1076
doi:10.1017/S1431927622004561
Microsc. Microanal. 28 (Suppl 1), 2022

Electron Counting and Phasing in MicroED

Johan Hattne[1,2], Michael W. Martynowycz[1,2], Max T.B. Clabbers[2] and Tamir Gonen[1,2,3*]

[1.] Howard Hughes Medical Institute, University of California, Los Angeles, CA, United States.
[2.] Department of Biological Chemistry, University of California, Los Angeles, CA, United States.
[3.] Department of Physiology, University of California, Los Angeles, CA, United States.
* Corresponding author: tgonen@g.ucla.edu

Microcrystal electron diffraction (MicroED) is a method in electron cryo-microscopy (cryo-EM) for determining high-resolution structures from minuscule crystals of a wide variety of molecules [1]. In MicroED, the electron microscope is operated in diffraction mode, where the recorded images correspond to a curved slice of reciprocal space rather than a projection of the sample in the beam. For periodic objects such as small, three-dimensional crystals, reciprocal space is comprised of a lattice of discrete, complex-valued peaks. Data collection aims at recording reciprocal space as accurately and completely as possible, whereas the goal of data processing is to reconstruct real space from these measurements, including the crystallographic phases of the Bragg peaks, which cannot be measured during data collection.

The phase problem in crystallography refers to recovering the lost phases of the Bragg spots in a dataset; there is no general solution. For macromolecular MicroED, phases can be calculated from an existing, homologous structure by molecular replacement [2]. Where no sufficiently similar search model can be found, the phase problem turns into a roadblock on the path towards a structure. For small-molecule or peptide MicroED data, prior chemical knowledge can introduce sufficient constraints to derive initial phase estimates from amplitudes alone [3]. Such direct methods are computationally infeasible for large proteins and are expected to fail unless the sample diffracts to about 1.2 Å or better. While molecular replacement is relatively robust with respect to the quality of the data, other phasing methods have more stringent requirements on data quality. In the absence of a suitable search model, it is of utmost importance that data collection yields the most accurate intensities.

Here, a hybrid approach was used to phase MicroED datasets from two macromolecules: lysozyme and proteinase K. First, small idealized helical fragments (a single three-residue fragment in the case of lysozyme, four identical 14-long helices for proteinase K) were placed by molecular replacement. The helix is a structural feature found in virtually all biomolecules, and helical fragments can be matched without any knowledge of the underlying protein. The initial map was then subjected to density modification. While often used for experimental phasing in X-ray crystallography, it has not resulted in interpretable maps for MicroED data before [4]. Given only the amino acid sequence, automated iterations of chain tracing and density modification eventually resulted in a complete protein model.

Molecular replacement with small fragments as well as density modification are critically dependent on the quality of the data, which is determined by both the sample and the measurement. To optimize the sample, crystals initially about 10 μm thick, were thinned with a focused ion-beam (FIB) to about the inelastic mean free path (approximately 300 nm for 300 kV electrons) [5]. FIB milling reproducibly provides lamellae that are thick enough to diffract and simultaneously reduces the probability that an electron interacts with the sample more than once on its way from the electron gun to the detector.

The electron-counting mode on a Falcon 4 direct electron detector was used to collect high-quality MicroED data. Apart from regular reset frames, there is essentially no dead time, which makes this camera well-suited for data collection under continuous rotation of the sample stage. In electron-counting mode it can measure very low electron counts with near-optimal quantum efficiency. Unlike integrating cameras, an electron-counting device counts the individual electrons, such that even weak signals can be accurately integrated. This implies that the flux, and hence the dose and consequently the damage to the sample, can be reduced without compromising the integrity of the measurement. Lowering the flux also reduces the probability that that two electrons arrive on the same pixel during a readout cycle. Because at most one electron can be counted on each frame, electron pileup causes undercounting, and the concomitant loss of accuracy is a major concern when using electron-counting devices for MicroED. In the present measurements, the exposure totals between 0.64 and 1 e^{-} $Å^{-2}$ for any crystal (about 4× less than what was used for measurements on integrating cameras) and essentially all pixel values, even those at low resolution where the probability of pileup is the greatest [6], fall within the linear range of the detector.

The readout noise from electron-counting devices can be effectively ignored and there is no penalty for reading out many frames at high frequency over collecting fewer frames at higher exposures [7]. While the total rotation range in the data obtained from a single crystal in this work (between 63° and 84°) is comparable to previous measurements, the number of frames per dataset is two orders of magnitude greater (well over 14k frames per crystal with the Falcon 4 at 250 Hz, compared to hundred or so images with integrating cameras at ~1 Hz) [8]. Correspondingly, the wedge of reciprocal space recorded on a frame in these electron-counted data is two orders of magnitude narrower than the wedge on a frame from an integrating camera. Such rapid fine slicing of reciprocal space captures the scattering process at much higher resolution in both time and space and has the potential to enable more accurate modelling of spot profiles [9]. However, the high frame rate leads to large datasets which are cumbersome to manage. Even though the sparse distribution of electron events tends to make electron-counted data amenable to data compression (here: ~98% compression ratio), the camera system nevertheless sums frames to overcome internal bandwidth limitations. This is equivalent to increasing the exposure time and widening the wedge recorded on each frame, but also increases the range of valid pixel values on a summed frame. Since most data processing pipelines developed for X-ray crystallography struggle with electron-diffraction images containing single-pixel spots stretched over many consecutive frames, the data were further summed to slices between 0.075° and 0.2° in width. The pixel values were also multiplied by 32 before rounding to integer: once the post-counting gain has been applied, this will preserve five fractional bits of the real-valued rasters. This has been found to be an acceptable compromise between precision and overflow.

Despite the low exposures, data from both lysozyme and proteinase K could be accurately integrated to high resolution (0.87 Å for lysozyme, 1.5 Å for proteinase K). To further increase accuracy and completeness, data from several isomorphous crystals were merged (lysozyme: 16 crystals, 88% complete; proteinase K: 2 crystals, 99% complete). Combined with refined methods for sample preparation, this enabled the fragment-based phasing approach, blending molecular replacement with density modification, but without using any previously known homologous structures.

https://doi.org/10.1017/S1431927622004561 Published online by Cambridge University Press

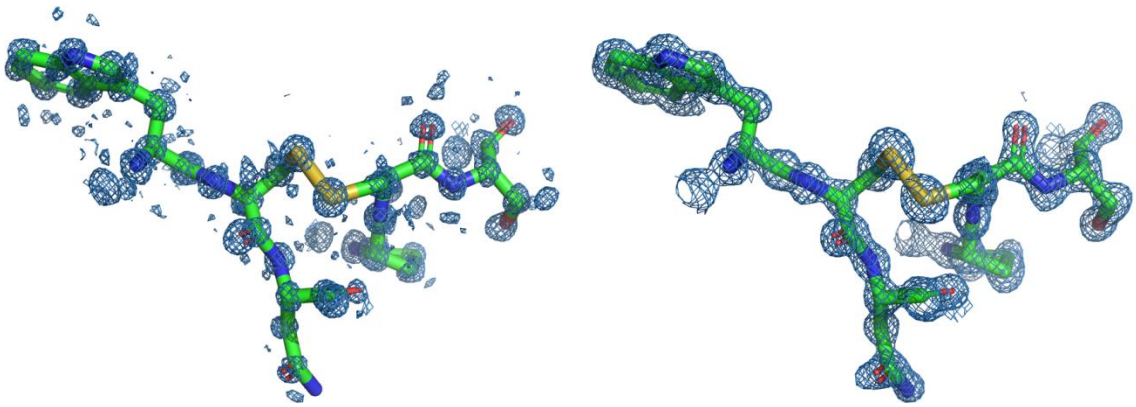

Figure 1. Densities around one of the four disulfide bonds in lysozyme. The meshes are contoured at 1.5 σ above the mean and carved to 2 Å around the selected atoms. **Left:** E_{calc} map after placing a 3-residue helical fragment with *Phaser* and density modification in *ACORN*. This map was automatically interpreted by *Buccaneer*. **Right:** The $2mF_o$-DF_c map after refinement in *REFMAC5*.

References:

[1] D Shi, BL Nannenga, MG Iadanza, T Gonen, eLife. **2** (2013), p. e01345. doi:10.7554/elife.01345
[2] BL Nannenga, T Gonen, Nat Methods. **16** (2019), p. 369. doi:10.1038/s41592-019-0395-x
[3] MR Sawaya et al., Proc Natl Acad Sci. **113** (2016), p. 11232. doi:10.1073/pnas.1606287113
[4] LS Richards et al., Acta Crystallogr Sect Struct Biol. **76** (2020), p. 703. doi:10.1107/S2059798320008049
[5] MW Martynowycz et al., Proc Natl Acad Sci. **118** (2021), p. e2108884118. doi:10.1073/pnas.2108884118
[6] F Leonarski et al., Nat Methods. **15** (2018), p. 799. doi:10.1038/s41592-018-0143-7
[7] M Battaglia et al., Nucl Instrum Methods Phys Res Sect Accel Spectrometers Detect Assoc Equip. **598** (2009), p. 642. doi:10.1016/j.nima.2008.09.029
[8] J Hattne, MW Martynowycz, PA Penczek, T Gonen, IUCrJ. **6** (2019), p. 921. doi:10.1107/S2052252519010583
[9] JW Pflugrath, Acta Crystallogr D Biol Crystallogr. **55** (1999), p. 1718. doi:10.1107/S090744499900935X

https://doi.org/10.1017/S1431927622004561 Published online by Cambridge University Press