

## **An Introduction to the Gold Standard for Macromolecular Crystallography Diffraction Data**

**H Bernstein<sup>1</sup>**

**<sup>1</sup>Ronin Institute for Independent Scholarship, c/o NSLS-II, Brookhaven National Lab, Bldg 745,  
Bellport, NY  
yayahjb@gmail.com**

In a major step forward for MX data management, a large portion of the research community concerned with High Data-Rate Macromolecular Crystallography (HDRMX) has now agreed to an updated specification of data and metadata for diffraction images produced at synchrotron light sources and X-ray free electron lasers (XFELs) [Herbert J. Bernstein, Andreas Förster, Asmit Bhowmick, Aaron S. Brewster, Sandor Brockhauser, Luca Gelisio, David R. Hall, Filip Leonarski, Valerio Mariani, Gianluca Santoni, Clemens Vornrhein, Graeme Winter, "Gold Standard for Macromolecular Crystallography Diffraction Data", IUCrJ submitted 2020]. This talk is an introduction to the new "Gold Standard" with use case examples for synchrotron and XFEL data. "Macromolecular crystallography (MX) is the dominant means of determining the three-dimensional structures of biological macromolecules. Over the last few decades, most MX data have been collected at synchrotron beamlines using a large number of different detectors produced by various manufacturers and taking advantage of various protocols and goniometry. These data came in their own formats, sometimes proprietary, sometimes open. The associated metadata rarely reached the degree of completeness required for data management according to Findability, Accessibility, Interoperability, and Reusability (FAIR) principles. Efforts to reuse old data by other investigators or even by the original investigators some time later were often frustrated. "In the culmination of an effort dating back more than two decades, a large portion of the research community concerned with High Data-Rate Macromolecular Crystallography (HDRMX) has now agreed to an updated specification of data and metadata for diffraction images produced at synchrotron light sources and X-ray free electron lasers (XFELs). This "Gold Standard" will facilitate processing of datasets independent of the facility at which they were collected and enable data archiving according to FAIR principles, with a particular focus on interoperability and reusability. This agreed standard builds on the NeXus/HDF5 NXmx application definition and the International Union of Crystallography (IUCr) imgCIF/CBF dictionary, and it is compatible with major data processing programs and pipelines. Just as with the IUCr CBF/imgCIF standard from which it arose and to which it is tied, the NeXus/HDF5 NXmx Gold Standard application definition is intended to be applicable to all detectors used for crystallography, and all hardware and software developers in the field are encouraged to adopt and contribute to the standard."