Development and Deployment of Enabling Technologies at the UCLA-DOE Institute

Jose A. Rodriguez,[1] David Eisenberg,[1] Marcus Gallagher-Jones[1*](marcusgj@chem.ucla.edu) Logan Richards,[1] Ambarneil Saha,[1] Matthew Agdanowski,[1] Roger Castells-Graells,[1] and **Todd O. Yeates**[1]

[1]UCLA-DOE Institute for Genomics and Proteomics; University of California, Los Angeles. Los Angeles, CA 90095

https://www.doe-mbi.ucla.edu/ucla-doe-institute-missions/

Project Goals:

Research in the UCLA-DOE Institute for Genomics and Proteomics (IGP) includes major efforts in the area of imaging science, where we are advancing the understanding of plant and microbial biosystems, their genomics and molecular biology. Our team is pioneering new enabling capabilities that facilitate the discovery of molecular structural features affecting protein function and specificity, to better our understanding of bioenergy crops and microbes. These capabilities span the broad areas of X-ray diffraction, electron microscopy, and micro-electron diffraction (MicroED), along with protein engineering and selection methods designed to advance those techniques. Emerging MicroED techniques present challenges in processing, refinement and phasing of diffraction data that our team is tackling in part by enabling rapid access to robust public-facing tools (webservers) that facilitate ED structure determination. Our team is also making critical advances to resolve protein structures smaller than about 50 kDa in size by cryo-EM. We have developed the first example of a working molecular scaffold that can image such small proteins.

Abstract:

Our efforts in imaging science and protein characterization bridge a number of technological areas to address pressing problems in protein structure and function.

Breakthroughs in cryo-EM –
Numerous technical advances have made cryo-EM an attractive method for atomic structure determination. Cryo-EM is ideally suited for very large structures; symmetrical structures like viruses are especially amenable. However, problems of low-signal-to-noise in imaging small proteins makes it practically impossible to determine structures smaller than about 50 kDa, leaving a great many cellular proteins and enzymes (and nucleic acid molecules) outside the reach of this important structural technique. In recent work, our DOE-UCLA IGP team has broken through this barrier by engineering novel scaffolds with sufficient rigidity and modularity to achieve resolution useful for interpreting atomic structure. A 3.8 Å resolution for a 26 kDa protein provides the leading benchmark for cryo-EM scaffolding (Fig. 1). Our ongoing efforts aim to further rigidify designed scaffolds, apply them to microbial and plant protein targets, and further extend the scaffolding approaches to nucleic acid molecules.

Enabling micro-ED methods -
A broad array of atomic structures has now been determined by MicroED; they include naturally occurring peptides, synthetic protein fragments and peptide-based natural products. However, *de*

novo structure determination by MicroED remains problematic for all but ideal crystals. Automated, fragment-based approaches to structure determination eliminate the need for atomic resolution diffraction, instead enforcing stereochemical constraints through libraries of small model fragments. We have demonstrated the application of fragment-based phasing on various macromolecular structures including some for which all traditional phasing methods have failed (Fig. 2). During refinement of MicroED structures, potential maps can reveal a wealth of typically untapped information about charged states. This information remains out of reach largely because of inadequate parameterizations of ionic electron scattering factors. To rectify this, we have developed a public web server accessible at: https://srv.mbi.ucla.edu/faes/. This resource enables Gaussian parameterization of elastic electron scattering factors in a form amenable to refinement in the program Phenix. We are exploring its application to a number of published cryo–EM structures that previously neglected charge, which plays an important role in electron scattering.

Universal tools for structure validation -

The veracity of validation efforts by the broad structure community has been vital in creating a robust PDB. Similar efforts are now developing in connection with structures determined by cryo-EM. Among the leading concerns is the issue of accurate determination of absolute length scale – typically embodied by refinement of the EM detector 'pixel size'. Discrepancies in this determination lead to structures suffering from overall stretching or compression, in some cases approaching 1%. This problem recalls parallel challenges addressed in earlier days of x-ray crystallography. By revisiting our earlier treatments of that problem, now in the context of cryo-EM, we have developed algorithms useful to the broad x-ray and EM communities for reliable post-facto analysis of absolute length scale in refined structures. A systematic study highlights the need for vigilance in calibrating widely used experimental facilities.

Collectively, the enabling capabilities we are developing will broadly facilitate the determination and refinement of unknown macromolecular structures with importance for bioenergy.

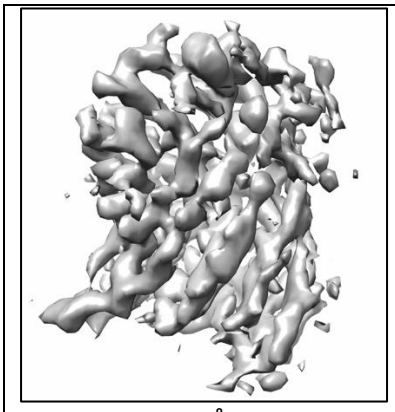

Figure 1. A 3.8 Å cryo-EM map of a small protein (26 kDa GFP).

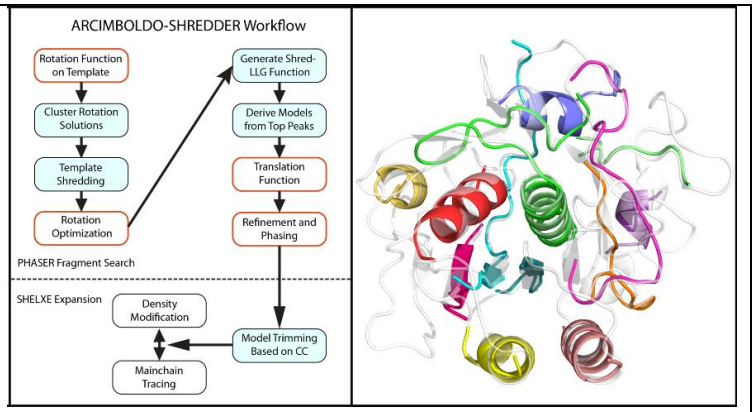

Figure 2. An automated fragment-based phasing pipeline for structure determination from electron diffraction datasets.

References

1. M.C. Thompson, T.O. Yeates, J.A. Rodriguez (2020). Advances in methods for atomic resolution macromolecular structure determination. https://f1000research.com/articles/9-667

This work is supported by DOE Office of Science, BER grant DE-FC02-02ER63421.