

Small Molecule Microcrystal Electron Diffraction (MicroED) for the Pharmaceutical Industry – Results from Examining Over Fifty Samples

Jessica Bruhn¹, Giovanna Scapin², Anchi Cheng¹, Thejusvi Ganesh¹, Sargis Dallakyan¹, Brandon Read¹, Travis Nieuwsma³, Kyle Lucier³, Megan Mayer³, Nicole Chiang³, Nicole Poweleit³, Philip McGilvray³, Timothy Wilson³, Michael Mashore³, Camille Hennessy³, Sean Thomson³, Clinton Potter³ and Bridget Carragher⁴

¹NanoImaging Services, San Diego, California, United States, ²Nanoimaging Services, New York, New York, United States, ³NanoImaging Services, United States, ⁴New York Structural Biology Center, New York, New York, United States

Abstract

The emerging field of microcrystal electron diffraction (MicroED) is of great interest to industrial researchers working in the drug discovery and drug development space. The promise of being able to routinely solve high-resolution crystal structures without the need to grow large crystals is very appealing. Despite MicroED's exciting potential, adoption across the pharmaceutical industry has been slow, primarily owing to a lack of access to specialized equipment and expertise. Here we present our experience building a small molecule MicroED service pipeline for the pharmaceutical industry. In the past year, we have examined more than fifty small molecule samples submitted by our clients, the majority of which have yielded data suitable for structure solution. This experience has led us to conclude that small molecule MicroED adoption will continue to grow within the pharmaceutical industry where it is able to rapidly provide structures inaccessible by other methods.

Introduction

Access to accurate small molecule structural information in a timely manner is important at all stages of drug development. Medicinal chemists have a variety of tools at their disposal to access such structural information. Techniques like mass spectrometry and nuclear magnetic resonance are routinely employed to aid in structure determination, but these techniques often produce somewhat ambiguous results, especially for larger compounds and those with chiral centers. Many groups are forced to turn to structural biologists to determine structures of the compound of interest bound to a protein target, but this approach can be quite challenging depending on the target. Single crystal X-ray diffraction of small molecule crystals is an attractive option as it generally produces structures at extremely high resolution ($< 1 \text{ \AA}$) and routinely provides the absolute configuration (handedness of each stereocenter in the compound). However, many compounds of pharmaceutical interest do not readily form crystals large enough for X-ray diffraction. Recently, MicroED has emerged as an option to fill this gap, producing high-resolution structures from crystals orders of magnitude smaller than those required for X-ray studies (1,2). In many cases, small molecules spontaneously form microcrystals suitable for MicroED data collection, eliminating the need for crystallization trials.

NanoImaging Services, a contract research organization, launched small molecule MicroED data collection and structure determination services in January of 2020. Here we describe our experience validating our pipeline and collecting data from over fifty pharmaceutical small molecule samples.

Materials and Methods

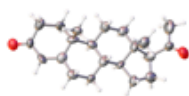
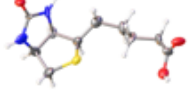
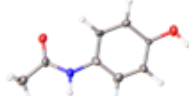
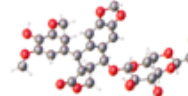
Grids for small molecule MicroED data collection were typically prepared by directly applying powdered compounds to a continuous carbon grid. Electron microscopy was performed using a Thermo Fisher Scientific Glacios Cryo Transmission Electron Microscope operated at 200kV and at -192°C . The

microscope was equipped with a Thermo Fisher Scientific CETA-D detector. Diffraction datasets were collected under parallel illumination conditions with a very low dose rate and a small condenser aperture (20 μm). Data collection was carried out with Legimon, which has been adapted to collect MicroED data (3). Crystal targets are manually selected at low magnification and added to a queue for automated data collection. Once the queue has been submitted, Legimon moves to and centers each crystal in imaging mode. Then the crystal is continuously rotated (typically from -60° to $+60^\circ$) with the microscope in diffraction mode and the camera collecting continuously in rolling shutter mode. The images are converted to SMV format with the appropriate metadata needed for data processing in standard crystallographic programs. Data reduction is carried out in DIALS (4) and structure solution is carried out in SHELXT, SHELXD and/or PHASER (5-7).

Results and Discussion

To validate our pipeline, we determined the structures of three compounds that had previously been determined by X-ray and electron diffraction methods: progesterone, biotin and paracetamol (1,2). Our structures produced encouraging data reduction and refinement statistics, especially compared to other MicroED structures in the literature (Figure 1). To further test our pipeline, we set out to solve the structure of a large organic compound that had not been previously solved by single crystal X-ray diffraction. We chose teniposide, an anti-cancer drug that has been on the market for almost two decades. We speculate that the lack of a crystal structure of this compound may have been due to difficulties growing large enough crystals for single-crystal X-ray diffraction measurements. We purchased this compound and were able to quickly determine the MicroED structure to 1.2 \AA and unambiguously assign all heavy atoms (Figure 1).

Since beginning our MicroED efforts, we have collected data from over fifty sample submitted by our clients. The vast majority of samples (51/60) produced data that we expect is amenable to structure solution (Figure 2A). Nearly 80% of small molecule data collected at NIS achieved 1.1 \AA resolution or higher according to $\text{CC1/2} \geq 0.33$ (Figure 2B), which is very important as it is difficult to solve structures at lower resolution using direct methods. Additionally, our pipeline produced data with at least 84% completeness for all samples examined, with more than half achieving 100% completeness. This is important as MicroED data collection is limited by the tilt range of the microscope stage. Despite this limitation, our pipeline can still produce highly complete data by collecting data from multiple crystals randomly oriented on the TEM grid. Taken together, our experience gives us great optimism for the role MicroED has played and will continue to play in the pharmaceutical industry.

	Progesterone C ₂₁ H ₃₀ O ₂ 	Biotin C ₁₀ H ₁₆ N ₂ O ₃ S 	Paracetamol C ₈ H ₉ NO ₂ 	Teniposide C ₃₂ H ₃₂ O ₁₃ S 
Data Collection				
Oscillation per frame (°)	0.9	0.89	0.89	0.89
Camera length (mm)	1186.5	1065.7	955.8	1065.7
Datasets combined	4	4	4	6
Temperature (°C)	-192	-192	-192	-192
Accelerating Voltage (kV)	200	200	200	200
Wavelength (Å)	0.02501	0.02501	0.02501	0.02501
Dose (e ⁻ /Å ² /°)	~0.045	~0.022	~0.022	~0.022
Data Reduction				
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ /n	P2 ₁ 2 ₁ 2 ₁
Unit cell a, b, c (Å)	10.28, 12.55, 13.50	5.14, 10.36, 20.80	7.07, 9.19, 11.49	9.27, 15.42, 20.50
Unit Cell α, β, γ (°)	90, 90, 90	90, 90, 90	90, 98.64, 90	90, 90, 90
Resolution (Å) ^{&}	13.60 – 1.05 (1.07 – 1.05)	20.83 – 0.85 (0.87 – 0.85)	11.42 – 0.80 (0.81 – 0.80)	20.49 – 1.20 (1.22 – 1.20)
Total reflections ^{&}	10,900 (421)	11,275 (231)	10,031 (331)	15,834 (792)
Unique reflections ^{&}	949 (43)	1,125 (47)	1,350 (67)	1,074 (51)
Completeness (%) ^{&}	100.0 (100.0)	97.9 (82.5)	84.0 (82.7)	100.0 (100.0)
Multiplicity	11.5 (9.8)	10.0 (4.9)	7.4 (4.9)	14.7 (15.5)
Mean I/σI ^{&}	8.5 (1.2)	5.4 (0.2)	3.8 (0.4)	6.7 (1.0)
R _{meas} ^{&}	0.25 (1.25)	0.26 (1.62)	0.30 (1.00)	0.26 (1.35)
R _{pim} ^{&}	0.07 (0.38)	0.08 (0.47)	0.10 (0.43)	0.07 (0.33)
CC _{1/2} ^{&}	0.993 (0.675)	0.988 (0.471)	0.975 (0.527)	0.995 (0.733)
Refinement				
Z / Z'	4 / 1	4 / 1	4 / 1	4 / 1
Total Reflections	10,209	10,626	9,049	14,062
R _{int} / R _{sigma}	0.22 / 0.16	0.23 / 0.38	0.26 / 0.32	0.25 / 0.20
Data / Restraints / Parameters	1,572 / 0 / 212	1825 / 0 / 147	1240 / 66 / 103	1,776 / 329 / 397
R ₁ / wR ₂ (all data)	0.13 / 0.22	0.18 / 0.26	0.20 / 0.31	0.16 / 0.26
R ₁ / wR ₂ (I ≥ 2σ(I))	0.09 / 0.20	0.10 / 0.24	0.12 / 0.29	0.10 / 0.23
Goof	0.952	1.000	1.104	0.938

[&] Numbers in parentheses refer to the highest resolution shell

Figure 1. Summary of the data collection, reduction, and refinement statistics for four structures determined by MicroED at NanoImaging Services Inc. An image of each crystal structure with atoms represented as thermal ellipsoids is included.

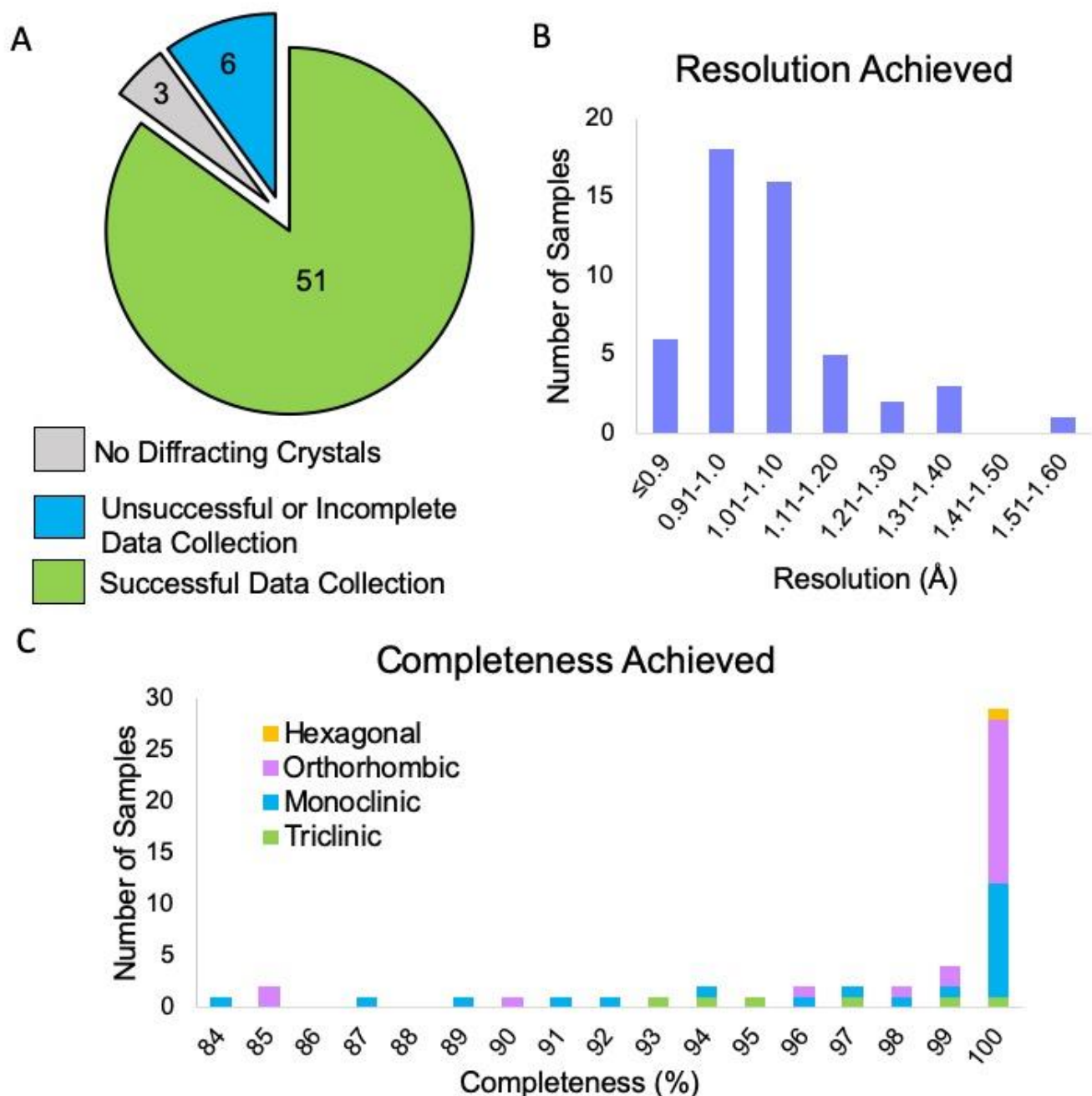


Figure 2. A) The outcome of MicroED data collection at NanoImaging Services for client and internal samples (progesterone, biotin, paracetamol and teniposide). Three samples did not contain diffracting crystals, but we note that X-ray powder diffraction (XRPD) pre-screening was not performed for these samples prior to submission. Six samples did not yield enough data suitable for structure solution and likely need additional microscope time possibly with an alternative grid preparation method. Fifty-one samples produced data that should be sufficient for structure solution. B) For the 51 successful data collections, the approximate resolution limits are shown. This resolution estimate is based on automated processing results from DIALS for the highest resolution crystal. This value is based on $CC1/2 \geq 33\%$. We note that the final resolution after processing by a trained crystallographer may deviate from these values. C) Of these data collections, the overall completeness is shown broken down by Bravais lattice

type. Automated processing in DIALS was carried out in the space group suggested by dials.cosym with all crystals for which integration was successful. Overall completeness is reported to the resolution limit of the crystals as determined in (B). The single hexagonal crystal was in space group P64.

References

1. Gruene, T., J.T.C. Wennmacher, C. Zaubitzer, J.J. Holstein, J. Heidler, A. Fecteau-Lefebvre, S. De Carlo, E. Muller, K.N. Goldie, I. Regeni, T. Li, G. Santiso-Quinones, G. Steinfeld, S. Handschin, E. van Genderen, J.A. van Bokhoven, G.H. Clever, and R. Pantelic. 2018. Rapid Structure Determination of Microcrystalline Molecular Compounds Using Electron Diffraction. *Angew Chem Int Ed Engl.* 57:16313-16317.
2. Jones, C.G., M.W. Martynowycz, J. Hattne, T.J. Fulton, B.M. Stoltz, J.A. Rodriguez, H.M. Nelson, and T. Gonen. 2018. The CryoEM Method MicroED as a Powerful Tool for Small Molecule Structure Determination. *ACS Central Science.* 4:1587-1592.
3. Cheng, A., C. Negro, J.F. Bruhn, W.J. Rice, S. Dallakyan, E.T. Eng, D.G. Waterman, C.S. Potter, and B. Carragher. 2021. Leginon: New features and applications. *Protein Sci.* 30:136-150.
4. Clabbers, M.T.B., T. Gruene, J.M. Parkhurst, J.P. Abrahams, and D.G. Waterman. 2018. Electron diffraction data processing with DIALS. *Acta Crystallographica Section D: Structural Biology.* 74:506-518.
5. Sheldrick, G.M. 2015b. SHELXT - integrated space-group and crystal-structure determination. *Acta Crystallogr A Found Adv.* 71:3-8.
6. Schneider, T.R., and G.M. Sheldrick. 2002. Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr.* 58:1772-1779.
7. McCoy, A.J., R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, and R.J. Read. 2007. Phaser crystallographic software. *J Appl Crystallogr.* 40:658-674.