Synchrotron Serial Crystallography with Multi-stage Merging of 1000's of Images

Herbert J. Bernstein[1], Lawrence C. Andrews[2], Kaden Badalian[3], Alexei Soares[4], Wuxian Shi[5], Jean Jakoncic[4], Martin Fuchs[4], Dieter Schneider[4], Robert M. Sweet[4], Sean McSweeney[4], John Skinner[4], Yusuke Yamada[6]

[1]Rochester Institute of Technology, Rochester, NY, USA;
[2]Retired, Kirkland, WA, USA;
[3]Binghamton University, Binghamton, NY, USA;
[4]Brookhaven National Laboratory, Upton, NY, USA;
[5]Case Western Reserve University, Cleveland, Ohio;
[6]Photon Factory, Tsukuba,Ibaraki, Japan

KAMO [1] and Blend [2] provide particularly effective tools to automatically manage the merging of large numbers of data sets from synchrotron serial crystallography. The requirement for manual intervention in the process can be reduced by extending Blend to support additional clustering options to improve the sensitivity to differences in unit cell parameters and to allow for clustering of nearly complete datasets on the basis of intensity differences. One applies KAMO twice, first using cell parameters. In this step either the simple cell vector distance of the original Blend is used, or the more sensitive NCDist [3], to find clusters to merge to at higher levels of completeness. In the second step one uses KAMO again using the correlation between the intensities at the common HKLs [4] to merge clusters in a way sensitive to structural differences that may not perturb the cell parameters sufficiently to make meaningful clusters.

[1] https://github.com/keitaroyam/yamtbx/

[2] Foadi, J., Aller, P., Alguel, Y., Cameron, A., Axford, D., Owen, R.L., Armour, W., Waterman, D.G., Iwata, S. and Evans, G. (2013). Clustering procedures for the optimal selection of data sets from multiple crystals in macromolecular crystallography. Acta Crystallographica Section D: Biological Crystallography, 69(8), 1617-1632.

[3] Andrews, L. C., & Bernstein, H. J. (2014). The geometry of Niggli reduction: BGAOL–embedding Niggli reduction and analysis of boundaries. Journal of applied crystallography, 47(1), 346-359.

[4] Assmann, G., Brehm, W., & Diederichs, K. (2016). Identification of rogue datasets in serial crystallography. Journal of applied crystallography, 49(3).