Protein structure determination by MicroED
Brent L Nannenga and Tamir Gonen

In this review we discuss the current advances relating to structure determination from protein microcrystals with special emphasis on the newly developed method called MicroED. This method uses a transmission electron cryo-microscope to collect electron diffraction data from extremely small 3-dimensional (3D) crystals. MicroED has been used to solve the 3D structure of the model protein lysozyme to 2.9 Å resolution. As the method further matures, MicroED promises to offer a unique and widely applicable approach to protein crystallography using nanocrystals.

Addresses
Janelia Farm Research Campus, Howard Hughes Medical Institute, 19700 Helix Drive, Ashburn, VA 20147, USA

Corresponding author: Gonen, Tamir (gonent@janelia.hhmi.org)

Introduction
The 3-dimensional (3D) structures of proteins and other biomolecules are an important source of insight into the function and mechanism of biological processes. Through the ability to visualize the structures of proteins and other biomolecules, structural biology has greatly enhanced our understanding of the biochemical processes of life. Of the techniques used to gather structural information, X-ray crystallography is by far the most successful and highly utilized technique, a fact highlighted by the more than 80,000 structures solved by this technique (www.pdb.org [1]). Despite the great advances in X-ray crystallography, the requirement for large well-ordered crystals remains a formidable barrier. For difficult targets such as membrane proteins and protein complexes, the optimization of initial small crystals found during screening can take significant time and resources and may never yield the large well-ordered crystals needed for traditional X-ray crystallography [2].

While continued development of microfocus beamlines have allowed researchers to obtain data from smaller and smaller crystals [3–7], the inherent problem of radiation damage caused by X-rays is one that is not easily overcome. In this review we discuss new methods for protein crystallography, with special focus on the recently developed MicroED method for diffraction of microcrystals and nanocrystals using a transmission electron microscope (TEM). MicroED can bypass the hurdles of crystal size and radiation damage allowing for the use of extremely small crystals for protein structure determination at atomic resolution [8**].

Bypassing the microcrystal radiation damage barrier
Even when diffraction experiments are conducted at cryogenic temperatures, the radiation damage experienced by a protein crystal during data collection will ultimately lead to a reduction in data quality as the experiment progresses [9–12]. Because the crystal must be sufficiently large to withstand the negative effects of radiation damage, there is a lower limit in size of the smallest crystals that will provide useful data by traditional X-ray crystallography. Therefore, there is a great need to find methods that could potentially overcome this obstacle and facilitate data collection from smaller and potentially more locally well-ordered crystals [13].

In an attempt to alleviate this problem, researchers have turned to femtosecond X-ray crystallography [14**,15*], which can use small microcrystals. The basis of this technique is ‘diffact-before-destruction’ where the diffracting X-ray pulse is so intense and quick that the data are obtained on a time scale faster than that of the radiation damage [16,17]. Since its inception the technique has continually been developed and optimized. Recently the method has been improved to the level that the integrated data was of high enough quality to determine phases experimentally [18]. While this technique shows great promise and has been used to determine the structures of several proteins (e.g. Photosystem I [14**], Lysozyme [15*,18], and Cathepsin B [19]), it requires a large number of crystals and access to X-ray lasers, which are relatively new and not widely available.

Another approach for reducing the effects of radiation damage while obtaining diffraction data is to use electrons instead of X-rays. The wavelength of electrons used for diffraction is approximately 50× shorter than the wavelengths used for X-ray diffraction. While this has effects on the Ewald sphere and what the resulting diffraction patterns look like (Figure 1), the principles are essentially the same between electron diffraction and X-ray diffraction. A key difference between the two diffraction techniques is that electrons deposit 2–3 orders of magnitude less energy into a crystal per useful scattering event [20*]. Additionally, electrons interact with matter much more strongly compared to X-rays; therefore, the requirement...
Comparison of diffraction data obtained from lysozyme crystals by electron diffraction and X-ray diffraction. Because the wavelength of the diffracting electrons is so short, the resulting Ewald sphere (left, red line) is essentially a plane when compared to the Ewald sphere for X-ray diffraction (right, red line). Diffraction only occurs when the Ewald sphere contacts a reflection in reciprocal space (top panels, white circles represent reflections in reciprocal space). Therefore, because the Ewald sphere is so flat, the patterns produced from electron diffraction (bottom left) appear as planar 2-dimensional slices through the 3-dimensional volume of reflections, whereas the patterns from X-ray diffraction (bottom right) appear as circular 2-dimensional projections of the sphere on the detector.

Electron diffraction of 3D protein crystals in the TEM has been attempted over the years but none yielded a refined structure [25–28]. One of the major hurdles was the beam damage associated with data collection, which typically allowed only one diffraction pattern to be collected per crystal [28]. By only having one pattern per crystal it is difficult to properly index the reflections and determine the crystallographic orientation especially as the data is affected by the shape function.

MicroED is a new method that was used successfully to determine protein structure by electron diffraction from microcrystals [8**]. In MicroED an extremely low electron dose is used to collect multiple electron diffraction patterns from each crystal at varying angles (Figures 2A and 3). Using this approach, the first complete high-resolution structure of a protein was reported from 3D microcrystals in a TEM. The remaining sections of this review will be devoted to MicroED and discussions on the future directions of this technique.

**MicroED sample preparation and data collection**

As with X-ray crystallography, conditions must be optimized in MicroED for sample preparation, cryo protection, and data collection strategy.

In order to prepare samples for MicroED, microcrystals are taken directly from the drops where they were formed and deposited on a carbon coated electron microscopy grid. The sample must have the excess liquid removed by blotting with filter paper, as the electron beam is unable to penetrate thick samples. Therefore, some optimization of blotting is necessary to strike the right balance between leaving enough solution so that the crystals are well-hydrated and preserved, but not so much that the sample is too thick for electron diffraction. Conditions for sample preparation can be screened by negative stain EM [29], however this only provides a rough starting point as
MicroED data collection and microcrystal visualization. (A) Schematic of MicroED data collection by tilting the sample stage between subsequent exposures. Each exposure is of relatively low dose (∼0.01 e⁻/Å²/s) which allows the collection of multiple diffraction patterns from a single crystal that are combined into a single data set. Each data set consists of up to 90 still frames taken at 0.1°–1° intervals. (B,C) Visualizing microcrystals (arrows) in the TEM prior to data collection. The dimensions of microcrystals suitable for MicroED range from approximately 1–10 μm in length and width and 0.1–1 μm in thickness. The crystals of a membrane transporter (B) visualized by negative stain EM, whereas the crystals of a novel designed protein (C) are seen with cryo-EM in over-focused diffraction mode.

conditions can deviate significantly when using stains and will need to be further optimized for cryoEM.

Once the sample is loaded into the microscope, the grid can be screened for microcrystals (Figure 2B and C). When a microcrystal is located, a test diffraction pattern is first collected to determine the quality of the crystal. Factors that are considered when assessing diffraction quality include, but are not limited to, the intensity and sharpness of the recorded reflections and maximum resolution obtained [30**]. Examples of high quality diffraction patterns can be seen in Figure 3. Once a promising microcrystal of good quality is found, a diffraction data set is collected by exposing the crystal with extremely low dose (∼0.01 e⁻/Å²/s) followed by tilting the crystal using the microscope stage and repeating this cycle of exposure and tilting (Figure 2A) until the crystal succumbs to the effects of radiation damage. The crystal is tilted by a constant value (typically 0.1°–1.0°) between exposures, which creates a data set of still diffraction patterns taken at defined intervals of rotation. Because it is clear how each still pattern relates to all other diffraction patterns within the data set, the orientation of the crystal can be determined and the data set correctly indexed and integrated [31**].

The stage of the electron microscope limits the attainable tilt angle to ±70°. This means that a single crystal can yield a 140° wedge of data. Depending on crystal symmetry, the degree of tilting performed, and the crystal’s orientation in the microscope, there may be missing regions of data, which will need to be filled in by more data sets originating from additional crystals (Figure 4). Each crystal would have to be oriented differently on the grid to allow ample sampling in reciprocal space. Following data integration it is possible to identify the orientation of the crystal and find out how much of the reciprocal space of the crystal was covered during data collection. By using this knowledge, data sets that complement each other can be combined, effectively filling in the data by removing missing wedges and improving completeness. Once a complete data set is obtained, the structure can be phased and refined using standard programs for X-ray crystallography such as PHENIX [32], CNS [33] and the CCP4 program suite [34]. As a proof of principle, this entire process was performed with lysozyme microcrystals giving rise to a refined structure of lysozyme at 2.9 Å resolution (Figure 4) and an Rwork/Rfree of 25.5/27.8 [8**]. Diffraction was recorded to 1.7 Å, however the data was truncated to 2.9 Å to ease processing with in house developed programs. In the future we expect that the resolution will be extended by integrating MicroED data with standard crystallography software and optimization of data collection strategies.
Examples of tilted diffraction data of Lysozyme collected by MicroED. Representative crystal images (left) and diffraction patterns (right) taken at tilts of 20°, 0°, and −20°. All images and diffraction patterns are from the same crystal. The MicroED diffraction data shows distinct reflections with a maximum resolution of 1.7 Å.
Combining data sets and solving the final structure. (A) The consequences of data completeness (left side of panels) on the resulting density maps (right side of panels). When a single data set from lysozyme having a large missing wedge of data is used (shown as orange in the (0 1 0) single data set panel), the data quality in the direction containing complete data (along (0 0 1)) shows reasonable density, whereas the density in the direction of the missing data (along (0 1 0)) is extremely poor. By combining additional data sets that include the missing reflections from the original single data set, the quality of the density can be recovered and a complete 3D map obtained. The protein backbone of lysozyme (Cα trace) is shown in blue in the density maps. (B) Stereo view of the final 2.9 Å structure of lysozyme solved by MicroED. The 2F_{obs}−F_{calc} map (contoured at 1.5σ) of the representative region shown displays well-defined density around the side chains and nearby waters.
Crystal size in MicroED
As was mentioned above, electrons interact with matter relatively strongly and this places limitations on the upper size limit of 3D crystals that can be used by MicroED. The effects of diffuse scattering and dynamic scattering within the crystal both increase as the crystal thickness increases, and while these are generally assumed to be negligible in X-ray crystallography, their effects cannot be completely ignored in electron diffraction. Diffuse scattering, which is caused by partial disorder within the crystal as well as inelastic scattering, leads to increased background noise and errors in the measurement of reflection intensities [35]. Dynamic scattering occurs when an electron that has been scattered elastically undergoes a second elastic scattering event as it exits the sample. Because both dynamic scattering events follow Bragg’s law, the effect is that a portion of the electrons destined to be counted as intensity for a specific reflection are redistributed to other reflections in the diffraction data set leading to inaccuracies in the resulting integrated data [36,37].

We found that Lysozyme crystals approximately 500 nm or thinner were not completely overcome by the effects described above and could provide usable data, whereas thicker crystals were essentially unusable. It is likely that different crystals will have different thickness limits as the density and the packing within the crystals will affect the mean free path of the electrons through the sample. Therefore, the usable range of crystal size will need to be determined for each microcrystal sample experimentally.

Data processing
In the current implementation of MicroED, the diffraction patterns are collected as still exposures rather than the oscillation method commonly used in X-ray crystallography. Because of this, the recorded reflections within a data set are typically partial reflections. To work with a data set containing only partial intensities, in-house programs and scripts were written to index the data and group all symmetry related intensities [31]. One of the main assumptions used to handle MicroED data is that the maximum intensity within a symmetry related group would originate from a diffraction pattern where the Ewald sphere passes through the largest cross-section of that given reflection. Therefore, only the maximum intensity within each symmetry related group was kept and taken to represent the most accurate measurement of the full intensity. By using this assumption, along with a few others, the data sets are concatenated and reflection files are generated. The final reflection files are then processed using standard refinement programs for X-ray crystallography as described above.

Future directions for MicroED
MicroED is a relatively new method and has room for much optimization with future work focusing on sample preparation, data collection strategies and data processing and phasing. Specifically, one area of work that we are currently focusing on is moving the technique away from the still diffraction patterns towards the collection of data from crystals that are oscillated. The obvious benefit is that the data sets collected this way would contain the measurement of full intensities as opposed to the partial intensities that still diffraction patterns provide.

Another method of collecting full intensities, which could be implemented in MicroED, is beam precession. This method relies on slightly tilting the beam and rotating it during data collection. Precession has been shown to provide fuller, more accurate intensities and reduce the effects of dynamic scattering, leading to greatly improved data quality [38–40].

MicroED promises to advance the field of structural biology by offering a parallel path to structure determination that has so far been dominated by X-ray crystallography. While there is much work to be done improving and optimizing the technique, this is an exciting time for electron crystallographers and those who would like to use micro and nanocrystals. While there will always be a place for traditional X-ray crystallography, one day — with the help of MicroED, microfocus X-ray beam lines, and femtosecond crystallography — many protein structures may be solved from extremely small crystals which have traditionally been discarded as undesirable.

Acknowledgements
We thank Dan Shi and Matthew G. Iadanza (JFRC) helpful insights and discussion. This work is supported by the Howard Hughes Medical Institute.

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest


8. Shi D, Nannenga BL, Iadanza MG, Gonen T: Three-dimensional electron crystallography of protein microcrystals. Elife 2013:2. This paper lays the foundation for MicroED and describes the proof of concept principles for the determination of the structure of lysozyme to 2.9 Å by electron diffraction of 3-dimensional protein microcrystals.


First high resolution structure (lysozyme to 1.9 Å) using XFELs and femtosecond X-ray crystallography.


Insightful overview of the benefits and shortcomings of neutrons, electrons and X-rays for structure determination. Comparisons between the effects of radiation damage between electrons and X-rays are particularly pertinent to this MicroED.


This study presents the first structure of a membrane protein which was solved by 2-dimensional electron crystallography. This paper helped to set the stage for the field of 2-dimensional crystallography which has solved the structures of a number of membrane proteins within a lipid membrane.


This work provides a thorough description of the steps necessary to collect diffraction data with a TEM.


This paper describes the programs and methodology which were initially used for processing MicroED data.


