Processing simultaneously-collected MAD data from two closely-spaced (90 eV) wavelengths measured at an X-ray free electron laser

Derek Mendez[a], William Weis[b], Axel Brunger[b,c], Soichi Wakatsuki[b], and Nicholas Sauter[a]

[a] Lawrence Berkeley National Laboratory, Berkeley, CA 94720, dermen@lbl.gov, nksauter@lbl.gov
[b] Department of Structural Biology, Stanford University, Stanford, CA 94305 USA, bill.weis@stanford.edu, soichi.wakatsuki@stanford.edu
[c] Department of Molecular and Cellular Physiology, Stanford University, Stanford, CA 94305 USA, brunger@stanford.edu,

Here we reveal a novel analysis pipeline for processing MAD XFEL crystallography data where each image is a superposition of measurements from two closely-spaced wavelengths. The crystal molecular units were Ytterbium derivatives of chicken-egg lysozyme, and the two wavelengths chosen were both near the Ytterbium L-III absorption edge. Specifically, the measured diffraction patterns comprised many overlapping reflections, leading to unique data analysis challenges and solutions.

XFELs are a great utility for investigating samples at room-temperature, and due to their ultra-short and bright X-ray pulses they remain a leading instrument for time-resolved structural studies. The vast majority of published XFEL structures are solved using molecular replacement, with an exception being [1], where two-wavelength MAD data was collected simultaneously from many lysozyme crystals. In [1], crucially, a large photon energy separation was utilized (~2000 eV) such that all measured spots were non-overlapping and the superimposed diffractions patterns were easily separated and processed independently. While conventional analysis tools and techniques were used to extract the structure factor magnitudes from the data and solve for the phases, it was revealed that multiple wavelength data significantly improves de novo phasing accuracy at XFELs (compared to single wavelength SAD phasing).

We present our efforts developing a novel approach where unit-cell structure factors may be extracted using a global fit of the two-wavelength data to the Karle-Hendrickson MAD equations [2]. This approach makes use of rapid simulation of the two-wavelength partialities and lattice contributions to each measured spot intensity in order to extract unit cell contributions. The energy separation is small (~90 eV), leading to many overlapping or nearly-overlapping reflections in the low and intermediate resolution regime where the MAD signal is expected to be large, and therefore standard methods of XFEL data processing do not apply.

[1] Gorel, A. et al (2017), *"Multi-wavelength anomalous diffraction de novo phasing using a two-colour X-ray free-electron laser with wide tunability",* Nat. Comm. (8)
[2] Karle, J. (1980) "*Some Developments in Anomalous Dispersion for the Structural Investigation of Macromolecular Systems in Biology*", Int. J. Quantum Chem. Symp. (7)