Using the pair angle distribution function for analysing protein structure

**P. L. R Adams, J. Binns, T. L. Greaves, A. V. Martin**

School of Science, RMIT University, Melbourne, Victoria 3000, Australia

S3826109@student.rmit.edu.au, andrew.martin@rmit.edu.au

X-Ray Free Electron Lasers provide a means of conducting crystallography experiments with remarkable time and spatial resolution. These methods can directly recover the electron density of materials. However, there are stringent requirements such as crystal size, number density per exposure, and the crystal order which are required for reconstruction. Membrane proteins, which do not readily crystallise or meet these requirements [1], are particularly interesting to study as they comprise up to 50% of drug targets [2], but less than 10% of the protein structures in the Protein Data Bank [3].

The Pair Angle Distribution Function (PADF) describes the three and four body correlations of the electron density in a sample, and can be recovered from X-ray cross-correlation analysis (XCCA) [4]. Although PADF analysis does not recover the electron density directly, it still contains significant information about the local three dimensional structure of the material. PADF analysis also has the potential to relax the stringent crystal requirements of current single crystal experiments.

We discuss the sensitivity of the PADF to different protein structures [5], and the correlations generated at different length scales; from atomic bonding to tertiary structure. Our aim is to further develop PADF analysis to recover crystal structure factors using X-ray cross-correlation analysis.
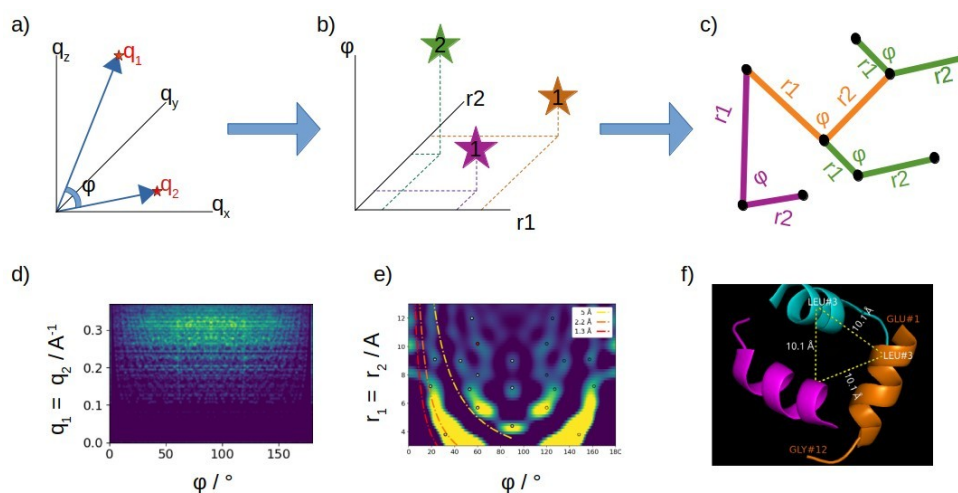
Figure 1. (a) Two scattering vectors in reciprocal space are correlated to produce the XCCA, which can be transformed into the PADF. (b) The PADF encodes three and four body spatial correlations of the material (c).   (d) The XCCA data from protein crystal diffraction [5]. (e) The PADF produced from the XCCA data in (d) [5]. (f) Intermolecular structure of protein 1AL1 in crystal, with matching distance in PADF, indicated by the red dot in e) [5].

[1] Johansson, L.C.; Arnlund, D.; White, T.A.; Katona, G.; DePonte, D.P.; Weierstall, U.; Doak, R.B.; Shoeman, R.L.; Lomb, L.; Malmerberg, E.; et al. Lipidic phase membrane protein serial femtosecond crystallography. Nat. Methods 2012, 9, 263–265.

[2] Cournia, Z.; Allen, T.W.; Andricioaei, I.; Antonny, B.; Baum, D.; Brannigan, G.; Buchete, N.V.; Deckman, J.T.; Delemotte, L.; del Val, C.; et al. Membrane protein structure, function, and dynamics: A perspective from experiments and theory. J. Membr. Biol. 2015, 248, 611–640.

[3] Berman, H.M.; Battistuz, T.; Bhat, T.N.; Bluhm, W.F.; Bourne, P.E.; Burkhardt, K.; Feng, Z.; Gilliland, G.L.; Iype, L.; Jain, S.; et al. The Protein Data Bank. Acta Crystallogr. Sect. D Biol. Crystallogr. 2002, 58, 899–907.

[4] Martin, A.V. Orientational order of liquids and glasses via fluctuation diffraction. IUCrJ 2017, 4, 24–36.

[5] Adams, Patrick, et al. "The Sensitivity of the Pair-Angle Distribution Function to Protein Structure." Crystals 10.9 (2020): 724.

Keywords: Novel Method; XFEL; Fluctuation Scattering