Published in final edited form as:
Drug Discov Today. 2014 October ; 19(10): 1505–1507. doi:10.1016/j.drudis.2014.07.002.

Progress in macromolecular crystallography depends on further miniaturization of crystallization experiments

Fabrice Gorrec

MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge Biomedical Campus, Cambridge CB2 0QH, UK, Desk: +44 (0) 1223 26 7807, Fax: +44 (0) 1223 26 8305, fgorrec@mrc-lmb.cam.ac.uk

X-ray crystallography is currently by far the most successful technique for determining the structures of biological macromolecules, their complexes and assemblies to high-resolution (better than ~ 3.5 Å). These structures have been proven to assist the rational design of new drug-like molecules and they are invaluable in helping us to understand life at the atomic level. Nowadays, X-ray crystallography is used in synergy with other biophysical techniques, especially electron microscopy, nuclear magnetic resonance and mass spectrometry.

Macromolecular crystallography has benefitted from a number of distinct advances over its recent history, for example recombinant technologies, more powerful synchrotron X-ray sources, automation and liquid handling of small volumes. The result has been an exponential increase in structures deposited at the Protein Data Bank (PDB) over the last two decades. Currently, the PDB holds more than 100,000 entries, 90 % of which have been determined by X-ray crystallography and most are proteins (there are also many RNA- and DNA-containing structures). As a result of the advances, crystallographers aim at solving increasingly larger and more complex structures that could not be tackled only a few years ago. For example, in 2006, the first complete, high-resolution structure of an entire bacterial ribosome was solved by X-ray crystallography [1]. The multitude of ribosome structures now available aids the quest for new antibiotics that target the ribosome. Another example is GPCRs, important trans-membrane proteins that for many decades proved recalcitrant to our efforts to crystallise. Again, a number of GPCR structures have already enlightened relevant areas of pharmacology [2].

Given the large number of reported structures, one might think that the structure determination process is straightforward. Unfortunately, this is far from the truth. For example, it is estimated that 21 % of *Escherichia coli* genes (*E. coli* has around 900 genes) and 26 % of human genes (~ 5539 genes) encode for membrane proteins while at present there are only about 350 unique membrane structures deposited in the PDB (265 structures in 2010 [3]). Maybe more importantly, some structures still require monumental efforts for their determination. We have only started to look at all the interactions of proteins, which all need to be investigated structurally if we want to understand them at the atomic level. In addition, a particular biological system often requires many different structures to be solved in order to answer questions asked about its mechanisms and evolution. This is demanding when it takes a few years, sometimes decades, for a particular state or structure to become

Europe PMC Funders Author Manuscripts

available through continuous trial and error. In other words, a vast amount of work is yet to be done in the field of macromolecular crystallography. Standardisations and process enhancements are urgently needed in order to make the methods accessible and workable on a large number of remaining problems.

One of the underlying issues in the structure determination process is the number of steps required, each of these steps being a potential dead end. In fact, large-scale statistics compiled from leading Structural Genomics Centers and Protein Science Initiatives (PSI) worldwide show that roughly two structures are solved for every 100 proteins originally investigated [4]. Sample production, usually by heterologous expression in the case of proteins is the first major bottleneck that must be tackled to obtain milligram quantities necessary for a successful structural biology project [5]. In most cases these days, an approach involving a multitude of assays run in parallel on multiple different samples is employed to minimise risk. These can include different constructs, complexes, nucleic acid sequences, source organisms, *etc*, which not only increases the probability of success but they may also result in structures of different biologically relevant states of the protein that are informative in understanding biological function.

Crystallization is the second major bottleneck in the process. It is often half-joked that “crystallizability” is inversely proportional to biological interest. For successful structure determination, crystals of diffraction quality are required, which means they have to be reproducible and ordered in all crystal directions such that diffraction data extend to at least 3.5 Å. Crystals also have to be large enough, albeit this is becoming less of a problem through increasingly more powerful X-ray sources. Since large macromolecular complexes are often rather flexible, it can be very hard for them to form a regular crystal lattice with the qualities required for solving their structure. The corresponding samples often exhibit poor stability and homogeneity that compounds this problem [5]. If and when crystals eventually grow, they often have unfavourable characteristics such as small size, high mosaicity, high solvent content, limited resolution, mechanical weakness, bad morphology, large number of molecules in unit cell, twinning and limited shelf life. All of these problems occur frequently and make the structure determination process demanding.

There are an almost infinite number of parameters that can be altered during crystallisation trials and it is often not clear which combinations to vary or not. As a result, a novel crystal structure generally results from thousands - if not tens to hundreds of thousands – individual crystallization experiments. Subsequently, macromolecular crystallization has been driven towards miniaturized and automated approaches that make those numbers manageable and reasonably cost effective. Over the past decade, the most common approach for crystallization - if such a thing exists - is the use of the vapour diffusion technique with 100 nanoliter droplets for each experiment in specialised 96-well trays (with standard foot-print and well-spacing) and commercially-available robotic liquid handlers [6]. This has meant that the initial screening for hits from each available sample is a standardised and routine process that consumes up to a few milligrams of sample while the price of one experiment (a single droplet) can be as low as 0.25 GBP. The size of initial crystals eventually produced using this technology is typically in the range of 10-100 micrometers.

Currently, problems arise when less than a milligram of sample is available and, unfortunately, this is certainly a more and more frequent occurrence. In this case, the standard screen cannot be run entirely and chances of success are reduced considerably (or worse, one can simply not proceed with crystallization). Subsequently, I would argue that the size of routine crystallization assays needs to be further reduced by at least an order of magnitude (to 1-10 nanoliters). Actually, advances in liquid-handling using sonic dispensing [7] and microfluidic chip technology [8] already enable such reduction. However, cost and variability increase dramatically when working with volumes below 100 nanoliters, hence the current trend is to stick to that range. For example, despite the much higher density achievable in microfluidic chips compared to standard plastic trays for vapour diffusion, the latter remain far cheaper to produce and integrate. This has meant that until now microfluidics have played only a very small role. Equally, sonic dispensing is currently an expensive technology, which has limited its uptake. We can envisage, however, that new technologies reducing the volumes required for the automated setup of crystallization experiments will in the future enable us to move to much larger screens - including duplicates - and also container technology that will enable screening of crystals in X-ray beams at high frequencies. Ideally, the tedious optical inspections of assays should be minimized.

In this context it is worth mentioning that recent investments in new X-ray sources and beamlines mean crystal size is becoming much less limiting. Structures can now be solved with data collected from small regions of typical crystals (i.e. below 10 μm) using microfocus X-ray beamlines [9] and this trend will continue as X-ray focusing technology gets better and synchrotrons produce smaller source foci. Even more exciting, data collected from crystals as small as 200 nm using an X-ray free electron laser (XFEL) resulted in a successful crystal structure determination by molecular replacement of photosystem I [10], one of the largest membrane protein whose structure has been determined by X-ray crystallography. In addition, handling many very small crystals can be facilitated by screening crystals directly either in special trays [11], in a stream of liquid using a flow-jet [10] or with acoustic technology [7]. It is clear that a concerted development of crystallization with other enabling technologies - such as systems for crystal detection and alignment, and merging of datasets from multiple crystals - is necessary to progress the field of X-ray crystallography [11]. However, novel approaches must be cost-effective to become wide spread (as the sonic dispensing and microfluidic technology cases have proven).

With further miniaturization of the crystallization assays, it will be possible to proceed with crystallization experiments even before sample production is fully optimized. Also, a much larger set of initial crystallization conditions could be used: this will increase chances of obtaining diffraction-quality crystals while reducing the need for later optimizations [12]. In addition experiments may be run in duplicate, avoiding the problem of single experiment failures. Finally, the behaviour of samples during crystallization assays could be systematically investigated to rationally guide subsequent experiments [13].

Abbreviations

XFEL X-ray free electron laser

GPCR	G protein-coupled receptor
MRC-LMB	Medical Research Council Laboratory of Molecular Biology
PDB	Protein Data Bank
PSI	Protein Structure Initiative

References

1. Selmer M, et al. Structure of the 70S ribosome complexed with mRNA and tRNA. Science. 2006; 313:1935–1942. [PubMed: 16959973]
2. Congreve M, Marshall F. The impact of GPCR structures on pharmacology and structure-based drug design. Br J Pharmacol. 2010; 159:986–996. [PubMed: 19912230]
3. Fagerberg L, et al. Prediction of the human membrane proteome. Proteomics. 2010; 10:1141–1149. [PubMed: 20175080]
4. Khurshid S, et al. Porous nucleating agents for protein crystallization. Nat Protoc. 2014; 9:1621–1633. [PubMed: 24922271]
5. Bill RM, et al. Overcoming barriers to membrane protein structure determination. Nat Biotechnol. 2011; 29:335–340. [PubMed: 21478852]
6. Stock D, et al. Robotic nanolitre protein crystallisation at the MRC Laboratory of Molecular Biology. Prog Biophys Mol Biol. 2005; 88:311–327. [PubMed: 15652247]
7. Soares AS, et al. Acoustically mounted microcrystals yield high-resolution X-ray structures. Biochemistry. 2011; 50:4399–4401. [PubMed: 21542590]
8. Zhu Y, et al. Nanoliter-scale protein crystallization and screening with a microfluidic droplet robot. Sci Rep. 2014; 4:5046. [PubMed: 24854085]
9. Riekel C, et al. Protein crystallography microdiffraction. Curr Opin Struct Biol. 2005; 15:556–562. [PubMed: 16168633]
10. Chapman HN, et al. Femtosecond X-ray protein nanocrystallography. Nature. 2011; 470:73–77. [PubMed: 21293373]
11. Axford D, et al. In situ macromolecular crystallography using microbeams. Acta Cryst. 2012; D68:592–600.
12. Gorrec F. The current approach to initial crystallization screening of proteins is under-sampled. J Appl Crystallogr. 2013; 46:795–797. [PubMed: 23682195]
13. Anderson MJ, et al. Phase knowledge enables rational screens for protein crystallization. Proc Natl Acad Sci U S A. 2006; 103:16746–16751. [PubMed: 17075056]