CryoEM in industry: challenges and opportunities

Giovanna Scapin

Nanoimaging Services, New York, New York, United States

Transmission electron microscopy has been an important research tool for many decades: the first transmission electron microscope was developed in 1931 by Max Knoll and Ernst Ruska, and in 1981 Jacques Dubochet and Alasdair McDowall introduced the concept of cryo-cooling individual molecules allowing them to retain their native shape and protecting them from environmental and radiation damage[1]. Since the "Resolution Revolution"[2] cryo-EM has revolutionized structural biology by enabling high-resolution structure determination of systems previously inaccessible.[3–6] The Biopharma community quickly understood that cryoEM could potentially transform the thinking about structures and biology[7–10]. But while cryoEM has proven to be effective in the early stages of drug development, and its applicability to small molecule lead identification and optimization has also been recently demonstrated[11,12], barriers still remain that prevent cryoEM from being a routinely successful tool in drug discovery.

Infrastructure requirements. The costs to install a microscope and the necessary ancillary equipment are in the multi-million-dollar range. Microscope installation requires dedicated space with specific environmental controls, and high capacity computational infrastructures. This investment is not easily available, and models to share the costs and the associated risks have been proposed[13]. For academic users, large-scale facilities have been established with funding by National agencies[14], or centered around pre-existing National facilities[15], but they remain largely unavailable to industrial users. CryoEM adoption by industry comes with unique requirements that have be partially met by establishing internal facilities or by forming industrial consortia (such as the UK Cambridge Pharmaceutical CryoEM Consortium[16]). Alternative options include partnering with specialized CROs, but the growth in this field has been slow.

A pipeline of structures. A successful SBDD project relies on a fast supply of structural data to inform subsequent cycles of design. This could correspond to the determination of several atomic-level structures a week. Though not yet routine in cryoEM, this is achievable with the current hardware and software: a 6K image data set can be collected overnight, and if a data processing pipeline is run concurrently[17,18], 3D reconstructions within 24h are feasible. The rate limiting step then becomes the reproducible generation of grids suitable for good data collections. Unfortunately, reproducibility with the current vitrification robots is difficult to achieve; thus, screening for optimal grids before each collection remains necessary, even if the sample has been previously characterized. Novel vitrification devices[19,20] are poised to improve the reproducibility and the quality of the grids, but they still require more extensive testing, particularly in SBDD pipelines.

Sample requirement and preparation. Even further upstream than grid preparation, but critically important for the success of a project, is the quality of the sample itself. Having a reliable source of good quality protein is as necessary in cryoEM as it is in any other structural approach. CryoEM offer multiple advantages over crystallography, especially in that there is more tolerance for discrete compositional and/or conformational heterogeneity. For example, raggedy ends on Fabs (generated by proteolysis) do not create problems in cryoEM imaging; conformationally variable molecules can be impossible to crystallize but have been successfully imaged by cryoEM (for example, see [21–23]). Nevertheless, if the disorder is a continuum (for example, a domain sampling different position with respect to the rest of the macromolecular assembly), it presents a level of difficulty that cannot be easily solved with the tools

currently available, although some progress has been made[24,25]. Another common misconception is that sample requirements for cryoEM are lower than what is necessary for crystallography. Although it is true that the sample needed for initial grid screening is less than the amount necessary to set up initial crystallization screenings, a full cryoEM optimization may end up demanding almost the same amount of total material. Lastly, the size of the macromolecule is important: although there have been reports of ~50kDa proteins successfully imaged by cryoEM[26–28], obtaining atomic level resolution for molecules smaller than 150-200 kDa is still challenging. A common strategy is to artificially increase the size of the macromolecule of interest by complexing it with Fab(s)[29,30] or known binding partners. This approach comes with a caveat. For a successful cryoEM high resolution structure determination, the "ordered" mass that can be visualized is important, not the absolute mass of the particle. If the "helper" molecule ends up assuming different orientations with respect to the target molecule, it will not increase the mass of the static portion of the particle. That is also the reason why protein expression partners such as GFP or GST usually are not helpful in cryoEM. Furthermore, when working with complexes, it is important to consider the relative affinity of the components with respect to the concentrations used in designing the experiments, and possibly confirm complex formation via SEC chromatography or mass photometry[31].

Conclusions. CryoEM is increasingly becoming the technique of choice to structurally enable drug discovery processes. In order to fully meet pharmaceutical needs of reproducibility, throughput, and cost, some challenges still remain, including sample preparation and instrument accessibility. Despite these current bottlenecks, cryoEM has already clearly demonstrated to be essential in expanding the space of structurally enabled targets and the breadth of structure-based drug design.

References
1. Dubochet, J., et al. J Microscopy 128, 219–237 (1982).
2. Kühlbrandt, W. Science 343, (2014).
3. Lau, C., et al. J Physiol. 596, 1107 (2018).
4. Safdari, H. A., et al. Trends Cell Bio 28, 591 (2018).
5. Ognjenović, J., et al. Ann Rev Biomed Engineering 21, 395 (2019).
6. Nogales, E. &Scheres, S. H. W. Mol Cell 58 677 (2015).
7. Scapin, G., et al. Cell Chem Bio 25, 1318 (2018).
8. Ceska, T., et al. Biochem Soc Trans 47, 281 (2019).
9. Renaud, J. P. et al. Nature Rev Drug Disc 17, 471 (2018).
10. García-Nafría, J. & Tate, C. G. Ann Rev PharmacolToxicol 60, 51 (2020).
11. Saur, M. et al. Drug Discovery Today 25, 485-490 (2020)
12. Zhang, X. et al. bioRxiv 2021.01.11.426276 (2021)
13. Stuart, D. I et al. Nature Methods 13, 607 (2016).
14. Democratizing cryo-EM: Broadening access to an expanding field | Science | AAAS. https://www.sciencemag.org/features/2020/03/democratizing-cryo-em-broadening-access-expanding-field.
15. Clare, D. K. et al. Acta Cryst Section D 73, 488 (2017).
16. Sader, K et al. Acta Cryst Section D 76, 313 (2020).
17. Tegunov, D. & Cramer, P. Nature Methods 16, 1146 (2019).
18. cryoSPARC Live | Real-time SPA Cryo-EM Software Systems. https://cryosparc.com/live.
19. Ravelli, R. B. G. et al. Nature Communications 11, 1 (2020).
20. Dandey, V. P. et al. J Struct Bio 202, 161 (2018).
21. Roh, S. H. et al. PNAS USA 114, 8259 (2017).
22. Zhao, Y. & Chiu, W. Microscopy and Microanalysis 25, 1006 (2019).

23. Sobti, M. et al. eLife 8, (2019).
24. Maji, S. et al. J Chem Info Mod 60, 2484 (2020).
25. Nakane, T K et al. eLife 7, (2018).
26. Fan, X. et al. Nature Communications 10, 2386 (2019).
27. Herzik, M. A., et a. Nature Communications 10, 1032 (2019).
28. Khoshouei, M., et al. Nature Communications 8, 16099 (2017).
29. Nygaard, R., et al. CurrOpin Struct Bio 64, 26 (2020).
30. Wu, S. et al. Structure 20, 582 (2012).
31. Soltermann, F. et al. bioRxiv 2020.01.31.925156 (2020).