



# Small Molecule Microcrystal Electron Diffraction for the Pharmaceutical Industry—Lessons Learned From Examining Over Fifty Samples

## OPEN ACCESS

### Edited by:

Dan Shi,  
Center for Cancer Research (NCI),  
United States

### Reviewed by:

Wei Zhao,  
California Institute of Technology,  
United States  
Brent L. Nannenga,  
Arizona State University, United States

### \*Correspondence:

Bridget Carragher  
bcarr@nanomimagingsservices.com  
Jessica F. Bruhn  
jbruhn@Nanolmagingsservices.com

### † Present address:

Megan L. Mayer,  
The Harvard Cryo-Electron  
Microscopy Center for Structural  
Biology, Harvard Medical School,  
Boston, MA, United States

### Specialty section:

This article was submitted to  
Structural Biology,  
a section of the journal  
Frontiers in Molecular Biosciences

**Received:** 19 January 2021

**Accepted:** 16 April 2021

**Published:** 12 July 2021

### Citation:

Bruhn JF, Scapin G, Cheng A,  
Mercado BQ, Waterman DG,  
Ganesh T, Dallakyan S, Read BN,  
Nieusma T, Lucier KW, Mayer ML,  
Chiang NJ, Poweleit N, McGilvray PT,  
Wilson TS, Mashore M, Hennessy C,  
Thomson S, Wang B, Potter CS and  
Carragher B (2021) Small Molecule  
Microcrystal Electron Diffraction  
for the Pharmaceutical  
Industry—Lessons Learned From  
Examining Over Fifty Samples.  
*Front. Mol. Biosci.* 8:648603.  
doi: 10.3389/fmolb.2021.648603

Jessica F. Bruhn<sup>1\*</sup>, Giovanna Scapin<sup>1</sup>, Anchi Cheng<sup>1</sup>, Brandon Q. Mercado<sup>2</sup>,  
David G. Waterman<sup>3,4</sup>, Thejusvi Ganesh<sup>1</sup>, Sargis Dallakyan<sup>1</sup>, Brandon N. Read<sup>1</sup>,  
Travis Nieusma<sup>1</sup>, Kyle W. Lucier<sup>1</sup>, Megan L. Mayer<sup>†</sup>, Nicole J. Chiang<sup>1</sup>, Nicole Poweleit<sup>1</sup>,  
Philip T. McGilvray<sup>1</sup>, Timothy S. Wilson<sup>1</sup>, Michael Mashore<sup>1</sup>, Camille Hennessy<sup>1</sup>,  
Sean Thomson<sup>1</sup>, Bo Wang<sup>5</sup>, Clinton S. Potter<sup>1</sup> and Bridget Carragher<sup>1\*</sup>

<sup>1</sup> Nanomimaging Services, San Diego, CA, United States, <sup>2</sup> Department of Chemistry, Yale University, New Haven, CT, United States, <sup>3</sup> UKRI STFC Rutherford Appleton Laboratory, Didcot, United Kingdom, <sup>4</sup> CCP4, Research Complex at Harwell, Rutherford Appleton Laboratory, Didcot, United Kingdom, <sup>5</sup> Biogen, Cambridge, MA, United States

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 members of 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. We also detail our experience determining small molecule MicroED structures of pharmaceutical interest and offer some insights into the typical experimental outcomes. 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.

**Keywords:** crystallography, transmission electron microscopy, electron diffraction (ED), MicroED, drug development, structure, small molecule, medicinal chemistry

## INTRODUCTION

The three-dimensional structure of a molecule provides detailed information regarding relative atom positions, bonding and intra-molecular interactions, which conversely informs stability, reactivity, solubility, and ultimately function. For example, in catalytic chemistry, information regarding the catalyst structure and grain morphology greatly helps in mechanism characterization and process optimization (Wennmacher et al., 2020). Knowledge of the active enantiomeric form of pharmaceutical compounds has become essential in recent years, as generally only one enantiomer is active, and the other(s) may be inactive or even unsafe (Ariëns, 1984). Determination of the 3D structure of synthetic organic reaction products or by-products can also allow for

the understanding of the reaction mechanism, and possibly optimization of the reaction itself. Structural characterization of small molecules has thus far primarily been carried out using nuclear magnetic resonance (NMR), mass spectrometry (MS), and X-ray diffraction (XRD), with single-crystal XRD being the preferred method when a high degree of confidence in the 3D model is required.

In the development of small molecule pharmaceuticals, knowledge of the crystalline solid form is extremely important, beyond confirmation of the proposed 2D structure. The crystal form adopted by an active pharmaceutical ingredient (API) and the lattice interactions that hold it together can have important ramifications for stability, tableting properties, solubility and dissolution rates, ultimately affecting bioavailability, potency, and even toxicity (Lu and Rohani, 2009; Censi and Di Martino, 2015). It is therefore critically important to understand the solid form(s) of APIs and the crystal lattice interactions that underly them. The majority of APIs are estimated to be able to exist in more than one polymorphic form (Raza et al., 2014). Polymorphs are alternative crystal forms of the same compound with different crystal lattice properties, which can give rise to drastically different drug properties. Selecting the optimal polymorph or other crystal form (salt, solvate, hydrate, or co-crystal) is therefore very important for drug formulation. Additionally, as the vast majority of drug substances currently being developed exhibit solubility issues, formulation chemists are increasingly hard pressed to find crystal forms that enhance solubility. It is therefore highly desirable to be able to easily determine the structure of the many crystal forms an API can adopt in order to understand the underlying lattice properties and better engineer the optimal crystal form for development.

Thus far, the vast majority of small molecule crystal structures have been determined by single-crystal XRD, but a major bottleneck for this method is that it requires the generation of large, well-ordered crystals. For synchrotron radiation sources, the smallest suitable crystals are typically around 5–10  $\mu\text{m}$ , and in-house X-ray sources require crystals about ten times larger than this (Gruene et al., 2018). When suitably sized crystals cannot be produced, X-ray powder diffraction (XRPD) (Thakral et al., 2018) can sometimes be used to determine a three-dimensional structure of the compound. While technical advances over the last 20 years have made it feasible to use XRPD to determine the crystal structures of materials of moderate complexity (Harris, 2012), determining crystal structures from powder diffraction remains significantly more challenging than single crystal diffraction data, particularly in the case of larger organic molecules for which there can be many overlapping peaks in the powder diffraction pattern.

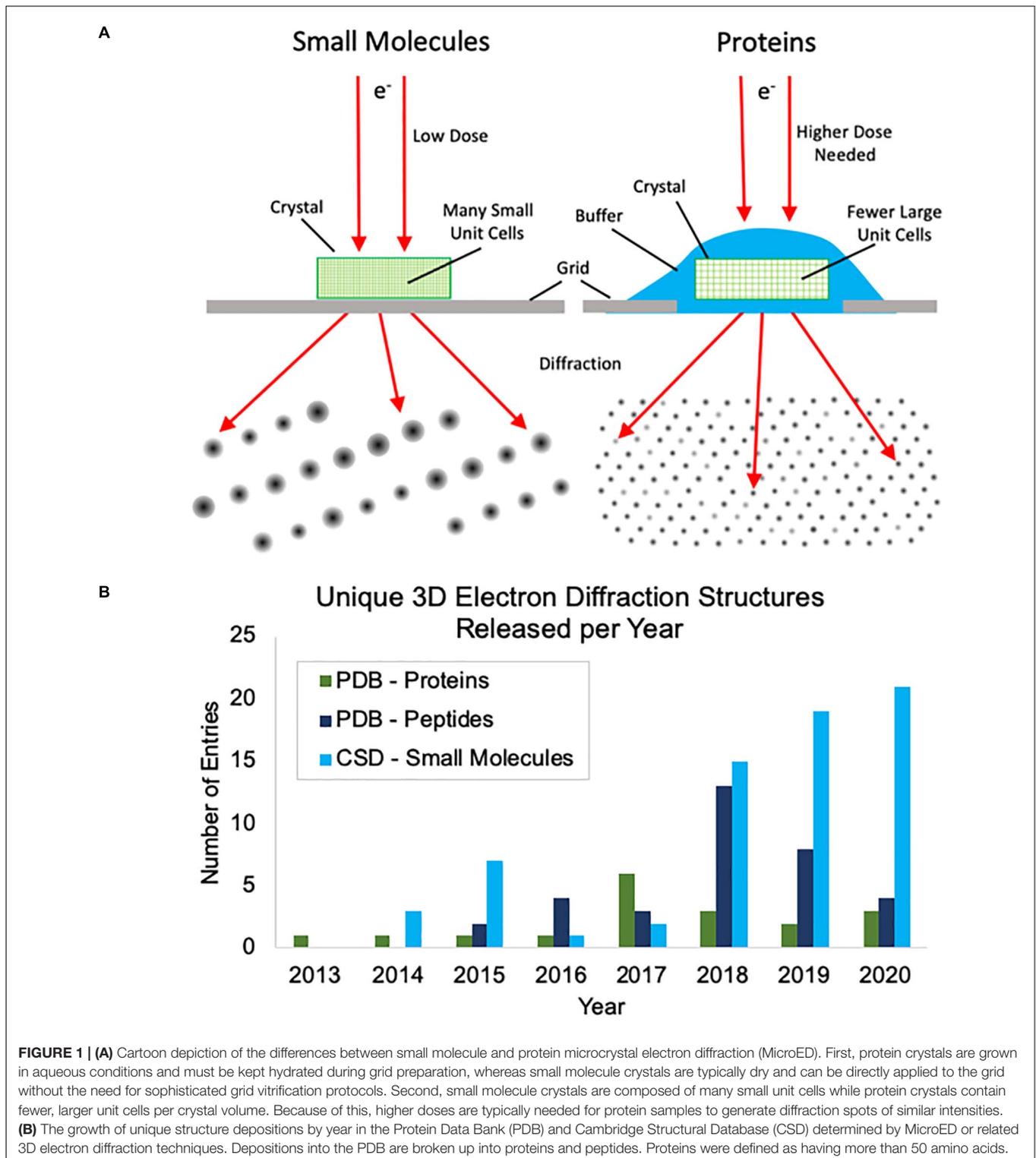
Very recently, microcrystal electron diffraction (MicroED) or 3D electron diffraction using the continuous rotation method has emerged as an attractive alternative technique for structure determination of proteins, peptides, inorganic compounds, and small organic molecules (Shi et al., 2013; Clabbers et al., 2018; Gruene et al., 2018; Jones et al., 2018). Electrons interact very strongly with matter, allowing for data collection from sub-micron sized crystals. This dramatically expands the possibilities for determining structures of crystal forms which do not readily

grow large crystals, such as the several thousand APIs for which only crystalline powders are available (Gruene et al., 2018). In many cases, screening for crystallization conditions can be eliminated entirely when MicroED is employed as many compounds readily form crystals during purification, thus saving time and material.

Thus far, most of the academic interest in, and development of, MicroED has centered around protein samples where there is a great need for methods to address samples that cannot produce large crystals. Protein MicroED presents unique challenges primarily in the realm of crystal screening, sample grid preparation and data collection. Hydrated crystals are very sensitive to changes in their environment and are generally applied to grids along with the crystallization solution or other buffer. This buffer must then be removed or sufficiently thinned to allow for penetration by the electron beam, without damaging the crystals, and this can be challenging to achieve. Many groups have resorted to using focused ion beams to mill through buffer coated crystals, adding yet another very expensive and labor-intensive step to the process (Duyvesteyn et al., 2018; Martynowycz et al., 2019; Beale et al., 2020). Small molecule crystals fortunately do not pose the same challenges for grid preparation, because they are generally dry powders that can easily be broken into suitably small crystals and can be directly applied to grids (**Figure 1A**). Additionally, protein crystals generally produce lower intensity diffractions spots at the same electron dose relative to small molecule crystals (**Figure 1A**) and this results in more radiation damage as the doses that must be used for protein crystal diffraction are generally slightly higher than for small molecules. Taken together, these factors make small molecule MicroED much easier to practically implement than protein MicroED and have likely contributed to the faster growth of small molecule structures being solved by MicroED and other, related 3D electron diffraction techniques. This trend can be seen when comparing the unique protein structures determined by MicroED/3D electron diffraction deposited in the Protein Data Bank (PDB) with the unique peptide structures deposited in the PDB and small molecule structures deposited in the Cambridge Structural Database (CSD) in recent years (**Figure 1B**).

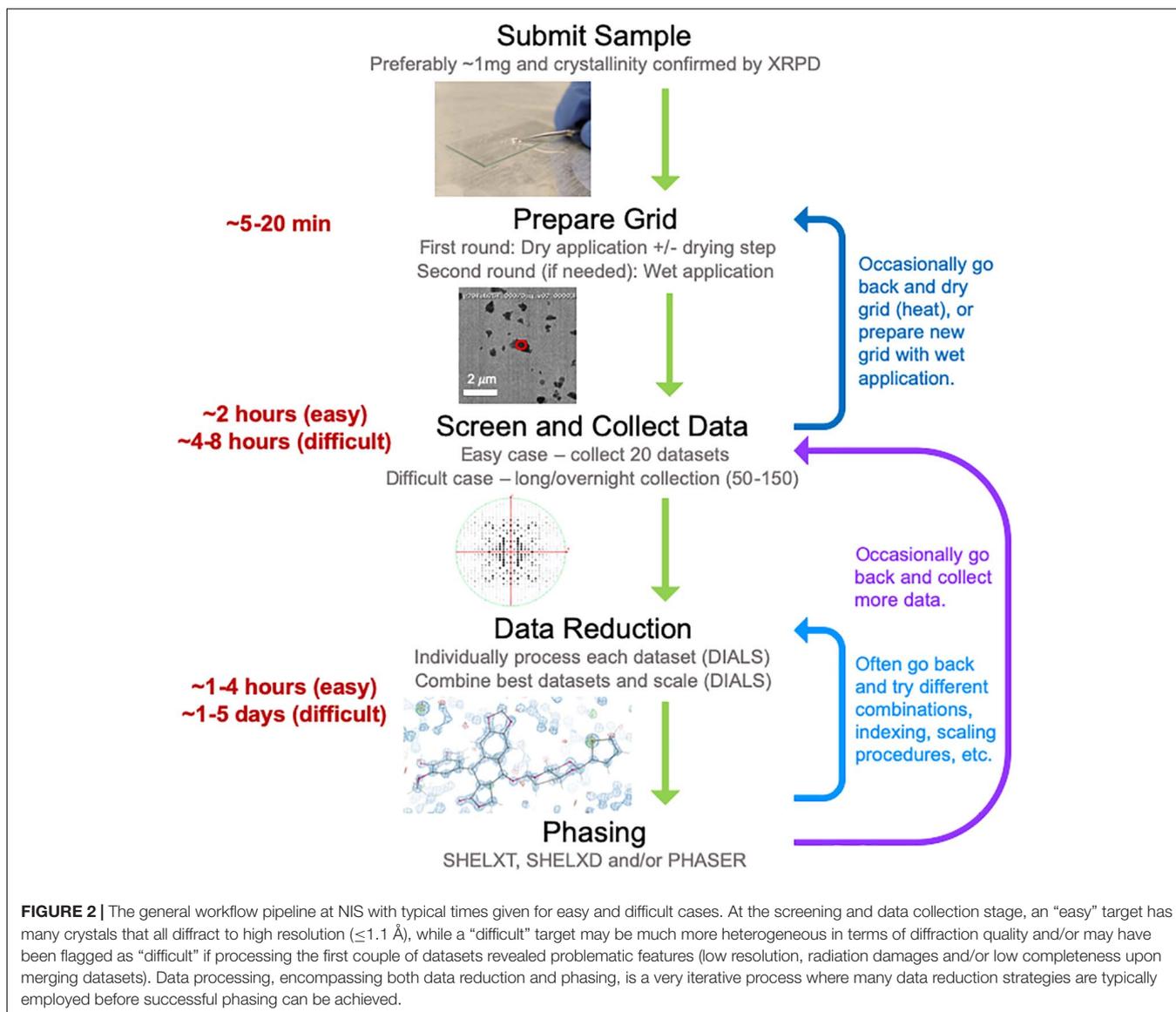
Despite the rapid growth of the field, small molecule MicroED adoption has been slow, especially in industrial settings. Although many research groups recognize the great potential of small molecule MicroED, widescale adoption is inhibited by steep requirements for instrumentation, infrastructure, and expertise. This is especially the case if only a limited number of projects are contemplated by a single group, making it hard to justify the required capital investments. Improvements in access to service facilities, in speed and automation of data collection and processing, and in the ease with which structure solution and refinement can be completed, will help to make small molecule MicroED far more broadly accessible, especially for researchers in the pharmaceutical industry.

Here we describe our experience establishing a MicroED data collection and structure determination service pipeline (**Figure 2**) at NanoImaging Services (NIS), a commercial service provider to the biopharmaceutical and biotechnology industries. To establish this pipeline, we adapted our existing data collection



software system, Legimon (Suloway et al., 2005), enabling the collection of continuous-rotation, electron-diffraction data with the camera in rolling-shutter mode, incorporating rapid crystal preview evaluation, and fully automating data collection of a queue of pre-selected crystal targets (Cheng et al., 2021).

We also established a robust data processing pipeline built on the DIALS (Clabbers et al., 2018) and SHELX (Schneider and Sheldrick, 2002; Sheldrick, 2015b) suites. Using this pipeline, we determined the structures of three small molecule validation targets (progesterone, biotin, and paracetamol) and



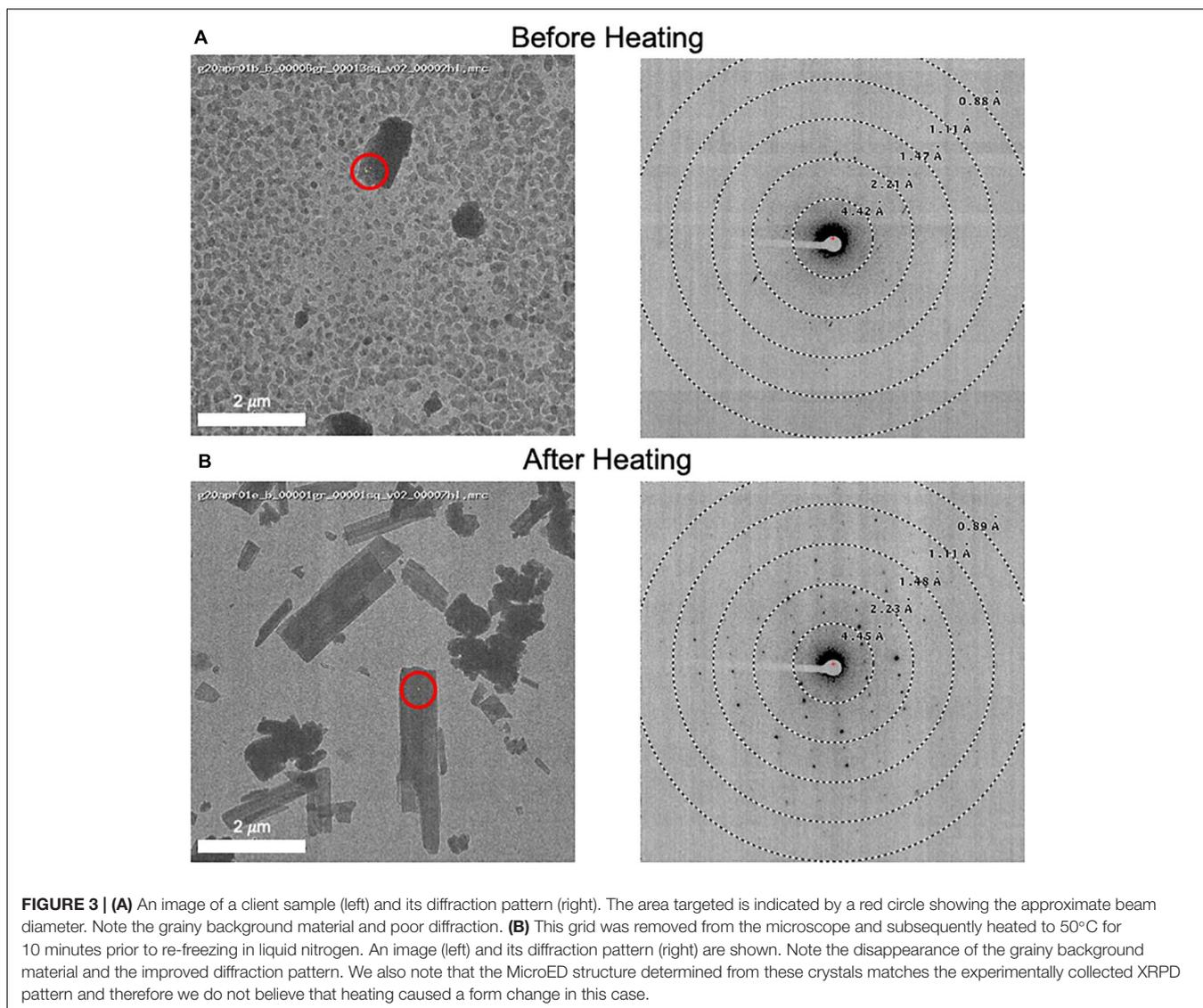
one novel, proof of concept structure (teniposide), all of which showed excellent refinement statistics relative to other structures determined by electron diffraction methods. Additionally, we have now examined over fifty small molecule samples submitted by our clients. We have also assisted in structure solution for a subset of these samples as requested by our clients. For the vast majority of these pharmaceutical samples, growing crystals large enough for single-crystal XRD was not possible despite the efforts of experienced small molecule crystallographers, and other structural techniques, such as NMR and MS, provided inconclusive results. MicroED is exceptionally well suited to provide structures for important and interesting compounds when other structural techniques are too challenging or provide incomplete information. Based on our experience providing MicroED services to the pharmaceutical industry over the past year, we are confident that MicroED is poised to join XRD, NMR, and MS as a routine tool for small molecule structure

determination, especially when targets pose challenges for the traditional, established methods.

## MATERIALS AND METHODS

### Sample Preparation

Grids for MicroED data acquisition were prepared as follows: (i), a small amount of dry sample was first ground between two glass slides; (ii) a pure carbon grid (Ted Pella 01840) that had not been plasma-cleaned or glow-discharged was dragged across the powder; (iii) the grid was gently tapped to remove excess powder; and (iv) the grid was clipped at room temperature and transferred to the transmission electron microscope for imaging at cryogenic temperature. In one case, when it was observed that the sample did not readily stick to the grid surface, a glow discharged grid was used instead, solving the problem. In many cases, the grid



was heated in a vacuum oven for a short period of time (15 min at 40°C) prior to plunging in liquid nitrogen in order to remove water and/or other volatile contaminants (**Figure 3**). Generally, only one grid was prepared per sample, and this was sufficient for data collection.

In a few instances, an alternative preparation method involving resuspending the sample in solvent, sonicating, and applying this crystal slurry to the grid was utilized. After drying, such grids were transferred to cryogenic temperatures. This method was employed when suitable crystals were not found using the dry application, but crystals were expected based on XRPD pre-screening results.

## Data Collection

Electron microscopy was performed using a Thermo Fisher Scientific (Hillsboro, OR, United States) Glacios Cryo Transmission Electron Microscope operated at 200 kV with the cryostage at  $-192^{\circ}\text{C}$ . The microscope was equipped with

a Thermo Fisher Scientific CETA-D detector. The camera length was calibrated using gold-palladium diffraction patterns collected using Grating Replica grids with Crossed Lines (EMS catalog #80051). Diffraction datasets were collected under parallel illumination conditions with a very low dose of  $\sim 0.1 \text{ e}^{-}/\text{Å}^2/\text{s}$ . A 20  $\mu\text{m}$  condenser aperture was used during data collection, resulting in a  $\sim 0.6 \mu\text{m}$  diameter beam on the sample. High gun lens (7.1) and spot size (10) settings were used to reduce the dose rate. The Legion software package, adapted to collect MicroED data as described previously (Cheng et al., 2021), was used for all data collection. Briefly, crystal selection based on low magnification imaging mode images was carried out manually by the operator. Legion provides the option to briefly test the diffraction of selected targets while queuing up targets for automated data collection. Automated data collection involves moving to and centering each target in imaging mode followed by recording the diffraction tilt series through the TEM User Interface with the camera set to

record continuously in rolling shutter mode with  $2 \times 2$  binning. This acquisition was synchronized with a continuous stage tilting function.

Diffraction datasets were then uploaded into the Legion database with an approximate tilt angle assignment and could be viewed from anywhere in real time using the web-based Image Viewer. The MRC format files (\*.mrc) were converted to SMV format files (\*.img) with appropriate header information needed for data processing. An offset was applied to remove negative pixel values and is reported in the image header as LEGINON\_OFFSET. A suggested pedestal value (IMAGE\_PEDESTAL) is given in the image header and corresponds to the largest minimum pixel values across the dataset after applying the offset. A DIALS specific format class can then be used to optimally handle the images<sup>1</sup>.

## Data Processing

All datasets were initially processed with DIALS using default settings appropriate for electron diffraction data in a semi-automated fashion (Clabbers et al., 2018; Winter et al., 2018). The best datasets were selected based on various quality indicators (unit cell refinement,  $CC_{1/2}$ ,  $R_{meas}$ ,  $I/\sigma I$ , etc.) for manual reprocessing. During manual processing, frames lacking reflections or exhibiting obvious radiation damage were removed from each dataset. These manually processed datasets were then fed into dials.cosym and scaled together in DIALS (Clabbers et al., 2018; Winter et al., 2018). Various combinations of datasets were evaluated with the goal of maximizing completeness and maintaining high redundancy while keeping favorable merging statistics. In most cases, these scaled, unmerged intensities were then trimmed to 0.9 Å regardless of merging statistics and then examined in XPREP (Sheldrick, 2008) before passing into SHELXT (Sheldrick, 2015b) and/or SHELXD (Schneider and Sheldrick, 2002) for phasing. Challenging cases typically required SHELXD with a high number of global tries. In many cases, it was necessary to attempt phasing from different combinations of datasets or in a number of possible space groups. When phasing was not possible in SHELXD, molecular replacement (MR) was employed. For MR, data were merged in AIMLESS (Evans, 2006; Winn et al., 2011) and then PHASER (McCoy et al., 2007) was employed with manually generated search models.

Refinement was carried out with SHELXL (Sheldrick, 2008) in Olex2 (Dolomanov et al., 2009) or REFMAC5 (Murshudov et al., 2011) and Coot (Emsley et al., 2010), depending on the phasing method. For structures refined with SHELXL, electron scattering factors were employed and riding hydrogens were added with the interatomic distances used for neutron scattering data (Gruene et al., 2018). The final resolution of the structure was established based on  $R$  factors and Goodness of Fit (GoF), and higher resolution reflections were excluded at this stage of data processing when appropriate. For all structures and values reported in this paper, the extinction parameter was not refined, which may be contributing to the higher  $R$ -factors relative to X-ray derived structures at the same resolution.

<sup>1</sup>[https://raw.githubusercontent.com/dials/dxtbx\\_ED\\_formats/master/FormatSMVCetaD\\_TUI.py](https://raw.githubusercontent.com/dials/dxtbx_ED_formats/master/FormatSMVCetaD_TUI.py)

## Progesterone

Progesterone (**Figure 4A**) was obtained from Sigma-Aldrich (MDL Number: MFCD00003658). One grid was prepared using the dry application method as described above. Four diffraction datasets were collected with a camera length of 1,186.5 mm. All datasets were processed and combined as described above and phasing was carried out in SHELXT. **Table 1** reports the final statistics. Raw data is available on Zenodo (10.5281/zenodo.3905397) and the final structure can be accessed from the CCDC (2084744).

## Biotin

Biotin (**Figure 4B**) was obtained from TCI (MDL Number: MFCD00005541). One grid was prepared using the dry application method as described above. Twelve datasets were collected with a camera length of 1,065.7 mm and data were processed as described above. Data from the best four crystals were combined and phasing was carried out in SHELXT. **Table 1** reports the final statistics. Raw data is available on Zenodo (10.3389/fmolb.2021.648603) and the final structure can be accessed from the CCDC (2083844).

## Paracetamol

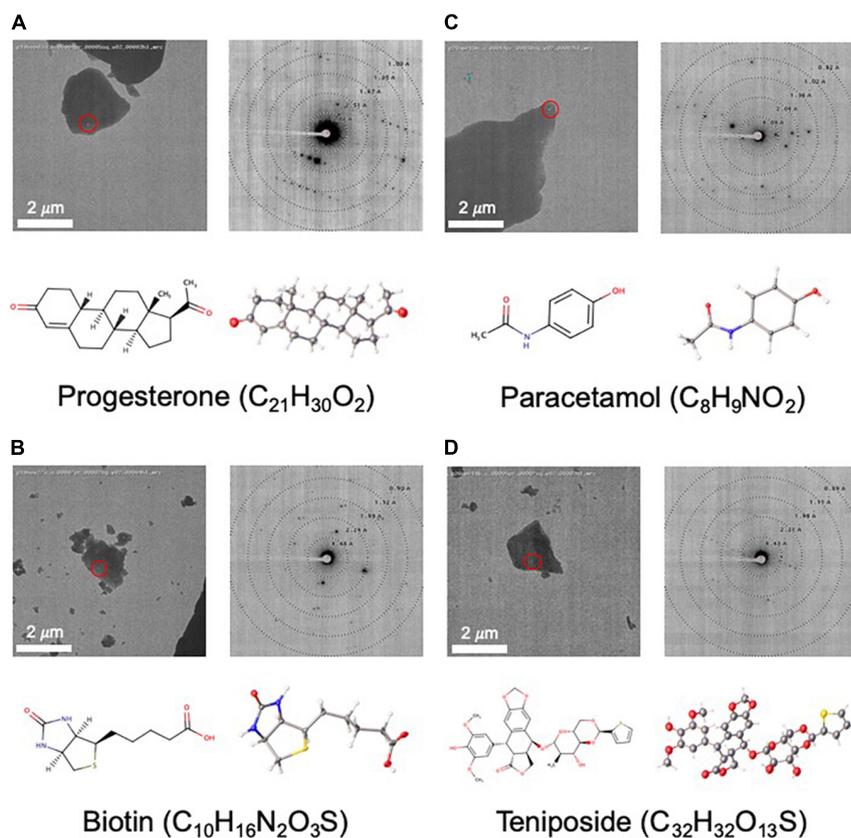
Paracetamol (**Figure 4C**) was extracted from Tylenol® using hot acetone and filtration. Once the crystals were dry, one grid was prepared using the dry application method as described above. Forty-nine particles/crystals were assessed for diffraction quality using Legion Preview Mode. Seventeen datasets were collected with a camera length of 955.8 mm and data were processed as described above. Eleven datasets failed at the indexing step and were discarded. Of the remaining six datasets, only four had the expected unit cell. These four datasets were combined and phased with SHELXT. **Table 1** reports the final statistics. Raw data is available on Zenodo (10.5281/zenodo.4737931) and the final structure can be accessed from the CCDC (2084732).

## Teniposide

Teniposide (**Figure 4D**) was obtained from Enzo (CAS Number: 29767-20-2). One grid was prepared using the dry application method as described above. Eighty-eight crystals were assessed with Legion Preview Mode and from these forty-four datasets were collected. Data were processed as described above. Ten datasets failed at the indexing stage and were discarded. Three indexed in a different crystal system and were also discarded. Of the remaining thirty-one data sets, the six best were combined. These data could not be phased by SHELXT and instead were solved by SHELXD (Sheldrick, 2008; Usón and Sheldrick, 2018) using all data to 0.9 Å. At the refinement stage, the final resolution was trimmed to 1.2 Å. **Table 1** reports the final statistics. Raw data is available on Zenodo (10.5281/zenodo.3937739) and the final structure can be accessed from the CCDC (2015361).

## Pharmaceutical Samples Submitted by Clients

Proprietary material was received from clients in powdered form. We requested that clients send at least 1 mg of material and



**FIGURE 4** | Structures determined by NIS as internal controls (A–C) and as a proof of concept for a more challenging sample (D). For each panel, an image of a targeted crystal is shown with the beam diameter shown as a red circle, along with an exemplary diffraction image with overlaid resolution rings. Below this, the 2D structure is shown alongside the MicroED structure with atoms displayed as thermal ellipsoids. Note that no distinction is made between single or double bonds in the thermal ellipsoid image. The common name and chemical formula are also given.

that they pre-screen the sample for crystallinity using XRPD if possible. Many, but not all, samples were pre-screened by XRPD. Grids were prepared using the dry application method or using a liquid resuspension step, and occasionally a drying step was applied to remove water or other volatile contaminants. Diffraction screening and data collection was carried out for fifty-six samples with a camera length of 1,065.7 mm. Automated data reduction was carried out for all samples as described above to assess data quality and preliminary space group assignments were made using *dials.cosym*. Most clients preferred to process and phase their own data, but occasionally when a client lacked internal expertise or when a client had difficulty processing and phasing their data, NIS proceeded with phasing. Phasing was carried out for fourteen client samples by NIS. Seven of these structures were fully refined by NIS to publication quality, again to assist clients lacking internal expertise.

## Figure Images

For **Figure 1B**, electron diffraction structures from the PDB were found using their online search tools and manually curated. The CSD structures were found using a combination of the CSD Python API (Groom et al., 2016) and the online search

tools, again paired with manual curation. The reciprocal space image in **Figure 2** was generated using ViewHKL in CCP4 (Winn et al., 2011) and the density map was generated in Coot (Emsley et al., 2010). The 2D chemical diagrams in **Figure 4** were generated in MarvinSketch (Marvin 20.4, ChemAxon<sup>2</sup>). Diffraction patterns with resolution rings in **Figure 4** were generated with Adxv Version 1.9.14<sup>3</sup>. ORTEP diagrams in **Figure 4** were generated using SHELXL in Olex2 (Sheldrick, 2008; Dolomanov et al., 2009).

## RESULTS

### Pipeline Validation (Structures of Progesterone, Biotin, and Paracetamol)

As part of our pipeline validation process, we determined the structures of three compounds, progesterone, biotin, and paracetamol (**Figures 4A–C**), that had already been solved by single crystal X-ray crystallography (Naumov et al., 1998; Shikii et al., 2004; Altaf and Stoeckli-Evans, 2013) and MicroED

<sup>2</sup><https://www.chemaxon.com>

<sup>3</sup>Copyright© 1994–2019 Andrew Arvai, The Scripps Research Institute.

**TABLE 1** | Summary of the data collection, reduction, and refinement statistics for four structures determined by NIS.

	<b>Progesterone</b> <b>C<sub>21</sub>H<sub>30</sub>O<sub>2</sub></b>	<b>Biotin</b> <b>C<sub>10</sub>H<sub>16</sub>N<sub>2</sub>O<sub>3</sub>S</b>	<b>Paracetamol</b> <b>C<sub>8</sub>H<sub>9</sub>NO<sub>2</sub></b>	<b>Teniposide</b> <b>C<sub>32</sub>H<sub>32</sub>O<sub>13</sub>S</b>
<b>Data collection</b>				
Oscillation per frame (°)	0.9	0.89	0.89	0.89
Camera length (mm)	1,186.5	1,065.7	955.8	1,065.7
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 <sup>-</sup> /Å <sup>2</sup> /s)	~0.045	~0.022	~0.022	~0.022
<b>Data reduction</b>				
Space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> / <i>n</i>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
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) <sup>a</sup>	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 <sup>a</sup>	10,900 (421)	11,275 (231)	10,031 (331)	15,834 (792)
Unique reflections <sup>a</sup>	949 (43)	1,125 (47)	1,350 (67)	1,074 (51)
Completeness (%) <sup>a</sup>	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/σ <sup>a</sup>	8.5 (1.2)	5.4 (0.2)	3.8 (0.4)	6.7 (1.0)
R <sub>meas</sub> <sup>a</sup>	0.25 (1.25)	0.26 (1.62)	0.30 (1.00)	0.26 (1.35)
R <sub>pim</sub> <sup>a</sup>	0.07 (0.38)	0.08 (0.47)	0.10 (0.43)	0.07 (0.33)
CC <sub>1/2</sub> <sup>a</sup>	0.993 (0.675)	0.988 (0.471)	0.975 (0.527)	0.995 (0.733)
<b>Refinement</b>				
Z/Z'	4/1	4/1	4/1	4/1
Total Reflections	10,209	10,626	9,049	14,062
R <sub>int</sub> /R <sub>sigma</sub>	0.22/0.16	0.23/0.38	0.26/0.32	0.25/0.20
Data/restraints/parameters	1,572/156/211	1,825/102/146	1,240/141/102	1,776/329/396
R <sub>1</sub> /wR <sub>2</sub> (all data)	0.16/0.29	0.22/0.37	0.29/0.46	0.17/0.28
R <sub>1</sub> /wR <sub>2</sub> [I ≥ 2σ(I)]	0.12/0.27	0.14/0.35	0.22/0.44	0.11/0.26
Goof	0.991	1.030	1.165	0.973

<sup>a</sup>Numbers in parentheses refer to the highest resolution shell.

(Gruene et al., 2018; Jones et al., 2018). In each of these cases, data collection took less than 2 h, and structure solution was possible in less than 1 day using only SHELXT. Compared to the client submitted small molecule structures solved at NIS, these validation targets were very easy to solve. Additionally, these structures showed highly encouraging statistics for MicroED structures (low R-factor and GooF values close to 1).

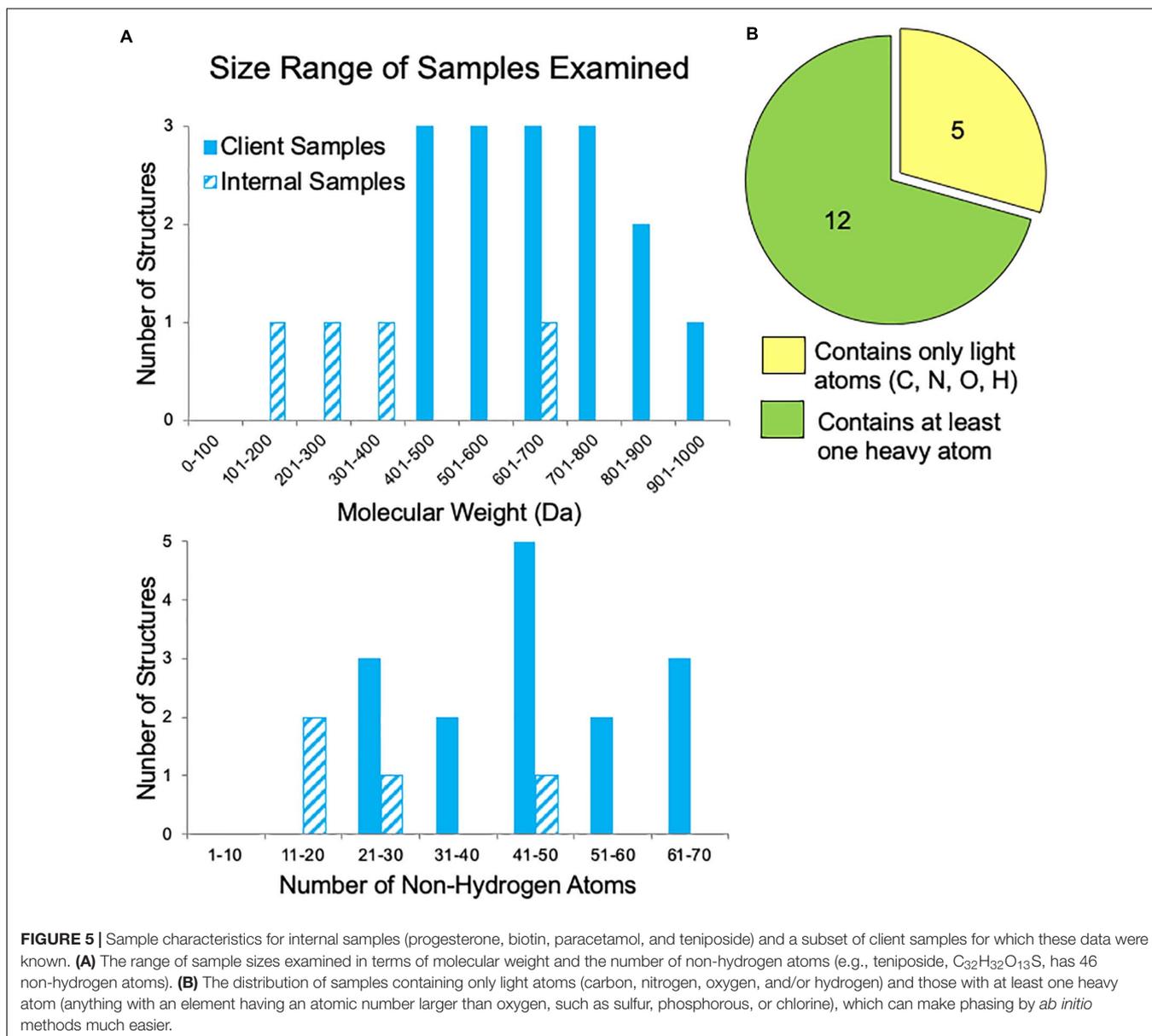
## Teniposide Structure

As a proof of concept for a more challenging target, we used MicroED to determine the novel structure of teniposide (**Figure 4D**), an anti-cancer drug that has been on the US market for almost two decades. Crystal diffraction quality varied greatly from crystal to crystal, so Legion Diffraction Preview Mode was very useful for identifying higher quality crystals. Manual targeting and diffraction screening of almost one hundred crystals took about 3 h and automated data collection of approximately fifty datasets took about 1.5 h. Manual data processing for this sample proved to be fairly challenging. These crystals exhibited radiation damage, making it necessary to trim out affected frames and necessitating combining data from

multiple crystals (six). Structure solution using SHELXT was unsuccessful and ultimately SHELXD was needed to successfully phase this data. While, structure solution was possible in 1 day, the data processing was not trivial and required a highly skilled crystallographer. The final structure has R<sub>1</sub>/wR<sub>2</sub> values for all data of 0.17/0.28 [0.11/0.26 for data with I ≥ 2σ(I)] and a GooF of 0.973.

## Data Collection for Pharmaceutical Samples

In our first year of offering small molecule MicroED data collection and structure solution services to pharmaceutical clients, we examined fifty-six samples. The majority of these samples failed to grow sufficiently large crystals for XRD studies despite attempts by experienced small molecule crystallographers. Additionally, for many samples, various NMR and MS experiments had already been performed, yet questions regarding the 2D and/or 3D structure still remained. Samples submitted by clients for which the molecular weight was disclosed ranged in size from 400 to 1,000 Da, corresponding to 20–70 non-hydrogen atoms (**Figure 5A**). These are fairly large



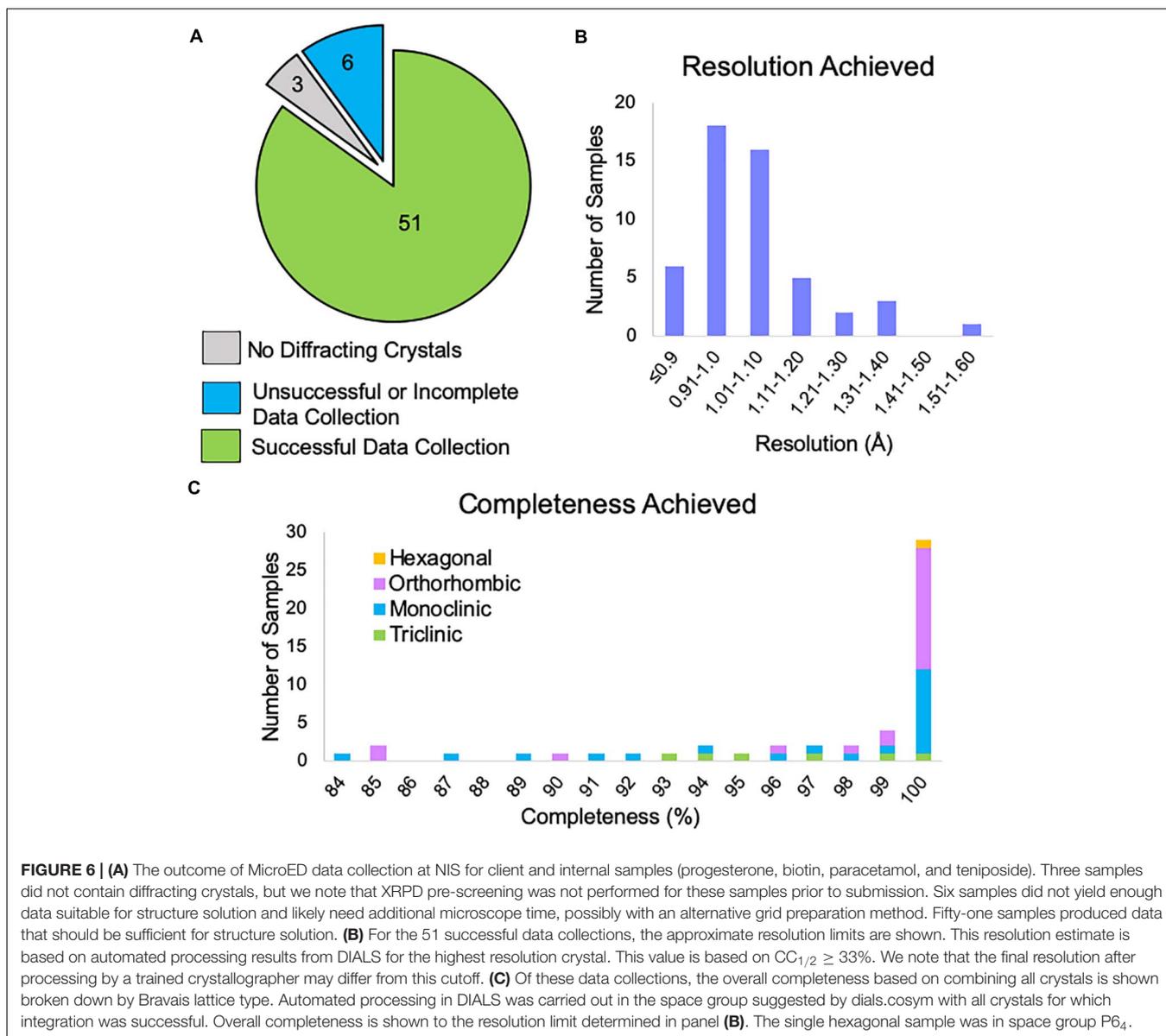
“small molecules”, which possibly contributed to their inability to be solved by traditional structural methods. Additionally, about one third of the samples contained only light atoms (carbon, nitrogen, oxygen, and hydrogen), which can make phasing by *ab initio* methods (SHELXT and SHELXD) more challenging if the resolution is poor or if there are other data quality issues (Figure 5B). Though anomalous signal is thought to be very minimal in electron diffraction, the presence of heavier atoms such as sulfur, chlorine, phosphorus, or bromine within a small molecule can be quite helpful for phasing. Heavy atom peaks, even without anomalous signal, are much easier to correctly locate by *ab initio* methods and can be used to bootstrap to a complete solution.

For samples considered “easy cases,” about 1 hour was spent finding crystals, screening them for diffraction quality and adding

targets to the queue, and about 1 hour was spent automatically collecting 20 datasets (Figure 2). For more challenging cases, the grid might be removed for a heating step prior to re-insertion or the wet application grid preparation protocol might be followed after the dry application failed to generate suitable crystals (Figure 2). In cases presenting data processing challenges, long (four or more hours) or overnight data collections might be set up, collecting 50–150 datasets to ensure enough high-quality data was collected for structure solution (Figure 2).

### Aggregate Data Collection Results

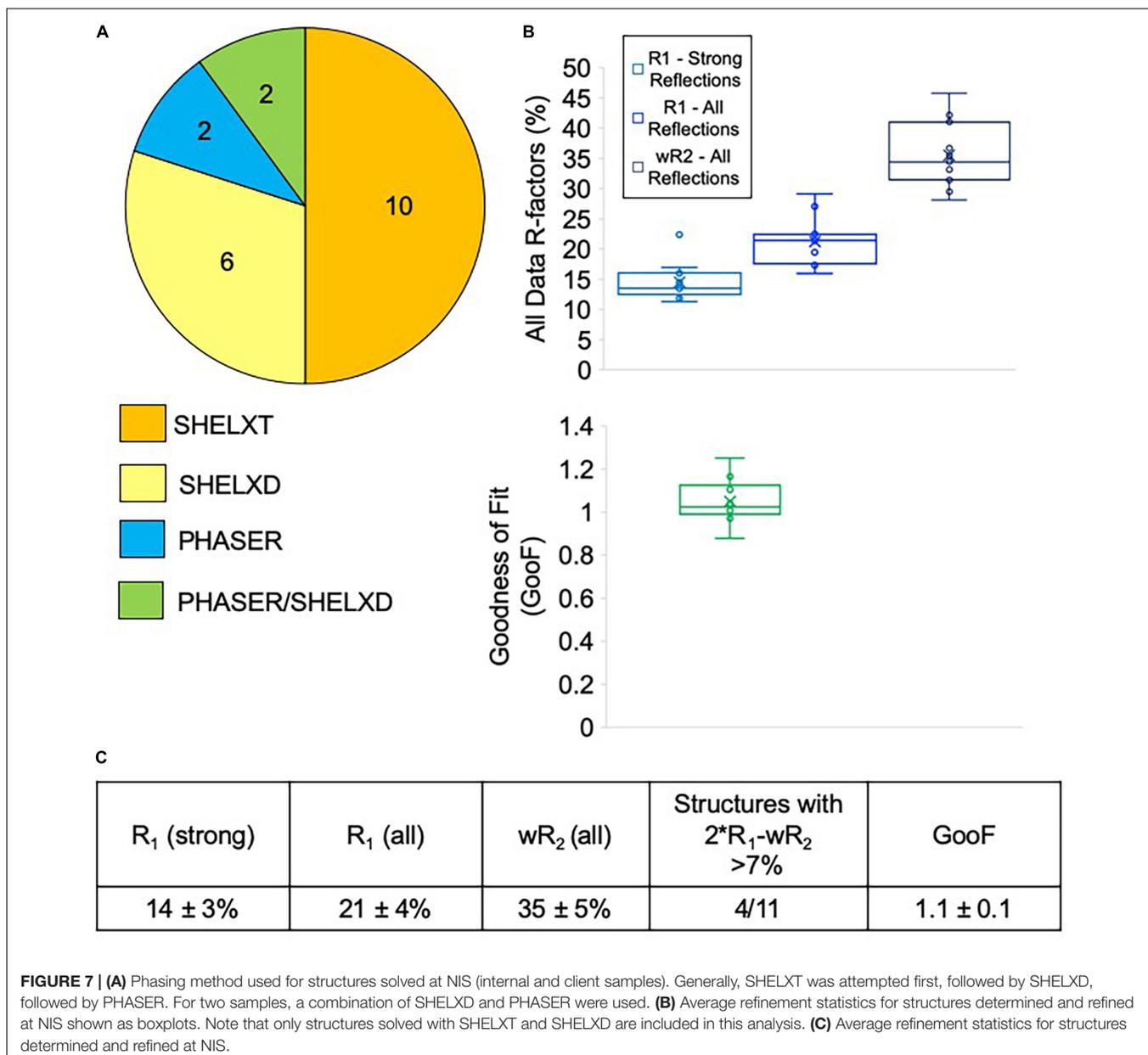
Of the samples examined by NIS (both internal and client samples), 85% produced data expected to be suitable for structure solution (Figure 6A). The remaining samples either showed no diffraction or had data collection curtailed due to



microscope time constraints. There were three samples that did not appear to contain crystals and after consultation with the client we found that XRPD was not collected from these samples prior to submission, leading us to suspect that these samples were not in fact crystalline. We note that, all samples for which XRPD was used as a pre-screening tool contained diffracting crystals. One of the six “unsuccessful” samples had obvious water contamination issues (frozen puddles of water partially dissolving the crystals). We have previously rescued such samples by heating the sample in a vacuum oven for a brief period of time prior to plunging into liquid nitrogen, and it is likely that this sample could also be rescued in this way. Another of the “unsuccessful” samples did not contain any diffracting crystals following our typical dry application to the TEM grid. We subsequently found that resuspension

of this sample in a solvent, applying the suspension to the TEM grid and allowing it to dry resulted in diffracting crystals. The remaining four samples contained diffracting crystals with a reasonable unit cell large enough to accommodate the compound of interest, but collection of a sufficiently large number of datasets was not possible during the client’s allotted microscope time. Many of our clients book single microscope days in which we try to collect data from as many samples as possible, which can lead to incomplete data collection for the last sample in the queue. Of course, these remaining samples can easily be revisited during the next scheduled microscope time.

Data reduction for all samples that produced viable data revealed that 78% of samples achieved resolution better than 1.1 Å based on a metric of  $CC_{1/2} \geq 33\%$  (**Figure 6B**).



After combining data from multiple crystals, 78% of samples reached completeness  $>95\%$ , while 57% of samples achieved 100% completeness (**Figure 6C**). We note that low completeness did not hinder structure solution for any of the samples phased at NIS. For lower completeness samples, we suspect that this is due to preferred orientation of the crystal on the grid and if so, it is doubtful that additional data collection from these grids would produce more complete data.

## Aggregate Phasing and Refinement Results

Most of the data collected at NIS was processed externally by our clients, and we do not have data on how many of these datasets were successfully phased and refined. We did, however, process

data from sixteen client samples, all of which we were able to successfully phase. Data processing difficulties varied widely for client samples, with some taking an hour or two to solve, while others required several rounds of iterative effort over multiple days (**Figure 7A**). Of the twenty internal and client samples determined by NIS, ten samples were considered easy to process and could be readily solved with SHELXT. The remaining ten samples had to be solved with SHELXD (six samples) or PHASER (two samples), or a combination of both PHASER and SHELXD (two samples). Structures requiring MR in PHASER generally were the result of low resolution (worse than  $1.1 \text{ \AA}$ ) and/or a poor data-to-parameter ratio (combination of low completeness and high  $Z'$ -number of molecules in the asymmetric unit). For the more challenging cases, we found that it helped to include

data to 0.9 Å, regardless of the actual resolution cutoff suggested by  $CC_{1/2}$ , and increase the cycles in SHELXT/D. Additionally, reducing the symmetry to P1 and phasing with SHELXD was extremely helpful for samples suffering from pseudo translational symmetry issues that might or might not be apparent during data processing. Once the structure was solved in P1, the actual crystal symmetry could be deduced by examining the model, and the data could then be re-solved in the correct space group. For PHASER, we found that success generally relied on having a very accurate search model (extremely low RMSD with the final refined structure) that represented the majority of the compound by mass.

Seven of the client structures determined by NIS were also fully refined by NIS at the request of clients lacking this expertise. Taken together with the four structures detailed in **Table 1**, structures collected and refined by NIS appear to achieve very encouraging refinement statistics when compared with other published MicroED structures. NIS structures achieved an average  $R_1/wR_2$  for all data of  $14 \pm 3\%/35 \pm 5\%$  and an average GooF value of  $1.1 \pm 0.1$  (**Figure 7B**). Ideally,  $wR_2$  should be about the same magnitude as two times  $R_1$  and deviations from this can indicate problems with the data or with the data handling. For the samples included in this analysis, only four out of eleven produced larger than 7% difference between  $2 \cdot R_1 - wR_2$  (**Figure 7C**). Overall, we are very encouraged by these refinement statistics given the resolution of our data and the current challenges posed by processing electron diffraction data.

## DISCUSSION

Microcrystal electron diffraction is becoming an attractive alternative structure determination technique for proteins (Shi et al., 2016) and small molecules (Gruene et al., 2018; Jones et al., 2018). Since electron diffraction allows for data collection from crystals orders of magnitude smaller than those required for single-crystal XRD, this technique can deliver structures of challenging targets for which large crystals are not available. This is especially exciting for small molecules, many of which readily form microcrystals, especially during the purification process. In such cases, MicroED often eliminates the need for crystallization screening, greatly reducing the sample requirements. Taken together, MicroED is poised to assist in the drug development pipeline at every step along the way: from the discovery phase, where sample quantities are generally too limited for crystallization trials, to process chemistry where structures of reaction products and by-products can guide synthesis strategies, and to formulation chemistry where the crystallization space of the API can be explored, better understood and engineered.

In medicinal and process chemistry, MicroED is poised to transform workflows making accurate structure determination a routine and essential tool, rather than an exotic, “nice-to-have” tool. Nevertheless, MicroED adoption has been slow, especially in industrial settings, where widespread adoption has been inhibited by steep requirements for instrumentation, infrastructure and expertise, especially when the number of projects may not

justify the initial investment. Improvements in access to facilities, in speed, reliability and automation for data collection and processing, and in structure solution and refinement, are needed for MicroED to fulfill its potential of becoming a routine tool in small molecule characterization. We have built and demonstrated an efficient, semi-automated MicroED pipeline that can be used to image multiple samples per day and solve their structures on a timeline of 1 day to several days, depending on the complexity of the sample data.

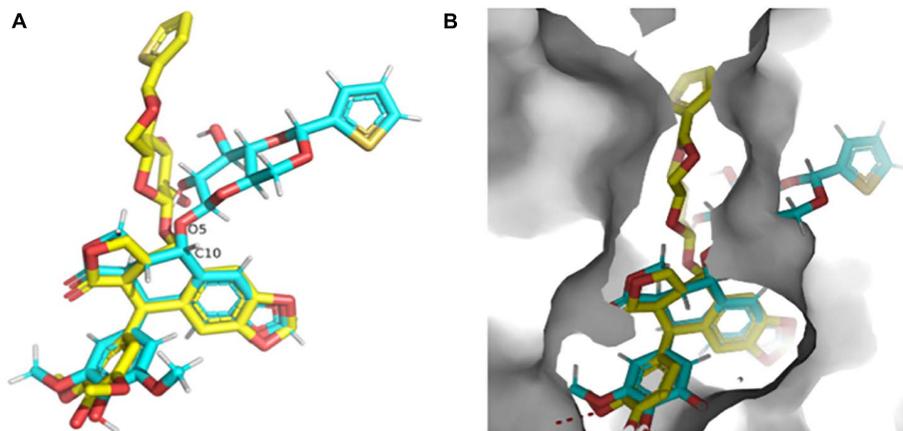
## Progesterone, Biotin, and Paracetamol Structures

To validate our MicroED setup, we determined the structures of three compounds that had already been solved by single crystal XRD and MicroED: progesterone, biotin and paracetamol (**Figures 4A–C**). Data for all three compounds achieved similar resolution and data reduction statistics compared to previous MicroED efforts (Gruene et al., 2018; Jones et al., 2018). The structures of all three compounds were solved with SHELXT (Sheldrick, 2008, 2015b) and refined with SHELXL (Sheldrick, 2015a) as implemented in Olex2 (Dolomanov et al., 2009), similarly to what was already reported. Statistics for the refined structures are fairly comparable to previously reported MicroED structures, with one notable difference being the value of the GooF parameter which in all the NIS structures is closer to the ideal value of 1 (**Figure 7C** and **Table 1**), whereas previous groups have obtained values closer to 2 (Gruene et al., 2018; Jones et al., 2018).

## Teniposide Structure

Teniposide [epipodophyllotoxin, 4'-demethyl-, 9-(4,6-*O*-2-thenylidene- $\beta$ -D-glucopyranoside)] is a semi-synthetic derivative of podophyllotoxin. This organic heterotetracyclic compound is found in the roots and rhizomes of *Podophyllum* species. It has been shown to exhibit antitumor activity via inhibition of DNA synthesis, achieved by forming a complex with topoisomerase II and DNA (Bh, 1992). This complex induces breaks in double-stranded DNA and prevents repair by topoisomerase II. Accumulated breaks in DNA prevent cells from entering into the mitotic phase of the cell cycle, and lead to cell death. Teniposide acts primarily in the G2 and S phases of the cell cycle. In the United States, teniposide is currently used in the treatment of children with Acute Lymphocytic Leukemia (ALL) that has not improved or that has worsened after treatment with other medications (Dahl et al., 1985).

Despite being on the market for almost two decades, there is no single-crystal X-ray structure of the solid form of teniposide deposited in the CCDC. It is possible that such a structure has been determined, but not deposited, by industrial researchers. We speculate that the lack of such a structure may be due to difficulties growing sufficiently large crystals for X-ray studies. For this reason, teniposide was chosen as a proof of concept target for our MicroED pipeline. Given the high cost and technical challenges associated with MicroED, we wanted to demonstrate that this technique could solve structures not easily accessible



**FIGURE 8 | (A)** Overlay of the teniposide MicroED structure determined in this paper (cyan carbon atoms, CCDC KUXJUL) and the structure of teniposide bound to human serum albumin (HSA) from PDB 4L9Q (yellow carbon atoms). **(B)** Overlay of the two teniposide molecules (same coloring scheme as in panel **(A)**) within the HSA binding pocket demonstrating that the conformation observed in solid state form is distinct from the conformation required for protein binding. UCSF Chimera was used to generate these images (Pettersen et al., 2004).

to XRD. Teniposide was also an attractive target given its size and complexity, especially considering that most of the early small molecule MicroED structures published in the literature had molecular weights of less than 400 Da, which generally makes structure solution easier. In contrast, all of the samples submitted by our industrial clients were of compounds with molecular weights more than 400 Da (**Figure 5A**). Thus, it was important for us to demonstrate that MicroED is a viable option for these larger compounds. While data collection and data processing were both more challenging for teniposide compared to progesterone, biotin, and paracetamol, we have shown that structure solution was still possible in a single day.

The high-resolution structure of the solid form of teniposide solved by MicroED is chemically consistent with a structure of teniposide bound to human serum albumin (HSA) (PDB entry 4L9Q; **Figure 8A**). The biggest difference is in the torsional angle at the C10-O5 bond ( $-134^\circ$  in 4L9Q versus  $-69^\circ$  in the structure reported here, **Figure 8A**). This bond is expected to be able to freely rotate and the change in torsional angle is likely the result of the constraints induced by the surrounding protein atoms in the HSA-bound form versus the constraints imposed in the solid form of the compound (**Figure 8B**).

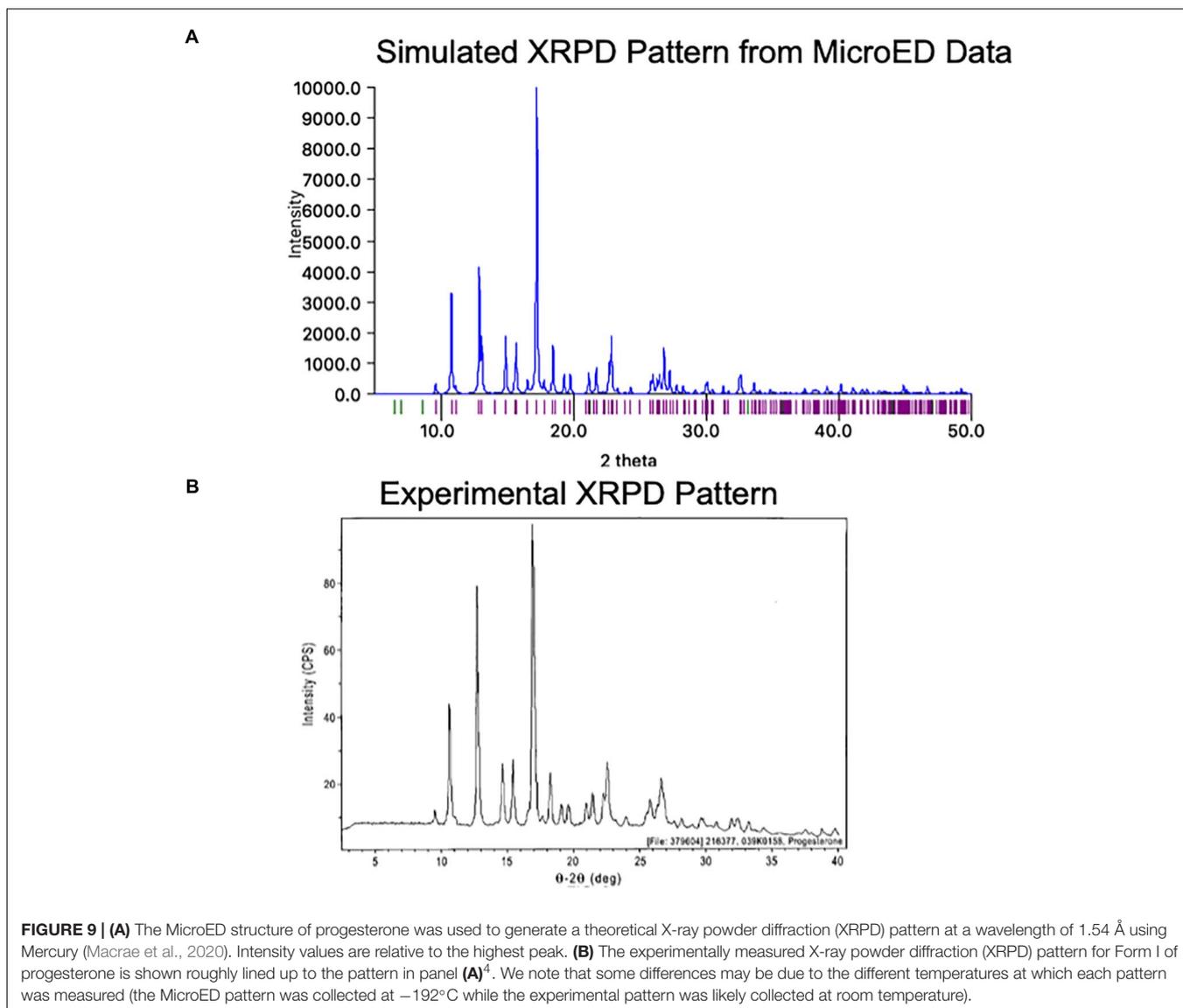
## Aggregate Data Collection and Processing Outcomes

At NIS, we were able to collect data suitable for structure solution for the majority of samples submitted (**Figure 6A**) and for data we have been asked to process we have had a 100% success rate for structure solution. This is very encouraging given the large size of many of these samples (**Figure 5A**) and the lack of heavy atoms (sulfur, chlorine, etc.) in 30% of the samples examined (**Figure 5B**), which is thought to make phasing by *ab initio* methods more challenging. Additionally, nearly 80% of samples collected at NIS achieved at least 1.1 Å resolution (**Figure 6B**). This is important given that 1.1 Å is generally

considered the point beyond which phasing by direct and dual space methods becomes exceedingly difficult (Sheldrick et al., 1993). Low completeness in many MicroED datasets is often cited as a concern for this method. We found that our pipeline produced at least 84% completeness for all samples for which data was collected, with more than half achieving 100% completeness (**Figure 6C**). We also note that low completeness (as low as 84%) did not hinder structure solution or map interpretation for data processed internally. The data for paracetamol were only 84% complete overall, yet the map for this compound was reasonably isotropic and easily interpreted. We were also encouraged by our average refinement statistics, which compare very favorably with MicroED structures in the literature (**Figures 7B,C**). Given these results we are satisfied that our MicroED pipeline is quite robust. We anticipate that for any small molecule sample proven to be crystalline by XRPD data, we should be able to solve a single crystal MicroED structure using 2–4 h of microscope time and one to several days of data processing time.

## Leveraging the Relationship Between MicroED and XRPD

One exciting feature of MicroED data is that the final coordinates produced are practically indistinguishable from those produced by X-ray methods. Because of this, a structure generated by MicroED can be readily used to generate a predicted XRPD pattern for its crystal form at a given X-ray wavelength, effectively generating an XRPD fingerprint. This fingerprint can be very useful as it can be compared to experimentally measured XRPD data (**Figure 9**). Comparing an experimentally measured XRPD spectra to the known MicroED-based spectra can quickly confirm the identity of a reaction product, such as might be produced during synthesis optimization. Additionally, XRPD from a mixture of crystals can be resolved into the relative ratios of these crystal forms in the mixture, as long as the expected XRPD fingerprints for each compound is known. In this way,



XRPD screening based on MicroED data for a low abundance contaminant can be very helpful in optimizing synthesis and purification strategies.

## Practical Notes for Establishing a MicroED Pipeline

There are many individuals interested in establishing their own MicroED pipeline, especially one centered around existing microscope equipment. While evaluating and comparing hardware requirements for MicroED is outside the scope of this paper, we do have some practical advice for newcomers to the field. It is important to keep in mind that compounds of interest to the pharmaceutical industry that fail to be solved by XRD often pose significant challenges for MicroED data processing.

<sup>4</sup> <https://patentimages.storage.googleapis.com/b3/6d/09/db2cfd1e296ab/US20170101433A1.pdf>

While examples of structures solved using lower end microscopes and cameras exist in the literature, we anticipate that many large organic crystals inaccessible by XRD might be impossible to solve with lower quality MicroED data. Data collected at NIS using the CETA camera prior to upgrading to the CETA-D camera was of lower quality, and while we were able to use it to solve the structure of paracetamol, we are doubtful that data from such a camera would have been sufficient to solve a more challenging crystal structure like teniposide using *ab initio* methods. We suggest that those interested in setting up their own MicroED facility first collect data from easy to solve crystals such as paracetamol or biotin but then try to solve a more difficult case, such as teniposide, to get a sense of whether or not their equipment and data processing pipeline can in practice deliver structures that require MicroED techniques. At this point in time, MicroED is relatively expensive and often poses significant data processing challenges. Until these two issues are addressed,

we suggest that MicroED be reserved for samples that absolutely require this technique because they fail to produce large crystals or because the sample quantity is limiting.

## Future of the Field

Our experience collecting data from over fifty pharmaceutical small molecule samples provides high optimism about the role MicroED will play in the future for drug discovery and development. We have demonstrated that structure determination by a MicroED service provider can be viable way for industrial researchers to gain access to this cutting-edge technology and leverage it to expedite research endeavors. We hope that future efforts in the field will yield better data processing tools for electron diffraction data, making it easier for newcomers to the field to process their own data which should in turn expedite results and reduce costs.

## DATA AVAILABILITY STATEMENT

All raw datasets contributing to the structures of progesterone, biotin, paracetamol, and teniposide have been deposited to the Zenodo, the open access public data repository as part of the 3DED/MicroED datasets community. The final structures (coordinates, reflections, and structure factors) of progesterone, biotin, paracetamol, and teniposide have been deposited in the Cambridge Crystallographic Data Center (CCDC), a part of the Cambridge Structural Database (CSD), a public data repository. Zenodo DOI and CCDC numbers are listed in the text.

## AUTHOR CONTRIBUTIONS

JB, GS, and BC conceptualized and wrote the manuscript. AC, JB, and TN established data collection strategies. JB, KL,

MMy, BR, TN, NC, NP, PM, TW, MMs, and CH prepared the samples and collected the data. JB, GS, BM, and BR processed the data. AC, TG, SD, BR, JB, DW, and ST automated data handling, processing and analysis. BW provided some samples, including the sample shown in **Figure 3**. CP helped secure the funding. All authors contributed to the article and approved the submitted version.

## FUNDING

Some of this work was funded by NIH/NIGMS grant number 1R44GM140666 awarded to GS and JB. The funders had no role in study design, data collection and interpretation, or the decision to submit this work for publication.

## ACKNOWLEDGMENTS

We would like to thank our clients for providing interesting samples, learning with us, and sharing how impactful these structures were for their research. We would like to thank the entire NIS team for supporting this effort in direct and indirect ways. We would like to thank Tim Grüne at University of Vienna for incredibly helpful advice regarding data processing, sample preparation, and for insightful discussions about the field of electron diffraction. We would like to thank Tamir Gonen at University of California Los Angeles for hosting MicroED workshops and helpful advice. We would like to thank José Rodríguez (UCLA), Lee Daniels (Cambridge Structural Database), Gérard Bricogne (Global Phasing Limited), and Paul Emsley (MRC Laboratory of Molecular Biology) for insightful discussions.

## REFERENCES

- Altaf, M., and Stoeckli-Evans, H. (2013). Chiral one- and two-dimensional silver(I)-biotin coordination polymers. *Acta Crystallogr. C* 69, 127–137. doi: 10.1107/s0108270113000322
- Ariëns, E. J. (1984). Stereochemistry, a basis for sophisticated nonsense in pharmacokinetics and clinical pharmacology. *Eur. J. Clin. Pharmacol.* 26, 663–668. doi: 10.1007/bf00541922
- Beale, E. V., Waterman, D. G., Hecksel, C., van Rooyen, J., Gilchrist, J. B., Parkhurst, J. M., et al. (2020). A workflow for protein structure determination from thin crystal lamella by micro-electron diffraction. *Front. Mol. Biosci.* 7:179. doi: 10.3389/fmolb.2020.00179
- Bh, L. (1992). Mechanisms of action of teniposide (VM-26) and comparison with etoposide (VP-16). *Semin. Oncol.* 19(2 Suppl. 6), 3–19.
- Censi, R., and Di Martino, P. (2015). Polymorph impact on the bioavailability and stability of poorly soluble drugs. *Molecules* 20, 18759–18776. doi: 10.3390/molecules201018759
- Cheng, A., Negro, C., Bruhn, J. F., Rice, W. J., Dallakyan, S., Eng, E. T., et al. (2021). Leginon: new features and applications. *Protein Sci.* 30, 136–150. doi: 10.1002/pro.3967
- Clabbers, M. T. B., Gruene, T., Parkhurst, J. M., Abrahams, J. P., and Waterman, D. G. (2018). Electron diffraction data processing with DIALS. *Acta Crystallogr. D Struct. Biol.* 74, 506–518. doi: 10.1107/S2059798318007726
- Dahl, G. V., Rivera, G., Pui, C. H., Mirro, J. Jr., Ochs, J., Kalwinsky, D. K., et al. (1985). A novel treatment of childhood lymphoblastic non-Hodgkin's lymphoma: early and intermittent use of teniposide plus cytarabine. *Blood* 66, 1110–1114. doi: 10.1182/blood.v66.5.1110.bloodjournal6651110
- Dolomanov, O. V., Bourhis, L. J., Gildea, R. J., Howard, J. A. K., and Puschmann, H. (2009). OLEX2: a complete structure solution, refinement and analysis program. *J. Appl. Crystallogr.* 42, 339–341. doi: 10.1107/s0021889808042726
- Duyvesteyn, H. M. E., Kotecha, A., Ginn, H. M., Hecksel, C. W., Beale, E. V., de Haas, F., et al. (2018). Machining protein microcrystals for structure determination by electron diffraction. *Proc. Natl. Acad. Sci. U.S.A.* 115, 9569–9573. doi: 10.1073/pnas.1809978115
- Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501. doi: 10.1107/S0907444910007493
- Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* 62, 72–82. doi: 10.1107/S0907444905036693
- Groom, C. R., Bruno, I. J., Lightfoot, M. P., and Ward, S. C. (2016). The cambridge structural database. *Acta Crystallogr. B Struct. Sci. Cryst. Eng. Mater.* 72, 171–179. doi: 10.1107/S2052520616003954
- Gruene, T., Wennmacher, J. T. C., Zaubitzer, C., Holstein, J. J., Heidler, J., Fecteau-Lefebvre, A., et al. (2018). Rapid structure determination of microcrystalline molecular compounds using electron diffraction. *Angew. Chem. Int. Ed. Engl.* 57, 16313–16317. doi: 10.1002/anie.201811318

- Harris, K. D. M. (2012). Powder diffraction crystallography of molecular solids. *Top. Curr. Chem.* 315, 133–178. doi: 10.1007/128\_2011\_251
- Jones, C. G., Martynowycz, M. W., Hattné, J., Fulton, T. J., Stoltz, B. M., Rodriguez, J. A., et al. (2018). The CryoEM method MicroED as a powerful tool for small molecule structure determination. *ACS Cent. Sci.* 4, 1587–1592. doi: 10.1021/acscentsci.8b00760
- Lu, J., and Rohani, S. (2009). Polymorphism and crystallization of active pharmaceutical ingredients (APIs). *Curr. Med. Chem.* 16, 884–905. doi: 10.2174/092986709787549299
- Macrae, C. F., Sovago, I., Cottrell, S. J., Galek, P. T. A., McCabe, P., Pidcock, E., et al. (2020). Mercury 4.0: from visualization to analysis, design and prediction. *J. Appl. Crystallogr.* 53, 226–235. doi: 10.1107/s1600576719014092
- Martynowycz, M. W., Zhao, W., Hattné, J., Jensen, G. J., and Gonen, T. (2019). Collection of continuous rotation MicroED data from ion beam-milled crystals of any size. *Structure* 27, 545–548.e2. doi: 10.1016/j.str.2018.12.003
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* 40, 658–674. doi: 10.1107/S0021889807021206
- Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., et al. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* 67, 355–367. doi: 10.1107/s0907444911001314
- Naumov, D. Y., Vasilchenko, M. A., and Howard, J. A. K. (1998). The monoclinic form of acetaminophen at 150K. *Acta Crystallogr. C* 54, 653–655. doi: 10.1107/s0108270197018386
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., et al. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612. doi: 10.1002/jcc.20084
- Raza, K., Kumar, P., Ratan, S., Malik, R., and Arora, S. (2014). Polymorphism: the phenomenon affecting the performance of drugs. *SOJ Pharm. Pharm. Sci.* 1, 1–10. doi: 10.15226/2374-6866/1/2/00111
- Schneider, T. R., and Sheldrick, G. M. (2002). Substructure solution with SHELXD. *Acta Crystallogr. D Biol. Crystallogr.* 58, 1772–1779.
- Sheldrick, G. M. (2008). A short history of SHELX. *Int. Union Crystallogr.* 64, 112–122.
- Sheldrick, G. M. (2015a). Crystal structure refinement with SHELXL. *Acta Crystallogr. C Struct. Chem.* 71, 3–8.
- Sheldrick, G. M. (2015b). SHELXT—integrated space-group and crystal-structure determination. *Acta Crystallogr. A Found. Adv.* 71, 3–8. doi: 10.1107/s2053273314026370
- Sheldrick, G. M., Dauter, Z., Wilson, K. S., Hope, H., and Sieker, L. C. (1993). The application of direct methods and Patterson interpretation to high-resolution native protein data. *Acta Crystallogr. D Biol. Crystallogr.* 49, 18–23. doi: 10.1107/s0907444992007364
- Shi, D., Nannenga, B. L., de la Cruz, M. J., Liu, J., Sawtelle, S., Calero, G., et al. (2016). The collection of MicroED data for macromolecular crystallography. *Nat. Protoc.* 11, 895–904. doi: 10.1038/nprot.2016.046
- Shi, D., Nannenga, B. L., Iadanza, M. G., and Gonen, T. (2013). Three-dimensional electron crystallography of protein microcrystals. *ELife* 2:e01345.
- Shikii, K., Sakamoto, S., Seki, H., Utsumi, H., and Yamaguchi, K. (2004). Narcissistic aggregation of steroid compounds in diluted solution elucidated by CSI-MS, PFG NMR and X-ray analysis. *Tetrahedron* 60, 3487–3492. doi: 10.1016/j.tet.2004.02.030
- Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., et al. (2005). Automated molecular microscopy: the new Legimon system. *J. Struct. Biol.* 151, 41–60. doi: 10.1016/j.jsb.2005.03.010
- Thakral, N. K., Zanon, R. L., Kelly, R. C., and Thakral, S. (2018). Applications of powder X-Ray diffraction in small molecule pharmaceuticals: achievements and aspirations. *J. Pharm. Sci.* 107, 2969–2982. doi: 10.1016/j.xphs.2018.08.010
- Usón, I., and Sheldrick, G. M. (2018). An introduction to experimental phasing of macromolecules illustrated by SHELX; new autotracing features. *Acta Crystallogr. D Struct. Biol.* 74, 106–116. doi: 10.1107/s2059798317015121
- Wennmacher, J. T. C., Li, T., Zaubitzer, C., Gemmi, M., Mugnaioli, E., Gruene, T., et al. (2020). Heterogeneity of nano-sized zeolite crystals. *Microporous Mesoporous Mater.* 294, 109897–109897. doi: 10.1016/j.micromeso.2019.109897
- Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., et al. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* 67, 235–242.
- Winter, G., Waterman, D. G., Parkhurst, J. M., Brewster, A. S., Gildea, R. J., Gerstel, M., et al. (2018). DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr. D Struct. Biol.* 74, 85–97. doi: 10.1107/s2059798317017235

**Conflict of Interest:** JB, GS, AC, TG, SD, BR, TN, KL, MMy, NC, NP, PM, TW, MMs, CH, ST, CP, and BC are or have been employed by Nanomaging Services, a commercial supplier of electron microscopy services to the biopharmaceutical and biotechnology industries.

BW is employed by the company Biogen.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Bruhn, Scapin, Cheng, Mercado, Waterman, Ganesh, Dallakyan, Read, Nieusma, Lucier, Mayer, Chiang, Poweleit, McGilvray, Wilson, Mashore, Hennessy, Thomson, Wang, Potter and Carragher. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.