Ice in Biomolecular Crystallography

D Moreau[1], H Atakisi[2], R Thorne[3]

[1]Cornell University [2]Cornell University, [3]Physics Dept, Cornell Univ

dwm265@cornell.edu

Diffraction images acquired from cryocooled protein crystals often include powder diffraction rings from ice. Distinguishing ice diffraction from the desired protein diffraction is non-trivial and often results in incorrectly measured protein structure factors (Parkhurst 2017). Building on the work of Thorn and coworkers (Thorn 2017), a revised metric, Pice, is defined for detecting ice from deposited protein structure-factor data. This metric is validated using full-frame diffraction data from the Integrated Resource for Reproducibility in Macromolecular Crystallography. Pice is a p-value testing the null hypothesis that the measured structure factors are not biased by ice diffraction. Pice is a condensed and familiar statistical metric that we believe could be used as a standard marker of data quality. Using Pice, an analysis of structure-factor data from a random sample of 89,827 Protein Data Bank (PDB) entries collected at cryogenic temperatures indicates that roughly 16% show evidence of ice contamination, and that this fraction increases with increasing solvent content and maximum solvent-cavity size. By examining the ice diffraction-peak positions at which structure-factor perturbations are observed, roughly 25% of these ice-contaminated entries have ice with a primarily hexagonal character, and that the absolute fraction of PDB entries exhibiting hexagonal ice has grown steadily over the last 20 years, indicating that inadequate cooling rates and/or cryoprotectant concentrations were used. The remaining 75% of ice contaminated entries show ice with a stacking-disordered or cubic character, arising from internal or cryoprotected surface solvent.

References:

- Parkhurst, J. M. et al. (2017). IUCrJ, 4, 626–638.

- Thorn et al. (2017), Acta Cryst. D73, 729–727,