Nothing Beats Good Data - Lessons Learned from Native-SAD Data Collection Can Give the Best Crystal Structure from Alphafold Molecular Replacement Models

Dr. John P Rose[1], Dr. Dayong Zhou[1], Dr. Lirong Chen[1], Dr. Bi-Cheng Wang[1]
[1]*University of Georgia*
jprose@uga.edu

Molecular Replacement using AlphaFold-generated search models is quickly becoming the first-choice method for the de novo determination of macromolecular crystal structures. The technique is simple, no heavy atom derivatives, no selenomethionine incorporation, just collect a data set on the native crystal, generate an AlphaFold search model based on the protein's sequence and solve the structure via Molecular Replacement. However, this is not as simple as it appears since structure quality is based on data quality regardless of the phasing method used. Thus, one must always strive to collect the best data possible.

Data quality is always important for Sulfur/Native SAD structure determination. Native-SAD phasing uses the anomalous scattering signal of light atoms in the native crystal collected using single-wavelength X-ray diffraction experiments. These atoms include sodium, magnesium, phosphorus, sulfur, chlorine, potassium, and calcium. The anomalous scattering signal (ΔF") for these light atoms is weak, for example, (ΔF") for sulfur ranges from 0.24e- at the selenium adsorption edge to 0.52e- using Cu Kα X-rays as compared to selenium's anomalous scattering signal (ΔF" = 3.82) at its absorption edge. Thus, Native- SAD phasing is challenging and is critically dependent on collecting accurate data since the small anomalous scattering signals of these light atoms will otherwise be lost in the noise. In short, lessons learned from Native-SAD Data Collection can give the best crystal structure using AlphaFold Molecular Replacement models.

The presentation will focus on the limitations of the AlphaFold model and using Native-SAD data collection methods and techniques to improve data accuracy, the quality of the resulting AlphaFold Molecular Replacement solution, and the final refined crystal structure.

Work is partially supported with funds from the Georgia Research Alliance, the National Institutes of Health (GM62407, RR028976 & FP00014437), and the University of Georgia Research Foundation.