MS3-P2 Decision making in automated structure solution pipelines

Melanie Vollmar[1], David Waterman[2], Irakli Sikharulidze[1], Graeme Winter[1], Dave Hall[1], Gwyndaf Evans[1]

1. Diamond Light Source Ltd, Harwell Science and Innovation Campus, Didcot, Oxfordshire, OX11 0DE, United Kingdom
2. Science and Technology Facilities Council, Rutherford-Appleton Laboratory, Harwell Science and Innovation Campus, Didcot, OX11 0FA, United Kingdom

email: melanie.vollmar@diamond.ac.uk

In recent years data collection, especially at third generation synchrotrons, has become so fast that in many cases scientists choose to shoot first and then ask the question later. This approach may not always be the most efficient and the resulting data set may be insufficient for structure determination due to a number of reasons, e.g. radiation damage leading to incomplete data or insufficient anomalous signal. Additionally, especially when using remote access, data collection is not always carried out by scientists with appropriate training and experiments may not be planned well enough.

Many tools and pipelines already employed at Diamond or within CCP4 help scientists to analyse their data and assist in decision-making while still having access to the crystal at the beamline. Databases such as ISpyB or beamline control software like GDA allow the user to provide additional information like protein sequence, crystallisation condition and atomic scatterer for derivative crystals prior to the experiment or while at the beamline. In the past two decades a multitude of quality metrics (e.g. R_{merge}, $R_{p.i.m.}$, I/σI, $CC_{½}$, CC_{ano}, CC_{map}) have been proposed, accepted, declined and modified to assist scientists in decision-making through the entire structure solution process.

Here we present how structure determination pipelines at Diamond will change over the next few years. In collaboration with CCP4 the existing pipeline infrastructure is to be transferred into a complex network for structure solution. Currently existing user data has been analysed with established software and pipelines and judged based on known quality metrics. In the long run this will serve as the basis to train a structure solution network in decision-making, which may involve defining additional new quality metrics or even allow scientists to provide biological data, results from mutation studies and sequence analysis, to increase the chances of solvability of a protein structure.

Weiss (2001) J Appl Cryst 34 130-135

Mueller-Dieckmann *et al.* (2005) Acta Cryst D61 1263-1272

Mueller-Dieckmann *et al.* (2004) Acta Cryst D60 28-38

Cianci et al (2008) Acta Cryst D64 1196-1209

Diederichs/Karpuls; (2013) Acta Cryst D69 1215-1222

Delageniere, S. *et al.* (2011). Bioinformatics, 27, 3186-3192

Beteva, A. *et al.* (2006). *Acta Cryst.* D62, 1162-1169

Keywords: structure solution, automated pipelines, decision making

MS3-P3 New metrics for anomalous data quality

Selina L.S. Storm[1], Guillaume Pompidor[1], Fabio Dall'Antonia[1], Gleb Bourenkov[1], Thomas R. Schneider[1]

1. European Molecular Biology Laboratory, Notkestr. 85, 22607 Hamburg, Germany

email: selina.storm@embl-hamburg.de

Experimental phasing is oftentimes hindered by radiation damage. Several metrics are available to judge data quality and the anomalous signal by taking all measured data into account [1-4].

For application in practice, it would be useful to have metrics with which one can easily identify a subset of data which are least radiation damaged and thereby increase then chance of successful phasing.

In absence of radiation damage, all Bijvoet differences would be the same. With increasing radiation damage, subsequently recorded Bijvoet mates and anomalous differences vary. By monitoring the behavior of subsequently recorded Bijvoet pairs, it should be possible to identify the point at which adding the next measurement, more noise than signal is introduced into a resulting data set.

We will present an evaluation of different metrics in terms of their ability to estimate the usefulness of Bijvoet differences as a function of dose deposited in a crystal.

References: [1] M. Weiss, J. Appl. Cryst. (2001): 130-135

[2] C. Mueller-Dieckmann et al., Acta Cryst. D (2005): 1263–1272

[3] P. Karplus, & K. Diederichs, Science (2012): 1030-1033

[4] G. Bunkóczi et al., Nature methods (2014): 126-130

Keywords: radiation damage, experimental phasing