Molecular Cell

CellPress

Technology review

Label-free visual proteomics: Coupling MS- and EM-based approaches in structural biology

Oleg Klykov,[1,2] Mykhailo Kopylov,[1] Bridget Carragher,[1,2] Albert J.R. Heck,[3,4] Alex J. Noble,[1,*] and Richard A. Scheltema[3,4,*]

[1]National Center for In-situ Tomographic Ultramicroscopy, Simons Electron Microscopy Center, New York Structural Biology Center, New York, NY, USA

[2]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY, USA

[3]Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, University of Utrecht, 3584 CH Utrecht, the Netherlands

[4]Netherlands Proteomics Center, 3584 CH Utrecht, the Netherlands

*Correspondence: anoble@nysbc.org (A.J.N.), r.a.scheltema@uu.nl (R.A.S.)

https://doi.org/10.1016/j.molcel.2021.12.027

SUMMARY

Combining diverse experimental structural and interactomic methods allows for the construction of comprehensible molecular encyclopedias of biological systems. Typically, this involves merging several independent approaches that provide complementary structural and functional information from multiple perspectives and at different resolution ranges. A particularly potent combination lies in coupling structural information from cryoelectron microscopy or tomography (cryo-EM or cryo-ET) with interactomic and structural information from mass spectrometry (MS)-based structural proteomics. Cryo-EM/ET allows for sub-nanometer visualization of biological specimens in purified and near-native states, while MS provides bioanalytical information for proteins and protein complexes without introducing additional labels. Here we highlight recent achievements in protein structure and interactome determination using cryo-EM/ET that benefit from additional MS analysis. We also give our perspective on how combining cryo-EM/ET and MS will continue bridging gaps between molecular and cellular studies by capturing and describing 3D snapshots of proteomes and interactomes.

INTRODUCTION

Evolution has produced crowded, complex environments inside of cells. Myriads of protein complexes, organelles, and membranes form interaction networks, or interactomes, whose aggregate behaviors constitute life. Visual proteomics is the study of cell interactomes, the three-dimensional (3D) organization of cell constituents, and, optionally, interactions between cells in tissues. The field of visual proteomics is just beginning to burgeon. Recently, millions of dollars in grant funding have been awarded to over a dozen labs to build hardware and software methods toward sub-nanometer visual proteomics of cells (refer to chanzuckerberg.com for details). Several methods for studying cellular life in molecular detail while retaining native or near-native contextual information have taken hold in recent years, including scanning electron microscopy (SEM), light microscopy, cryoelectron microscopy and tomography (cryo-EM/ET), and mass spectrometry (MS). In this review, we report on the confluence of direct 3D visualization of cellular interiors with cryo-EM/ET and information on protein complexes and their interactions obtained with MS. We discuss the combination of cryo-EM/ET and MS, representing the highest-resolution label-free approach currently possible—a combination we call label-free visual proteomics. Furthermore, we discuss prospective advances in label-free visual proteomics that may allow this interdisciplinary field to answer fundamental biological questions and to directly contribute to medicine.

Several additional methods in light microscopy and *in situ* labeling that should also advance the field of visual proteomics are being developed. Currently, commercial cryo-fluorescent light microscopes (cryo-FLMs) can localize fluorescent signals inside biological specimens with resolutions on the order of hundreds of nanometers. Several efforts toward cryo-super-resolution light microscopes (cryo-SRLMs) with precision on the order of 10 nm or better are in development (for review, see Bauerlein and Baumeister, 2021). Perhaps the most promising development in SRLM is MINFLUX, which may localize proteins with 1–3 nm precision (Gwosch et al., 2020) if devitrification issues with frozen specimens are first overcome. Multiple non-metal EM tags for labeling and identifying proteins of interest vitrified *in situ* have been proposed (Silvester et al., 2021; Tan et al., 2021), yet are not in mainstream use and may disrupt native protein conditions. In this review, we focus on label-free MS approaches for detecting protein quantities, their modifications, protein-protein interactions (PPIs), and protein complex components in conjunction with cryo-EM/ET.

MS is a well-established analytical method for analyzing ionized biomolecules, providing ion signal intensities as a function of the mass-to-charge ratio (Aebersold and Mann, 2003). This method identifies and relatively quantifies proteins together

with pinpointing their post-translational modifications (PTMs). MS can also retrieve interactomic and low/medium-resolution structural information (10–30 Å) from biological samples with complexities ranging from purified protein samples to cellular specimens. The second method we focus on is cryo-EM. Cryo-EM can routinely produce high-resolution structures (2.5–4.5 Å) of purified proteins and protein complexes by averaging together images of thousands to millions of copies of the same protein target. Cryo-ET can work with various *in situ* specimens including intact cells to produce individual, low-resolution 3D reconstructions, called tomograms (typically tens of angstroms). Like cryo-EM, medium/high-resolution structures (3.5–20 Å) may be obtained with cryo-ET by averaging thousands to hundreds of thousands of copies of the same protein target. Cryo-EM/ET and MS may be performed on samples from the same specimen and without labels, making them highly complementary. In the literature, their combination has led to striking novel insights into a variety of biological questions. Here we provide a brief overview of cryo-EM/ET, followed by a description of the current state of structural MS methods most relevant to their combination with EM. Next, we highlight the most recent accomplishments achieved by combining cryo-EM and MS, focusing on results that would have been unattainable using only one method. We provide an update on the previously reviewed complementarity of XL-MS- and EM-based methods (Schmidt and Urlaub, 2017) and highlight studies that employ other structural MS techniques for aiding cryo-EM interpretation. Additionally, we discuss novel approaches that incorporate MS-based techniques into cryo-EM structural characterization of complex samples. Next, we describe the benefits and limitations of performing cryo-ET and XL-MS on *in situ* cellular specimens. Then we discuss prospects of *in situ* cellular cryo-ET and XL-MS and explain why this combination is a particularly attractive method for integrative system-wide studies investigating snapshots of living cells. Last, we suggest several future directions and frontiers that may be uniquely attainable by combining MS- and EM-based methodologies, along with improvements that are required to make visual proteomics with cryo-ET more accessible and routine. We envision that implementation of these approaches will provide a comprehensive basis for coupling label-free EM and MS methods as a link between structural, cellular, and systems biology.

Overview of cryo-EM approaches

During the past several years, cryo-EM has emerged as an essential method for protein structure determination (Figure 1, top). Cryo-EM is reviewed in this issue by by Saibil, 2022, so here we only briefly introduce the core methods and concepts most relevant for this review. Currently, cryo-EM encompasses two major approaches: single-particle analysis (SPA) and cryoelectron tomography (cryo-ET), which are used for determining the structure of proteins, protein complexes, and morphologies of more complex biological objects such as cells and tissues. For both SPA and cryo-ET, the sample must be vitrified to preserve it in its native or near-native state and protect it from the vacuum of the electron microscope. SPA is primarily useful when the sample is a solution of a purified protein with minimal heterogeneity (for review, see Wu and Lander, 2020). The sample is applied to an EM grid, reduced to a thin film, and vitrified by plunging into a cryogen (most often liquid ethane), trapping individual protein complexes in a variety of orientations in a thin layer of vitrified ice (Figure 1, top). Typically, hundreds to thousands of high-magnification images of the vitrified sample are acquired using a cryo-transmission electron microscope. Individual 2D protein projection images are called ''particles,'' which typically number in the thousands to hundreds of thousands. Large quantities of data are required because the incident electron beam onto the sample quickly deteriorates the proteins' chemical bonds, so exposures must be acquired using very low doses of electrons, which results in images with very low signal-to-noise ratio (SNR). This issue can be overcome by averaging sufficient numbers of low-SNR particle images with sufficient coverage of protein projection angles. Particles are selected and computationally combined using the Fourier slice theorem (Sigworth, 2016) to produce a 3D Coulomb potential density map from each homogeneous protein subset. This process is routinely performed with well-established software packages like CryoSparc (Punjani et al., 2017) and Relion (Scheres, 2012, 2020). The resolution of the 3D map depends on several factors, most critically the structural rigidity and homogeneity of the sample itself. The most well-behaved sample, apoferritin, has been used to demonstrate that cryo-EM can produce structures with resolutions up to 1.2 Å (Nakane et al., 2020; Yip et al., 2020); more typical resolutions are in the range of 2.5–4.5 Å (Lawson et al., 2011).

Complex and unique environments, such as cell interiors, are typically not amenable to SPA cryo-EM due to much higher degrees of heterogeneity. Instead, cryo-ET is performed by collecting sequential cryo-EM images of the area of interest while tilting the specimen stage in the microscope (for review, see Turk and Baumeister, 2020). Due to electron beam damage of the sample, each 2D image projection is collected with a fraction of the electron dose typically used in SPA. These series of 2D image projections, or tilt-series, are then computationally aligned and combined to produce 3D reconstructions, called tomograms. After plunge freezing, tilt-series for cellular specimens thinner than ~500 nm may be collected as described above. However, to visualize the *in situ* structure of interiors of cells thicker than ~500 nm, more complicated approaches are required due to fundamental electron transparency limits in cryo-TEMs (for review, see Bauerlein and Baumeister, 2021). While cellular specimens up to ~5 μm thick can be successfully vitrified by plunge freezing (Mahamid et al., 2016), for thicker specimens a high-pressure freezer (HPF) is required. HPF freezing is done under near-instantaneous high pressure, which provides vitrification throughout the whole volume of specimens up to 100–200 μm (Harapin et al., 2015). Subsequently, the vitrified specimen is transferred to a cryo-focused ion beam scanning electron microscope (cryo-FIB/SEM) where large portions of the material are ablated (FIB-milled) *in situ* to create thin volumes, called lamellae (typically 100–300 nm thick). Lamellae prepared from plunge-frozen cells are typically several microns in each cross-sectional direction, whereas lamellae prepared from high-pressure-frozen grids using either the cryo-liftout (Schaffer et al., 2019) or waffle method (Kelley et al., 2020) may be tens of microns in each cross-sectional direction. The resulting lamellae are

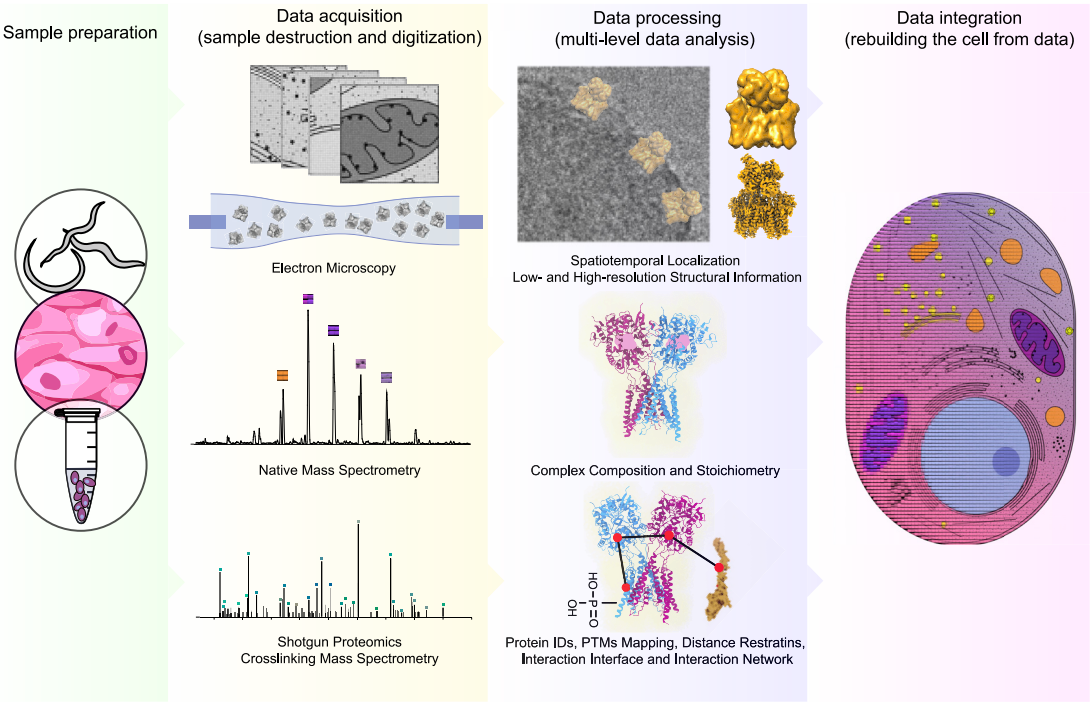


Figure 1. Schematic overview of EM and MS approaches described in this review, and illustrated by two domains of AMPA protein in complex with auxiliary subunit γ5

Shown is PDB: 7RYY. The initial sample is an intact organism, tissue, or cells. SPA cryo-EM and nMS require purification of the protein or protein complex of interest, while *ex situ* XL-MS and shotgun proteomics work with both purified proteins or protein complexes and highly heterogeneous samples like whole-cell lysates. *In situ* XL-MS and cryo-ET can be applied to intact samples; however, for samples thicker than ~500 nm, thinning with FIB-milling must be performed prior to cryo-ET imaging. Each method provides complementary information about the biological system of interest, which is combined later to answer a variety of biological questions.

cross-sections of the biological system in both space and time due to the specimen being frozen. The FIB-milled sample is then transferred to the cryo-TEM, where tilt-series may be collected and reconstructed into tomograms, allowing spatial visualization and interpretation of biological systems of interest *in situ*. Individual objects of interest, including proteins, protein complexes, and cellular membranes, are then identified and segmented out of the 3D tomograms to improve the analyses. Alignment and classification of repeating objects, called sub-tomogram processing, is used to achieve higher-resolution structures, which may subsequently be mapped back into the 3D tomograms for higher-fidelity visualization. After every step in these cryo-ET sample preparation and collection workflows, cryo-FLM may be performed to help localize objects of interest or confirm the collection from the areas with objects of interest. The sample preparation (for review, see Turk and Baumeister, 2020) and data analysis workflows (for review, see Leigh et al., 2019; Pyle and Zanetti, 2021) are continually under development. As we describe herein, multiple MS approaches can be used concurrently to critically supplement and help with interpretation in cryo-ET analysis of cellular specimens.

Current state of MS-based structural proteomics

Structural proteomics has been used as an umbrella term describing a variety of approaches, including those that use structural mass spectrometric analyses of protein samples. Among them, several can provide structural information on protein- and system-wide levels. In this review we focus on shotgun proteomics, native mass spectrometry, and crosslinking mass spectrometry, as these techniques are most used for structural studies in combination with cryo-EM (Britt et al., 2021). Several other approaches, including thermal proteome profiling (TPP, also referred to as cellular thermal shift assay MS [CETSA-MS]), top-down MS (TD-MS or denaturing MS), hydrogen-deuterium exchange MS (HDX-MS), covalent labeling, complexome profiling, and limited proteolysis MS (LiP-MS), are not discussed here, although we recognize their potential in integrative structural biology; for reviews of structural MS-based approaches, see Lössl et al., 2016; Britt et al., 2021; Wittig and Malacarne, 2021.

The first approach we will discuss is shotgun proteomics (Figure 1, bottom), a well-established technique that examines complex protein mixtures by digesting the proteins into peptides and analyzing the peptide mixture by a combination of high-performance liquid chromatography (HPLC) and MS. This technique is used for both qualitative (protein identification) and quantitative (protein abundance) characterization of proteins, which can be extended to include mappings of PTMs. Shotgun proteomics can reliably identify proteins as well as provide their relative and absolute quantities (Pan et al., 2009) and is now extensively optimized and automated across laboratories and instruments. As an input, it requires several millions of cells or

several hundred nanograms of purified protein material per run. With the current state-of-the-art setups, ~80% of proteins expressed in a single cell type can be characterized within 2 days (Nagaraj et al., 2012). In general, sample preparation starts with cell lysis, followed by protein denaturation through addition of chaotropic agents. Next, reduction and alkylation reagents are applied to break disulfide bridges, followed by the addition of proteases to digest the proteins into peptides. The resulting peptide mixture is then desalted and injected into a mass spectrometer coupled to an online HPLC peptide separation platform, which is used to separate potentially hundreds of thousands of peptides over a liquid chromatography (LC) separation gradient. The LC gradient set-up time typically ranges from 15 min for protein complexes to several hours for cell lysates. Subsequently, peptides are ionized through electrospray ionization (ESI), then the resulting peptide ions are guided into the mass spectrometer. Data acquisition uses so-called tandem MS (MS2, MS/MS) strategies, in which each cycle consists of a survey scan to record the precursor masses of the peptides eluting at that point followed by mass selection and fragmentation of precursors of interest. For the analysis of specific PTMs, the sample preparation workflow can be complemented with an enrichment or sample-fractionation step as well as alternative data-acquisition strategies or fragmentation methods. The acquired data are analyzed with various fully automated software platforms (for review, see Chen et al., 2020). Another elegant technique based on shotgun proteomics is affinity purification mass spectrometry (AP-MS), where typically an organism is genetically modified to express one protein (the bait protein) with an enrichment handle (Gingras et al., 2007). If the interaction is sufficiently stable, interacting partners (preys) will be co-purified with the bait. As the enrichment strategy can give rise to background proteins, a control is employed to distinguish the background from the true interactors (Mellacheruvu et al., 2013). High-speed AP-MS has been developed where over 100 samples per day can be measured (Hosp et al., 2015), allowing this method to serve as an extensive source of protein interaction information.

A second approach is native MS (nMS, Figure 1, middle), an MS-based technique that directly works on intact proteins and protein complexes by circumventing the denaturing step, thereby providing a bridge between interactomics and structural biology (for review, see Bolla et al., 2019; Vimer et al., 2020; Tamara et al., 2021). In nMS, proteins or protein complexes are first buffer-exchanged into a MS-compatible buffer that minimally disturbs their native structural organization, including inter- and intra-molecular connections. Sample in the MS-compatible buffer is then injected directly into the mass spectrometer to measure the mass of the intact protein complex. For a single nMS measurement, a few microliters of micromolar sample concentration, which are typically made in-house, are injected through the static needle. From the detected intact mass and when the masses of complex components are known, conclusions can be made about the stoichiometry of each subunit in the complex. In addition, the ionized assemblies can be mass-selected and subjected to fragmentation conditions to investigate the relative strength of the binding of the different subunits as well as to obtain masses of each individual component. Furthermore, this approach can also be coupled to ion mobility to approximate the cross-section and thus size and shape of the assembly. There is no theoretical restriction on how big the assembly of interest could be, with reports as high as tens of megadaltons (Miller et al., 2021). However, increased sample heterogeneity, as for instance introduced by extensive protein glycosylation, interferes with unambiguous assignments of stoichiometries. Despite that, nMS provides an ideal method to test potential heterogeneity in samples as it can easily identify and quantify different co-occurring protein complex states (Snijder et al., 2017).

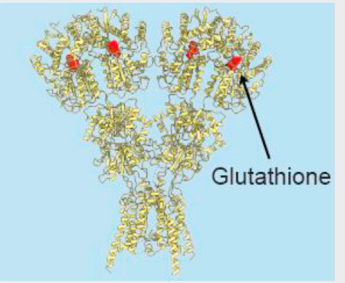
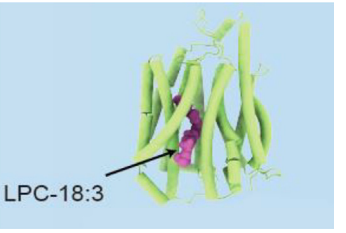
A third approach that can bridge the gap between EM and MS is crosslinking MS (XL-MS) (Figure 1, bottom). XL-MS has gained substantial traction in the past decade and has been reviewed extensively by Leitner et al., 2014; Yu and Huang, 2017; O'Reilly and Rappsilber, 2018. XL-MS supplies distance and orientation constraints for specific residues at low/medium-resolution (10–30 Å) (Rappsilber, 2011). The workflow is similar to the standard shotgun proteomics pipeline but typically requires at least 10 μg of protein complex or several millions of cells. Another difference from shotgun proteomics is the chemical crosslinking reactions and fractionation or enrichment steps that increase the number of identified crosslinked peptides. In the crosslinking reaction, small chemical reagents are introduced that contain a spacer arm with an amine reactive group on both ends (Belsom and Rappsilber, 2021). During this step, the crosslinking reagent covalently links amino acids in proximity both within and between proteins. Next, the standard shotgun proteomics steps and additional fractionation or enrichment are performed, resulting in covalently linked peptide pair products that can be identified by MS. The final output of an XL-MS experiment are distance constraints within and between proteins defining, for example, PPI interfaces (Schmidt and Urlaub, 2017). XL-MS data acquisition strategies are typically similar to those used in the shotgun proteomics workflows but can vary depending on the crosslinking reagent used (Stieger et al., 2019).

INTEGRATIVE STRUCTURAL BIOLOGY: COMBINING OUTPUT DATA FROM MS AND CRYO-EM

Shotgun MS-based proteomics and cryo-EM

The ability for shotgun proteomics to identify all proteins in a sample, complete with disparate PTMs, is useful for cryo-EM data interpretation (Table 1, top). At typical resolutions in cryo-EM, densities in protein regions with small, heterogeneous, or flexible PTMs are often not resolved well enough to comprehensively characterize the PTMs' compositions. At the same time, these unresolved regions do suggest the presence of PTMs, which can be aligned to those extracted by shotgun proteomics even though MS methods typically analyze the mixture of all proteoforms present in the sample. For example, in a study by Green et al., 2021, the cryo-EM density map of the plant ion channel GLR3.4 contained a bulky density that was hypothesized to be glutathionylation (GSH). As this PTM is rare, shotgun proteomics was used to conclusively show its presence. PTMs can also have a direct effect on protein complex assemblies, e.g., when monomers may assemble into oligomers. Comprehensive mapping of PTMs on cryo-EM structures of brain-derived tau filaments highlighted distinct PTM profiles depending on tauopathy

Table 1. Combining EM and MS output data in integrative structural biology

Approach	Key features	Sample amounts	Time requirements	Examples	
EM + shotgun proteomics	●*ex situ* ●peptide-level information ●requires genetic modifications (AP-MS) ●high throughput	●~100 ng for purified proteins and protein complexes ●from several cells in advanced single-cell proteomics pipelines up to millions of cells for standard pipelines	●sample preparation: overnight digestion, several hours for fractionation ●MS data acquisition: 3 h runs for lysates, 2–3 h per fraction for fractionated lysates, 60–90 min for purified protein or protein complexes	●PTM mapping (Green et al., 2021; Arakhamia et al., 2020) ●PTM sequencing (Hurdiss et al., 2020) ●protein identification from heterogeneous mixtures (Verbeke et al., 2018)	 based on PDB: 7LZH (Green et al., 2021)
EM + native MS	●*ex situ* ●protein identification is limited by sample heterogeneity ●requires MS-compatible buffer ●low throughput	●several microliters of micromolar protein solutions per run (2–3 microliters of ~5 micromolar solution)	●sample preparation: 3–5 h for buffer exchange (multiple samples in parallel) ●MS data acquisition: up to 10–20 min per sample	●protein complex composition and stoichiometry (Snijder et al., 2017; Ahdash et al., 2017) ●small molecule binding (Cater et al., 2021) ●cryo-EM sample screening (Olinares et al., 2021)	 based on PDB: 7MJS (Cater et al., 2021)
EM + XL-MS	●*ex situ*, *in situ* ●structural information at low and intermediate resolution ●peptide-level information ●high throughput	●*ex situ*: concentrations of 1 mg/mL are optimal but can be as low as 0.2 mg/mL; minimum 10 μg of protein or protein complex, but 100 μg is optimal ●*in situ*: varies based on cell type; tens of millions of mammalian cells is generally enough	same as in shotgun proteomics but with 30–60 min for crosslinking reaction and mandatory fractionation	●protein complex stabilization (Kenner et al., 2019; Huang et al., 2020) ●model validation (Yan et al., 2019; Dijkman et al., 2021) ●model building (Shakeel et al., 2019; Albanese et al., 2020) ●validation of interaction interface (Fagerlund et al., 2017; Meyer et al., 2019) ●structural details for SMC-kleisin complexes (Bürmann et al., 2019; Taschner et al., 2021; Yu et al., 2021b)	 based on PDB: 6QLE (Yan et al., 2019)

Overview of the key features of each described MS-based technique, sample amounts, time requirements, and highlighted reports of combinations of EM and MS output data. Example panels are based on PDB: 7LZH (Green et al., 2021), PDB: 7MJS (Cater et al., 2021), and PDB: 6QLE (Yan et al., 2019).

CellPress

(Arakhamia et al., 2020). In this study, a combination of cryo-EM and shotgun proteomics pinpointed the role of ubiquitination as the mediator of interaction interfaces between protofilaments, where different ubiquitination positions led to a wide diversity of tau filament structures. Glycosylation is another PTM that is readily visible in cryo-EM densities. In a study by Hurdiss et al., 2020, the cryo-EM structure of coronavirus-HKU1 hemagglutinin esterase was extensively complemented by shotgun proteomics data focused on the analysis of glycosylated peptides. Taken together, glycoproteomics and cryo-EM resolved the details of glycans covering the surface of the virus.

While MS-based proteomics is unable to directly provide protein structural information, its quantitative capabilities are useful in investigating protein complexes from whole-cell lysates. To achieve this, the complexity must first be reduced with near-native pre-fractionation at the protein level by, for example, size-exclusion chromatography (SEC) (Foster et al., 2006). Proteins associated as a protein complex will elute together, which is also clearly visible in the detected abundance traces over the fractions. This approach was extended by Verbeke et al., 2018 by including negative-stain EM analysis of each SEC fraction, enabling the structural characterization of the identified complexes. Termed shotgun-EM, this integrative method allows for more heterogeneous samples than in standard SPA to be performed on purified protein targets. At the same time, shotgun-EM has an advantage of preserving near-native protein interaction landscapes. Based on the MS output data, hundreds of protein complexes from fractionated HEK293 cell lysate were cross-correlated with the top three EM densities derived from the negative-stain EM data. While two densities were identified as two different versions of proteasomes, the third one was left unassigned, highlighting the need for more complex experimental designs including supplementary sources of structural information.

Based on the protein abundance derived from shotgun proteomics data, the stoichiometry of the individual proteins in a complex of known partners can be estimated (Hosp et al., 2015). For discovering the new protein complexes on a system-wide level at near-native conditions (Krogan et al., 2006), a modified version of shotgun proteomics, termed AP-MS, is applied. One of the advantages of AP-MS is the integration of a genetically incorporated enrichment handle on the protein of interest, which simultaneously can be used for the purification of the corresponding protein complex for SPA cryo-EM. AP-MS in its current state is a powerful tool that pinpoints thousands of PPIs (Huttlin et al., 2021; Michaelis et al., 2021). The potential combination of AP-MS and EM has recently been reviewed in more detail by Chowdhury et al., 2020.

nMS and cryo-EM

nMS has progressed rapidly over the past few years (Britt et al., 2021; Tamara et al., 2021). Most importantly, the method has become accessible to a wider community due to technological advances, software developments, and optimization of sample preparation and separation approaches. From the perspective of combining nMS output data with cryo-EM (Table 1, middle), nMS can be used as an initial assessment of the targets of interest, which is particularly useful for heterogeneous systems. Snijder et al., 2017 have demonstrated the power of nMS combined with SPA cryo-EM by applying them to a highly dynamic circadian oscillator system in cyanobacteria composed of the proteins KaiC, KaiB, and KaiA. nMS determined complex formation details and established that structural heterogeneity could be reduced by assembling the complex *in vitro* at lower temperatures. These optimized conditions were used in cryo-EM that completed the characterization and provided the structure of the fully assembled KaiCBA complex at 4.7 Å resolution. In the case where structures of complex components are available, nMS can be used to extend the structure of existing oligomeric assemblies. Ahdash et al., 2017 applied nMS to demonstrate that HerA is primarily represented by a heptamer, and after validation with 2D class averages generated by cryo-EM, the model of the hexamer was extended to a heptamer by means of molecular docking.

As high-resolution nMS is capable of measuring mass differences on the order of a single amino acid, it is commonly applied to analyze the binding of small molecules including drugs and lipids (Bennett et al., 2021; Keener et al., 2021). In the case of membrane protein transporters, both the lipid environment of the membrane regions and the lipid cargo itself can be defined with high precision and confidence (Cater et al., 2021). We anticipate that as the resolution of SPA cryo-EM increases for membrane proteins (already typically below 3.5 Å), more structures will allow lipid densities in membrane regions to be resolved. For membrane proteins, nMS indicates precisely how many lipid molecules and which lipids are directly involved in membrane protein functionality (Bolla et al., 2020; Gault et al., 2020). Combined with cryo-EM, nMS is poised to become an attractive approach for comprehensive characterization of membrane protein-bound lipid environments (Chorev et al., 2018).

In SPA cryo-EM, sample integrity and overall quality depend in part on the buffer conditions used to prepare the sample just prior to freezing it on the EM grid. As opposed to negative-stain EM sample screening, nMS does not introduce artifacts and allows for the identification of appropriate buffer conditions at which protein complexes are less heterogeneous, more stable, and represent the desired complex composition. While nMS requires a dedicated, experienced operator, it can be done within 5–6 h for several samples and may be applied to screen sample buffer conditions prior to vitrification instead of preparing and screening multiple cryo-EM grids, which, with data analysis, takes about 1 day per sample. Olinares et al., 2021 have demonstrated this nMS screening approach on several examples, including capturing the correct complex state for 3 out of 4 distinct protein complex targets. In the fourth case, nMS showed a single intact protein complex peak, but after classification of particles picked from EM micrographs, only 20% of these particles represented the same complex composition detected with nMS. Most particles found by cryo-EM represented the more complete version of the protein complex, indicating its partial dissociation during nMS analysis. It is important to emphasize that despite the relative ionic strength and pH being maintained in the MS-compatible buffer, in some cases protein stability and solubility issues may arise due to the buffer-exchange step as well as due to consequences of the electrospray process and the transition from solution into gas phase. Nonetheless, nMS screening of cryo-EM samples has several advantages, such

as relative simplicity, accessibility, and low sample consumption (on the order of micrograms). Also, the recently introduced setup for online buffer exchange for nMS analysis (VanAernum et al., 2020) allows for the buffer-exchange step into MS-compatible buffers to occur within the instrument, instead of buffer exchanging for hours. With multiple MS-compatible buffers introduced, this modification significantly increases the screening speed and sample throughput, resulting in assessment of the sample quality within minutes.

XL-MS and cryo-EM

XL-MS witnessed a strong uptake in use, as this technique most naturally aligns with cryo-EM approaches (Table 1, bottom). XL-MS is capable of extracting distance constraints from highly flexible regions, making it highly complementary to cryo-EM. In addition, XL-MS is compatible with the same sample conditions, with the exception of Tris-containing buffers when N-hydroxysuccinimide (NHS) chemistry is used, for which the buffer usually can be exchanged (e.g., to HEPES). XL-MS also requires low sample amounts, in the order of 100 μg of protein-purified complex per experiment. With advances in sample preparation techniques, MS instrumentation, and data analysis platforms, XL-MS has progressed from a niche application to a method of choice since 2014 (for review, see Schmidt and Urlaub, 2017). XL-MS has been applied to study PPIs for *ex situ* samples including purified proteins and protein complexes, cell lysates, and extracellular vesicles as well as *in situ* specimens including intact isolated organelles, prokaryotic cells, biopolymers, and tissues. In general, XL-MS studies are conducted separately from the cryo-EM pipeline, then output data from each approach are combined. Highlighting their compatibility even further, crosslinking reagents have also seen direct use in cryo-EM to prevent the disruption of unstable or transient protein complexes during vitrification, resulting in an effect similar to the gradient fixation approach (GraFix, Stark, 2010). Among many examples, when only the crosslinking reaction is performed without further identification of the crosslinked peptides, this approach helped, e.g., to obtain an EM density of the multi-subunit eIF2B:eIF2 and eIF2B:eIF2α systems (Kenner et al., 2019). It is, however, beneficial to perform the full XL-MS pipeline as the provided distance constraints can be used to guide and validate structures generated from cryo-EM densities of the stabilized protein complex (Huang et al., 2020).

Validation of structures derived from cryo-EM data is the most straightforward integration of XL-MS. Both pipelines can run in parallel, and after building the model based on the cryo-EM data, XL-MS restraints can be mapped onto the model to validate its quality (Yan et al., 2019; Dijkman et al., 2021). For regions that are not resolved at high resolution due to physiologically relevant conformational dynamics, XL-MS characterization can additionally guide model building and highlight domain movements, including, for example, highly flexible N-terminal tails (Shakeel et al., 2019; Albanese et al., 2020). Another integration method is to employ the distance restraints from XL-MS to guide placement of components of a protein complex within a negative stain or low/medium-resolution cryo-EM density (8.1 Å) and to confirm protein interaction sites (Fagerlund et al., 2017; Lee et al., 2020). This is especially useful for cryo-ET, which commonly produces low/medium-resolution structures (Meyer et al., 2019). This complementarity of low/medium-resolution EM/ET maps with XL-MS distance restraints for modeling protein complexes has also been demonstrated for low-resolution cryo-EM by Bullock et al., 2018. With a sophisticated modeling approach applied to a set of known protein complexes, the authors show that complex structures can be modeled more precisely when both EM and XL-MS information are used together with their scoring function as opposed to structures modeled with each method separately.

An area where the combination of XL-MS and cryo-EM particularly shines is for protein structures of elongated complexes with dynamic coiled-coil regions. Bürmann et al., 2019 first reported this approach for structural maintenance of chromosomes (SMC)-kleisin complexes from yeast. In this study, negative-stain EM and 2D classes obtained with cryo-EM were used to access the complex morphology. Additional analysis of MukBEF and cohesin with XL-MS provided information about flexible regions for each of the components based on the positions of the crosslinked residues. This allowed for reliable localization of the head, arm, midpoints, and hinge regions of the protein complex. Both Taschner et al., 2021 and Yu et al., 2021b utilized a similar approach for the SMC5/6 complex from yeast in the presence of the "non-SMC elements" subunits.

Structural proteomics and cryo-EM in integrative modeling

In structural biology, integrative modeling refers to the combination of multiple sources of input data for molecular docking that result in a comprehensible 3D model of the structure of interest. An illustrative, tour-de-force example was the deciphering of the nuclear pore complex (NPC) structure using MS-based proteomics, XL-MS, nMS, cryo-ET, and the integrative modeling platform (IMP) (Kim et al., 2018). The NPC is a large macromolecular machine that mediates transport between the nucleus and cytoplasm with a mass of ~52 MDa when isolated or ~87 MDa including the membrane, cargo, and nuclear transport factors (NTFs). To assess the complex stoichiometry and composition, nMS and quantitative shotgun proteomics were used alongside phospholipid analyses and calibrated imaging analysis of GFP-tagged nucleoporin family proteins. XL-MS on the purified crosslinked complex resulted in more than 3,000 unique distance restraints derived from two parallel experiments with different crosslinking reagents. Cryo-ET provided low-resolution EM maps (~28 Å) that depicted the overall complex morphology. Taken together, the data produced by these approaches were combined with previously published information and used in a sophisticated modeling pipeline based on IMP (Russel et al., 2012), which resulted in the intact structure of the yeast NPC with sub-nanometer precision for interaction interfaces. Recently, the yeast NPC structure has been solved at much higher resolutions (7–11 Å), with a root mean square deviation between scaffold proteins of both structures of ~8 Å for main Cα chains (Akey et al., 2021).

Integrative pipelines are not limited to purified protein complexes. For example, the *C. thermophilum* interactome has been extensively studied by a combination of negative-stain EM, cryo-EM, MS-based proteomics, XL-MS, and

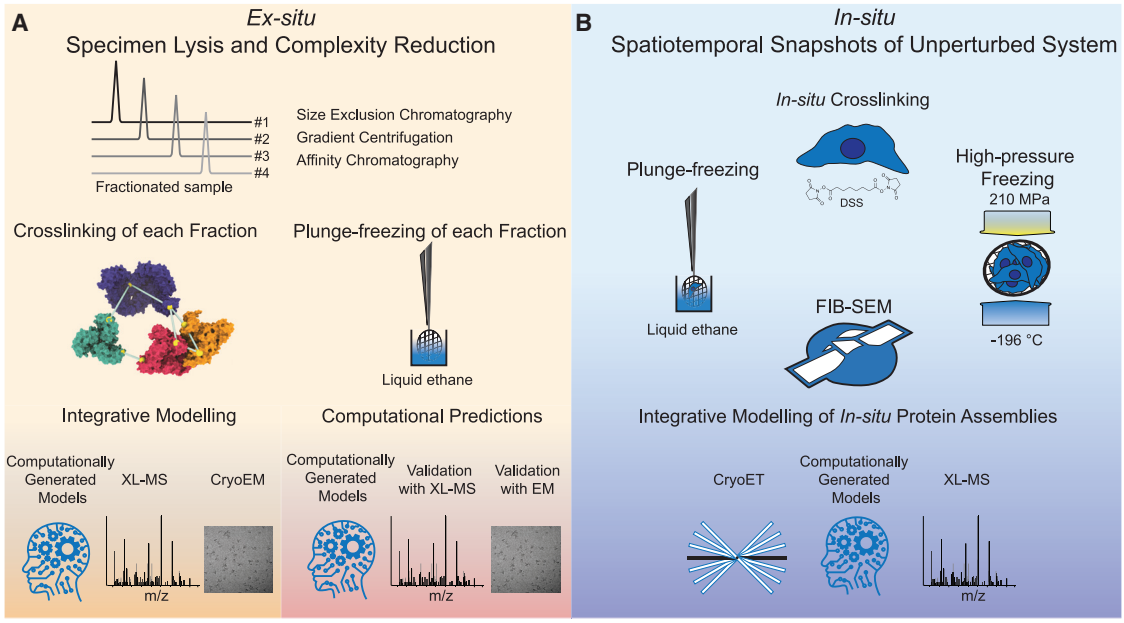


Figure 2. Data workflows for generating molecular encyclopedias

(A) *Ex situ* data generation workflow for samples prepared from whole-cell lysates. First, samples are lysed and fractionated to reduce the sample complexity. Next, each fraction is subjected to XL-MS or shotgun-EM in parallel. From the generated data, structural models of protein complexes are generated in two ways. In the case of integrative structural modeling, structural models of individual proteins are generated computationally, and then models of protein complexes are assembled with input from XL-MS and shotgun-EM. Alternatively, structural models of protein complexes are generated computationally, and then the best models are chosen based on XL-MS and shotgun-EM data.

(B) *In situ* data generation workflow for creating molecular encyclopedias. Samples are first crosslinked with a membrane-permeable reagent and processed as in the proteome-wide XL-MS workflow. For acquiring *in situ* cryo-ET data, samples are vitrified in their intact state either by plunge-freezing or HPF. In the case of thin-edge plunge-frozen cell experiments, cryo-ET data are collected from the thin cellular areas where FIB-milling is not required. For all other cryo-ET experiments, including HPF samples and thick cellular regions, an additional FIB-milling step is performed to obtain thin lamellae.

restraints-driven docking (Kastritis et al., 2017). The whole organism was lysed, fractionated by SEC, and then subjected to both MS and EM pipelines in parallel (similar to the shotgun-EM approach discussed above). To increase the amount of structural information, XL-MS was added to the standard proteomics pipeline and the cryo-EM data supplemented with the negative-stain EM. Combined with computational predictions, this analysis characterized the abundance and interaction interfaces of 108 protein complexes, comprising one-third of the whole *C. thermophilum* proteome. Data derived from EM and XL-MS were used to create and validate a computationally predicted interaction network. At the same time, both pipelines provided biological insights on their own: shotgun-cryo-EM allowed for structural determination of otherwise elusive intermediate proteins in fatty acid synthesis, while the constraints derived from XL-MS were used to predict interaction interfaces by means of restraints-guided protein-protein docking with HADDOCK (van Zundert et al., 2016). This study serves as an example of how modified shotgun-EM combined with XL-MS can significantly increase the amount and depth of information one can obtain for cell systems of interest.

We anticipate that based on the pipeline suggested by Kastritis et al., 2017, it is possible to build molecular encyclopedias for cells and biological systems consisting of proteome-wide interaction interface models created by integrative structural modeling (Figure 2). Also, as the system of interest is lysed prior to analysis, many diverse samples can be studied with the proposed approach described below (Figure 2A). First, the samples are lysed, then fractionated, and the structural information from each fraction is obtained by XL-MS and modified shotgun-cryo-EM. Next, the initial structures of individual proteins are generated computationally with, for example, AlphaFold 2 (Jumper et al., 2021) or RoseTTAfold (Baek et al., 2021), which have recently emerged as promising methods for structural model prediction. Computationally predicted structures still must be validated biologically, biochemically, and biophysically, especially in cases when the protein target has no close homology models, physiologically relevant conformational states, or intrinsically disordered regions. Nonetheless, both methods have already been applied to the proteome-wide structural interactome investigations of yeast (Humphreys et al., 2021) and human proteomes in combination with XL-MS data. (Burke et al., 2021). Then, generated models of individual proteins can be combined with the system-wide XL-MS and EM data and subjected to integrative structural modeling with IMP (Russel et al., 2012) or Assembline (Rantos et al., 2021). With the introduction of AlphaFold-Multimer (Evans et al., 2021), structures of protein complexes can also be predicted. In this case, low/medium-resolution densities derived from shotgun-EM and restraints information from XL-MS can be used for several purposes: (1) validation of protein complex models, (2) differentiation between possible structures from a variety of the modeling predictions,

and (3) adjusting the positions of the components within the obtained protein complex structures. Similarly, if the approximate interaction interface regions are known, defining flexibility for those known regions at the docking step has the potential to account for the proteins that undergo significant structural rearrangement when forming a complex. Ultimately, with cryo-ET reaching higher resolutions at a system-wide level and with further development of proteome-wide *in situ* XL-MS, computational predictions could be combined with these more native *in situ* approaches (Figure 2B), which we extensively discuss below. A similar approach has been utilized in deciphering the structure of the human NPC based on computational predictions and cryo-ET data (Mosalaganti et al., 2021).

NOVEL PIPELINES

Integration of MS-based approaches into cryo-EM

The next logical step after combining EM and MS output data but before combining hardware would be to incorporate MS tools into the cryo-EM pipeline. Here we describe in detail two novel solutions: (1) dedicated cryoID software (Ho et al., 2020) and (2) build and retrieve, or BaR (Su et al., 2021). Both approaches are complementary to each other and have the potential to extend the structural biologists' toolbox.

Analogous to bottom-up proteomics, in which protein identifications are built up from peptides as the smallest building blocks, cryoID can be viewed as a bottom-up structural proteomics tool. The 20 amino acids are combined into 6 smaller groups, with each group made of amino acids with similar cryo-EM densities for the side chains. For maps better than 4 Å, cryoID assigns protein residue side chain identities in the form of 6-letter codes based on structural features. Predicted sequences are then eventually used to screen across a list of potential candidates identified by MS-based proteomics performed on the corresponding fractions. In the proof-of-concept study (Ho et al., 2020), cryoID was first successfully tested on EM maps from the Electron Microscopy Data Bank (EMDB) (Lawson et al., 2011). The authors then tested cryoID on new cryo-EM data from sucrose gradient fractions of *P. falciparum* lysate wherein sets of cryo-EM densities were processed to the highest possible resolution using cryoSPARC and Relion. In addition to EMDB structures and several previously known protein complexes, cryoID has been successfully applied to elucidate a near-native structure of the RhopH complex, which was not possible to model with conventional structural biology approaches (Ho et al., 2021). It is important to emphasize that the reported cryoID approach is not an alternative to *in situ* cryo-ET pipelines that capture systems of interest in their near-native states in their biological contexts. Current literature using cryoID explores structures obtained from lysates of single organisms and suggests that native interactomes are affected by a loss of transient or weak interactions. The molecular weight (MW) of proteins reported using cryoID varies in the range of 435 kDa to 788 kDa, leaving low-MW and megadalton complexes unexplored. Also, the reported pipeline only addresses targets in the soluble parts of the proteome.

BaR is suitable for, but not limited to, the characterization of low-abundance proteins and membrane proteins. Cellular samples are lysed, then solubilized in a detergent followed by incorporation into nanodiscs and SEC enrichment. After extensive cryo-EM imaging, data analysis is performed with BaR as follows: initial cryo-EM maps are used as starting models for independent 3D heterogeneous classification and refinement after particle stacks are cleaned with 2D classification. Simultaneously, nMS and shotgun proteomics are used to confirm the presence of the target protein complexes in the sample. The BaR pipeline is exemplified in *E. coli* K12 crude cell membranes and raw cell lysates producing EM maps with resolutions between 2.2 and 3.6 Å for the proteins in the MW range of 111 kDa to 363 kDa, both for cytosolic and membrane targets. One of the advantages of BaR is the omission of protein purification, which for some targets is the only available approach. In another example, BaR was applied to the *E. coli* BL21(DE3)ΔacrB cells to overexpress the otherwise elusive 93 kDa membrane protein, *Burkholderia pseudomallei* hopanoid biosynthesis-associated RND (BpHpN, from the family of resistance-nodulation-cell division or RND transporters). Due to its low expression levels, previous attempts at purifying this target have failed. With BaR, an EM density of BpHpN was obtained at ~3.6 Å resolution. BaR is comparable to the cryoID pipeline for soluble proteins for near-native protein structure characterization. For membrane proteins investigated with BaR, extraction with detergent might alter the native protein structure and higher-order structural assemblies despite subsequent incorporation into nanodiscs. Like cryoID, BaR does not capture transient or weak interactions and thus far has only been applied to complexes in the lower-MW range, with a maximum MW of 363 kDa. Another important limitation is the large number of cryo-EM micrographs required to achieve high resolution.

For both cryoID and BaR, nMS and shotgun proteomics are essential for identification purposes, which significantly simplifies the assignment of cryo-EM densities from the heterogeneous sample. Combining cryoID and BaR with XL-MS on lysates will lead to stabilizing transient and weak interactions, allowing for previously unreported parts of the interactome to be resolved. XL-MS will also stabilize higher-order protein organization, and identification of distance constraints will help guide the integrative modeling to uncover higher-order protein complex assemblies. Ultimately, comprehensive XL-MS characterization could pave the road toward identifying densities produced with BaR and cryoID sample preparation that are not readily assignable to known protein complexes. BaR and cryoID target different parts of the proteome and various MW ranges, suggesting a need for their simultaneous application, or even integrating both methods into a single pipeline for a more complete characterization of biological systems of interest.

FROM STRUCTURES TO SYSTEMS

Snapshots of living cells: The need for *in situ* cryo-ET cellular structural interactomics

Most biological processes within living cells require a higher order of functional organization than a single molecule. Signal transfer within the cell, cellular regulation, and protein synthesis followed by their transport across membranes are just a few examples of crucial processes for healthy cellular function that are regulated and mediated by various protein-protein and protein-lipid interactions (for review, see Westermarck et al., 2013).

To perform these cellular functions, proteins form assemblies, resulting in highly complex molecular machines. Analogous to the terms genotype and phenotype, the term proteotype, which describes the active, real-time state of the proteome in a cell considering both the protein complement and organization, was introduced recently (Röst et al., 2015). Cellular environments provide unique conditions for proteotype investigations, and purification of certain macromolecular complexes inevitably leads to their decomposition, deformation, or complete loss due to low expression levels. *In situ* approaches circumvent purification and are often the only approaches for obtaining insights into the morphologies and behaviors of molecular assemblies inside the cell (Watanabe et al., 2020; Burbaum et al., 2021; Peukes et al., 2021). Similarly, protein networks within cells and on cell surfaces require a bird's-eye view to analyze their functionality and higher-order organization (Mahamid et al., 2016; Brasch et al., 2019).

The number of functional PPIs within a human cell is estimated at ~200,000 (Hein et al., 2015). As of now, no single approach can address such complexity within a single experiment, and therefore the development of multifaceted and comprehensive system-wide strategies involving complementary methods are needed. Proteome-wide screens (Krogan et al., 2006; Havugimana et al., 2012), a combination of AP-MS and proximity-dependent biotinylation identifications (BioID) called multiple approaches combined (MAC) (Liu et al., 2018), and multiple XL-MS datasets have already successfully been generated in various interactomics studies, establishing a basis for several public databases: CORUM (Giurgiu et al., 2019), proteomics.fi (Liu et al., 2018), and XlinkDB (Schweppe et al., 2016). However, purification approaches do not preserve the native protein interaction landscape, particularly in the case of weak or transient interactions. With the capability of cryo-ET to visualize all interactions directly inside the system, for large-MW proteins it is currently possible to capture otherwise elusive (Guo et al., 2018; O'Reilly et al., 2020) or, hypothetically, previously unknown protein assemblies without the use of labels by integrating cryo-ET with MS proteomics analyses.

In situ cryo-ET captures a spatiotemporal snapshot of near-native protein systems at unprecedented resolution. The spatial context provides functional information through localizing protein complexes within cellular/biological environments (Gupta et al., 2021). Additionally, observations of intermediate complex states without introducing bias from labels or purification by biochemical or genetic manipulation is only possible with *in situ* approaches (Erdmann et al., 2021). Another advantage of *in situ* approaches is the characterization of membrane proteins within their native lipid environment. Investigating samples at near-native conditions allows for more complete PPI profiles in membrane regions (Yu et al., 2021a). An intermediate solution is to recreate native lipid environments through vesicle formation (Chen and Kudryashev, 2020). This approach was successfully used to give unique insights into ryanodine receptor 1 (RyR1) activation, which filled structural gaps in models obtained with SPA.

System-wide *in situ* XL-MS and its limitations

From a single crosslinking experiment on lysed human cells, thousands of crosslinks can routinely be detected (Klykov et al., 2018). *In situ* XL-MS experiments, however, would be more appealing as these can be performed directly in the cell without disrupting protein complexes. This approach so far has successfully been applied to intact organelles with large transport systems located on the membrane, including isolated nuclei and mitochondria (Liu et al., 2017; Schweppe et al., 2017; Fasci et al., 2018). Beyond organelles, *in situ* XL-MS has been applied to biopolymers (Klykov et al., 2020), tissues (Chavez et al., 2018), thylakoid membranes (Albanese et al., 2020), and mammalian sperm cells (Leung et al., 2021). In the case of other types of intact cells, the reports in the literature remain controversial because all currently available and most widely used crosslinking reagents, except for disuccinimidyl suberate (DSS), carry charges, preventing their passage through the membrane. Several studies have attempted to perform these experiments with, for example, prokaryotes (O'Reilly et al., 2020) and eukaryotes (Slavin et al., 2021), although other reports have highlighted that the crosslinked peptides solely form in the extracellular space (Armony et al., 2021).

In addition to the limitations described above, the efficiency of the crosslinking reaction itself is not optimal with current reagents in terms of reaction conditions in buffer solution for cell lysates and *in situ*, which highlights the need to explore further chemistry. A decade ago, the number of detectable crosslinked peptides from cell lysates was estimated to be a maximum of 4%–6% from the total amount of potentially crosslinked residues (Leitner et al., 2010). Even after numerous advances in the field, a large part of the proteome remains uncovered (Klykov et al., 2018). High-density crosslinking (Belsom et al., 2016, 2017; Mintseris and Gygi, 2020) boosts the number of identifications through either alternative reagent chemistry or combinations of reagents. An arguably more elegant approach is to create novel crosslinking reagents that incorporate an enrichable handle for extraction of crosslinked peptides. The two most successful versions of this concept are the protein interaction reporter (PIR) linker, which incorporates a biotin handle (Zhong et al., 2017), and PhoX, which carries a handle for an immobilized metal affinity chromatography (IMAC) enrichment (Steigenberger et al., 2019). By pulling the crosslinked peptide pairs from the background, significantly higher depth of analysis can be achieved. Further optimizations are currently being explored with, for example, extensive fractionation (Lenz et al., 2021), advanced MS setups (Steigenberger et al., 2020), tailored data acquisition approaches (Hauri et al., 2019; Giese et al., 2021), and more robust reporting in terms of false-positive identifications (Lenz et al., 2021). A combination of these modified reagents, technological advancements, and optimization of reagents' membrane permeability will undoubtedly further improve proteome-wide *in situ* crosslinking results.

LABEL-FREE VISUAL PROTEOMICS

Challenges in visual proteomics

Currently, the cryo-FIB/SEM-ET workflow from sample preparation to tomogram analysis has several challenging aspects that are continually being developed and improved, including (1) object of interest identification/localization during cryo-FIB-milling and in cryo-ET tomograms, (2) throughput of cryo-FIB-milling,

tilt-series collection, and tomogram-processing workflows, and (3) resolving repeating proteins smaller than several hundred kilodaltons to high resolutions.

Apart from quantitative and stoichiometric information of protein and protein complex abundances in cells, locating objects of interest at the subcellular level may be supplemented by FLM, cryo-FLM, and eventually cryo-SRLM (Bauerlein and Baumeister, 2021) of the specimen on the EM grid, which often require fluorescent tags as labels. To retain more native, label-free visual proteomics information, a separate labeled specimen may be prepared to confirm presence of the objects of interest, localize those objects relative to organelles as landmarks, and obtain potential initial models of objects of interest. Then this additional knowledge can help guide localization during cryo-FIB-milling, cryo-ET collection, and cryo-ET processing. Additional labels for protein localization in tomograms, including metal and nonmetal tags (Silvester et al., 2021; Tan et al., 2021), may be biochemically added to the labeled specimen to help identify and understand the morphologies of the structures of interest in the corresponding label-free specimen.

To address cryo-FIB-milling throughput, software automation has recently been introduced both commercially and as open source packages (Klumpe et al., 2021), new FIB-milling modalities have been developed to increase the size of the lamellae (Schaffer et al., 2019; Kelley et al., 2020), and commercial companies are currently exploring using plasma sources for FIB-milling instead of gallium, which is expected to increase throughput by several factors. Faster and more robust lamellae production will require an increase in cryo-ET tilt-series collection throughput, which may be overcome by the BISECT method (Bouvette et al., 2021) for parallel tilt-series collection. As cryo-ET tilt-series are collected more rapidly, the tomogram post-processing pipeline will need to become more automated to account for the increased throughput. Multiple software projects are heading in this direction, including a full pipeline for tilt-series to high-resolution structure in EMAN2 (Chen et al., 2019), a full community-based tomogram-processing pipeline using several software packages (Burt et al., 2021), and a standalone fully automated tomogram-processing pipeline by the Kudryashev lab (refer to https://github.com/KudryashevLab/tomoBEAR).

The last of the challenges requires first localizing each individual object in the size range of the protein of interest, which is especially difficult in the crowded and noisy tomogram environments. Increasing the contrast of cryo-ET tomograms by collecting with a Volta-phase plate (Danev and Baumeister, 2016) or laser-phase plate (Schwartz et al., 2019) in the cryo-TEM or by applying deep learning denoising models (Buchholz et al., 2018; Bepler et al., 2020; Tegunov et al., 2021) to the tomograms will help, but likely will not allow these objects to be cleanly differentiated. A new method for creating deep learning models to significantly reduce tomogram anisotropicity, implemented in IsoNet (Liu et al., 2021), will also help but not solve this challenge. Once proteins of interest are localized, high-resolution sub-tilt-series refinement pipelines (Himes and Zhang, 2018; Chen et al., 2019; Tegunov and Cramer, 2019; Tegunov et al., 2021; Kimanius et al., 2021) may be used to maximize the resolution of each segmented object. An alternative method for cases where models of proteins or components exist may be to use a new high-resolution template-matching software to directly and confidently localize proteins of interest (Rickgauer et al., 2017). Once all structures of interest are determined in a cryo-ET dataset, those higher-resolution structures may be mapped back into the tomogram scenes, with denoising and isotropic correction, then combined with additional MS data from the specimen. Such visual molecular encyclopedias will help researchers identify and quantify observed proteins and interactions versus those predicted by MS and XL-MS of corresponding lysates. Subsequently, proteins and interactions that exist in MS studies but are missing or in improper abundance in cryo-ET may be more easily identified after the re-mapping procedure by computationally subtracting known objects. With the ever-expanding BioImage Archive (Ellenberg et al., 2018), a public repository for biological imaging and associated data, annotated datasets with multiple data sources may be used to train future deep learning models to discover signals and patterns (Mund et al., 2021) within and across datasets that humans and other analysis methods may miss.

Comprehensive structural characterization of cells and tissues by MS and cryo-ET

According to the classical definition of visual proteomics (Nickell et al., 2006), one of its goals is to complement and extend MS-based inventories and to retrieve quantitative information of the PPIs inside the cell. The first successful implementations of this approach have been done in investigations of the spatial proteome of *T. acidophilium* (Sun et al., 2007) and *L. interrogans* (Beck et al., 2009). Visual proteomics together with *in situ* cryo-ET requires that individual macromolecular complexes are annotated within tomograms (Pyle and Zanetti, 2021). Although it is possible to get functional information of certain targets with sophisticated visual proteomics approaches (Erdmann et al., 2021; Martinez-Sanchez et al., 2020), *in situ* cellular cryo-ET datasets are inherently crowded, complex, and heterogeneous. Moreover, raw cryo-ET data are unlabeled. While cells can be made to express fluorescent tags through genetic engineering, current mainstream approaches to visualize the expressed tags are limited to resolutions of several hundred nanometers. Current methods and reports are only scraping the surface of the wealth of information generated by *in situ* cryo-ET. We postulate that to achieve the goals of visual proteomics to characterize any cell of any size and every macromolecular assembly, we should employ a modern variant of visual proteomics wherein cryo-ET is complemented and extended with the MS-based inventories. Here, we define the combination of MS and cryo-ET as *in situ* label-free visual proteomics, along with any future label-free extensions. The clear advantage of *in situ* label-free visual proteomics is that the specimens are perturbed minimally from their native states during structural and interactome investigations, with the only exception being crosslinking for XL-MS studies.

The first layer of visual proteomics information of the cell would be provided by shotgun MS (Figure 3A). In addition to qualitative characterizations of samples, MS-based proteomics reports relative or absolute amounts of the proteins present. Complementary quantitative information may then be retrieved from cellular cryo-ET data in the form of localization of proteins and

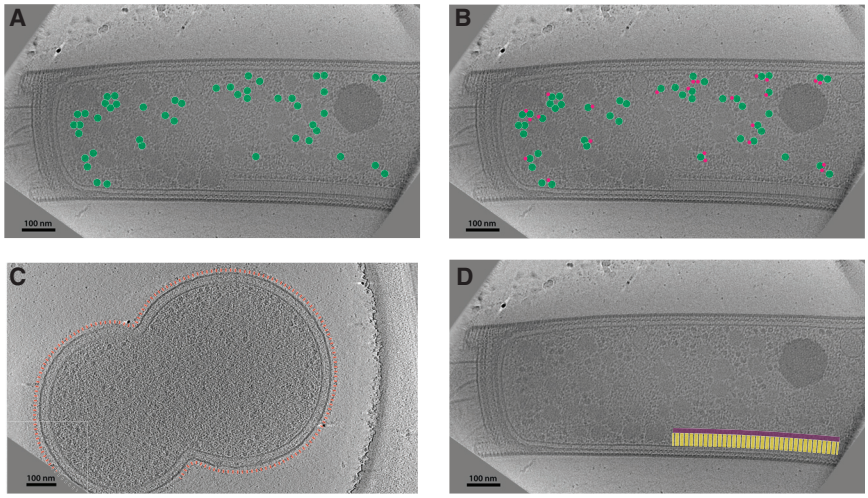


Figure 3. Label-free visual proteomics with *in situ* shotgun proteomics, XL-MS, and cryo-ET

(A) Cross-section of a cellular tomogram with ribosome densities highlighted in green. In this case, ribosome quantification performed with shotgun proteomics would identify the relative amount of ribosomes present in the sample as opposed to quantification by means of visual proteomics, which would omit the ribosomes not present in the field of view (analogous to Albert et al., 2019). Identification of the most abundant proteins with shotgun MS allows for correlation of the most-often-detected protein in the tomogram.

(B) Cross-section of a cellular tomogram with ribosome highlighted in green and unidentified ribosome-associated densities highlighted in red. In this case, XL-MS would pinpoint the most probable ribosome interaction partner and help to identify those densities (analogous to O'Reilly et al., 2020).

(C) Cross-section of a cellular tomogram highlighting the heterogeneous gingipain proteins in orange. Additional conformational stabilization of gingipains by XL-MS reagents would reduce flexibility and help in the high-resolution cryo-ET pipeline (analogous to Ke et al., 2020).

(D) Cross-section of a cellular tomogram with membrane-bound chemosensory arrays and unidentified associated densities highlighted in yellow and purple, respectively. In this case, XL-MS along with quantitative shotgun MS data generated in parallel would help to identify unassigned densities, protein complex orientation, and higher-order protein organization through protein self-links (analogous to Leung et al., 2021).

The cross-sections through the cellular tomograms are taken from the publicly available Atlas of Bacterial and Archaeal Cell Structure (https://www.cellstructureatlas.org) with the protein densities assigned based on the information from the atlas. Figures 3A, 3B, and 3D represent data from *M. hungatei* (10.22002/D1.1357) and Figure 3C are acquired on *P. gingivalis* (10.22002/D1.1578).

their interactions relative to the cell interior. Both workflows indeed correlate between each other very well, as has been demonstrated by Albert et al., 2019. In this study, the authors applied quantitative MS-based proteomics of cell lysates and cryo-ET of lamellae generated from unicellular algae *C. reinhardtii*, then assessed and compared the concentrations of ribosomes, proteasomes, and Cdc48 protein. The ribosome concentration was found to be identical in both methods, and the concentrations of Cdc48 differed insubstantially. For proteasomes, visual inspection of tomograms reported slightly lower values due to the localization of proteasomes in nuclei, which was not present on the tomograms. This study demonstrates the power of label-free visual proteomics in protein quantification by cryo-ET, yet it shows the need for additional proteomics validation by MS for the targets located in various subcellular compartments that are not yet identifiable in tomograms.

The possibility of performing *in situ* XL-MS alongside cryo-ET appears to be the most attractive and potent combination of methods for visualizing cellular proteomes (Figure 3B). The landmark study by O'Reilly et al., (2020) reporting the structure of the expressome complex in action demonstrates the potential of combining these approaches. Due to its relatively high sensitivity, *in situ* XL-MS led the investigators to discover the presence of the active expressome complex, which was further structurally characterized by cryo-ET. A total of 352 cryo-ET tilt-series were collected from the control sample, and only 2.8% of the total number of ribosomal particles comprised the active expressome. While the extensive fractionation and data acquisition for fractionated crosslinked samples generally takes longer than cryo-ET acquisition, the straightforwardness and robustness of XL-MS with the relatively well-established data analysis pipeline makes it a powerful tool for capturing the *in situ* conformations of complexes prior to cryo-ET experiments. XL-MS may also be used to screen for abundant protein complexes without requiring massive amounts of data or extensive fractionation, and the whole pipeline can be performed in a matter of hours. Together, the study by O'Reilly et al. showed the power of label-free visual proteomics by combining cryo-ET and XL-MS. At the same time, it showed the need for improvements in the pipeline, particularly with cryo-ET data processing.

While XL-MS of cellular specimens describes interactomes unbiasedly, low-resolution information from cryo-ET provides a precise landscape onto which information from XL-MS data may be mapped. Due to its nominal low resolution, however, correct assignments of proteins within cryo-ET tomograms are often problematic when few, if any, proteins are identified by cryo-ET alone. XL-MS can shed light on interaction partners when at least one component of a complex is assigned and can provide an encyclopedia of protein complexes present in the sample from which candidate complexes can be identified, for example, by approximating size. When most or all of the components of a given protein complex are known, *in situ* XL-MS and cryo-ET are technically capable of deciphering the interaction interfaces. While cryo-ET alone can produce EM densities for larger (several hundred kilodaltons) proteins with high abundance within the cell at medium/high resolutions sufficient for interaction studies, currently medium/high-resolution cryo-ET is limited to very few cases (Schur et al., 2016; Tegunov et al., 2021; Pyle and Zanetti, 2021), and integrative modeling rather than *de novo* model building must be used to obtain the final complex structure. When only low-resolution structures are available for the low-abundant or transient protein complexes, XL-MS restraints could be used alongside cryo-ET densities for integrative modeling (O'Reilly et al., 2020). Alternatively,

XL-MS distance restraints could help with the validation of the interaction interfaces for the cryo-ET models, which has already been done for purified targets (Meyer et al., 2019).

Like purified protein complexes *ex situ*, XL-MS is able to stabilize protein complexes *in situ* prior to downstream analysis (Hevler et al., 2021; Figure 3C). This approach has already been combined with cryo-ET, though with a low-selectivity 4% formaldehyde fixation reagent (Ke et al., 2020). It is important to realize that crosslinking by fixation does not produce new conformations but potentially shifts the equilibrium toward certain conformational states. This property can be used to target specific states in a conformational range. Moreover, with UV-activatable crosslinking reagents, timed snapshots of cellular systems of interest can be obtained. The UV moiety reacts within the millisecond range and may be triggered at any moment with UV light. This could (1) provide temporal context after a cellular event or addition of a drug to the system and (2) lock the system in a certain time-dependent conformational state that will increase the number of particles in cryo-ET and therefore increase the possibility of getting high-resolution EM density maps. While the heterobifunctional UV-reactive reagents used in high-density XL-MS are much more efficient than monofunctional UV reagents, their application *in situ* is so far not possible due to the massive number of generated crosslinked peptides. In this case, a targeted approach with much lower throughput may find an application (Yang et al., 2017).

The most important impact that XL-MS information may have on cryo-ET data in the near future is to provide insights into the biological function of large macromolecular assemblies. Leung et al., 2021 show the potential for MS-based methods to supplement cryo-ET data in terms of protein assignment and localization, resulting in valuable biological insights. First, based on low-resolution cryo-ET density maps, periodicity of the arrays, and pore sizes, it was possible to identify the components of repeating protein arrays in tomograms as voltage-dependent anion channels (VDACs). Label-free visual proteomics quantification detected VDAC2 and VDAC3 as the most abundant proteins in the sample, and the crystal structures of VDAC2 fit the cryo-ET densities. Next, from the interaction partners of VDAC2/VDAC3 identified by *in situ* XL-MS, it was possible to pinpoint glycerol kinase (GK) for fitting into one of the cryo-ET densities. This visual proteomics approach allowed for integrative model building, which was accomplished by first rigid-body fitting of homology models of GK and VDAC2 with the correct orientations of the GK-VDAC complexes chosen by the XL-MS restraints (Figure 3D). Based on the results of this study, it was possible to confirm the direct link between GK protein arrays and submitochondrial reticulum.

While XL-MS can be performed label-free on various samples without additional sample preparation steps, for cryo-ET, the sample thickness is a limiting factor. O'Reilly et al., 2020 used the *M. pneumonia* organism, which is thin enough (below 150 nm) for the electron beam to pass through, but in the case of Leung et al., 2021, cryo-FIB-milling was required to create TEM-transparent thin sections of their mammalian sperm cells. Cryo-FIB-milling can create lamellae from almost any type of biological sample, including whole cells, vitrified organisms, and tissues. While practically difficult, cryo-FIB-milling is being rapidly developed with the introduction of automated milling software solutions and new instrumentation (Klumpe et al., 2021; Tacke et al., 2021) as well as advanced vitrification and milling techniques (Schaffer et al., 2019; Kelley et al., 2020). Due to its ability to create near-native cross-sections of biological specimens that allow for the highest-resolution *in situ* 3D visualization by cryo-ET tomograms, cryo-FIB-milling is rapidly being turned to by biologists and biomedical scientists. We envision that, similar to O'Reilly et al., 2020 and Leung et al., 2021, combining XL-MS, cryo-FIB, and cryo-ET will lead to complementary *in situ* characterizations of a wide variety of biological objects.

Analyzing samples with multiple techniques within the same instrument

EM and MS are highly complementary approaches that are intertwined in terms of the biological questions they can answer. In a long-term perspective, deeper integration of these methods into a single rapid and efficient EM-MS pipeline would prove beneficial for the structural biology research community. Here we describe two integration concepts: (1) analyzing the same sample by MS and EM with separate instruments and (2) coupling the mass spectrometer into the TEM or SEM instruments.

Hardware integration of MS approaches into the EM pipeline already exists. One of the first examples of such coupling was reported by Mikhailov et al., 2014 based on nMS instrumentation. In this setup, a mass spectrometer was employed to separate ions on the fly followed by their soft landing on an EM grid for subsequent negative-stain EM. Termed native electrospray ion-beam deposition (ES-IBD), the approach possesses all the limitations of both nMS and negative-staining EM, namely sample purity, potential perturbation of the native state of proteins in vacuum, and the low resolution of negative-stain EM. Recently, the approach was modified and implemented with current nMS instruments for both negative-stain (Westphall et al., 2021) and frozen (Esser et al., 2021) grid preparation. Frozen EM grid preparation with ES-IBD is especially interesting, as this would potentially allow for obtaining high-resolution structures. First, however, the issues of protein structure perturbation due to exposure to vacuum, transfer through the instrument, dehydration, and landing on the EM grid must be solved. Coupling ES-IBD with the online buffer-exchange setup discussed above (VanAernum et al., 2020) would facilitate the speed of analysis and improve the quality of EM grids by decreasing sample heterogeneity. Unfortunately, ES-IBD is not applicable for *in situ* studies, as it is based on nMS, which works with purified samples.

Another pioneering integrative approach is deep visual proteomics (DVP; Mund et al., 2021). DVP requires immunofluorescence labeling (for cells) or immunohistochemical staining (for tissues) and employs high-resolution confocal imaging, ultrasensitive proteomics pipelines, and deep learning image analysis for single-cell phenotyping and isolation. To excise the regions of interest (ROIs), automated laser microdissection is performed, excising ROIs with a precision of 200 nm. The proof-of-concept study demonstrates utility for both cellular and tissue samples. For cells, DVP correlates cellular phenotypes, heterogeneity, and changes on a proteome level, whereas for tissues, DVP can pinpoint disease-specific protein targets across single cell types and quantify the spatial diversity of those targets. The primary

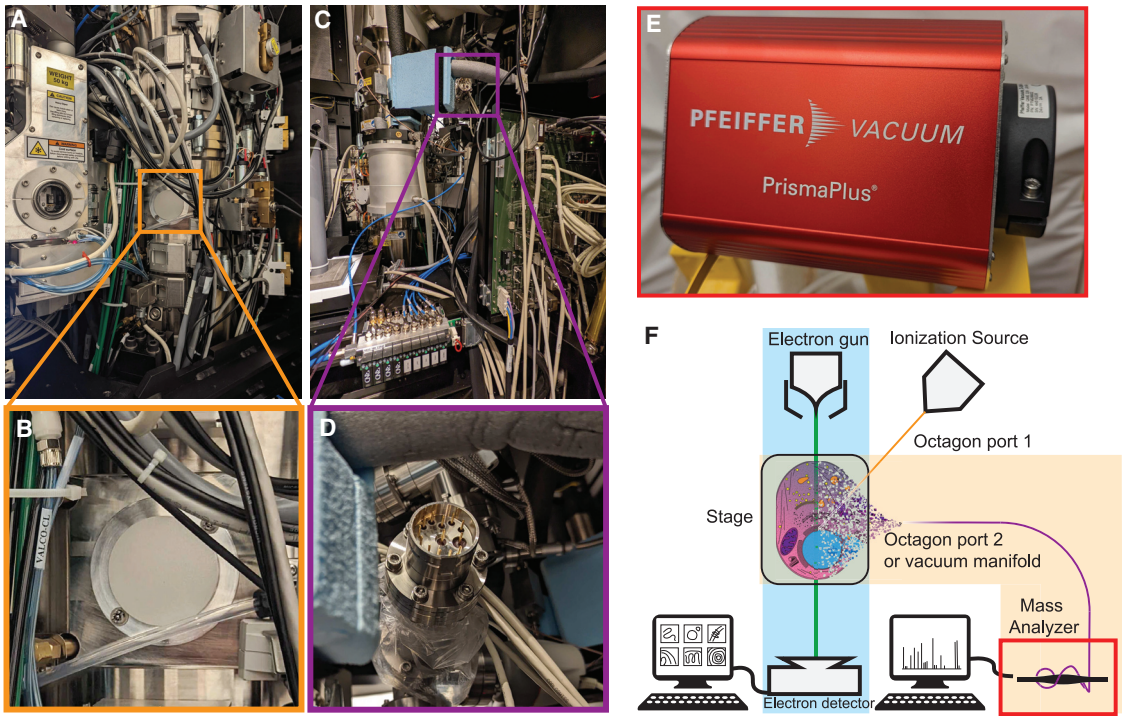


Figure 4. Integration of EM and MS hardware

(A–D) The front view of a G3 Titan Krios column showing an available octagon port (A, B) and a back view of the same column showing an adaptor for a residual gas analyzer (RGA) that is connected to the vacuum manifold (C, D). There is an extra octagon port available in the front of a G3 Titan Krios TEM column with ample space within the enclosure (not shown).

(E) Image of the RGA itself, which allows for measuring of low-MW contaminants in the TEM column. Instrument parts containing mass analyzers are highlighted with red boxes.

(F) Schematic representation of a combined EM and MS instrument for real-time localization of ROIs. The sample is ionized in the area close to the ROI with high precision, and then ions are guided toward the advanced mass analyzer, which is schematically represented as an orbitrap. After confirming the presence of the protein target in the vicinity, the ROI is either imaged by TEM or subjected to FIB-milling.

limiting factor in DVP is the requirement for room-temperature samples, which places the higher-resolution *in situ* cryo-ET methods out of reach. A modified DVP pipeline with integrated cryo-ET imaging could open frontiers well beyond our current boundaries, including simultaneous structural and visual cellular-wide characterization of proteins and protein complexes on nanometer and sub-nanometer scales. Integration of cryo-EM/ET pipelines into DVP will allow retrieval in medium/high-resolution of structural information from *in situ* samples supplemented with MS identification and quantification of the targets of interest. One way to combine DVP and cryo-EM/ET would involve thawing of the sample after imaging. Cells or tissues typically maintain their native properties while in a vitreous state, and in theory, intact cells with thin edges ~200 nm or thinner may be imaged in cryo-ET in a vitreous state, then thawed and localized with single-cell isolation and laser microdissection. The amount of protein material in this case is severely limited even for advanced single-cell proteomics (scP) MS approaches (Specht et al., 2021) and represents the major challenge of the proposed workflow. A possible alternative may be to decrease the correlation accuracy of EM visualization and MS identifications from a single imaged cell area to whole grid square, which for cell thin-edge experiments contains a handful of cells and will partially circumvent the sensitivity issue. Another way to increase the amount of protein material is to apply modified DVP to a specimen prepared with the waffle method (Kelley et al., 2020) or cryo-liftout (Schaffer et al., 2019), where the amount of sample in the grid square is even higher than in a thin-edge cell experiment. For such cryo-FIB-milled samples and precise correlation of imaged areas on lamellae, a thawing strategy would require isolating previously imaged lamellae prior to MS analysis, which may be possible with the cryo-liftout capabilities that are already integrated into modern cryo-FIB-milling instruments. A challenge for all post-cryo-EM/ET MS analyses will be overcoming the radiation damage that has already accrued onto the sample during cryo-TEM collection. Performing MS experiments on the exact areas after TEM imaging could provide insight into the radiation damage through quantifying the accrued chemical modifications. In the end, we anticipate that a reliable method will be developed to combine cellular *in situ* cryo-ET spatial information with MS to assign specific proteins and interactions to specific tomogram densities inside the cell.

Protein assignment in the rich *in situ* cellular environment is a complicated task. The identification process would be significantly facilitated when MS is directly coupled or integrated inside the TEM instrument. As MS methods are destructive by their nature, the only possible combination is to perform MS analysis on the sample after EM imaging is done. The sensitivity limitation would be less pronounced as the sample would be directly

transferred into the MS instrument. Cryo-TEM instrumentation equipped with a vacuum port (Figures 4A–4D) can be coupled to quadrupole-based residual gas analyzers (RGA, Figure 4E) for identifying low-MW contaminants present in the microscope column. Extending the quadrupoles with advanced ion traps, orbitraps, or time-of-flight (TOF) detectors would be the next step toward identifying proteins and peptides. While several technical parameters regarding hardware placements and connections between the instruments have yet to be established, the most difficult part is the ionization of the vitrified sample. Advanced extraction techniques like liquid extraction surface analysis (LESA, Griffiths et al., 2019) allow intact protein sampling from glass slides and even thin tissue sections of thawed intact cells. While ionization directly from vitreous ice is currently not feasible, protein sampling from cryo-FIB-milled and then thawed tissue lamellae may be possible. LESA coupled to nMS is so far limited to maximum 800 kDa proteins ionized from a glass slide, or maximum 80 kDa for thin tissue slices. Currently, native LESA-MS does not have a limitation on heterogeneity due to the small number of proteins ionized, but in the future further digestion into peptides followed by shotgun proteomics approaches might be necessary to identify all the proteins present. If this is the case, online protein digestion (Villegas et al., 2019) together with microreactors and immobilized proteases must also be integrated into the coupled EM-MS instrument.

While MS can be performed post-TEM imaging for identification purposes, real-time integration of MS into SEM or TEM instruments would help target ROIs analogous to cryo-FLMs, but without labeling requirements. Cryo-FLM currently suffers from low-resolution (typically several hundred nanometers), and even with the most advanced imaging solutions, localizing the ROIs and guiding cryo-FIB-milling and cryo-TEM imaging are not straightforward. With a mass spectrometer inside the cryo-SEM/TEM instrument, real-time navigation of FIB-milling and TEM imaging may be successfully combined with cryo-FLM navigation for better precision (Figure 4F). Additionally, due to the nature of MS analyses and destroying samples during ionization, the exact ROI should be kept intact, making this approach viable only if the ionization precision is better than cryo-FLM precision. TOF mass analyzers are already integrated into FIB-SEM instruments and are used in material science (Rickard et al., 2020), leaving the sample ionization from within the vitreous ice while avoiding the sample devitrification as the major challenge.

CONCLUDING REMARKS

The tremendous advances in cryo-EM/ET and structural MS fields with regards to resolution in cryo-EM/ET and depth of analysis in MS have enabled the detection of protein-protein interactions on a system-wide scale with molecular precision. Label-free visual proteomics, described here as the combination of cryo-EM/ET and MS on *in situ* cell/tissue specimens, produces near-native structures combined with interaction networks that allow for more accurate and comprehensive cell/tissue specimen descriptions. Supplementing label-free specimens of interest with fluorescently labeled counterparts for additional localizability may further increase the number of biological and biomedical questions that may be answered, particularly with future developments in cryo-SRLMs. Further development of deep learning approaches to data analysis, such as AlphaFold, DVP, and cryo-ET tomogram unsupervised segmentation and classification, will continually reduce the data-processing bottleneck of these highly dense datasets while simultaneously improving data interpretability. We envision that label-free visual proteomics will be widely applied for comprehensive, high-resolution, near-native structural interactomic studies of cells and tissues for the foreseeable future.

ACKNOWLEDGMENTS

A.J.R.H. and R.A.S. acknowledge support by the research program NWO TA project number 741.018.201, which is partly financed by the Dutch Research Council (NWO); additional support for A.J.R.H. and R.A.S came through the European Union Horizon 2020 program INFRAIA project Epic-XS (project 823839). A.J.N., M.K., and B.C. acknowledge support from the Simons Foundation (SF349247) and NIH (GM103310, U24GM129539, U24GM139171).

AUTHOR CONTRIBUTIONS

O.K., M.K., B.C., A.J.R.H., A.J.N., and R.A.S. conceived, wrote, and edited the manuscript. For cryo-EM/ET/FIB-SEM, contact A.J.N. at anoble@nysbc.org; for MS, contact R.A.S. at r.a.scheltema@uu.nl.

DECLARATION OF INTERESTS

The authors declare no conflicts of interest.

REFERENCES

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. Nature *422*, 198–207.

Ahdash, Z., Lau, A.M., Byrne, R.T., Lammens, K., Stüetzer, A., Urlaub, H., Booth, P.J., Reading, E., Hopfner, K.P., and Politis, A. (2017). Mechanistic insight into the assembly of the HerA-NurA helicase-nuclease DNA end resection complex. Nucleic Acids Res. *45*, 12025–12038.

Akey, C.W., Singh, D., Ouch, C., Echeverria, I., Nudelman, I., Varberg, J.M., Yu, Z., Fang, F., Shi, Y., Wang, J., et al. (2021). Comprehensive Structure and Functional Adaptations of the Yeast Nuclear Pore Complex. bioRxiv. Published November 3, 2021. https://doi.org/10.1101/2021.10.29.466335.

Albanese, P., Tamara, S., Saracco, G., Scheltema, R.A., and Pagliano, C. (2020). How paired PSII-LHCII supercomplexes mediate the stacking of plant thylakoid membranes unveiled by structural mass-spectrometry. Nat. Commun. *11*, 1361.

Albert, S., Wietrzynski, W., Lee, C.-W., Schaffer, M., Beck, F., Schuller, J.M., Salomé, P.A., Plitzko, J.M., Baumeister, W., and Engel, B.D. (2019). Direct visualization of degradation microcompartments at the ER membrane. Proc. Nat. Acad. Sci. USA *117*, 1069–1080.

Arakhamia, T., Lee, C.E., Carlomagno, Y., Duong, D.M., Kundinger, S.R., Wang, K., Williams, D., DeTure, M., Dickson, D.W., Cook, C.N., et al. (2020). Posttranslational Modifications Mediate the Structural Diversity of Tauopathy Strains. Cell *180*, 633–644.e12.

Armony, G., Heck, A.J.R., and Wu, W. (2021). Extracellular crosslinking mass spectrometry reveals HLA class I - HLA class II interactions on the cell surface. Mol. Immunol. *136*, 16–25.

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G.R., Wang, J., Cong, Q., Kinch, L.N., Schaeffer, R.D., et al. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. Science *373*, 871–876.

Bauerlein, F.J.B., and Baumeister, W. (2021). Towards Visual Proteomics at High Resolution. Journal of Molecular Biology *433*, 167187.

Beck, M., Malmström, J.A., Lange, V., Schmidt, A., Deutsch, E.W., and Aebersold, R. (2009). Visual proteomics of the human pathogen Leptospira interrogans. Nat. Methods *6*, 817–823.

Belsom, A., and Rappsilber, J. (2021). Anatomy of a crosslinker. Curr. Opin. Chem. Biol. *60*, 39–46.

Belsom, A., Schneider, M., Fischer, L., Brock, O., and Rappsilber, J. (2016). Serum Albumin Domain Structures in Human Blood Serum by Mass Spectrometry and Computational Biology. Mol. Cell. Proteomics *15*, 1105–1116.

Belsom, A., Mudd, G., Giese, S., Auer, M., and Rappsilber, J. (2017). Complementary Benzophenone Cross-Linking/Mass Spectrometry Photochemistry. Anal. Chem. *89*, 5319–5324.

Bennett, J.L., Nguyen, G.T.H., and Donald, W.A. (2021). Protein–Small Molecule Interactions in Native Mass Spectrometry. Chem. Rev. Published online August 27, 2021. acs.chemrev.1c00293.

Bepler, T., Kelley, K., Noble, A.J., and Berger, B. (2020). Topaz-Denoise: general deep denoising models for cryoEM and cryoET. Nat. Commun. *11*, 5208.

Bolla, J.R., Agasid, M.T., Mehmood, S., and Robinson, C.V. (2019). Membrane Protein-Lipid Interactions Probed Using Mass Spectrometry. Annu. Rev. Biochem. *88*, 85–111.

Bolla, J.R., Corey, R.A., Sahin, C., Gault, J., Hummer, A., Hopper, J.T.S., Lane, D.P., Drew, D., Allison, T.M., Stansfeld, P.J., et al. (2020). A Mass-Spectrometry-Based Approach to Distinguish Annular and Specific Lipid Binding to Membrane Proteins. Angew. Chem. Int. Ed. Engl. *59*, 3523–3528.

Bouvette, J., Liu, H.-F., Du, X., Zhou, Y., Sikkema, A.P., da Fonseca Rezende E Mello, J., Klemm, B.P., Huang, R., Schaaper, R.M., Borgnia, M.J., and Bartesaghi, A. (2021). Beam image-shift accelerated data acquisition for near-atomic resolution single-particle cryo-electron tomography. Nat. Commun. *12*, 1957.

Brasch, J., Goodman, K.M., Noble, A.J., Rapp, M., Mannepalli, S., Bahna, F., Dandey, V.P., Bepler, T., Berger, B., Maniatis, T., et al. (2019). Visualization of clustered protocadherin neuronal self-recognition complexes. Nature *569*, 280–283.

Britt, H.M., Cragnolini, T., and Thalassinos, K. (2021). Integration of Mass Spectrometry Data for Structural Biology. Chem. Rev. Published online September 10, 2021. acs.chemrev.1c00356.

Buchholz, T.-O., Jordan, M., Pigino, G., and Jug, F. (2018). Cryo-CARE: Content-Aware Image Restoration for Cryo-Transmission Electron Microscopy Data. ArXiv181005420 Cs. https://arxiv.org/abs/1810.05420.

Bullock, J.M.A., Sen, N., Thalassinos, K., and Topf, M. (2018). Modeling Protein Complexes Using Restraints from Crosslinking Mass Spectrometry. Structure *26*, 1015–1024.e2.

Burbaum, L., Schneider, J., Scholze, S., Böttcher, R.T., Baumeister, W., Schwille, P., Plitzko, J.M., and Jasnin, M. (2021). Molecular-scale visualization of sarcomere contraction within native cardiomyocytes. Nat. Commun. *12*, 4086.

Burke, D.F., Bryant, P., Barrio-Hernandez, I., Memon, D., Pozzati, G., Shenoy, A., Zhu, W., Dunham, A.S., Albanese, P., Keller, A., et al. (2021). Towards a structurally resolved human protein interaction network. bioRxiv. https://doi.org/10.1101/2021.11.08.467664.

Bürmann, F., Lee, B.G., Than, T., Sinn, L., O'Reilly, F.J., Yatskevich, S., Rappsilber, J., Hu, B., Nasmyth, K., and Löwe, J. (2019). A folded conformation of MukBEF and cohesin. Nat. Struct. Mol. Biol. *26*, 227–236.

Burt, A., Gaifas, L., Dendooven, T., and Gutsche, I. (2021). A flexible framework for multi-particle refinement in cryo-electron tomography. PLoS Biol. *19*, e3001319, e3001319.

Cater, R.J., Lin Chua, G., Erramilli, S.K., Keener, J.E., Choy, B.C., Tokarz, P., Chin, C.F., Quek, D.Q.Y., Kloss, B., Pepe, J.G., et al. (2021). Structural basis of omega-3 fatty acid transport across the blood-brain barrier. Nature *595*, 315–319.

Chavez, J.D., Lee, C.F., Caudal, A., Keller, A., Tian, R., and Bruce, J.E. (2018). Chemical Crosslinking Mass Spectrometry Analysis of Protein Conformations and Supercomplexes in Heart Tissue. Cell Syst. *6*, 136–141.e5.

Chen, W., and Kudryashev, M. (2020). Structure of RyR1 in native membranes. EMBO Rep. *21*, e49891.

Chen, M., Bell, J.M., Shi, X., Sun, S.Y., Wang, Z., and Ludtke, S.J. (2019). A complete data processing workflow for cryo-ET and subtomogram averaging. Nat. Methods *16*, 1161–1168.

Chen, C., Hou, J., Tanner, J.J., and Cheng, J. (2020). Bioinformatics Methods for Mass Spectrometry-Based Proteomics Data Analysis. Int. J. Mol. Sci. *21*, 2873.

Chorev, D.S., Baker, L.A., Wu, D., Beilsten-Edmands, V., Rouse, S.L., Zeev-Ben-Mordehai, T., Jiko, C., Samsudin, F., Gerle, C., Khalid, S., et al. (2018). Protein assemblies ejected directly from native membranes yield complexes for mass spectrometry. Science *362*, 829–834.

Chowdhury, S., Happonen, L., Khakzad, H., Malmström, L., and Malmström, J. (2020). Structural proteomics, electron cryo-microscopy and structural modeling approaches in bacteria-human protein interactions. Med. Microbiol. Immunol. (Berl.) *209*, 265–275.

Danev, R., and Baumeister, W. (2016). Cryo-EM single particle analysis with the Volta phase plate. eLife *5*, e13046.

Dijkman, P.M., Marzluf, T., Zhang, Y., Chang, S.-Y.S., Helm, D., Lanzer, M., Bujard, H., and Kudryashev, M. (2021). Structure of the merozoite surface protein 1 from Plasmodium falciparum. Science Advances. https://doi.org/10.1126/sciadv.abg0465.

Ellenberg, J., Swedlow, J.R., Barlow, M., Cook, C.E., Sarkans, U., Patwardhan, A., Brazma, A., and Birney, E. (2018). A call for public archives for biological image data. Nat. Methods *15*, 849–854.

Erdmann, P.S., Hou, Z., Klumpe, S., Khavnekar, S., Beck, F., Wilfling, F., Plitzko, J.M., and Baumeister, W. (2021). In situ cryo-electron tomography reveals gradient organization of ribosome biogenesis in intact nucleoli. Nat. Commun. *12*, 5364.

Esser, T.K., Boehning, J., Fremdling, P., Agasid, M.T., Costin, A., Fort, K., Konijnenberg, A., Bahm, A., Makarov, A., Robinson, C.V., et al. (2021). Mass-selective and ice-free cryo-EM protein sample preparation via native electrospray ion-beam deposition. bioRxiv. https://doi.org/10.1101/2021.10.18.464782.

Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J., et al. (2021). Protein complex prediction with AlphaFold-Multimer. bioRxiv. https://doi.org/10.1101/2021.10.04.463034.

Fagerlund, R.D.R.D., Wilkinson, M.E.M.E., Klykov, O., Barendregt, A., Pearce, F.G.G., Kieper, S.N.S.N., Maxwell, H.W.R., Capolupo, A., Heck, A.J.R., Krause, K.L.K.L., et al. (2017). Spacer capture and integration by a type I-F Cas1-Cas2-3 CRISPR adaptation complex. Proc. Natl. Acad. Sci. USA *114*, E5122–E5128.

Fasci, D., van Ingen, H., Scheltema, R.A., and Heck, A.J.R. (2018). Histone interaction landscapes visualized by crosslinking mass spectrometry in intact cell nuclei. Mol. Cell. Proteomics *17*, mcp.RA118.000924-mcp.RA118.000924.

Foster, L.J., de Hoog, C.L., Zhang, Y., Zhang, Y., Xie, X., Mootha, V.K., and Mann, M. (2006). A mammalian organelle map by protein correlation profiling. Cell *125*, 187–199.

Gault, J., Liko, I., Landreh, M., Shutin, D., Bolla, J.R., Jefferies, D., Agasid, M., Yen, H.-Y., Ladds, M.J.G.W., Lane, D.P., et al. (2020). Combining native and 'omics' mass spectrometry to identify endogenous ligands bound to membrane proteins. Nat. Methods *17*, 505–508.

Giese, S.H., Sinn, L.R., Wegner, F., and Rappsilber, J. (2021). Retention time prediction using neural networks increases identifications in crosslinking mass spectrometry. Nat. Commun. *12*, 3237.

Gingras, A.-C., Gstaiger, M., Raught, B., and Aebersold, R. (2007). Analysis of protein complexes using mass spectrometry. Nat. Rev. Mol. Cell Biol. *8*, 645–654.

Giurgiu, M., Reinhard, J., Brauner, B., Dunger-Kaltenbach, I., Fobo, G., Frishman, G., Montrone, C., and Ruepp, A. (2019). CORUM: the comprehensive resource of mammalian protein complexes-2019. Nucleic Acids Res. *47* (D1), D559–D563.

Green, M.N., Gangwar, S.P., Michard, E., Simon, A.A., Portes, M.T., Barbosa-Caro, J., Wudick, M.M., Lizzio, M.A., Klykov, O., Yelshanskaya, M.V., et al.

(2021). Structure of the Arabidopsis thaliana glutamate receptor-like channel GLR3.4. Mol. Cell *81*, 3216–3226.e8.

Griffiths, R.L., Konijnenberg, A., Viner, R., and Cooper, H.J. (2019). Direct Mass Spectrometry Analysis of Protein Complexes and Intact Proteins up to >70 kDa from Tissue. Anal. Chem. *91*, 6962–6966.

Guo, Q., Lehmer, C., Martínez-Sánchez, A., Rudack, T., Beck, F., Hartmann, H., Pérez-Berlanga, M., Frottin, F., Hipp, M.S., Hartl, F.U., et al. (2018). In Situ Structure of Neuronal C9orf72 Poly-GA Aggregates Reveals Proteasome Recruitment. Cell *172*, 696–705.e12.

Gupta, T.K., Klumpe, S., Gries, K., Heinz, S., Wietrzynski, W., Ohnishi, N., Niemeyer, J., Spaniol, B., Schaffer, M., Rast, A., et al. (2021). Structural basis for VIPP1 oligomerization and maintenance of thylakoid membrane integrity. Cell *184*, 3643–3659.e23.

Gwosch, K.C., Pape, J.K., Balzarotti, F., Hoess, P., Ellenberg, J., Ries, J., and Hell, S.W. (2020). MINFLUX nanoscopy delivers 3D multicolor nanometer resolution in cells. Nat. Methods *17*, 217–224.

Harapin, J., Börmel, M., Sapra, K.T., Brunner, D., Kaech, A., and Medalia, O. (2015). Structural analysis of multicellular organisms with cryo-electron tomography. Nat. Methods *12*, 634–636.

Hauri, S., Khakzad, H., Happonen, L., Teleman, J., Malmström, J., and Malmström, L. (2019). Rapid determination of quaternary protein structures in complex biological samples. Nat. Commun. *10*, 192, 192.

Havugimana, P.C., Hart, G.T., Nepusz, T., Yang, H., Turinsky, A.L., Li, Z., Wang, P.I., Boutz, D.R., Fong, V., Phanse, S., et al. (2012). A census of human soluble protein complexes. Cell *150*, 1068–1081.

Hein, M.Y., Hubner, N.C., Poser, I., Cox, J., Nagaraj, N., Toyoda, Y., Gak, I.A., Weisswange, I., Mansfeld, J., Buchholz, F., et al. (2015). A human interactome in three quantitative dimensions organized by stoichiometries and abundances. Cell *163*, 712–723.

Hevler, J.F., Zenezeni Chiozzi, R., Cabrera-Orefice, A., Brandt, U., Arnold, S., and Heck, A.J.R. (2021). Molecular characterization of a complex of apoptosis-inducing factor 1 with cytochrome c oxidase of the mitochondrial respiratory chain. Proc. Natl. Acad. Sci. USA *118*, e2106950118.

Himes, B.A., and Zhang, P. (2018). emClarity: software for high-resolution cryo-electron tomography and subtomogram averaging. Nat. Methods *15*, 955–961.

Ho, C.-M., Li, X., Lai, M., Terwilliger, T.C., Beck, J.R., Wohlschlegel, J., Goldberg, D.E., Fitzpatrick, A.W.P., and Zhou, Z.H. (2020). Bottom-up structural proteomics: cryoEM of protein complexes enriched from the cellular milieu. Nat. Methods *17*, 79–85.

Ho, C.-M., Jih, J., Lai, M., Li, X., Goldberg, D.E., Beck, J.R., and Zhou, Z.H. (2021). Native structure of the RhopH complex, a key determinant of malaria parasite nutrient acquisition. Proc. Natl. Acad. Sci. USA *118*, e2100514118.

Hosp, F., Scheltema, R.A., Eberl, H.C., Kulak, N.A., Keilhauer, E.C., Mayr, K., and Mann, M. (2015). A Double-Barrel Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) System to Quantify 96 Interactomes per Day. Mol. Cell. Proteomics *14*, 2030–2041.

Huang, W., Masureel, M., Qu, Q., Janetzko, J., Inoue, A., Kato, H.E., Robertson, M.J., Nguyen, K.C., Glenn, J.S., Skiniotis, G., and Kobilka, B.K. (2020). Structure of the neurotensin receptor 1 in complex with β-arrestin 1. Nature *579*, 303–308.

Humphreys, I.R., Pei, J., Baek, M., Krishnakumar, A., Anishchenko, I., Ovchinnikov, S., Zhang, J., Ness, T.J., Banjade, S., Bagde, S.R., et al. (2021). Computed structures of core eukaryotic protein complexes. Science *374*, m4805.

Hurdiss, D.L., Drulyte, I., Lang, Y., Shamorkina, T.M., Pronker, M.F., van Kuppeveld, F.J.M., Snijder, J., and de Groot, R.J. (2020). Cryo-EM structure of coronavirus-HKU1 haemagglutinin esterase reveals architectural changes arising from prolonged circulation in humans. Nat. Commun. *11*, 4646.

Huttlin, E.L., Bruckner, R.J., Navarrete-Perea, J., Cannon, J.R., Baltier, K., Gebreab, F., Gygi, M.P., Thornock, A., Zarraga, G., Tam, S., et al. (2021). Dual proteome-scale networks reveal cell-specific remodeling of the human interactome. Cell *184*, 3022–3040.e28.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589.

Kastritis, P.L., O'Reilly, F.J., Bock, T., Li, Y., Rogon, M.Z., Buczak, K., Romanov, N., Betts, M.J., Bui, K.H., Hagen, W.J., et al. (2017). Capturing protein communities by structural proteomics in a thermophilic eukaryote. Mol. Syst. Biol. *13*, 936, 936.

Ke, Z., Oton, J., Qu, K., Cortese, M., Zila, V., McKeane, L., Nakane, T., Zivanov, J., Neufeldt, C.J., Cerikan, B., et al. (2020). Structures and distributions of SARS-CoV-2 spike proteins on intact virions. Nature *588*, 498–502.

Keener, J.E., Zhang, G., and Marty, M.T. (2021). Native Mass Spectrometry of Membrane Proteins. Anal. Chem. *93*, 583–597.

Kelley, K., Raczkowski, A.M., Klykov, O., Jaroenlak, P., Bobe, D., Kopylov, M., Eng, E.T., Bhabha, G., Potter, C.S., Carragher, B., and Noble, A.J. (2020). Waffle Method: A general and flexible approach for FIB-milling small and anisotropically oriented samples. bioRxiv. https://doi.org/10.1101/2020.10.28.359372.

Kenner, L.R., Anand, A.A., Nguyen, H.C., Myasnikov, A.G., Klose, C.J., McGeever, L.A., Tsai, J.C., Miller-Vedam, L.E., Walter, P., and Frost, A. (2019). eIF2B-catalyzed nucleotide exchange and phosphoregulation by the integrated stress response. Science *364*, 491–495.

Kim, S.J., Fernandez-Martinez, J., Nudelman, I., Shi, Y., Zhang, W., Raveh, B., Herricks, T., Slaughter, B.D., Hogan, J.A., Upla, P., et al. (2018). Integrative structure and functional anatomy of a nuclear pore complex. Nature *555*, 475–482.

Kimanius, D., Dong, L., Sharov, G., Nakane, T., and Scheres, S.H.W. (2021). New tools for automated cryo-EM single-particle analysis in RELION-4.0. Biochem. J. *478*, 4169–4185.

Klumpe, S., Fung, H.K.H., Goetz, S.K., Zagoriy, I., Hampoelz, B., Zhang, X., Erdmann, P.S., Baumbach, J., Müller, C.W., Beck, M., et al. (2021). A Modular Platform for Streamlining Automated Cryo-FIB Workflows. eLife. https://doi.org/10.7554/eLife.70506.

Klykov, O., Steigenberger, B., Pektaş, S., Fasci, D., Heck, A.J.R., and Scheltema, R.A. (2018). Efficient and robust proteome-wide approaches for cross-linking mass spectrometry. Nat. Protoc. *13*, 2964–2990.

Klykov, O., van der Zwaan, C., Heck, A.J.R., Meijer, A.B., and Scheltema, R.A. (2020). Missing regions within the molecular architecture of human fibrin clots structurally resolved by XL-MS and integrative structural modeling. Proc. Natl. Acad. Sci. USA *117*, 1976–1987.

Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P., et al. (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature *440*, 637–643.

Lawson, C.L., Baker, M.L., Best, C., Bi, C., Dougherty, M., Feng, P., van Ginkel, G., Devkota, B., Lagerstedt, I., Ludtke, S.J., et al. (2011). EMDataBank.org: unified data resource for CryoEM. Nucleic Acids Res. *39*, D456–D464.

Lee, B.-G., Merkel, F., Allegretti, M., Hassler, M., Cawood, C., Lecomte, L., O'Reilly, F.J., Sinn, L.R., Gutierrez-Escribano, P., Kschonsak, M., et al. (2020). Cryo-EM structures of holo condensin reveal a subunit flip-flop mechanism. Nat. Struct. Mol. Biol. *27*, 743–751.

Leigh, K.E., Navarro, P.P., Scaramuzza, S., Chen, W., Zhang, Y., Castaño-Díez, D., and Kudryashev, M. (2019). Subtomogram averaging from cryo-electron tomograms. In Methods in Cell Biology (Academic Press Inc.), pp. 217–259.

Leitner, A., Walzthoeni, T., Kahraman, A., Herzog, F., Rinner, O., Beck, M., and Aebersold, R. (2010). Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. Mol. Cell. Proteomics *9*, 1634–1649.

Leitner, A., Joachimiak, L.A., Unverdorben, P., Walzthoeni, T., Frydman, J., Förster, F., and Aebersold, R. (2014). Chemical cross-linking/mass spectrometry targeting acidic residues in proteins and protein complexes. Proc. Natl. Acad. Sci. USA *111*, 9455–9460.

Lenz, S., Sinn, L.R., O'Reilly, F.J., Fischer, L., Wegner, F., and Rappsilber, J. (2021). Reliable identification of protein-protein interactions by crosslinking mass spectrometry. Nat. Commun. *12*, 3564.

Leung, M.R., Chiozzi, R.Z., Roelofs, M.C., Hevler, J.F., Ravi, R.T., Maitan, P., Zhang, M., Henning, H., Bromfield, E.G., Howes, S.C., et al. (2021). In-cell structures of a conserved supramolecular array at the mitochondria-cytoskeleton interface in mammalian sperm. bioRxiv. https://doi.org/10.1101/2021.02.16.431372.

Liu, F., Lössl, P., Rabbitts, B.M., Balaban, R.S., and Heck, A.J.R. (2017). The interactome of intact mitochondria by cross-linking mass spectrometry provides evidence for co-existing respiratory supercomplexes. Mol. Cell. Proteomics *16*, RA117.000470-RA117.000470.

Liu, X., Salokas, K., Tamene, F., Jiu, Y., Weldatsadik, R.G., Öhman, T., and Varjosalo, M. (2018). An AP-MS- and BioID-compatible MAC-tag enables comprehensive mapping of protein interactions and subcellular localizations. Nat. Commun. *9*, 1188.

Liu, Y.-T., Zhang, H., Wang, H., Tao, C.-L., Bi, G.-Q., and Zhou, Z.H. (2021). Isotropic Reconstruction of Electron Tomograms with Deep Learning. biorXiv. https://doi.org/10.1101/2021.07.17.452128.

Lössl, P., van de Waterbeemd, M., and Heck, A.J.R. (2016). The diverse and expanding role of mass spectrometry in structural and molecular biology. EMBO J. *35*, 2634–2657.

Mahamid, J., Pfeffer, S., Schaffer, M., Villa, E., Danev, R., Cuellar, L.K., Förster, F., Hyman, A.A., Plitzko, J.M., and Baumeister, W. (2016). Visualizing the molecular sociology at the HeLa cell nuclear periphery. Science *351*, 969–972.

Martinez-Sanchez, A., Kochovski, Z., Laugks, U., Meyer zum Alten Borgloh, J., Chakraborty, S., Pfeffer, S., Baumeister, W., and Lučić, V. (2020). Template-free detection and classification of membrane-bound complexes in cryo-electron tomograms. Nat. Methods *17*, 209–216.

Mellacheruvu, D., Wright, Z., Couzens, A.L., Lambert, J.-P., St-Denis, N.A., Li, T., Miteva, Y.V., Hauri, S., Sardiu, M.E., Low, T.Y., et al. (2013). The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. Nat. Methods *10*, 730–736.

Meyer, N.L., Hu, G., Davulcu, O., Xie, Q., Noble, A.J., Yoshioka, C., Gingerich, D.S., Trzynka, A., David, L., Stagg, S.M., and Chapman, M.S. (2019). Structure of the gene therapy vector, adeno-associated virus with its cell receptor, AAVR. eLife *8*, e44707, e44707.

Michaelis, A.C., Brunner, A.-D., Zwiebel, M., Meier, F., Strauss, M.T., Bludau, I., and Mann, M. (2021). The social architecture of an in-depth cellular protein interactome. bioRxiv. https://doi.org/10.1101/2021.10.24.465633.

Mikhailov, V.A., Mize, T.H., Benesch, J.L.P., and Robinson, C.V. (2014). Mass-selective soft-landing of protein assemblies with controlled landing energies. Anal. Chem. *86*, 8321–8328.

Miller, L.M., Bond, K.M., Draper, B.E., and Jarrold, M.F. (2021). Characterization of Classical Vaccines by Charge Detection Mass Spectrometry. Anal. Chem. *93*, 11965–11972.

Mintseris, J., and Gygi, S.P. (2020). High-density chemical cross-linking for modeling protein interactions. Proc. Natl. Acad. Sci. USA *117*, 93–102.

Mosalaganti, S., Obarska-Kosinska, A., Siggel, M., Turonova, B., Zimmerli, C.E., Buczak, K., Schmidt, F.H., Margiotta, E., Mackmull, M.-T., Hagen, W., et al. (2021). Artificial intelligence reveals nuclear pore complexity. bioRxiv. https://doi.org/10.1101/2021.10.26.465776.

Mund, A., Coscia, F., Hollandi, R., Kovács, F., Kriston, A., Brunner, A.-D., Bzorek, M., Naimy, S., Rahbek Gjerdrum, L.M., Dyring-Andersen, B., et al. (2021). AI-driven Deep Visual Proteomics defines cell identity and heterogeneity. bioRxiv. https://doi.org/10.1101/2021.01.25.427969.

Nagaraj, N., Alexander Kulak, N., Cox, J., Neuhauser, N., Mayr, K., Hoerning, O., Vorm, O., and Mann, M. (2012). System-wide Perturbation Analysis with Nearly Complete Coverage of the Yeast Proteome by Single-shot Ultra HPLC Runs on a Bench Top Orbitrap. Mol. Cell. Proteomics *11*, M111.013722.

Nakane, T., Kotecha, A., Sente, A., McMullan, G., Masiulis, S., Brown, P.M.G.E., Grigoras, I.T., Malinauskaite, L., Malinauskas, T., Miehling, J., et al. (2020). Single-particle cryo-EM at atomic resolution. Nature *587*, 152–156.

Nickell, S., Kofler, C., Leis, A.P., and Baumeister, W. (2006). A visual approach to proteomics. Nat. Rev. Mol. Cell Biol. *7*, 225–230.

O'Reilly, F.J., and Rappsilber, J. (2018). Cross-linking mass spectrometry: methods and applications in structural, molecular and systems biology. Nat. Struct. Mol. Biol. *25*, 1000–1008.

O'Reilly, F.J., Xue, L., Graziadei, A., Sinn, L., Lenz, S., Tegunov, D., Blötz, C., Hagen, W.J.H., Cramer, P., Stülke, J., et al. (2020). In-cell architecture of an actively transcribing-translating expressome. Science *369*, 2020.02.28.970111-2020.02.28.970111.

Olinares, P.D.B., Kang, J.Y., Llewellyn, E., Chiu, C., Chen, J., Malone, B., Saecker, R.M., Campbell, E.A., Darst, S.A., and Chait, B.T. (2021). Native Mass Spectrometry-Based Screening for Optimal Sample Preparation in Single-Particle Cryo-EM. Structure *29*, 186–195.e6.

Pan, C., Kumar, C., Bohl, S., Klingmueller, U., and Mann, M. (2009). Comparative proteomic phenotyping of cell lines and primary cells to assess preservation of cell type-specific functions. Mol. Cell. Proteomics *8*, 443–450.

Peukes, J., Lovatt, M., Leistner, C., Boulanger, J., Morado, D.R., Kukulski, W., Zhu, F., Komiyama, N., Briggs, J.A.G., Grant, S.G.N., et al. (2021). The molecular infrastructure of glutamatergic synapses in the mammalian forebrain. bioRxiv. https://doi.org/10.1101/2021.02.19.432002.

Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290–296.

Pyle, E., and Zanetti, G. (2021). Current data processing strategies for cryo-electron tomography and subtomogram averaging. Biochem. J. *478*, 1827–1845.

Rantos, V., Karius, K., and Kosinski, J. (2021). Integrative structural modeling of macromolecular complexes using Assembline. Nat. Protoc. Published online November 29, 2021. 10.1038/s41596-021-00640-z.

Rappsilber, J. (2011). The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. J. Struct. Biol. *173*, 530–540.

Rickard, W.D.A., Reddy, S.M., Saxey, D.W., Fougerouse, D., Timms, N.E., Daly, L., Peterman, E., Cavosie, A.J., and Jourdan, F. (2020). Novel Applications of FIB-SEM-Based ToF-SIMS in Atom Probe Tomography Workflows. Microsc. Microanal. *26*, 750–757.

Rickgauer, J.P., Grigorieff, N., and Denk, W. (2017). Single-protein detection in crowded molecular environments in cryo-EM images. eLife *6*, e25648.

Röst, H.L., Malmström, L., and Aebersold, R. (2015). Reproducible quantitative proteotype data matrices for systems biology. Mol. Biol. Cell *26*, 3926–3931.

Russel, D., Lasker, K., Webb, B., Velázquez-Muriel, J., Tjioe, E., Schneidman-Duhovny, D., Peterson, B., and Sali, A. (2012). Putting the pieces together: integrative modeling platform software for structure determination of macromolecular assemblies. PLoS Biol. *10*, e1001244, e1001244.

Saibil, H.R. (2022). Cryo-EM in molecular and cellular biology. Mol. Cell, 274–284.

Schaffer, M., Pfeffer, S., Mahamid, J., Kleindiek, S., Laugks, T., Albert, S., Engel, B.D., Rummel, A., Smith, A.J., Baumeister, W., and Plitzko, J.M. (2019). A cryo-FIB lift-out technique enables molecular-resolution cryo-ET within native Caenorhabditis elegans tissue. Nat. Methods *16*, 757–762.

Scheres, S.H.W. (2012). A Bayesian view on cryo-EM structure determination. J. Mol. Biol. *415*, 406–418.

Scheres, S.H.W. (2020). Amyloid structure determination in RELION-3.1. Acta Crystallogr. D Struct. Biol. *76*, 94–101.

Schmidt, C., and Urlaub, H. (2017). Combining cryo-electron microscopy (cryo-EM) and cross-linking mass spectrometry (CX-MS) for structural elucidation of large protein assemblies. Curr. Opin. Struct. Biol. *46*, 157–168.

Schur, F.K.M., Obr, M., Hagen, W.J.H., Wan, W., Jakobi, A.J., Kirkpatrick, J.M., Sachse, C., Kräusslich, H.-G., and Briggs, J.A.G. (2016). An atomic model of HIV-1 capsid-SP1 reveals structures regulating assembly and maturation. Science *353*, 506–508.

Schwartz, O., Axelrod, J.J., Campbell, S.L., Turnbaugh, C., Glaeser, R.M., and Müller, H. (2019). Laser phase plate for transmission electron microscopy. Nat. Methods *16*, 1016–1020.

Schweppe, D.K., Zheng, C., Chavez, J.D., Navare, A.T., Wu, X., Eng, J.K., and Bruce, J.E. (2016). XLinkDB 2.0: integrated, large-scale structural analysis of protein crosslinking data. Bioinformatics *32*, 2716–2718.

Schweppe, D.K., Chavez, J.D., Lee, C.F., Caudal, A., Kruse, S.E., Stuppard, R., Marcinek, D.J., Shadel, G.S., Tian, R., and Bruce, J.E. (2017). Mitochondrial protein interactome elucidated by chemical cross-linking mass spectrometry. Proc. Natl. Acad. Sci. USA *114*, 1732–1737.

Shakeel, S., Rajendra, E., Alcón, P., O'Reilly, F., Chorev, D.S., Maslen, S., Degliesposti, G., Russo, C.J., He, S., Hill, C.H., et al. (2019). Structure of the Fanconi anaemia monoubiquitin ligase complex. Nature *575*, 234–237.

Sigworth, F.J. (2016). Principles of cryo-EM single-particle image processing. Microscopy (Oxf.) *65*, 57–67.

Silvester, E., Vollmer, B., Pražák, V., Vasishtan, D., Machala, E.A., Whittle, C., Black, S., Bath, J., Turberfield, A.J., Grünewald, K., and Baker, L.A. (2021). DNA origami signposts for identifying proteins on cell membranes by electron cryotomography. Cell *184*, 1110–1121.e16.

Slavin, M., Zamel, J., Zohar, K., Eliyahu, T., Braitbard, M., Brielle, E., Baraz, L., Stolovich-Rain, M., Friedman, A., Wolf, D.G., et al. (2021). Targeted in situ cross-linking mass spectrometry and integrative modeling reveal the architectures of three proteins from SARS-CoV-2. Proc. Natl. Acad. Sci. USA *118*, e2103554118.

Snijder, J., Schuller, J.M., Wiegard, A., Lössl, P., Schmelling, N., Axmann, I.M., Plitzko, J.M., Förster, F., and Heck, A.J.R. (2017). Structures of the cyanobacterial circadian oscillator frozen in a fully assembled state. Science *355*, 1181–1184.

Specht, H., Emmott, E., Petelski, A.A., Huffman, R.G., Perlman, D.H., Serra, M., Kharchenko, P., Koller, A., and Slavov, N. (2021). Single-cell proteomic and transcriptomic analysis of macrophage heterogeneity using SCoPE2. Genome Biol. *22*, 50.

Stark, H. (2010). GraFix: stabilization of fragile macromolecular complexes for single particle cryo-EM. Methods Enzymol. *481*, 109–126.

Steigenberger, B., Pieters, R.J., Heck, A.J.R., and Scheltema, R.A. (2019). PhoX: An IMAC-Enrichable Cross-Linking Reagent. ACS Cent. Sci. *5*, 1514–1522.

Steigenberger, B., van den Toorn, H.W.P., Bijl, E., Greisch, J.-F., Räther, O., Lubeck, M., Pieters, R.J., Heck, A.J.R., and Scheltema, R.A. (2020). Benefits of Collisional Cross Section Assisted Precursor Selection (caps-PASEF) for Cross-linking Mass Spectrometry. Mol. Cell. Proteomics *19*, 1677–1687.

Stieger, C.E., Doppler, P., and Mechtler, K. (2019). Optimized Fragmentation Improves the Identification of Peptides Cross-Linked by MS-Cleavable Reagents. J. Proteome Res. *18*, 1363–1370.

Su, C.C., Lyu, M., Morgan, C.E., Bolla, J.R., Robinson, C.V., and Yu, E.W. (2021). A 'Build and Retrieve' methodology to simultaneously solve cryo-EM structures of membrane proteins. Nat. Methods *18*, 69–75.

Sun, N., Beck, F., Knispel, R.W., Siedler, F., Scheffer, B., Nickell, S., Baumeister, W., and Nagy, I. (2007). Proteomics analysis of Thermoplasma acidophilum with a focus on protein complexes. Mol. Cell. Proteomics *6*, 492–502.

Tacke, S., Erdmann, P., Wang, Z., Klumpe, S., Grange, M., Plitzko, J., and Raunser, S. (2021). A streamlined workflow for automated cryo focused ion beam milling. J. Struct. Biol. *213*, 107743.

Tamara, S., den Boer, M.A., and Heck, A.J.R. (2021). High-Resolution Native Mass Spectrometry. Chem. Rev. Published online August 20, 2021. https://doi.org/10.1021/acs.chemrev.1c00212.

Tan, Z.Y., Cai, S., Noble, A.J., Chen, J.K., Shi, J., and Gan, L. (2021). Heterogeneous non-canonical nucleosomes predominate in yeast cells in situ. bioRxiv. Published online April 4, 2021. https://doi.org/10.1101/2021.04.04.438362.

Taschner, M., Basquin, J., Steigenberger, B., Schäfer, I.B., Soh, Y.-M., Basquin, C., Lorentzen, E., Räschle, M., Scheltema, R.A., and Gruber, S. (2021). Nse5/6 inhibits the Smc5/6 ATPase and modulates DNA substrate binding. EMBO J. *40*, e107807.

Tegunov, D., and Cramer, P. (2019). Real-time cryo-electron microscopy data preprocessing with Warp. Nat. Methods *16*, 1146–1152.

Tegunov, D., Xue, L., Dienemann, C., Cramer, P., and Mahamid, J. (2021). Multi-particle cryo-EM refinement with M visualizes ribosome-antibiotic complex at 3.5 Å in cells. Nat. Methods *18*, 186–193.

Turk, M., and Baumeister, W. (2020). The promise and the challenges of cryo-electron tomography. FEBS Lett. *594*, 3243–3261.

van Zundert, G.C.P., Rodrigues, J.P.G.L.M., Trellet, M., Schmitz, C., Kastritis, P.L., Karaca, E., Melquiond, A.S.J., van Dijk, M., de Vries, S.J., and Bonvin, A.M.J.J. (2016). The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. J. Mol. Biol. *428*, 720–725.

VanAernum, Z.L., Busch, F., Jones, B.J., Jia, M., Chen, Z., Boyken, S.E., Sahasrabuddhe, A., Baker, D., and Wysocki, V.H. (2020). Rapid online buffer exchange for screening of proteins, protein complexes and cell lysates by native mass spectrometry. Nat. Protoc. *15*, 1132–1157, Rapid online.

Verbeke, E.J., Mallam, A.L., Drew, K., Marcotte, E.M., and Taylor, D.W. (2018). Classification of Single Particles from Human Cell Extract Reveals Distinct Structures. Cell Rep. *24*, 259–268.e3.

Villegas, L., Pero-Gascon, R., Benavente, F., Barbosa, J., and Sanz-Nebot, V. (2019). On-line protein digestion by immobilized enzyme microreactor capillary electrophoresis-mass spectrometry. Talanta *199*, 116–123.

Vimer, S., Ben-Nissan, G., and Sharon, M. (2020). Mass Spectrometry Analysis of Intact Proteins from Crude Samples. Anal. Chem. *92*, 12741–12749.

Watanabe, R., Buschauer, R., Böhning, J., Audagnotto, M., Lasker, K., Lu, T.W., Boassa, D., Taylor, S., and Villa, E. (2020). The in situ structure of Parkinson's disease-linked LRRK2. Cell *182*, 1508–1518.e16.

Westermarck, J., Ivaska, J., and Corthals, G.L. (2013). Identification of protein interactions involved in cellular signaling. Mol. Cell. Proteomics *12*, 1752–1763.

Westphall, M.S., Lee, K.W., Salome, A.Z., Lodge, J., Grant, T., and Coon, J.J. (2021). 3D Structure Determination of Protein Complexes using Matrix-Landing Mass Spectrometry. bioRxiv. https://doi.org/10.1101/2021.10.13.464253.

Wittig, I., and Malacarne, P.F. (2021). Complexome Profiling: Assembly and Remodeling of Protein Complexes. Int. J. Mol. Sci. *22*, 7809.

Wu, M., and Lander, G.C. (2020). Present and Emerging Methodologies in Cryo-EM Single-Particle Analysis. Biophys. J. *119*, 1281–1289.

Yan, K., Yang, J., Zhang, Z., McLaughlin, S.H., Chang, L., Fasci, D., Ehrenhofer-Murray, A.E., Heck, A.J.R., and Barford, D. (2019). Structure of the inner kinetochore CCAN complex assembled onto a centromeric nucleosome. Nature *574*, 278–282.

Yang, Y., Song, H., He, D., Zhang, S., Dai, S., Xie, X., Lin, S., Hao, Z., Zheng, H., and Chen, P.R. (2017). Genetically encoded releasable photo-cross-linking strategies for studying protein-protein interactions in living cells. Nat. Protoc. *12*, 2147–2168.

Yip, K.M., Fischer, N., Paknia, E., Chari, A., and Stark, H. (2020). Atomic-resolution protein structure determination by cryo-EM. Nature *587*, 157–161.

Yu, C., and Huang, L. (2017). Cross-Linking Mass Spectrometry: An Emerging Technology for Interactomics and Structural Biology. Anal. Chem. *90*, 144–165.

Yu, J., Rao, P., Clark, S., Mitra, J., Ha, T., and Gouaux, E. (2021a). Hippocampal AMPA receptor assemblies and mechanism of allosteric inhibition. Nature *594*, 448–453.

Yu, Y., Li, S., Ser, Z., Sanyal, T., Choi, K., Wan, B., Kuang, H., Sali, A., Kentsis, A., Patel, D.J., and Zhao, X. (2021b). Integrative analysis reveals unique structural and functional features of the Smc5/6 complex. Proc. Natl. Acad. Sci. USA *118*, e2026844118.

Zhong, X., Navare, A.T., Chavez, J.D., Eng, J.K., Schweppe, D.K., and Bruce, J.E. (2017). Large-Scale and Targeted Quantitative Cross-Linking MS Using Isotope-Labeled Protein Interaction Reporter (PIR) Cross-Linkers. J. Proteome Res. *16*, 720–727.