

Special Issue: Future of Cell Biology

Review

Cryo-Electron Tomography: Can it Reveal the Molecular Sociology of Cells in Atomic Detail?

Martin Beck^{1,*} and Wolfgang Baumeister^{2,*}

Traditionally, macromolecular structure determination is performed *ex situ*, that is, with purified materials. But, there are strong incentives to develop approaches to study them *in situ* in their native functional context. In recent years, cryo-electron tomography (cryo-ET) has emerged as a powerful method for visualizing the molecular organization of unperturbed cellular landscapes with the potential to attain near-atomic resolution. Here, we review recent work on several macromolecular assemblies, demonstrating the power of *in situ* studies. We also highlight technical challenges and discuss ways to meet them.

Introduction

Structural studies of the macromolecular repertoire that makes up a living cell have traditionally taken a reductionist or ‘divide and conquer’ approach: cells are taken apart and fractionated and the purified components studied by one of the methods structural biology has at its disposal. X-ray crystallography, nuclear magnetic resonance (NMR), and increasingly, electron microscopy (EM) have contributed to the vast and rich portfolio of molecular structures deposited in databases. This, in turn, has had an impact on our understanding of the molecular biology of the cell in a profound way. As a result of the demands on sample homogeneity and concentration, and therefore, the need to isolate and purify the molecules of interest, these approaches have limitations. Only rarely are biological functions performed by individual molecules; most cellular functions arise from interactions between the many different molecular species inhabiting cells [1]. Weak or transient interactions are lost when the effects of macromolecular crowding are reduced. Moreover, many fundamentally important supramolecular assemblies are so deeply rooted in their cellular environments that it is impossible to isolate them without violation of their structural integrity. The depletion of essential metabolites or the absence of post-translational modifiers might lead to nonphysiological conformational or oligomeric states, and their spatiotemporal modulation, for example, during the cell cycle, remains elusive when taken out of context. To what extent the structure and function of a macromolecular assembly observed in isolation is relevant to the *in situ* scenario is case dependent and in practice, usually unknown.

Postreductionist structural biology [2] aims at understanding the structure and function not of individual molecules, but of functional modules and their interplay in a cellular context—the **molecular sociology** of cells [3] (see *Glossary*). This is made possible by structural studies performed *in situ*, that is, in unperturbed environments. Given the stochasticity of cellular structures, only methods that do not rely on deterministic and hence repetitive structures

Trends

Recent technological advances in cryo-electron microscopy have caused a leap forward in the structure determination of isolated macromolecular assemblies by single particle analysis. This is referred to as the ‘resolution revolution’.

We are just beginning to take advantage of these and other technical improvements in *in situ* studies by electron tomography. Studying molecular assemblies and machines in unperturbed cellular environments has a huge discovery potential.

¹Structural and Computational Biology Unit, European Molecular Biology Laboratory, 69117 Heidelberg, Germany

²Department of Molecular Structural Biology, Max Planck Institute of Biochemistry, D-82152 Martinsried (Planegg), Germany

*Correspondence:
martin.beck@embl.de (M. Beck) and
baumeist@biochem.mpg.de (W. Baumeister).

Box 1. Cryo-EM Imaging Modalities

Cryo-EM refers to the visualization of frozen-hydrated biological materials, molecules or cells. Vitrification, that is, freezing without ice crystal formation of native (not fixed and not stained) samples provides the best structural preservation that is physically possible. Given the sensitivity of ice-embedded biological materials to electron irradiation, images must be recorded in a low-dose mode to minimize radiation damage. As a consequence, they usually suffer from a poor signal to noise ratio (SNR), making it necessary to retrieve high-resolution information by averaging-based noise reduction.

Different imaging modalities can be used to study frozen-hydrated samples: electron crystallography [45], single particle analysis (SPA) [46], and tomography [47]. Electron crystallography requires that the molecule under study exists in the form of a regular two-dimensional (2D) array; either natural or synthetic. This facilitates averaging and three-dimensional (3D) reconstruction. Electron diffraction is used to collect Fourier amplitudes, whereas phases are obtained by calculating Fourier transforms of images. Electron crystallography has been particularly useful for membrane proteins, which can be studied in quasineutral lipid environments; often to atomic resolution. However, a limitation is the need to solubilize them with detergents first and then to persuade them to assemble into well-ordered one molecule thick 2D arrays.

SPA is arguably a misnomer; the particles under study are single only in the sense that they are not integrated in a regular array but free standing. Ideally, they are embedded in a thin layer of ice in random orientations. For each particle of a large (up to 10^6 or more) set of individual particles, the orientation with respect to the beam must be determined before combining the data to yield a 3D reconstruction. Ideally, all individual particles are identical in structure but the method is tolerant to some heterogeneity; image classification allows the performing of some purification *in silico*. As a result of recent developments in instrumentation and software, often cryo-EM SPA attains atomic or near-atomic resolution. There is indeed a strong trend towards becoming the method of choice—in preference to X-ray crystallography—in structural studies of molecular machines and assemblies. A variation of SPA is helical reconstruction [48], whereby the individual segments of a helical arrangement are considered as single particles that have to fulfill the overall geometry of the helix under investigation.

Unlike the aforementioned methods, which critically depend on the existence of repetitive structures, electron tomography (ET) can provide 3D reconstructions of nondeterministic and hence variable objects such as organelles or cells. In ET, a set of projection images is acquired at different angular orientations of the sample. The images of such a tilt series are then aligned and computationally merged to form a 3D image; the tomogram. This is more informative than the component images because it separates features superimposed in the projection images. Cryo-ET [49] has to reconcile the need to maximize the angular range and number of projections, which determine resolution, with the requirement that all images have sufficient contrast for precise alignment, while keeping the cumulative dose subcritical. The development of procedures for automated tilt series acquisition made cryo-ET of vitrified biological samples practical. Tomograms that contain repetitive molecular features can be extracted computationally and subjected to averaging and classification using methods akin to those used in SPA to improve SNR and resolution; albeit with some differences. In SPA, the so-called ‘angular assignment problem’ needs to be addressed, that is, the orientation of each particle needs to be determined, while tomograms are distorted along the z axis because of the ‘missing wedge’ resulting from technical limitations in collecting data over the full angular range. This must be taken into account in subtomogram averaging (Box 3).

can provide information at this crucial level of the structural hierarchy. Cryo-ET (Box 1) fulfills these criteria: it provides molecular resolution information of cells and organelles unadulterated by specimen preparation. Rapid freezing ensures the best structural preservation that is physically possible, and resolution can, under favorable conditions, attain the subnanometer range [4,5].

People have different views with regard to the meaning of structural biology *in situ*. Therefore, it may be useful to offer some clarification: *sensu stricto* the term should apply only to a scenario in which the cellular environment is preserved in its entirety, that is, all organelles and supramolecular assemblies are kept at their native place. *Sensu largo*, it may refer to situations where the local environment of a macromolecular assembly is either preserved or reconstituted such that some functionality is maintained.

In this review, we discuss several examples highlighting both aforementioned scenarios. We show how the application of *in situ* approaches has led to new insights into the function of macromolecular machines that had remained elusive to traditional structural biology. We discuss the technical challenges, potential solutions (Boxes 1–3), and the prospects for studying the molecular sociology of the cell in atomic detail.

Glossary

Molecular census: determination of the equilibrium of coexisting functional states of a given macromolecular assembly inside cells or of molecular populations (quantitative visual proteomics).

Molecular sociology: description and analysis of functional molecular neighborhoods in cells.

Structural analysis *in situ sensu largo*: refers to experimental conditions under which parts of the cellular environment are meaningfully preserved, such as in isolated organelles.

Structural analysis *in situ sensu stricto*: refers to conditions under which the cellular environment is preserved in its entirety.

Subtomogram averaging: a structure determination method that relies on combining the information contained in several copies of the macromolecular assembly of interest contained in cryo-electron tomograms, that is the subtomograms, by computational particle averaging.

Template matching: a computational method that enables visual proteomics. The structure of macromolecular assemblies (or a library of templates) is correlated with the respective signatures found in cryo-electron tomograms to determine their positions and orientations.

Visual proteomics: experimental determination of the spatial coordinates of macromolecular assemblies within intact cellular volumes.

Close to Native Structural Analysis of Membrane-Associated Complexes

The structural and functional analysis of membrane-associated and embedded complexes is notoriously difficult using *ex situ* approaches, because the native biophysical properties and the topology of membranes *in situ* are poorly reproduced by detergents *in vitro*. For systems of moderate complexity, the *in vitro* reconstitution of membrane systems based on artificial vesicles or nanodiscs has proven to be a viable approach. Although this bears the risk of missing out on yet-undiscovered components, it brings the structural analysis closer to the *in situ* scenario by mimicking more closely the native membrane environment. For more complex biological systems that rely on interplay of a multitude of components, *in vitro* reconstitution is challenging, if not impossible. New insights into such complexes were obtained using ***in situ sensu largo*** approaches, where organelles were isolated and structurally analyzed. Only recently, methods have been developed that allow for the structural analysis of such complexes ***in situ sensu stricto*** (Box 2). They have revealed remarkable new insights, although their application hitherto has been limited to a few case studies.

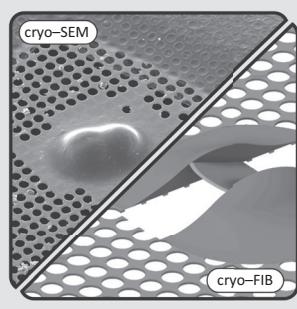
Coat Protein Complex I (COPI)

Vesicular trafficking is of importance for maintaining distinct compositions of various cellular organelles, for endocytosis and exocytosis, and for maintaining membrane homeostasis. Specialized complexes exist but they rely on coat proteins that facilitate membrane binding and fusion. COPI is a seven-membered protein complex that mediates retrograde transport [6,7]. It coats the membranes of the *cis* Golgi apparatus, allowing vesicles to bud off and travel to the endoplasmic reticulum or within the Golgi. Although high resolution structures of the components of COPI have been obtained by X-ray crystallography, and biochemical analysis had revealed how those form coatamers, the basic coating module that is recruited to membranes *en bloc*, it remained elusive how COPI oligomerizes to form a vesicle coat (reviewed in [8]). Recent work that investigated the structure of *in vitro* reconstituted COPI vesicles (Figure 1A) using cryo-ET and **subtomogram averaging** resolved this issue [8,9] (Boxes 1 and 3). Initial analysis at moderate resolution revealed that the basic structural element of COPI when bound to membranes is a triad of coatamers (Figure 1B). Strikingly, the triad forms various heterotypic interactions with its neighbors such that quaternary structure of the individual vesicles is variable, allowing vesicles of different sizes to form [8]. Nature flexibly reuses the same basic building block to fulfill multiple different structural tasks. Higher resolution structural analysis provided a sufficiently resolved envelope to fit into it the crystal structures of the individual proteins [9] (Figure 1C). The protein positioning is suggestive of specific interfaces that have been elusive to traditional approaches but could be confirmed by crosslinking mass spectrometry (MS) (Box 3) applied to reconstituted vesicles. Although the critical protein contacts have now been determined, it would be desirable to understand the COPI structure at an atomic level. To this end, the *in vitro* system might be further exploited at even higher resolution. Structural analysis *in situ sensu stricto*, even at moderate resolution, might reveal in which physiological context the different oligomeric states occur inside of cells (Figure 1A). They will also show how well reconstituted vesicles recapitulate the *in vivo* architecture and provide insights into their cargo.

Nuclear Pore Complexes (NPCs)

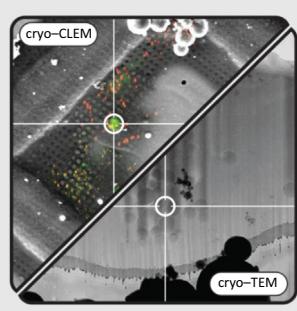
NPCs bridge the nuclear envelope to form an ~40 nm wide central channel that facilitates the exchange of macromolecules between the nucleoplasm and cytoplasm in all eukaryotes. Although NPCs are evolutionarily related to the vesicle coats, with which they share some components and domain folds [10], they are architecturally more complex. Whereas a vesicle coat binds to single layered, convex membranes, the NPC coat binds to a hole within the fused inner and outer nuclear membranes that display both convex and concave features. This biophysically more elaborate task is reflected by a more complex proteomic composition. NPCs comprise ~30 components called nucleoporins (Nups) that are organized in a modular manner. Certain subsets of Nups are preassembled into subcomplexes prior to NPC assembly *in vivo*;

Box 2. Technologies for Advancement of Cryo-ET

**FIB Micromachining**

A critical parameter in cellular ET is sample thickness. Cells thicker than 0.5–1 µm become nontransparent to medium voltage (<300 kV) electrons, often restricting structural studies to prokaryotes or thin peripheral regions or appendages of eukaryotic cells. Thicker samples must be thinned prior to their examination by EM, and the thinner the sample can be made, the higher the attainable resolution.

There are several options to achieve sample thinning: one is **cellular fractionation or laser microdissection for the isolation of organelles**. The disruption of the cellular context bears the risk of causing unwanted distortions of the structures under scrutiny. Cryo-sectioning, that is, physical sectioning of frozen hydrated samples [50] with a microtome preserves the cellular environment but requires highly specialized skills and is prone to artefacts such as substantial compressions (up to 40%) of cellular volumes. Focused ion beam (FIB) milling is becoming the method of choice for preparing thin (100–200 nm) samples of cells and tissues [51,52]. In this method, **sample thinning is achieved by ablating material with a beam of (gallium) ions**. Different milling geometries can be used, generating wedges with a thickness gradient or a lamella of uniform thickness. Although, presently, only cells cultured on grids can be processed, new developments are underway that will allow targeting cellular volumes in thicker tissues or multicellular organisms. FIB-based micromachining of frozen-hydrated samples is a gentle procedure ensuring an optimal preservation of the cellular landscape. In conjunction with cryo-fluorescence LM, it allows the precise targeting of cellular structures.

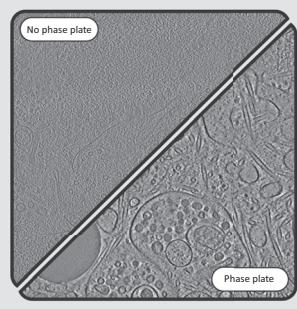
**Correlative LM-EM**

The large landscape of most eukaryote cells (typically several hundred micrometers) compared to the small fields of view of tomograms (typically one or a few micrometers) requires navigation to the features of interest. This can be achieved by combining LM for their identification and localization and then using their coordinates to guide examination in the EM or, if thinning is required, FIB micromachining. Correlative LM-EM (also referred to as cryo-CLEM) [53] targets features of interest with a precision of a few hundred nanometers. This will be of crucial importance in targeting rare or transient cellular events for structural analysis (see Concluding Remarks). An interesting perspective is the use of super-resolution LM in a correlative workflow, combining the power

of fluorescence microscopy in identifying molecular species with their visualization at high resolution in EM. The smaller the resolution gap, the more powerful correlative approaches will be. It is a distinct advantage of EM that the images do not merely represent the positions of a label in an otherwise invisible environment. EM image formation relies on the intrinsic spatial variation of density, therefore, the native molecular environment of the molecule of interest is visualized as well.

Detectors

Automated ET relies heavily on electronic devices for image recording, which provide direct access to digital images. Until a few years ago, only charge-coupled devices (CCDs) were available for data acquisition, limiting performance in SPA as well as in tomography. In CCDs, the impinging electrons must be converted into photons, whereas direct electron detectors based on complementary metal–oxide–semiconductor (CMOS) technology use electrons directly to generate image output. They outperform CCD cameras in terms of detective quantum efficiency (DQE) and in readout speed [54]. By operating in an electron counting mode, they correct for motions of the sample during exposure and thereby remove this often resolution-limiting effect [55]. The availability of commercial direct detectors has already changed the perspectives of cryo-EM in a profound way [56], and there is room for further improvement. Although it might not be possible to attain an ideal DQE of 1, with every electron contributing to the signal, a twofold improvement, for example, by optimizing the CMOS pixel design and the readout speed would be welcome, further reducing exposure and thereby minimizing radiation damage.



Phase Plates

Image formation of unstained biological materials relies on phase contrast. When images are recorded in focus, structural features remain almost invisible. Therefore, images are traditionally recorded slightly out of focus to generate defocus phase contrast. An undesirable effect of this mode of image formation is a phase contrast transfer function (CTF) with contrast reversals at some spatial frequencies and information loss at others; this makes it necessary to correct for the oscillations of CTF.

Phase plates overcome this problem. They are devices that produce phase contrast by introducing a phase shift between the scattered and unscattered waves at the diffraction plane of the microscope. They take images in focus with a continuous CTF across the entire spatial frequency spectrum

and with much improved contrast in the lower spatial frequency range. Consequently, the resulting images/tomograms can be interpreted directly and intuitively. Although earlier phase plate designs were cumbersome to use and had undesirable side effects such as 'fringing' of the images, the recently developed Volta phase plate [57] combines ease of use with durability. It holds promise to improve data quality in both SPA [58,59] and cryo-ET [22,60].

some of them can be biochemically isolated or reconstituted and structurally characterized *ex situ* [11]. Although many of the high-resolution structures of Nups have been resolved using X-ray crystallography, it remained subject to some debate how exactly Nup subcomplexes oligomerize to form the nuclear pore scaffold [12]. This has been primarily caused by the following experimental challenges: (i) hitherto nuclear pores can neither be reconstituted *in vitro* nor natively purified; and (ii) the inter-subcomplex contacts are often formed by short linear motifs contained in loops [13]. This nonrigid mode of interaction renders the higher order structure inherently inaccessible to structural studies relying on traditional methods. Only in the context of the fused inner and outer nuclear membranes do those structural elements accurately fall into place. Therefore the structurally rigid subcomplexes had to be reconstituted *in vitro* and analyzed by X-ray crystallography to provide high-resolution structures. The more flexible linkage between subcomplexes have been charted biochemically and by crosslinking MS *in vitro* (reviewed in [14]). Purified nuclear envelopes containing intact NPCs have been analyzed using cryo-ET with subsequent subtomogram averaging [15–18] (Box 1), complemented by super-resolution microscopy, fluorophore counting, and stoichiometric and proximity measurements by MS (reviewed in [14]) (Box 3). The convergence of this combined bottom up *in vitro* and top down *in situ sensu largo* approach has resulted in a pseudoatomic map of the nuclear pore scaffold [15,17–20] (Figure 2A). It revealed that Nups subcomplexes occur in multiple structural instances per asymmetric unit that conformationally adapt to their local environment. Similar to their COPI relatives, they form locally specific heteromeric interactions.

NPC structure has also been analyzed *in situ sensu stricto*, namely by cryo-ET within intact u2os [21] and HeLa [22] cells. A recent study [22] has implemented the direct electron detector, phase plate, and specimen thinning technologies (Box 2), and revealed the molecular sociology of NPCs in the context of nuclear membranes, the lamin meshwork, and cytoskeletal structures that have been partially lost in previous *in situ sensu largo* studies (Figure 2B). This was facilitated by an improved signal-to-noise-ratio (SNR) of the primary data (Figure 2C). From a relatively small data set, a low-resolution structure of the NPC could be obtained that is conformationally dilated as compared to previous analyses both *sensu stricto* [22] (not shown) and *sensu largo* [18] (Figure 2C); possibly facilitating the translocation of large cargo and explaining why NPC subcomplexes are flexibly connected. These findings highlight the enormous discovery potential of *in situ* structural biology. Furthermore, true atomic maps of the NPC would be highly desirable to understand mechanistically why alterations of the nucleocytoplasmic transport system are of human health relevance (reviewed in [23,24]).

Box 3. Retrieval of Information from cryo-electron tomograms**Image processing**

Because of dose limitations, tomograms of ice-embedded samples have a poor SNR. Image processing offers some remedy. Several approaches perform 'de-noising', that is, suppression of noise while preserving the signal. This is often sufficient to delineate features such as filaments or membranes by automated segmentation procedures [61,62].

The identification and localization of macromolecular structures can be achieved by template matching [63,64]. Templates are typically derived from high- or medium-resolution structures obtained from X-ray crystallography or SPA. They are used to perform a systematic search of tomographic volumes for matching structures—or alternatively, to identify repetitive patterns within the tomograms in an unbiased way. The size of structures that can be identified with high fidelity depends on the resolution of the tomograms. With improvements in resolution and large template libraries becoming available, it will be possible to perform a comprehensive mapping of molecular landscapes, the realization of visual proteomics [38].

When the approximate coordinates of macromolecules have been determined, higher-resolution structures can be obtained by subtomogram averaging [65,66]. In a procedure that is conceptually similar to SPA, 3D volumes are computationally extracted and subjected to classification and averaging. Iteratively, the angular orientation of the particle is refined and used to generate an improved structure. The alignment in subtomogram averaging is algorithmically and computationally more demanding because it is done in 3D and has to take into account distortions resulting from the missing wedge (Box 1). Under favorable conditions, resolutions in the subnanometer range can be attained *in situ*, and there is no fundamental limitation that would prevent attaining even higher (near-atomic) resolutions.

Data Integration

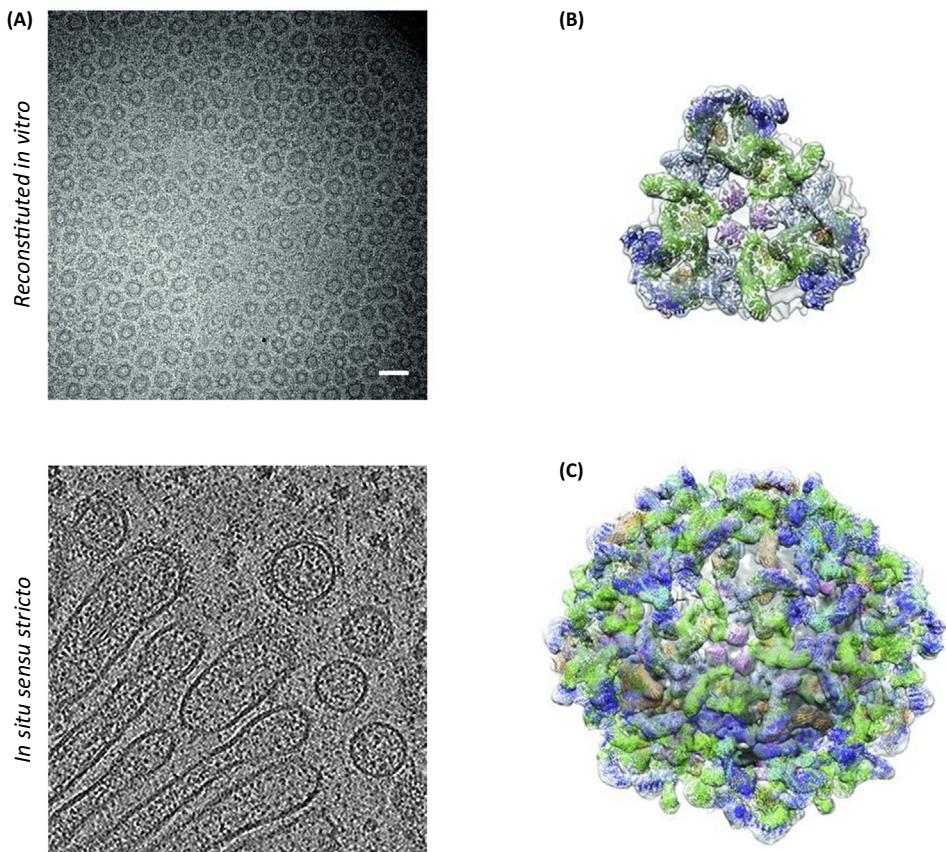
The fitting of X-ray structures into lower resolution EM maps has become common practice in SPA. Also the integration with MS methods has turned out to be powerful. Although the mass spectrometer measures signals arising from the combined bulk of molecular species, the obtained information is often synergistic to cryo-EM methods and can facilitate visual proteomics approaches. Quantitative MS [67] cannot only be used to identify the molecular inventory of compositionally complex, natively isolated specimen, but also to obtain accurate stoichiometries thereof. Crosslinking MS [68] identifies amino acids in close proximity to each other that can be translated into spatial restraints—similar to NMR, although more sparse. Such spatial restraints can be implemented into structural modeling procedures or alternatively, be used to validate those. Both the aforementioned MS methods can be carried out with relatively small amounts of sample. They are readily applicable to a complex specimen such as isolated organelles or even intact cells and thus ideally synergize with cryo-ET. Complementary information about protein stoichiometry, proximity, and conformational variability can be objectively integrated with structural data using molecular modeling frameworks that are reviewed elsewhere [69].

Cryo-ET as an OMICs Technology and Discovery Tool for Systems Biology?

In situ structural biology will need to cope with complex and big data sets in the near future that are challenging to interpret. As such, it is likely to benefit from experimental design and data analysis strategies that have been well established for OMICs technologies. For example, quantitative MS is often used to compare the abundance of proteins across different biological states, such as biological replicates of unperturbed cells subjected to some kind of stimulus. Statistical analysis of such data reveals which of the observed abundance changes are significant to facilitate their interpretation in the context of biological mechanisms. In structural biology *in situ*, protein abundances might be translated into the abundance of specific structural states such as protein complex association, proximity, or conformation and the same experimental design principles and statistical analysis methods could be used. Also, the combination with other OMICs methods that are often readily applicable to the same specimen could be powerful in studying biological mechanisms. Taken together, cryo-ET holds potential to become an unbiased discovery tool for systems biology similar to proteomics and genomics methods.

Ribosome-Translocation Machinery

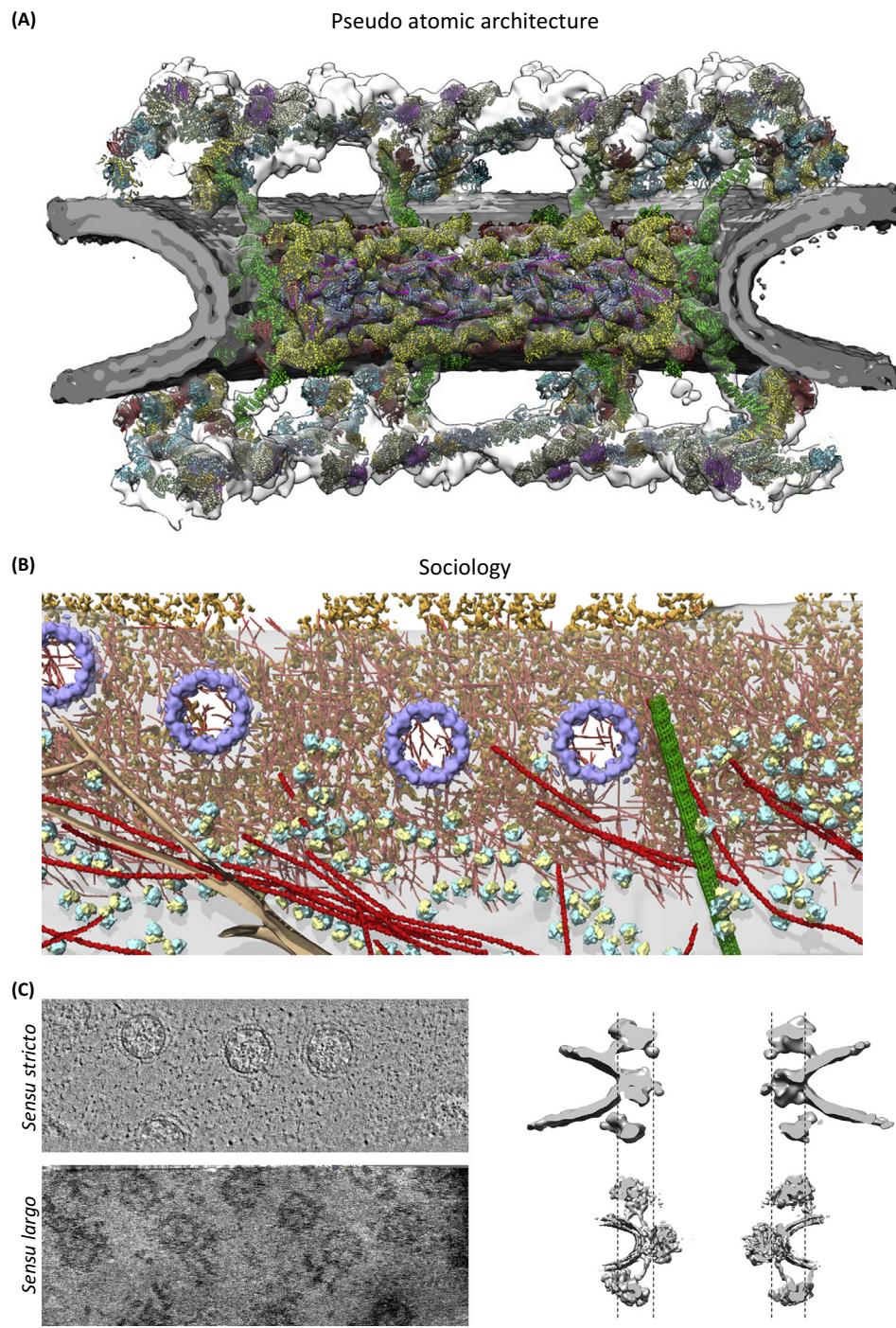
The translation machinery of the cell, the ribosome, exists in a free and in a membrane-bound form. Because of the cytoplasmic localization of ribosomes, proteins designated to end up in the lumen of the endoplasmic reticulum (ER) or in the Golgi apparatus must be translocated across or inserted into the ER membrane, respectively. To this end, the ribosome associates with Sec61, a heterotrimeric complex spanning the ER membrane (reviewed in [25]). Its core component, Sec61 α , facilitates the insertion of the native chain into the ER membrane. It had



Trends in Cell Biology

Figure 1. Structural Analysis of Vesicle Coats. (A) Electron micrograph of *in vitro* reconstituted vesicles (top) in comparison to trafficking vesicles as observed *in situ sensu stricto* (bottom, tomographic slice). (B) A triad of coatomer complexes is the basic structural element of the COPI coat and observed only in the context of membranes. Within the coat, it forms various contacts with its neighbors to build vesicles with a variable size; one of which is shown in (C). Images adapted, with permission, from [9], except for bottom left panel in (A) (courtesy of Ben Engel).

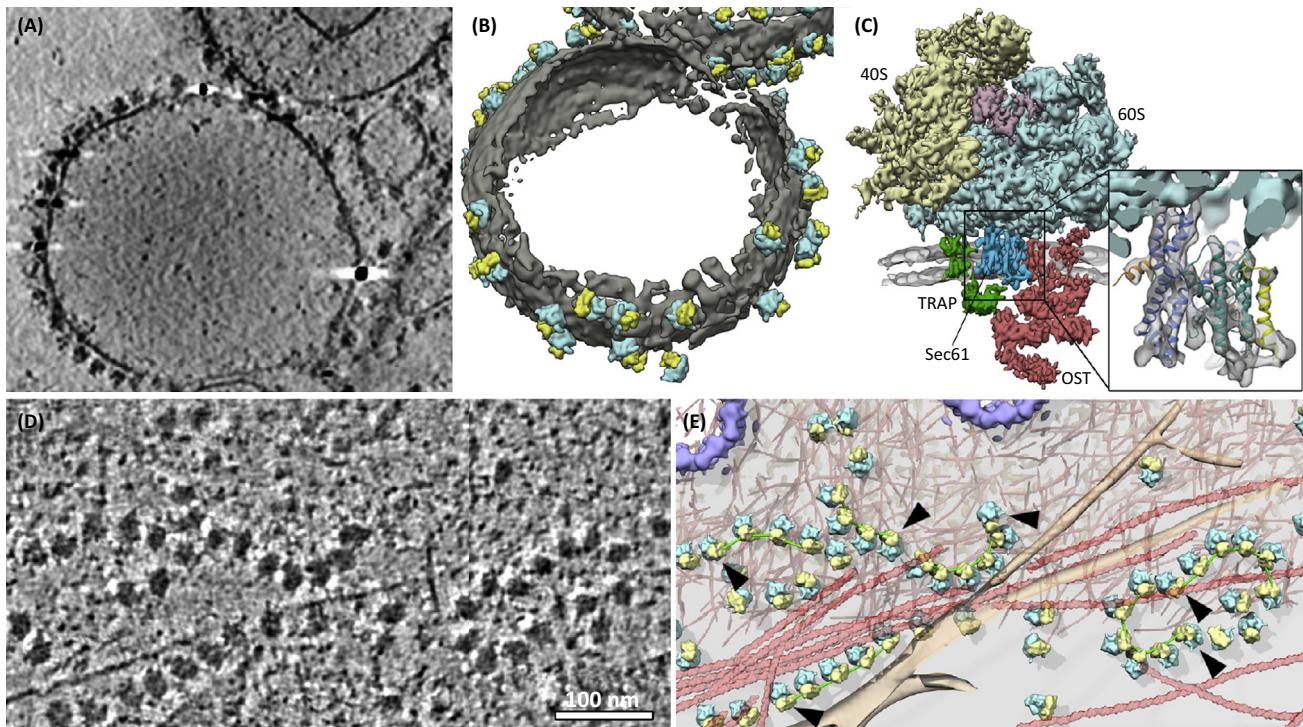
been previously proposed that Sec61 acts as a physical gate and can adopt either a closed or open conformation depending on the inserting state of the nascent chain [26,27]. This model was based on structural data obtained by single particle analysis (SPA) (Box 1) of isolated complexes, such as the ribosome–Sec61 complex purified in the presence of detergent [26,27]. A recent tomographic study investigated the structure of the ribosome–Sec61 complex in a close to native, membrane-embedded state, that is located within rough ER vesicles purified from dog pancreas (Figure 3A, B) [5]. Sec61 was found to adopt an open conformation even in the noninserting state (Figure 3C), suggesting that Sec61 assumes a closed conformation only in the absence of the ribosome [25]. These findings highlight the importance of structural analysis *in situ*. The ER-derived vesicles represent an *in situ sensu largo* scenario. Comparison to the ER membranes as observed inside intact cells (Figure 3D, E) reveals differences in membrane topology. While isolated vesicles relax into entropically favored spherical shapes, ER membranes display more elaborate topologies *in situ*. Due to their rigidity, size and cellular abundance, ribosomes are easily accessible to structural studies *in situ* by cryo-ET. It is therefore reasonable to assume that it will be possible to study the mechanism of translation and translocation in near atomic detail *in situ sensu stricto*.



Trends in Cell Biology

Figure 2. Structural Analysis of the NPC Requires *in situ* Approaches Preserving Nuclear Membrane Topology. (A) Pseudoatomic composite structure of the human NPC scaffold [17] obtained by the integration of *in vitro* and *in situ* structural biology. The tomographic map that provides the structural framework is shown as isosurface rendered (membranes in grey; proteinaceous regions transparent) and superimposed with high-resolution structures obtained by X-ray crystallography. (B) Sociology of NPCs (purple) within intact cells in the context of microtubules (green), actin filaments (dark red), ribosomes (yellow/green), the lamin meshwork (greyish red), and the nuclear interior (brown). (C) Comparison of *in situ* *sensu stricto* and *sensu largo* approaches. Slices through tomograms are shown on the left and slices through isosurface rendered, averaged structures are shown on the right. Although the SNR of the primary data is higher in the top

(Figure legend continued on the bottom of the next page.)



Trends in Cell Biology

Figure 3. Structural Analysis of the Translocon *in situ*. A tomographic slice (A) and isosurface rendering (B) showing isolated ER vesicles that assume a spherical topology in isolation (*in situ sensu largo*). The vesicles are decorated with ribosome–translocon complexes that were subjected to subtomogram averaging to obtain the structure shown in (C). (C) Cryo-EM structure of the translocon complex together with the translocon associated protein complex (TRAP, green) and the oligosaccharyltransferase (OST, red); large subunit in bright blue, small subunit in yellow (Inset: Sec61 assumes an open conformation). Tomographic slice (D) and isosurface rendering (E) of the outer nuclear membrane observed *in situ sensu strictu*. Ribosomes form clearly recognizable linear polysome chains on the low curvature membrane.

The Proteasome—a Giant Molecular Machine for Regulated Intracellular Protein Degradation

The ubiquitin–proteasome system (UPS) is the major route used by eukaryotic cells for the disposal of misfolded or damaged proteins and for controlling the lifespan of regulatory proteins. The 26S proteasome is a multisubunit complex of 2.5 MDa, which targets proteins marked for destruction by the covalent attachment of polyubiquitin chains. The holocomplex comprises two subcomplexes: the core complex also referred to as the 20S proteasome, and one or two regulatory complexes, the 19S caps, which bind to either one or both ends of the barrel-shaped core. The core complex executes degradation; its architecture allows the proteolytic action to be confined to a nanocompartment sequestered from the cellular environment it inhabits, the nucleus and the cytosol. The 19S regulatory particles recruit substrates marked for degradation and prepare them for translocation into the core particle. This includes their initial binding to ubiquitin receptors, their deubiquitylation and unfolding, and the opening of the gate, which controls access to the interior of the proteolytic core [28].

Although the structure of the 20S core could be determined in the 1990s by X-ray crystallography, the structure of the 26S holocomplex remained elusive [29]. The compositional and conformational heterogeneity of 26S preparations made its structural characterization

panel, the experimental throughput is limited and thus the resolution of the averages is lower as compared to the bottom panel. Broken lines in the right panel indicate the inner and outer diameter of the NPC inner ring, implying a massive conformational change.

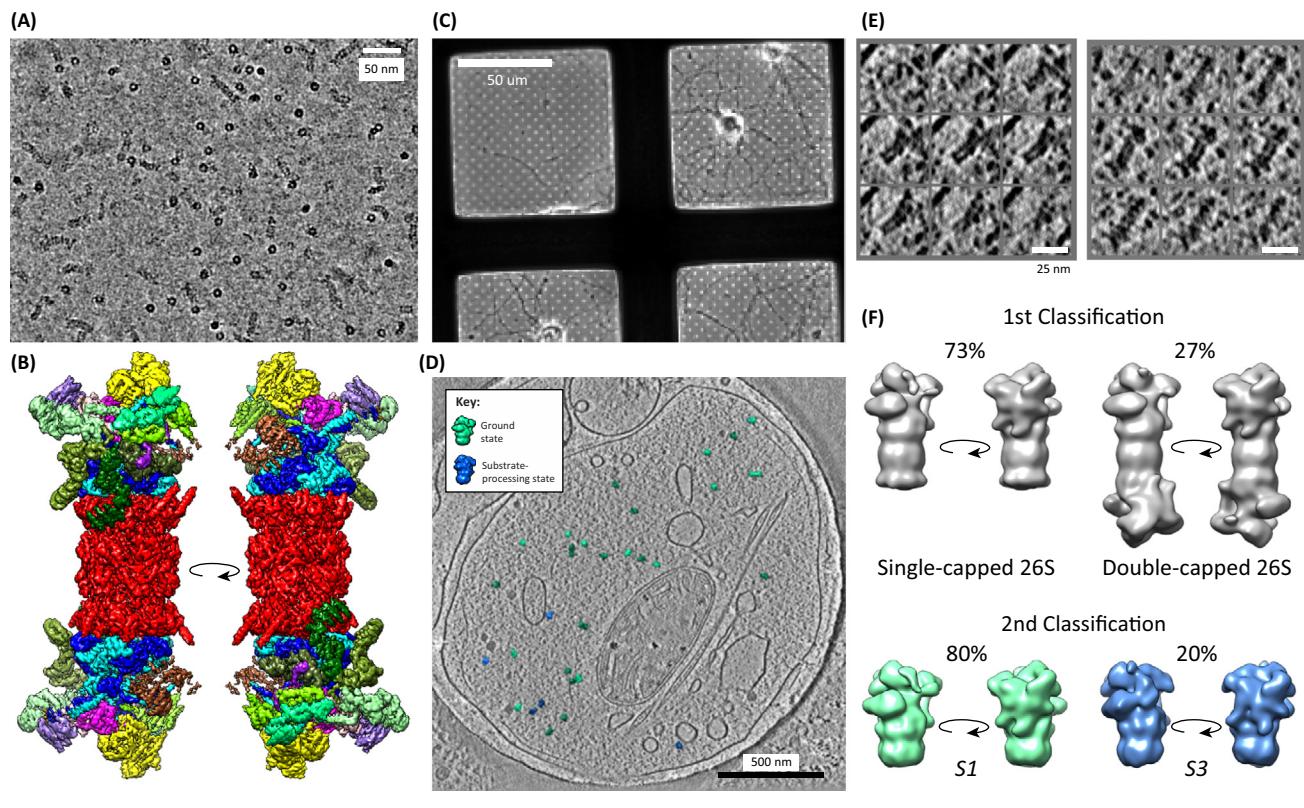


Figure 4. Ex situ and in situ Structural Studies of 26S Proteasome. (A) Cryo-EM image of isolated human 26S proteasomes. (B) 3D structure of human 26S proteasome at 3.9 Å resolution obtained by SPA. The subunits are color coded. Red: core particle; blue and light blue: AAA-ATPase heterohexamer; brown: Rpn1; yellow: Rpn2; green: proteasome–cyclosome–initiation factor (PCI) subunits (Rpn3, -5, -6, -7, -9, -12); magenta: Rpn8, -11; purple: Rpn10. (C) Primary cultured hippocampal neurons grown on gold EM grid. (D) Single-capped and double-capped 26S proteasomes displayed as green (ground state) and blue (substrate-processing state) isosurfaces overlaid on a slice from a cryo-electron tomogram of an intact neuron acquired with the Volta phase plate. (E) Consecutive slices of 26S individual proteasomes detected *in situ*. Left: single-capped 26S proteasome. Right: double-capped 26S proteasome. (f, top panel) first classification of the data set into single-capped and double-capped 26S proteasome. (f, bottom panel) second classification for 19S conformational states (S1 state and S3) [34]. (A) and (B) modified, with permission, from [33]. (D–F) modified, with permission, from [35].

challenging. In addition to the 33 canonical subunits, multiple proteins interact with 26S proteasome stoichiometrically such as deubiquitylating enzymes or shuttling ubiquitin receptors. During its functional cycle, the 19S regulatory particle undergoes large-scale conformational changes driven by the AAA-ATPase module, which forms the base of the regulatory particle. Substrates must be translocated from the initial binding sites to the ‘mouth’ of the ATPase for unfolding, and the resident deubiquitylase Rpn11 must be positioned such that it can cleave off polyubiquitin chains.

Only recently has the molecular architecture of the 26S holocomplex been elucidated using, cryo-ET, SPA (Box 1), and integrative approaches [30,31] (Box 3). SPA is more tolerant to structural heterogeneity than crystallographic methods; it allows coexisting states to be separated computationally and therefore became, in many cases, the method of choice for determining the structure of large macromolecular assemblies. With the yeast 26S proteasome, it was possible to generate a near-atomic model and to define its conformational space. More recently, in the wake of advances in technology (Box 2), the structure of the human 26S proteasome has been determined to a resolution of 3.9 Å (Figure 4A, B), providing deeper insights into the role of ATP binding and hydrolysis [32–34].

Structural studies *in situ sensu strictu* were performed initially with rat hippocampal neurons grown on EM grids [35] (Figure 4 C). The filiform nature of large parts of the neurons allows examining them *in toto* by cryo-ET, that is, without the need for thinning. Data recorded with the Volta phase plate (Box 2) render individual 26S proteasomes clearly visible in the tomograms and, given the size and distinct shape of the complex, **template matching** (Box 3) allows determining their positions in the cell with high fidelity and nanometer accuracy (Figure 4D, E). Subtomogram averaging and classification (Box 3) across the population of proteasomes has allowed one to distinguish between particles carrying one or two regulatory particles and to assess their conformational state (Figure 4F). Surprisingly, single-capped proteasomes are the predominant species. By correlation with the high-resolution SPA structures of the dominant conformations, one can infer which particle is in the ground state, that is, waiting for substrate, and which one is engaged in substrate processing. Such a **molecular census** provides insights into the capacity and usage of a key player of cellular proteostasis. Future studies will aim at a comprehensive mapping of proteasomes in all major cellular territories, including the nucleus, as well as improving resolution. By mapping other molecular species involved in intracellular protein degradation, such as tripeptidyl peptidase [36] or CDC48, their interaction networks can be revealed.

Concluding Remarks

The structures discussed in this review were not obtained by the traditional approaches to macromolecular structure determination. However they took advantage of *a priori* information about the components making up the supramolecular assemblies and this information was furnished by the established methods of structural biology. It is thus fair to state that traditional *ex situ* studies and *in situ* studies complement one another. Cryo-ET is the method of choice to create the structural framework for the integration of partial structures of supramolecular assemblies and allows putting them into a cell biological context (Box 3). The principal advantage of cryo-ET is that it allows coping with structures of high complexity that are pleomorphic in nature, such as cells or organelles. Its full potential still needs to be realized.

There is growing awareness that the concept of diffusion controlled, randomly colliding reactants does not hold for the interior of cells [37] but the current methodological toolbox limits further examination of this idea. For example, proteomic measurements based on MS (Box 3) are carried out on the combined lysates of multiple cells or organelles. Any spatial information—for example, about functional neighborhoods and associations in networks—is lost and cell-specific properties are averaged over the population of lysed cells. Electron tomograms of cells contain the structural signatures of all their components and reveal their spatial coordinates. However, to extract features of interest in a meaningful way is a nontrivial task because of the sheer wealth of information. It had been proposed that this could be achieved by matching the positions and orientations of the structural signatures of protein complexes into the tomograms—a concept referred to as **visual proteomics** [38] (Box 3). Early studies that have assessed the feasibility of the visual proteomics concept concluded that the accuracy of identifying complexes is critically dependent on the resolution and SNR and therefore only a rather limited set of large and abundant complexes could be targeted [39]. Recent technological advances (Boxes 2 and 3) have led to significant improvements of SNR, contrast, and resolution, and thus hold promise to render visual proteomics feasible for a larger set of macromolecular complexes than was anticipated a few years ago. Due to the abundance variation of molecular complexes exceeding seven orders of magnitude in mammalian cells [40,41] and three orders of magnitude in bacterial cells [42], only a certain fraction of the proteome will be accessible to visual proteomics (see Outstanding Questions). However, many of the major cellular functions involve large complexes. The generation of key metabolites, transcription, translation, and protein folding and degradation do rely on abundant functional modules [43], which can be studied *in situ*. Cryo-ET offers the possibility to investigate how their structures are modulated in the context of their functional

Outstanding Questions

The molecular machines discussed in this review conduct central cellular functions and are abundant in most cells. Other biological processes, such as DNA repair, are rare events and rely on less abundant molecules. It remains to be seen to what extent will the presently attainable throughput of *in situ* structural biology techniques will allow for the visualization of rare molecular species? Correlative light microscopy (LM)-EM might be helpful to guide the structural analysis to the volume of interest.

How can the method be more accessible? It is of importance that instrumentation facilitating the complex workflow required for performing *in situ* studies becomes commercially available making the method accessible to a wider community.

environment by studying the molecular sociology of cells [22]. However, even more simplistic outcomes, such as the determination of local concentrations of complexes, a ‘molecular census’ [35], can be expected to provide novel insights into metabolic and kinetic networks.

Will we ultimately be able to obtain subnanometer or even near atomic resolution maps of entire cellular landscapes—the holy grail of structural biology *in situ*? If this can be achieved, then the localization of macromolecules would become almost entirely unambiguous. Combined with the high-resolution structural repertoire contained in databases, pseudoatomic maps will be generated. It has been recently demonstrated that cryo-ET in conjunction with subsequent subtomogram averaging can, in principle, reach near atomic resolution [44], even with the present state-of-the art instrumentation and image analysis tools. Although an idealized *in vitro* sample with an essentially unlimited number of identical copies of the structure of interest was used in this study, it highlights the potential of the technology. It appears fair to conclude that, in the long term, the sociology of the cell, at least in part, will be revealed in atomic detail.

Acknowledgments

We thank Drs. Jürgen Plitzko and Friedrich Förster for critical reading of the manuscript. Drs. Jan Kosinski, Yoshiyuki Fukuda, Julia Mahamid, Jürgen Plitzko and Ben Engel are acknowledged for help with preparing the figures.

References

- Alberts, B. (1998) The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell* 92, 291–294
- Asano, S. *et al.* (2016) *In situ* cryo-electron tomography: a post-reductionist approach to structural biology. *J. Mol. Biol.* 428, 332–343
- Robinson, C.V. *et al.* (2007) The molecular sociology of the cell. *Nature* 450, 973–982
- Schur, F.K. *et al.* (2013) Determination of protein structure at 8.5 Å resolution using cryo-electron tomography and sub-tomogram averaging. *J. Struct. Biol.* 184, 394–400
- Pfeffer, S. *et al.* (2015) Structure of the native Sec61 protein-conducting channel. *Nat. Commun.* 6, 8403
- Letoumeur, F. *et al.* (1994) Coatomer is essential for retrieval of dityrosine-tagged proteins to the endoplasmic reticulum. *Cell* 79, 1199–1207
- Malhotra, V. *et al.* (1989) Purification of a novel class of coated vesicles mediating biosynthetic protein transport through the Golgi stack. *Cell* 58, 329–336
- Faini, M. *et al.* (2013) Vesicle coats: structure, function, and general principles of assembly. *Trends Cell Biol.* 23, 279–288
- Dodonova, S.O. *et al.* (2015) Vesicular transport. A structure of the COPI coat and the role of coat proteins in membrane vesicle assembly. *Science* 349, 195–198
- Devos, D. *et al.* (2004) Components of coated vesicles and nuclear pore complexes share a common molecular architecture. *PLoS Biol.* 2, e380
- Schwartz, T.U. (2005) Modularity within the architecture of the nuclear pore complex. *Curr. Opin. Struct. Biol.* 15, 221–226
- Hoelz, A. *et al.* (2011) The structure of the nuclear pore complex. *Annu. Rev. Biochem.* 80, 613–643
- Fischer, J. *et al.* (2015) Linker Nups connect the nuclear pore complex inner ring with the outer ring and transport channel. *Nat. Struct. Mol. Biol.* 22, 774–781
- Hurt, E. and Beck, M. (2015) Towards understanding nuclear pore complex architecture and dynamics in the age of integrative structural analysis. *Curr. Opin. Cell Biol.* 34, 31–38
- Bui, K.H. *et al.* (2013) Integrated structural analysis of the human nuclear pore complex scaffold. *Cell* 155, 1233–1243
- Eibauer, M. *et al.* (2015) Structure and gating of the nuclear pore complex. *Nat. Commun.* 6, 7532
- Kosinski, J. *et al.* (2016) Molecular architecture of the inner ring scaffold of the human nuclear pore complex. *Science* 352, 363–365
- von Appen, A. *et al.* (2015) *In situ* structural analysis of the human nuclear pore complex. *Nature* 526, 140–143
- Lin, D.H. *et al.* (2016) Architecture of the symmetric core of the nuclear pore. *Science* 352, <http://dx.doi.org/10.1126/science.aaf1015>
- Stuwe, T. *et al.* (2015) Nuclear pores. *Architecture of the nuclear pore complex coat*. *Science* 347, 1148–1152
- Maimon, T. *et al.* (2012) The human nuclear pore complex as revealed by cryo-electron tomography. *Structure* 20, 998–1006
- Mahamid, J. *et al.* (2016) Visualizing the molecular sociology at the HeLa cell nuclear periphery. *Science* 351, 969–972
- Kohler, A. and Hurt, E. (2010) Gene regulation by nucleoporins and links to cancer. *Mol. cell* 38, 6–15
- Raices, M. and D’Angelo, M.A. (2012) Nuclear pore complex composition: a new regulator of tissue-specific and developmental functions. *Nat. Rev. Mol. Cell Biol.* 13, 687–699
- Pfeffer, S. and Forster, F. (2016) Sec61: a static framework for membrane-protein insertion. *Channels* 10, 167–169
- Gogala, M. *et al.* (2014) Structures of the Sec61 complex engaged in nascent peptide translocation or membrane insertion. *Nature* 506, 107–110
- Voorhees, R.M. *et al.* (2014) Structure of the mammalian ribosome-Sec61 complex to 3.4 Å resolution. *Cell* 157, 1632–1643
- Baumeister, W. *et al.* (1998) The proteasome: paradigm of a self-compartmentalizing protease. *Cell* 92, 367–380
- Forster, F. *et al.* (2013) Unveiling the long-held secrets of the 26S proteasome. *Structure* 21, 1551–1562
- Lander, G.C. *et al.* (2012) Complete subunit architecture of the proteasome regulatory particle. *Nature* 482, 186–191
- Lasker, K. *et al.* (2012) Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. *Proc. Natl. Acad. Sci. U. S. A.* 109, 1380–1387
- Beck, F. *et al.* (2012) Near-atomic resolution structural model of the yeast 26S proteasome. *Proc. Natl. Acad. Sci. U. S. A.* 109, 14870–14875
- Schweitzer, A. *et al.* (2016) Structure of the human 26S proteasome at a resolution of 3.9 Å. *Proc. Natl. Acad. Sci. U. S. A.* 113, 7816–7821
- Unverdorben, P. *et al.* (2014) Deep classification of a large cryo-EM dataset defines the conformational landscape of the 26S proteasome. *Proc. Natl. Acad. Sci. U. S. A.* 111, 5544–5549
- Asano, S. *et al.* (2015) A molecular census of 26S proteasomes in intact neurons. *Science* 347, 439–442

36. Schonegele, A.M. *et al.* (2012) The structure of human tripeptidyl peptidase II as determined by a hybrid approach. *Structure* 20, 593–603
37. Hyman, A.A. *et al.* (2014) Liquid-liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* 30, 39–58
38. Nickell, S. *et al.* (2006) A visual approach to proteomics. *Nat. Rev. Mol. Cell Biol.* 7, 225–230
39. Forster, F. *et al.* (2010) Visual proteomics. *Methods Enzymol.* 483, 215–243
40. Beck, M. *et al.* (2011) The quantitative proteome of a human cell line. *Mol. Syst. Biol.* 7, 549
41. Nagaraj, N. *et al.* (2011) Deep proteome and transcriptome mapping of a human cancer cell line. *Mol. Syst. Biol.* 7, 548
42. Malmstrom, J. *et al.* (2009) Proteome-wide cellular protein concentrations of the human pathogen *Leptospira interrogans*. *Nature* 460, 762–765
43. Beck, M. *et al.* (2009) Visual proteomics of the human pathogen *Leptospira interrogans*. *Nat. Methods* 6, 817–823
44. Schur, F.K. *et al.* (2016) An atomic model of HIV-1 capsid-SP1 reveals structures regulating assembly and maturation. *Science* 353, 506–508
45. Stahlberg, H. *et al.* (2015) 3D reconstruction of two-dimensional crystals. *Arch. Biochem. Biophys.* 581, 68–77
46. Frank, J. (2002) Single-particle imaging of macromolecules by cryo-electron microscopy. *Annu. Rev. Biophysics Biomol. Struct.* 31, 303–319
47. Lucic, V. *et al.* (2005) Structural studies by electron tomography: from cells to molecules. *Annu. Rev. Biochem.* 74, 833–865
48. Fromm, S.A. and Sachse, C. (2016) Cryo-EM structure determination using segmented helical image reconstruction. *Methods Enzymol.* 579, 307–328
49. Mahamid, J. and Baumeister, W. (2012) Cryo-electron tomography: the realization of a vision. *Microsc. Anal.* 26, 45–48
50. Al-Amoudi, A. *et al.* (2004) Cryo-electron microscopy of vitreous sections. *EMBO J.* 23, 3583–3588
51. Rigort, A. *et al.* (2012) Focused ion beam micromachining of eukaryotic cells for cryo-electron tomography. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4449–4454
52. Villa, E. *et al.* (2013) Opening windows into the cell: focused-ion-beam milling for cryo-electron tomography. *Curr. Opin. Struct. Biol.* 23, 771–777
53. Plitzko, J.M. *et al.* (2009) Correlative cryo-light microscopy and cryo-electron tomography: from cellular territories to molecular landscapes. *Curr. Opin. Biotechnol.* 20, 83–89
54. Faruqi, A.R. and Henderson, R. (2007) Electronic detectors for electron microscopy. *Curr. Opin. Struct. Biol.* 17, 549–555
55. Li, X. *et al.* (2013) Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* 10, 584–590
56. Kuijper, M. *et al.* (2015) FEI's direct electron detector developments: Embarking on a revolution in cryo-TEM. *J. Struct. Biol.* 192, 179–187
57. Danev, R. *et al.* (2014) Volta potential phase plate for in-focus phase contrast transmission electron microscopy. *Proc. Natl. Acad. Sci. U. S. A.* 111, 15635–15640
58. Glaeser, R.M. (2016) Protein complexes in focus. *eLife* 5 e16156
59. Danev, R. and Baumeister, W. (2016) Cryo-EM single particle analysis with the Volta phase plate. *eLife* 5, e13046
60. Khoshouei, M. *et al.* (2016) Subtomogram analysis using the Volta phase plate. *J. Struct. Biol.* <http://dx.doi.org/10.1016/j.jsb.2016.05.009>
61. Fernandez, J.J. (2012) Computational methods for electron tomography. *Micron* 43, 1010–1030
62. Frangakis, A.S. and Forster, F. (2004) Computational exploration of structural information from cryo-electron tomograms. *Curr. Opin. Struct. Biol.* 14, 325–331
63. Bohm, J. *et al.* (2000) Toward detecting and identifying macromolecules in a cellular context: template matching applied to electron tomograms. *Proc. Natl. Acad. Sci. U. S. A.* 97, 14245–14250
64. Frangakis, A.S. *et al.* (2002) Identification of macromolecular complexes in cryo-electron tomograms of phantom cells. *Proc. Natl. Acad. Sci. U. S. A.* 99, 14153–14158
65. Forster, F. *et al.* (2005) Retrovirus envelope protein complex structure in situ studied by cryo-electron tomography. *Proc. Natl. Acad. Sci. U. S. A.* 102, 4729–4734
66. Wan, W. and Briggs, J.A.G. (2016) Cryo-electron tomography and subtomogram averaging. *Methods Enzymol.* 579, 329–367
67. Aebersold, R. and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* 422, 198–207
68. Leitner, A. *et al.* (2010) Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol. Cell Proteomics* 9, 1634–1649
69. Alber, F. *et al.* (2008) Integrating diverse data for structure determination of macromolecular assemblies. *Annu. Rev. Biochem.* 77, 443–477