as unearthed a number of diffraction datasets that may benefit from reprocessing. The resource provides a way to safeguard public availability of "orphan" data left behind by principal investigators closing their labs and/or extinct large structural biology projects.

References:

[1] Grabowski, M. et al. (2016) Acta Cryst D72, 1181-1193

[2] Grabowski, M. and Minor, W. (2017), IUCR J 4, 3-4

Keywords: [X-Ray diffraction](#), [reproducibility](#), [raw data](#)