Review

Sample preparation of biological macromolecular assemblies for the determination of high-resolution structures by cryo-electron microscopy

Holger Stark* and Ashwin Chari

Research Group of 3D Electron Cryomicroscopy, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, Göttingen D-37070, Germany

*To whom correspondence should be addressed: E-mail: hstark1@gwdg.de

Received 25 September 2015; Accepted 5 November 2015

Abstract

Single particle cryo-EM has recently developed into a powerful tool to determine the 3D structure of macromolecular complexes at near-atomic resolution, which allows structural biologists to build atomic models of proteins. All technical aspects of cryo-EM technology have been considerably improved over the last two decades, including electron microscopic hardware, image processing software and the ever growing speed of computers. This leads to a more widespread use of the technique, and it can be anticipated that further automation of electron microscopes and image processing tools will soon fully shift the focus away from the technological aspects, onto biological questions that can be answered. In single particle cryo-EM, no crystals of a macromolecule are required. In contrast to X-ray crystallography, this significantly facilitates structure determination by cryo-EM. Nevertheless, a relatively high level of biochemical control is still essential to obtain high-resolution structures by cryo-EM, and it can be anticipated that the success of the cryo-EM technology goes hand in hand with further developments of sample purification and preparation techniques. This will allow routine high-resolution structure determination of the many macromolecular complexes of the cell that until now represent evasive targets for X-ray crystallographers. Here we discuss the various biochemical tools that are currently available and the existing sample purification and preparation techniques for cryo-EM grid preparation that are needed to obtain high-resolution images for structure determination.

Key words: Cryo-EM, macromolecular complexes, high-resolution, biochemistry, sample optimization

Obtaining macromolecular complexes for high-resolution cryo-EM

Molecular machines or macromolecular complexes are supramolecular assemblies of biomolecules with a variety of functions. These may constitute soluble complexes in the cytoplasm, the nucleus or any other subcellular organelle, proteins associated with the intracellular or extracellular cytoskeletal frameworks or intrinsic membrane proteins with their associated proteins. Obtaining these complexes in a highly purified, yet intact state is often tedious owing to
their compositional complexity and the associated relative structural instability. However, the ability to purify macromolecular machines in an intact, stable manner is imperative to mechanistically interrogate their function within cells and to elucidate three-dimensional (3D) structures. Thus, the starting point in sample preparation for any successful structure determination endeavor, whether X-ray crystallography or cryo-EM, is a protocol for the biochemical purification of the samples of interest in an intact and stable manner (Fig. 1).

The recent years have seen the advent of many sophisticated recombinant technologies to enable the overexpression and purification of virtually any molecular machine from the various kingdoms of life. The extensive description of all these methods would far exceed the scope and size constraints of this review. We therefore draw the attention of the interested reader to a series of excellent recent reviews, which discuss these methods in detail [1–14]. In the following, we thus only highlight the most prominent, which have led to some high-resolution cryo-EM 3D structures in the recent months.

The assembly of macromolecular complexes, on first biophysical principles, depends on the diffusion-driven, random encounter of individual subunits to form the entire ensemble [15]. However, many individual subunits of a given molecular machine complete their folding process in the context of the final complex and as a consequence, expose large hydrophobic stretches in isolation [16,17]. These properties of individual macromolecular complex subunits make them susceptible to aggregation and instability, when overexpressed alone [18]. Multiprotein expression strategies are therefore often necessary to obtain complexes in a soluble and stable manner. This can either be achieved making use of the natural operon mode of gene expression in bacteria (e.g. E. coli), or the baculovirus expression system (BVES) by co-infecting cells with a multitude of viruses. A potential caveat of the former strategy is that bacteria do not encompass the machinery required to fold proteins from eukaryotic sources and additionally the need to establish efficient cloning strategies for a multitude of expression constructs. The co-infection strategy with BVES requires a fine titration of individual virus titers. This leads to concerns of reproducibility among different virus stocks and a high logistical demand to maintain stocks for large numbers of subunit-expressing viruses. Berger and colleagues have developed

![Fig. 1. Schematic overview of the sample preparation workflow for the determination of high-resolution structures by Cryo-EM.](https://academic.oup.com/jmicro/article-abstract/65/1/23/2579661)

The sample preparation workflow consists of three general modules: biochemistry (left panel), electron microscopy (middle panel) and 3D structure determination (right panel). The workflow starts with the biochemistry module. Macromolecular complexes can either be obtained employing recombinant expression schemes or from native sources and can be purified using standard biochemical procedures. At this stage, the quality of the biochemical preparation is verified by negative stain EM (middle panel, top). Depending on the quality of the negative stain EM preparation, the sample is subjected to optimization procedures (quality control; left panel, bottom part), or if the sample appears homogenous in negatively stained images one proceeds with the preparation of unstained cryo-EM grids (middle panel, bottom). Both negatively stained samples as well as cryo samples can yield initial 3D models of the sample under study (right panel). Noteworthy, to determine an initial 3D model, the random conical tilt method can be of enormous help (right panel, top). Cryo-EM specimens are capable of yielding high-resolution structures of macromolecular complexes (right panel, bottom part). Should determination of high-resolution structures be possible, the determined structures should be validated (right panel, bottom), models built and can aid in the interpretation of functional complexes (left panel, bottom). Knowledge about functional modules of the specimens under study can yet again aid in the biochemical optimization of macromolecular complexes (left panel), which might yield samples from which structures of higher resolution can be determined.
elegant solutions to tackle multiprotein expression not only in *E. coli* and BVES, but also in mammalian cell lines and the baker’s yeast *Saccharomyces cerevisiae* [19–25]. This so-called ACEMBL pipeline is a new robot-assisted concept for high-throughput multigene assembly by a technology called tandem recombiner (TR) and a fully automated pipeline for producing complex multiprotein specimens [25]. An innovative feature within the ACEMBL pipeline is to express multisubunit complexes as polyproteins, which compensate for the imbalance of expression levels of individual subunits [26]. This approach mirrors the strategy employed by many viruses to express their genomes. The genome is translated in the form of polyproteins, which is then processed by site-specific proteases to yield the individual proteins. When applied to recombinant multiprotein complexes, the individual open reading frames are interrupted by engineered TEV protease cleavage sites and the TEV protease is co-expressed as the terminal part of the polyprotein [26]. Polyproteins not only normalize the expression levels among different subunits but might also promote the accurate folding and assembly of different parts of the macromolecular complex by spatial proximity of the individual subunits. The latter feature is exemplified by the recent successful structure determination of the heterotrimeric influenza virus RNA polymerase, which was recalcitrant to recombinant production without the polyprotein strategy [27]. The engineered plasmids and biological materials required for all aspects of the ACEMBL pipeline are commercially available from Geneva Biotech. The strength of the ACEMBL pipeline for high-resolution structure determination by cryo-EM is documented by the fact that many recombinant specimens were generated using this technology.

Despite the strength of recombinant systems in generating multiprotein complexes for structural biology as discussed above, purification from native sources is not to be neglected [13]. This is often achieved in a straightforward manner by classical biochemical purification, when the macromolecular complex of interest is present in high copy numbers. In these cases, the source for macromolecular complexes can be tissues obtained from abattoirs, or large numbers of cells obtained by fermentation. Successful examples for this strategy include cytosolic and organelle ribosomes, chaperonins, eukaryotic multi-subunit RNA polymerases and respiratory chain complexes [13]. When classical biochemistry is difficult due to low copy numbers of the macromolecular machine of interest, they can be purified by employing affinity tagging strategies. Noteworthy in this regard is the TAP-tagging strategy originally introduced by Seraphin and coworkers [28,29]. Many TAP-tagged *S. cerevisiae* strains are commercially available for example from Cellzome. The strength of native complex purification lies in two primary facts. (i) Macromolecular complexes often use intricate biogenesis pathways in living cells to be assembled from parts [30–32]. This introduces a potential conundrum even with multiprotein expression strategies if one considers that the associated assembly-chaperones and -factors may not be present in the given overexpression host. Also overexpression can easily overload assembly pathways even if present in the expression host. In proof of this notion, the successful recombinant production of multisubunit eukaryotic RNA polymerases has yet to be reported [13,33]. (ii) Many interactions among subunits in macromolecular complexes rely on directed post-translational modification. It is common knowledge that overexpressed subunits/complexes are often post-translationally modified in a promiscuous manner, favoring native purification.

**Structural dynamics of macromolecular complexes**

After the successful crystallographic structure determination of ribosomal subunits in 2005, Moore and Steitz [34] wrote in a review article: ‘Crystallographically, the ribosome might turn out to be the lysozyme or chymotrypsin of large macromolecular assemblies’. After 2 years of successful high-resolution structure determination by cryo-EM, a similar picture emerges. Overall in the EM Databank (EMDB), entries below 4 Å resolutions are mostly characterized by inherently stable complexes, with 33% of entries comprising ribosomes and functional complexes thereof, 43% viruses and 10% cytoskeletal filaments. In 2014, ribosomal entries below 4 Å increased to a proportion above 55%. Even in 2005, Moore and Steitz [34] predicted: ‘The challenges posed by these [macromolecular complexes other than ribosomes] systems are enormous, but, in fact, they are almost entirely biochemical…’ Having purified the macromolecular complex of interest either using recombinant strategies or native purification, one would like to proceed with structure determination (Fig. 1). However in keeping with the prediction of Moore and Steitz [34], even after obtaining large amounts of biochemical specimens in high concentrations and apparently homogenous by SDS-PAGE, it is not uncommon to experience that negative stain EM grid preparation reveals a high degree of heterogeneity.

An important feature, which one has to keep in mind, is that most macromolecular complexes in the cell carry out their functional tasks by binding to various other protein factors and ligands during their functional reaction cycle. Cryo-EM studies of these complexes therefore inherently image molecules that may be in a variety of functional states, which may be associated with significant conformational changes of the molecules. The internal order of macromolecular complexes studied by cryo-EM will therefore be significantly lower compared with the order in a crystal in crystallographic studies. For some molecular machines such
as the ribosome, antibiotics are available, which allow the stalling of distinct functional states. This is a very elegant method to study the structure of a complex at high resolution by cryo-EM, which simultaneously answers questions about how these large macromolecular complexes perform their respective reactions [35]. However, there are numerous macromolecular complexes in the cell that might be difficult to arrest biochemically in a given functional state. On the computational side, great progress was made recently to tackle heterogeneity problems by means of image processing. However, it would always be desirable to have biochemical tools to arrest the molecules in interesting functional states that are required to enhance the functional understanding.

A computational approach to tackle the heterogeneity problem is to mask certain regions of interest during the image processing and restrict the structure determination to the most stable areas of the complex [36,37]. Whereas this is the fastest way to obtain high-resolution structures, one may not get a defined structure for the more mobile and structurally less well-defined parts of the complex. The real advantage of cryo-EM over X-ray crystallography is however that the method promises to get both, high-resolution structures and information about the stable and the dynamic parts of macromolecules alike [38]. To get there routinely, aside from improved strategies for the production and purification of macromolecular complexes, further methodological developments in biochemistry as well as in computational image processing are required essentially.

Proteoplex

The rationale for ProteoPlex stems from the longstanding observation that a stringent control of solvent pH is one of the strongest determinants to preserve protein quaternary structure [41]. This is easily evidenced by surveying the pH values used for the determination of X-ray crystal structures in the Protein Data Bank (PDB). The pH range in the PDB spans 4.5–9 with no difference between macromolecular complex entries and the full PDB (Fig. 2). When considering that successful crystallization can only occur if the sample is present as a thermodynamically stable, monodisperse species in solution, this observation indicates the pH optima of many purified molecular machines span a wide pH range often distinct from neutrality. It is therefore surprising that most structures deposited in the Electron Microscopy Databank (EMDB) were determined in a narrow pH range spanning between pH 7 and 8.5. The structures of many macromolecular complexes derived by single-particle cryo-EM, therefore, may have been determined far from their thermodynamic optima. Moreover, the difficulties in determining structures of some macromolecular machines could be a consequence of non-optimal buffer conditions. Considering the difficulties in purifying macromolecular complexes and the associated low yields, it is understandable that many researchers do not systematically explore buffer conditions, but make an arbitrary, ‘best-guess’ choice and move on to structure determination.

To systematically explore stabilizing buffer conditions, ProteoPlex, the first available high-throughput method to study the stability of complexes utilizes a sparse-matrix approach. The basis for measurement is the unfolding of samples in the presence of an environmentally sensitive fluorescent dye, as first implemented in differential scanning fluorimetry (DSF) [42]. An environmentally sensitive fluorescent dye is hereby defined by its property to be largely quenched in polar environments, but to become hyper-fluorescent when in apolar environments (such as the unfolded hydrophobic core of proteins). While the measurement principles among DSF and ProteoPlex are similar, the most important distinctions lie in the careful choice of instrumentation to accurately measure unfolding transitions of macromolecular complexes and the manner in which resulting unfolding curves are interpreted. In DSF, stabilization is characterized by the increase in the melting temperature (defined by determining the mid-point of the unfolding transition) of the sample in various buffer conditions. While this approach is suitable to find stabilizing conditions of single proteins and/or small complexes (two to three subunits) that behave as two-state folders, it is unsuitable for large macromolecular complexes (Fig. 2c). The failure of DSF to perform with multisubunit complexes lies in the fact

Techniques for the stabilization of macromolecular complexes for high-resolution cryo-EM

When cellular machines are extracted from their cellular environment, it is advisable to find buffering conditions, which aim at mimicking the cellular milieu, to obtain homogenous and stable samples for structural biology. Two critical parameters have to be considered when optimizing the stability of macromolecular complexes. (i) The buffering conditions in which multisubunit specimens are purified strongly influence their stability, by possibly reflecting their cellular environment. (ii) Many macromolecular complexes require osmolytes such as glycerol or sucrose for their stability. These additives, which mimic the crowded cellular environment, however, are incompatible with cryo-EM grid preparation, because they do cause a significant loss of contrast in cryo-EM images. In the following, we will present two techniques – ProteoPlex and GraFix – established in our own laboratory, which are directed to counteract the aforementioned two pitfalls to cryo-EM grid preparation [39,40].
that they do not unfold cooperatively (i.e. two-state unfolding) under destabilizing conditions, but solely under stabilizing ones [39]. Therefore, in order to predict stabilizing conditions with ProteoPlex, unfolding transitions need to be found, which resemble two-state unfolding. This is achieved by a theoretical thermodynamic framework, which is tailored to describe the unfolding of large complexes, and a hierarchical sorting scheme to efficiently find stabilizing conditions. When the pH stability maxima of the more than 80 different macromolecular complexes measured so far by ProteoPlex are plotted along with the pH values mined from the PDB and EMDB (Fig. 2a), the ProteoPlex stability

![Fig. 2. Biochemical sample optimization of macromolecular complexes by ProteoPlex.](image)

(a) Survey of pH value distributions of structures deposited in the PDB. Depicted are macromolecular complexes (orange dotted line) and all entries (green dotted line). Structures determined by single-particle cryo-EM in the EMDB are indicated with a black dotted line. The optimized pH values of over 80 complexes assayed with ProteoPlex are also shown (orange solid line). (b) Analysis of Biomphalaria glabrata hemoglobin complex (Bghb, 1.5 MDa native molecular weight) before (left) and after (right) buffer optimization imaged by negative stain EM (scale bar: 20 nm). Under standard purification conditions in Tris (pH 7.4; left panel), Bghb is found to be present as aggregated and dissociated particles in EM images. In imidazole (pH 5.8), the buffer determined as optimal by ProteoPlex, EM analysis reveals a monodisperse field of compact particles (right panel). (c) Comparison of melting curve evaluation by ProteoPlex and conventional DSF. Results obtained by both methods do perfectly agree (green) only for complexes, which are already optimal. In early stages of a project, conventional DSF-based data evaluation does not provide useful results and there is no overlap between DSF and ProteoPlex results. This reflects a lack of biochemical control in early stages of a given project, characterized by multi-state unfolding behavior of complexes, which is mis-interpreted by standard DSF. The accuracy of finding stabilizing conditions for any kind of macromolecular complex is strictly dependent on the correct analysis of unfolding states, as implemented in ProteoPlex. Close to two-state unfolding behavior of macromolecular complexes correlates with stabilization of the sample under study. Improvement in stability and solubility of complexes requires an iterative procedure. Notably, we have increasingly experienced that it is essential to repurify complexes in the optimal pH/buffer conditions from the previous step, before subjecting them to an additional round of optimization.
maxima resemble the broad pH distribution found in the PDB. As a further step in the optimization of macromolecular complexes, additional stability may be afforded by accessory and inhibitory ligands. A principle advantage of ProteoPlex in comparison to DSF is that the method allows not only for screening small molecule stabilizers, but also proteinaceous co-factors and binders such as nanobodies and/or Fab fragments. Among all 80 different macromolecular complexes we have measured, we have seen an improvement in homogeneity in negative stain EM without fail, moreover it is customary that purification yields are enhanced when performed in the optimal buffer determined by ProteoPlex.

**GraFix**

GraFix is a density gradient centrifugation procedure, which combines a density gradient for size fractionation and fixation of macromolecular complexes by mild chemical fixation [40]. In contrast to batch fixation with chemical crosslinkers [43], the sample subjected to GraFix successively encounters an increasing concentration of chemical fixative, which largely prevents intermolecular crosslinks and favors intramolecular ones. Even if intermolecular crosslinks occur, they are removed from the intramolecularly crosslinked species owing to the simultaneous size fractionation. Since its inception, GraFix has been widely adopted by the EM community and it has been employed to a variety of complexes unrelated in composition and function [44]. In our opinion, GraFix offers some principal advantages for (cryo-) EM specimen preparation, which we will highlight in the following: (i) the concentrations of macromolecular complexes required for the preparation of EM grids with useful distribution of particles are relatively low (<0.1 mg ml\(^{-1}\)). These low concentrations often are below the dissociation constants of intersubunit interactions, such that partial dissociation of subunits is commonly observed and contributes to sample heterogeneity. The mild chemical fixation in GraFix prevents this dissociation. (ii) EM supports such as continuous carbon films are relatively harsh surfaces, which can be disruptive to complexes upon binding. Yet again the fixative in GraFix is beneficial to sample integrity. (iii) As alluded to above, most complexes require osmolytes for stability. However, cryo-EM grid preparation is best performed in the absence of such osmolytes, making the removal of glycerol or sucrose essential. At this stage, we and others frequently observe that complexes dissociate and appear inhomogeneous. When buffer exchange is performed by the use of desalting columns after GraFix treatment, the disintegration is not observed in most cases. We attribute this favorable property of GraFix yet again to the mild chemical fixation.

**Membrane proteins**

Integral membrane proteins (IMPs) constitute ∼30% of the proteome and represent the majority of pharmaceutical targets [45,46]. The recent years have seen that cryo-EM is capable of elucidating 3D structures of IMPs at high resolution [47–55]. When considering that crystallization of IMPs is tedious and very difficult to achieve, cryo-EM might well turn out to be generally superior for IMP structure determination. While all the aforementioned biochemical methods which apply for macromolecular complexes in general hold true, additional obstacles have to be tackled for membrane proteins. A major challenge for structural biology of IMPs is to find a suitable means to extract them from their native environment – the membrane – and stabilize them in an aqueous environment. This is generally achieved by solubilization of IMPs in detergents at concentrations higher than the critical micellar concentration (cmc). However, this harbors some potential difficulties: the free detergent can easily be mistaken for protein particles in electron micrographs [56], high detergent concentrations complicate the controlled formation of thin ice in cryo-EM grid preparation [56] and the detergent micelle bound to IMPs affects image processing.

Accordingly several approaches have been taken to remove free detergent in cryo-EM grid preparation of IMPs. These include the utilization of amphiphilic polymers, amphipols, protein caged lipid bilayers, nano-discs or engineered beta-sheet peptides [57–59]. Among these it is noteworthy that amphipols have enabled the subnanometer structure elucidation of TRPV channels and the γ-secretase complex [54,55]. The ryanodine receptor structures were determined either after reconstitution into nano-discs, in detergent-lipid bicelles or in simple detergent solution [48,52,53]. A potential alternative, which combines the strengths of GraFix described above with simultaneous detergent removal, is GraDeR. It utilizes the slow-off rate of the detergent LMNG along with the size fractionation and chemical fixation provided by GraFix [56]. This procedure increased the success rate of thin ice film formation in cryo-EM grid preparation (Fig. 3) and provided micelle-free high-contrast cryo-EM images of three different membrane proteins [56].

**Small proteins**

In contrast to X-ray crystallography, structure determination by cryo-EM becomes increasingly more challenging for smaller proteins. The reason is that each image needs to provide sufficient contrast to accurately align the particles with high accuracy. Unfortunately, contrast directly correlates with the size of the protein complexes. Despite these
drawbacks with regard to size, the limits have already been shifted toward smaller proteins and two high-resolution structures were published recently of complexes that can be considered relatively small for cryo-EM. The 465 kDa enzyme β-galactosidase was solved at the highest resolution of 2.2 Å yet obtained by cryo-EM [60] and the 170 kDa

Fig. 3. Overview of the Gradient Detergent Removal (GraDeR) workflow. (a) IMPs are solubilized in lauryl maltose-neopentyl glycol (LMNG) either by direct extraction or after exchange from another detergent. Free LMNG detergent micelles and monomers are removed by gradient centrifugation. LMNG-stabilized IMP in the absence of free detergent micelles and monomers can be used for single-particle cryo-EM. Inset: Chemical structure of LMNG, which basically represents two fused dodecylmaltoside molecules. The combination of a lipid-like hydrophobic tail and a hydrophilic headgroup big enough to prevent the formation of bilayers, together with a central, stiff quaternary carbon bond, convey LMNG with remarkable chemical properties. (b) GraDeR considerably improves the contrast of cryo-EM images. Depicted are cryo-EM images of the Thermus thermophilus V-ATPase prepared in the presence of dodecylmaltoside (DDM) detergent (left panel) or after GraDeR (right panel). Note the improved image contrast of V-ATPase particles after GraDeR (Scale bars, 20 nm).
γ-secretase was determined at 3.4 Å resolution [61]. An elegant approach to study even smaller proteins (<100 kDa) was recently introduced by the Cheng group [49,62], making use of single or double Fab fragments bound to the protein of interest. Two Fab fragments add another ∼70 kDa to the protein, providing sufficient signal in cryo-EM images to determine structures at the 1 nm resolution level.

**EM grid preparation**

Generally, there are two major techniques to prepare purified macromolecular complexes for electron microscopic imaging. Both techniques have in common that they need to protect the molecules from the high vacuum of the microscope. In the classical negative stain imaging, this is normally achieved by embedding the molecules in a layer of heavy metal salt solution. When dried, the heavy metal salt surrounds the molecules. Heavy metals scatter electrons more strongly than biological objects leading to an inverted (negative) contrast. Most commonly used heavy metal salts are uranyl acetate and uranyl formate, which provide very high image contrast and additionally serve as chemical fixatives. However, they are soluble only at low pH and only heavy metal salts such as ammonium molybdate or phosphotungstic acid can be used at neutral pH, albeit with the result of reduced contrast. Normally, negatively stained specimens are imaged at room temperature, working with dried-out samples. Depending on the staining conditions, size and molecular architecture of the molecules, this can lead to substantial flattening of the molecules on the grid [63]. Generally, negative staining is limited in resolution to ~15 Å. Better resolution can be obtained by a technique called cryo-negative staining [64,65], where the heavy metal salt-treated grid is not dried but rapidly frozen and imaged under cryo conditions in the electron microscope. Negative stain EM is generally still a powerful technique to obtain initial 2D and 3D structural information of a given complex. Additionally, it is extremely helpful when used as a quality control instrument during sample optimization, because it provides information about the structural integrity of molecules and their potential tendency to dissociate and/or to aggregate (Fig. 1). Often, the latter two parameters are difficult to appreciate by classical biochemical techniques used during sample purification. Furthermore, according to our experience there is generally very good agreement between the 3D structures obtained by negative stain and cryo-EM at resolution levels of 15–25 Å (Fig. 4).

High-resolution structure determination of macromolecular complexes is nowadays exclusively done in vitrified buffer solutions, where crystalline ice formation is prevented by the rapid freezing process [66]. The extremely high cooling rate of liquid ethane is mostly used to rapidly plunge-freeze cryo-EM grids. The desired vitrified water layer should only be slightly thicker than the molecules to obtain the best possible signal quality and contrast. Most macromolecular complexes are in a size range of 10–100 nm, which is smaller than the mean-free path lengths of 300 kV electrons [67] and makes macromolecular complexes ideally suited to be studied by cryo-EM. Although cryo grid preparation was a manual task of experienced experimentalists for a long time, cryo grids can now be prepared in an almost automated manner and several robotic freeze-plunging machines are commercially available.

**Grid supports**

With the availability of direct electron detectors in cryo-EM, it has become possible to record images as a series of movie frames. Instead of an integrated total image, one now obtains several tens of image frames on new generation pixel detectors, which allows studying the differences between the image frames by computational processing [68,69]. By following the particles from one image frame to another, it is now possible to computationally compensate for image blurring, which may be the result of stage drift, beam-induced particle motion and charging. The alignment of image frames minimizes the blurring of particle images and thus restores high-resolution information that was previously lost in blurred integrated images [68,70]. In spite of this computational possibility to restore the image quality, it is desirable to have as little motion as possible in the images. In recent publications, the use of graphene and gold grids was demonstrated to substantially contribute to a reduction of image blurring [71,72]. It is interesting to note that until now the most dramatic blurring effects occur within the first three to five electrons per Å². Usually this cannot be corrected efficiently by computational means and therefore these electrons are currently discarded from analysis. However, exactly those electrons potentially carry the highest quality image information, because they carry the signal that is almost free of beam damage [73]. Further development of optimal grid supports in combination with image frame processing will hopefully provide access to these most valuable parts of the images in future. This could well be an essential future technological improvement that is needed to push the resolution limits of cryo-EM structures even further.

**Summary and conclusion**

Sample purification and preparation for cryo-EM plays an important – and often underestimated – role. Optimizing sample quality is not always straightforward, but still the
fastest way to successful high-resolution structure determination. Interestingly, we have observed for a number of macromolecular complexes that after successful optimization of all purification conditions, the complexes were not only optimized for cryo-EM imaging and structure determination, but we have also obtained initial crystals of such samples in our very first crystallization attempts \[39\]. This shows that the optimum conditions for cryo-EM and crystallization are probably very similar.

There are currently four examples of cryo-EM structures at \( \leq 3 \) Å resolution, for which structures have been elucidated by X-ray crystallography as well. A comparison between the EM and the corresponding X-ray structures reveals that the X-ray structures were solved at higher resolution with one exception \[35\]. Very recently, structures of more dynamic and flexible molecules such as complexes related to pre-mRNA splicing were determined at impressive resolutions \[74,75\]. Those complexes will remain difficult targets for crystallography, but also for cryo-EM, because it remains a challenging task in cryo-EM to determine the entire structure of such large and dynamic complexes at isotropic resolution. Recent computational tools were developed for local alignments of flexible regions within large complexes \[75\]. Ultimately, what is needed to understand functional aspects of these molecular machines are high-resolution structures of these macromolecular complexes in whole, with their dynamic domains shown in different conformational states. This will mainly require biochemical optimization, in combination with advanced computational image processing tools and will provide us with insights of macromolecular machines in motion at atomic or near-atomic resolution.

**Fig. 4.** Similarity of 3D structures determined by negative stain EM and cryo-EM. (a) Raw image of human anaphase-promoting complex/cyclosome (APC/C) after negative staining using uranyl formate (upper panel) and a 3D structure reconstructed by the single-particle method (bottom panel) at 18 Å resolution (scale bar, 20 nm). (b) Raw image of human APC/C obtained by cryo-EM (upper panel) and a 3D structure reconstructed by the single-particle method (bottom panel) at 18 Å resolution (scale bar, 20 nm). Note that the structures obtained from negatively stained and cryo images correspond well to each other (compare bottom panels of (a) and (b)).
Funding

Work in the authors’ laboratory is funded by the Max-Planck Society, by the German Research Foundation (DFG) grants (SFB860/TPA5) to H.S. and (CH 1098/1-1) to A.C. and an R&D grant by Instruct to H.S. and A.C., part of the European Strategy Forum on research Infrastructures (ESFRI) and supported by national member subscriptions.

Conflicts of interest

The authors declare competing financial interest. A.C. and H.S. have filed a patent application (WO2013034160 A1: ‘Methods for analyzing biological macromolecular complexes and use thereof’) and are currently developing a commercial product for ProteoPlex in cooperation with FEI Company.

References


