CHEMICAL REVIEWS

pubs.acs.org/CR Review

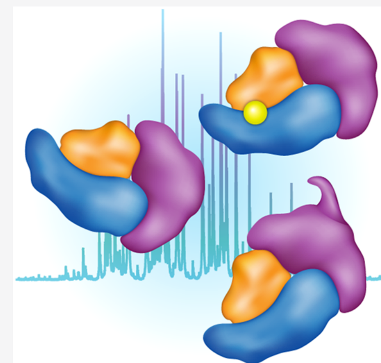
Characterizing Endogenous Protein Complexes with Biological Mass Spectrometry

Rivkah Rogawski and Michal Sharon*

Cite This: *Chem. Rev.* 2022, 122, 7386−7414 Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Biological mass spectrometry (MS) encompasses a range of methods for characterizing proteins and other biomolecules. MS is uniquely powerful for the structural analysis of endogenous protein complexes, which are often heterogeneous, poorly abundant, and refractive to characterization by other methods. Here, we focus on how biological MS can contribute to the study of endogenous protein complexes, which we define as complexes expressed in the physiological host and purified intact, as opposed to reconstituted complexes assembled from heterologously expressed components. Biological MS can yield information on complex stoichiometry, heterogeneity, topology, stability, activity, modes of regulation, and even structural dynamics. We begin with a review of methods for isolating endogenous complexes. We then describe the various biological MS approaches, focusing on the type of information that each method yields. We end with future directions and challenges for these MS-based methods.



CONTENTS

1. Introduction 7386
2. Preparation of Endogenous Protein Complexes for Mass Spectrometry 7388
2.1. General Considerations for MS Protein Complex Sample Preparation 7388
2.2. Biochemical Purification 7389
2.3. Direct Immunoprecipitation 7389
2.4. Affinity Tagging 7389
2.4.1. Methods for Generating Affinity Tagged Endogenous Complexes 7390
2.5. Cross-Linking to Preserve Noncovalent Interactions 7391
3. MS-Based Techniques for Studying Endogenous Complexes 7391
3.1. Biological MS Delivers Insights Across Multiple Levels of Complex Regulation 7391
3.2. Sample Separation and Fractionation 7392
3.2.1. Electrophoresis-Based Methods 7392
3.2.2. Liquid Chromatography 7392
3.2.3. Capillary Zone Electrophoresis 7393
3.2.4. Ion Mobility 7393
3.3. Bottom-Up Approaches 7393
3.3.1. Proteomics 7393
3.3.2. Protein Footprinting 7396
3.4. "Top-Down" Approaches 7400
3.4.1. Top-Down MS for PTM Analysis 7400
3.4.2. Native MS of Protein Complexes 7401
3.5. Integrative Approaches and Computational Modeling 7403
4. Conclusion and Future Directions 7403
Author Information 7404
Corresponding Author 7404
Author 7404
Notes 7404
Biographies 7404
Acknowledgments 7404
References 7404

1. INTRODUCTION

Proteins encoded by cellular DNA are the workhorses of the cell, carrying out diverse biochemical tasks that generate cellular phenotypes.[1] However, many proteins do not act alone but associate with other proteins to form functional complexes, greatly increasing the complexity of molecular species found in the cell. Moreover, complex formation can depend on protein location and post-translational modifications (PTMs), further increasing the diversity of cellular species. A major challenge in biology is to map the cellular protein complexes and determine how complex composition varies as a function of cellular state. 22% of the protein coding genes in humans are represented in CORUM, a repository of experimentally studied complexes

Special Issue: Mass Spectrometry Applications in Structural Biology

Received: March 15, 2021
Published: August 18, 2021

© 2021 The Authors. Published by American Chemical Society

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

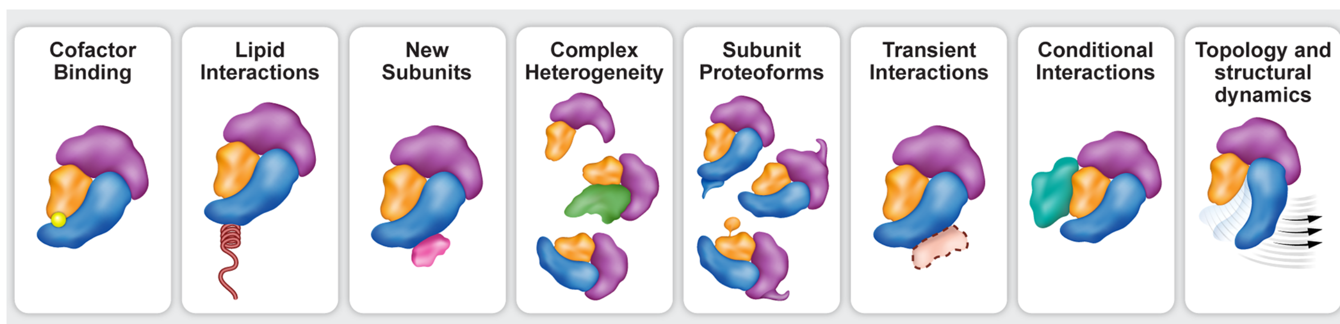


Figure 1. Biological MS yields a wide range of information about biological complexes. Depending on the specifics of sample preparation and the MS method chosen, biological MS can shed light on many properties of endogenous complexes. This includes determining binding of cofactors and lipids and their effects on complex composition, identification of new subunits, heterogeneity of complexes, post-translational modifications, and transient interactions or interactions that are only present under a distinct physiological condition. Structures and topologies, as well as structural dynamics, can also be revealed.

Overview of Different Biological MS Approaches

	Proteomics	HDX	Cross Linking	Intact/Top Down	Native MS
Sample Requirments (mol)	1 x 10^{-18} to 1x 10^{-15} for detection	100 x 10^{-12}	1 to 10 x 10^{-12} for purified; 5-10 mg raw sample for in situ crosslinking	0.5 x 10^{-12} (sensitivity decreases with MW)	10 x 10^{-12} (sensitivity decreases with MW)
Treatment	Enzymatic Digestion	H_2O → D_2O; D_2O Exchange + Enzymatic Digestion	Cross Linking + Enzymatic Digestion	Protein Separation	Complex Buffer Exchange
MS Analysis	LC-MS Peptide Analysis and Quantification	LC-MS Peptide Analysis and Quantification	LC-MS Peptide Analysis and Quantification	Mass of Protein Protein Fragmentation	Mass of Protein Dissociation of Non-covalent Interactions
Generated Information	Interactions Composition Abundances	Relative Flexibility Protein-protein Interfaces	Interactions 3D Distance Restraints	Composition Proteoform Identification	Composition Stoichiometry Ligand Binding Topology Shape

Figure 2. Overview of various biological MS techniques. Each biological MS method requires different amounts of protein (row 1: refs 161, 209, and 250 corroborate amounts for proteomic, in situ cross-linking, and HDX, respectively). For each method, the proteins must be treated in a different way (second row). While all include gas phase ion generation and MS analysis, the experimental setup differs slightly between samples, with bottom-up proteomic methods sequencing and quantifying peptides via LC/MS-MS and top-down and native MS measuring the mass of the whole protein and using tandem MS for protein fragmentation and dissociation of covalent and noncovalent interactions, respectively (third row). Each method yields different information, as detailed in section 3 and described in the last row.

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

from mammalian organisms,[2] and many more protein complexes likely exist.[3] The topology and structural dynamics of protein complexes should be determined in the endogenous state, as similar to that as in cells or tissues, for the findings to have maximal physiological relevance. However, because of their limited quantity and heterogeneity, the range of biochemical and structural techniques that can be applied to endogenous complexes is restricted.

Biological mass spectrometry (MS), a term that encompasses a wide range of MS-based techniques, has provided tremendous insights into the composition and structural features of endogenous protein complexes. MS-based techniques balance throughput with resolution, filling a gap between time-intensive, all-atom structural techniques such as NMR and X-ray crystallography and high-throughput molecular biology techniques that do not provide detailed structural information. Here we detail the different MS based methods, focusing on the insights they provide into endogenous complexes and comparing between them. Depending on the approach taken, biological MS can determine the interaction partners for a particular protein, describe complex topology and structure, and even provide insights into conformational fluctuations (Figure 1).

This review is intended as an overview of sample preparation techniques and biological MS approaches available for in-depth characterization of endogenous complexes. Rather than focusing on technical aspects of specific MS techniques, which are detailed in many excellent reviews cited here[4−6] and below, we focus on the compositional and structural insights that can be gained from biological MS and compare how different MS methods can yield this information. We also highlight recent exciting advances ranging from improvements in LC/MS technology to CRISPR/CAS9 technologies that enable epitope tagging of endogenous proteins, increasing relevance for a wider range of systems.

We begin by reviewing methods for isolating endogenous complexes, highlighting advantages and disadvantages of each. We define endogenous complexes as those composed of proteins expressed in the organism of origin, either under native promoters or under conditions as close to native promoters as possible. We then introduce the different biological MS methods for characterizing those complexes, pointing the reader to an updated technical review for each technology and detailing the advantages and disadvantages of each method. For each technology, we describe the biological information that it can provide, focusing on specific examples that highlight the power of biological MS and apologizing in advance for publications excluded due to lack of space. We end with a review of future directions and exciting developments that will open new frontiers in the study of protein complexes.

2. PREPARATION OF ENDOGENOUS PROTEIN COMPLEXES FOR MASS SPECTROMETRY

2.1. General Considerations for MS Protein Complex Sample Preparation

The information extracted from MS experiments is limited by the scale and quality of the input samples. Although many MS modalities can be applied to crude biological samples, as discussed below, typically assemblies must be extracted from cells or tissues and purified for MS analysis. It is crucial to remove as many contaminants, or copurifying proteins that are not part of the complex, as possible. In this section, we discuss general principles for the enrichment of protein complexes from different tissues and cell types. Once a suitable tissue and enrichment scheme has been selected, optimization of sample preparation protocol is a combination of art and science, and there is no substitute for screening many purification conditions. In fact, rapid screening platforms to test buffers and conditions in parallel, analogous to X-ray crystallography screens, have been used for MS analysis.[7]

Complex abundance will directly affect both the sample purification method as well as the choice of MS technique. Generally, bottom-up proteomics requires significantly less material than top-down MS methods (see section 3 and Figure 2). Therefore, while we focus primarily on purification strategies for proteins that are not overexpressed, we include a few strategies for ectopic overexpression with the caveat that overexpression might lead to nonphysiological complex formation. Moreover, for heteromeric complexes of unknown composition, not all subunits can be selected for overexpression. In some cases, overexpression of a single subunit may lead to proportionally altered levels of associated subunits such that subunit stoichiometry is preserved in the cell.[8]

Three main strategies exist for complex enrichment (Figure 3): biochemical purification, immunoprecipitation, and affinity tagging. All three strategies have yielded robust insight into biological complexes and their structural and compositional properties. For any strategy, it may be desirable to begin with subcellular fractionation to enable the retrieval of spatially defined complexes of interest from different cellular compartments such as the nuclei, cytosol, and mitochondria.[9,10] Because compound assembly state can depend on cofactor binding, it can be important to include cofactors during purification to prevent complex disassembly. For example, the 26S proteasome dissociates readily once removed from cells if ATP is not included in the buffers in the physiological range of 1 mM.[11] If the complex is known to have catalytic or enzymatic functions that can be reconstituted in vitro, it is worth developing in vitro assays that can be performed on fractions or purified complexes to validate that the complex is functional. For example, in our lab, 20S proteasome containing fractions are tested for degradation activity with fluorescent activity peptides.[12]

The desired final buffer and assembly state depend on the MS technique applied. For bottom-up proteomics analysis (see section 3.3.1), because proteins are digested and separated on a column, the complex does not need to be in an MS compatible buffer. In fact, the complex is typically denatured before digestion, and thus protein precipitation via organic solvents such as tricholoroacetic acid can be used to separate proteins from buffer components. After digestion, peptides are loaded onto a chromatography column that exchanges them into an MS compatible buffer. For chemical footprinting methods (see section 3.3.2), complexes should be in a native state for the foot-printing step but are then denatured and digested for MS analysis, relaxing complex purification requirements. However, buffers must be compatible with the chemistry of the label used. For example, for cross-linking MS with disuccinimidyl dibutyric urea, buffers containing primary amines should be avoided because they will react with the cross-linkers.[13] For native and top-down MS (see section 3.4), complexes should be eluted in an assembled, native state and exchanged into MS compatible buffers, as detailed in ref 14. In particular, high concentrations of salts can suppress signal intensity and require rounds of buffer exchange. A number of

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

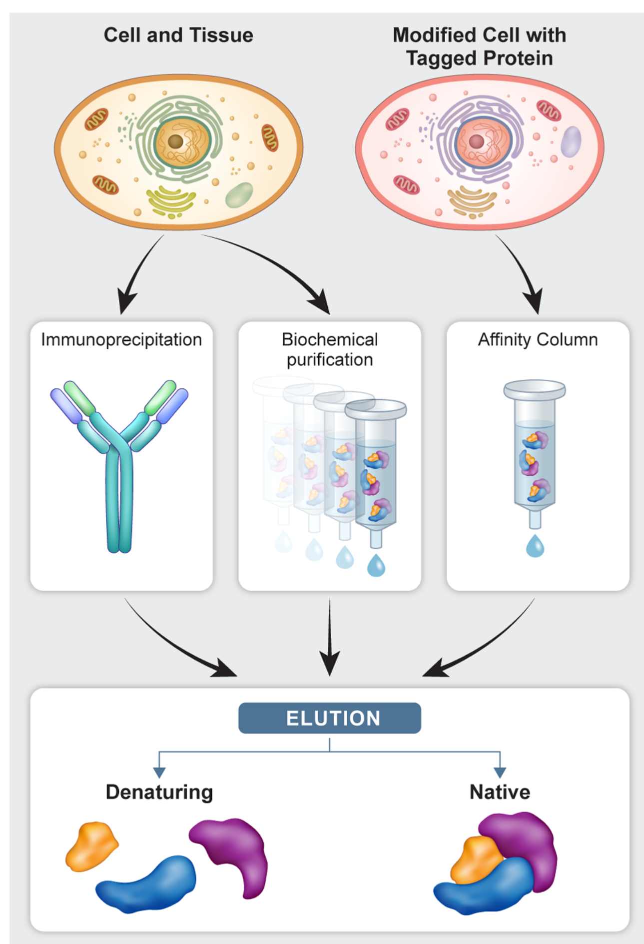


Figure 3. Strategies for protein complex isolation for biological MS. Purification of endogenous complexes can start either from unmodified cells and tissues, in which immunoprecipitation and biochemical purification can be applied, or by genetic manipulation of cells to express an epitope tagged protein, preferably at the endogenous locus (see section 2.4.1). At the end of the purification, complexes can either be eluted in a denatured state, in which case bottom-up proteomics or intact protein MS can be applied, or eluted in a native state, which will enable application of native-MS, HDX, or cross-linking MS.

salts that can maintain physiological ionic strength and pH while remaining compatible with MS, including ammonium acetate, ethylene diammonium diacetate buffer, and others.

2.2. Biochemical Purification

Biochemical purification strategies rely on differences in physicochemical properties between the complex of interest and its biological matrix. Starting from a homogenate of biological material, sequential centrifugation, precipitation, and/or chromatographic steps gradually yield solutions enriched in the complex of interest. Chromatographic purification steps can separate complexes based on their size (size exclusion chromatography), charge (ion exchange chromatography), or hydrophobicity (hydrophobic interaction chromatography). Protein complexes can also be enriched based on density via sucrose gradient centrifugation or via selective precipitation of some components, most popularly using ammonium sulfate. We refer the reader to guides for design of protein purification strategies for complete discussion of these techniques.[15]

The advantage of a biochemical purification is that it yields endogenous complexes that are unperturbed by the addition of non-native amino acid sequences or amplification of individual subunits. Complexes are also not bound to antibodies, which must be removed for downstream MS analysis. Disadvantages of biochemical purifications, and the reason that they have been primarily supplanted by affinity purification methodologies, is that they are lengthy, lossy, and must be optimized separately for each protein complex. The lengthiness can cause dissociation of transient interactions and the lossiness means that abundance of the complex must be quite high to provide enough starting material to survive the successive enrichment steps. Given these limitations, it is not surprising that biochemical purifications have been mainly applied to prepare stable and abundant complexes for MS analysis. These include the ribosome,[16] the proteasome,[12] the COP9 signalosome,[17] and varying complexes of vinculun/Arp[18] proteins.

Biochemical purification techniques are also used in cofractionation MS analysis of complexes, discussed below. In these experiments, complex biological samples are fractionated biochemically without tracking a particular component. Crude fractions are subjected to MS analysis, enabling identification of complexes either by proteomic colocalization or by directly observing the complex via native MS.[19] A wide range of biochemical fractionation techniques have been used for this purpose, including ion exchange chromatgraphy,[19−21] isoelectric focusing,[20−22] sucrose density gradient centrifugation,[20,21] size exclusion chromatography,[22,23] native gel-based electrophoresis strategies,[24] and capillary zone electrophoresis.[25]

2.3. Direct Immunoprecipitation

The ability to generate antibodies that bind specific epitopes in a protein of interest has accelerated many areas of biology. In protein purification, antibodies conjugated to sepharose or magnetic beads are used to immunoprecipitate (IP) protein complexes. Usually an antibody targeting one subunit in a complex is chosen, and IP under nondenaturing conditions will lead to isolation of the entire complex, provided that the epitope chosen is accessible in the complex. Reference 26 describes a protocol for choosing and optimizing a primary antibody for IP followed by MS analysis.

As with biochemical purifications, direct IP has the advantage that the protein does not need to be modified for purification. However, it requires robust antibodies that recognize the epitope. Antibodies can be conformer or proteoform specific, leading to a biased sampling of complex distribution. Typically, polyclonal antibodies are preferred over monoclonal antibodies because they can potentially recognize a range of epitopes and maintain population heterogeneity.

An additional consideration is dissociating the tightly bound protein/antibody complex after IP for MS analysis. This can be done by elution with high salt[27] or via denaturing methods including low pH glycine buffer or urea. However, as discussed above, high salts and detergents are not compatible with all MS methods, and quaternary structure is destroyed by denaturing elution. Additionally, contamination of the purified sample with large quantities of the primary antibody can lead to masking of the MS signals of proteins of interest.

2.4. Affinity Tagging

Arguably, the most popular method for purifying endogenous protein complexes is affinity tagging, in which additional bases are added to the DNA sequence coding for the protein of

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

Table 1. Properties of a Selection of Popular Tags for Affinity Purification Followed by MS Analysis

tag	nature	mass (kDa)	origin	elution method
HA, 3X HA	9 amino acids	1.102	human influenza hemagluttin[36]	low pH or HA peptide
C-myc	10 amino acids	1.202	C-myc gene product	low pH or myc peptide
GST	220 amino acids	26	enzyme glutathione-*S*-transferase	excess reduced glutathione
GFP	238 amino acids	27	green fluorescent protein	low pH or denaturing conditions
SpA	depends on how many repeats of IgG domain are used	1 engineered *Z*-domain: 6.6[37]	protein A from *Staphylococcus aureus*	SpA binding peptides, low pH, denaturing
FLAG, 3X-FLAG	8 amino acids	1.031	engineered[38]	FLAG-peptide, low pH, denaturing
CBP	26−28 amino acids	4	calmodulin binding peptide	EGTA and high salt
STREP	8 amino acids	1.058	genetic random library selection[39]	biotin or biotin derivative
His	6 amino acids	0.841	engineered[40]	imidazole or EDTA
V5	14 amino acids	1.421	simian virus 5	low pH or competitive elution with V5 peptide

interest. This sequence is transcribed along with the protein of interest, which then contains a tag that can be isolated using commercially available affinity resins for the tag of interest. Affinity methods provide a single-step purification in a high-throughput fashion that can be theoretically applied to any target of interest. The speed of purification can also preserve transient or weak biomolecular interactions.

A wide range of affinity tags are available, with novel tags introduced frequently (reviewed in refs 28−30). Table 1 summarizes the properties of some popular tags for purification of endogenous complexes. Tandem affinity tags for tandem affinity purification (TAP) are also available that combine multiple epitopes for coupling multiple affinity steps.[31−34] TAP approaches generally lead to better contaminant removal at the expense of dissociating transient interactions. The sensitivity of modern mass spectrometers combined with improved methods for contaminant detection makes the use of TAP tags less important.

When choosing a tag, its size, charge, and the method of removal from affinity resin is critical. As discussed above, this last property can be an extremely important variable for downstream MS analysis. If competitive elution strategies which involve addition of peptides or other small molecules interfere with downstream MS analysis because they limit the dynamic range, gel filtration or buffer exchange can reduce their concentration in the sample. Host organism is also an important consideration, with His tags being effective for purification of proteins from *E. coli* but significantly less useful from mammalian cells due to nonspecific copurification of contaminants.

Placement of the tag on the protein of interest is also crucial to avoid disrupting functional complex formation. Any known facts about protein structure and assembly should be incorporated. Tags can be placed at the N- or C-terminus of the protein, as well as internal to the sequence in a position accessible in the final folded protein structure.[35] It is not possible to empirically predict which location is better for a protein of interest. For N-terminally tagged proteins, polypeptides that are improperly terminated early will contain the affinity tag and copurify along with the desired protein; for C-terminally tagged proteins, alternative initiation sites can also lead to a mixture of different proteins with the C-terminal tag. Ideally, biochemical or biological assays should confirm that the affinity tag has not impacted the protein's endogenous fold, function, interactions, and localization. These can either be in vitro assays that assess enzymatic function or protein structural properties after tag addition, or *in cell* assays to confirm the location of the tagged protein and normal physiological state after tagging. However, it is always possible that the tag perturbs a completely unknown function or association that the protein engages in, and this should be considered as part of the experimental design and analysis.

2.4.1. Methods for Generating Affinity Tagged Endogenous Complexes. To use affinity tags to purify endogenous complexes, the DNA coding for at least one subunit of the complex must be modified with the affinity tag. The gold standard, which presumably produces complexes as similar to native as possible, is to tag the protein at the endogenous gene locus in the appropriate host. Endogenous tagging strategies generally capitalize on the principle of homologous recombination, reviewed in refs 41 and 42, in which homologous regions of DNA recombine and regions in between the homology patches can be swapped. Efficient homologous recombination relies on the generation of precise double-stranded DNA breaks (DSBs); the introduction of CRISPR/Cas9 technology,[43] which revolutionized the simple and specific generation of DSBs, has advanced affinity tagging in a range of organisms, including *Escherichia coli*[44,45] and yeast.[46] However, it is particularly powerful for endogenous tagging in mammalian cells, and several groups have presented tools and approaches for CRISPR based epitope tagging in mammalian cells.[47−50] Moreover, the use of CRISPR to generate epitope tagged whole model organisms, such as mice,[51] can significantly advance comparative purification of complexes from different tissues and under different physiological stresses. However, users of CRISPR should be aware that it can be subject to off-target effects.[52]

For *E. coli*,[41,53,54] and yeast,[55−59] robust tools for epitope tagging predate CRISPR/CAS9 technologies, and various libraries of epitope tagged yeast and *E. coli* strains are available, as well as yeast donor libraries that enable fast tagging of ORFs with any desired epitope.[60,61] For mammalian cells, epitope tagging pre-CRISPR was generally accomplished via transcription activator-like effector nuclease (TALENS) or zinc finger nucleases(ZNFs),[48] which are composed of protein domains that recognize specific DNA sequences fused to nonspecific DNA cleavage domains. These modalities are more difficult to scale than CRISPR technologies, because new protein constructs must be designed for each gene targeted.

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

Therefore, for mammalian cells in particular, when tagging of endogenous loci is not feasible or practical, other approaches exist that result in endogenous or near-endogenous levels of protein. Bacterial artificial chromosomes containing the protein of interest under control of the endogenous promoter can be stably transfected into mammalian cells[62] for protein expression. Ectopic promoters can also generate near-native levels of protein. For example, tetON systems, in which addition of tetracycline derivatives drive gene expression, are titratable, which allows researchers to control the level of ectopic expression.[63] These ectopic genes can be introduced into cells via transient transfection, random generation of stable clones, or episomal vectors, which are maintained in the nucleus in a nonintegrated state and replicate, allowing generation of a semistable cell line.[64] Lentivirus and adenoviruses can also be used to generate stable clones by integration into the mammalian genome; however, random integration or integration of multiple copies of the gene can lead to disturbances in cellular physiology and uncontrolled gene silencing. As a solution, ectopic genes can be targeted to "safe harbors" in the human genome,[65] such as AAVS1,[66] where integration of genes has been shown to minimally perturb cellular physiology or gene regulation.

Regardless of how the DNA coding for the tagged protein is introduced, some form of selection is typically required to generate cell populations containing the tagged protein. Most often this is done using antibiotic selection, where along with the desired mutations an antibiotic resistance cassette is introduced.[44,45,48,50] Fluorescent proteins can also be added which enable cell sorting for the desired tagged proteins.[67] Both of these methods do involve expression of exogenous proteins by the cell. Single cells can also be isolated and sequenced to cultivate clonal strains with a single mutation; this enables the addition of a tag without need for additional selection cassettes.[47]

2.5. Cross-Linking to Preserve Noncovalent Interactions

Depending on the mode of MS analysis chosen, it may be helpful to add a cross-linking step as part of the purification process. Cross-linking refers to covalently linking together reactive amino acids in proteins and protein complexes.[68−70] Cross-linking can be coupled to any of the three enrichment strategies discussed above (Figure 3) and enables samples to be purified stringently while retaining weak but physiological associations. Cross-linking generally increases the number of proteins that copurify[71,72] and is also used to generate distance constraints for cross-linking MS (see section 3.3.2.1).

A wide range of cross-linkers exist,[73] and the choice of cross-linker will depend on downstream application. Cross-linker concentration and reaction time should be optimized to ensure that only physiological interactions are captured. Formaldehyde and glutaraldehyde are two popular small, cell-permeable cross-linkers that are useful for simply preserving interactions, with glutaraldehyde functional at low temperatures.[74] Other cross-linkers are designed for subsequent cross-linking MS analyses, as reviewed in ref 73 and discussed below (section 3.3.2.1), many of which are also cell permeable. Some cross-linkers can be affinity purified from cell lysate via reactive handles, often based on click chemistry, for the enrichment of otherwise low-abundant cross-linked peptides from complex biological samples.[75]

Cross-linking can occur at many points throughout the purification procedure. Cell-permeable compounds can be added before cellular lysis. For example, in rapid immunoprecipitation mass spectrometry of endogenous protein (RIME), a protocol developed by Carrol and co-workers, in-cell formaldehyde cross-linking followed by immunoprecipitation and MS analysis is used to identify protein complexes.[76−78] Intriguingly, Fabre and co-workers[79] found that cross-linking before cellular fractionation prevented leakage of proteins between different fractions and preserved native localization of proteins after fractionation. Cross-linking can also take place in lysates,[80] in cell powders,[74] or after complexes are captured on affinity beads[81]

Chait and co-workers have engineered lysine-free anti-GFP nanobodies for on-bead cross-linking with lysine cross-linkers; the anti-GFP nanobodies will not be affected by cross-linking.[81] Cross-linking can also be performed in gel after native gel separation[82] of endogenous complexes. However, it is important to note that Zhang et al. found different cross-linking results between in-cell and in-lysate linking;[71] therefore, for optimal preservation of biological interactions, cross-linking should take place before cell lysis.

In addition to adding exogenous cross-linkers, photo-cross-linking amino acids can be incorporated into cellular proteins and used to capture protein−protein interactions in cell following light activation.[83] Photoleucine,[84] photomethionine,[84] and photolysine[85] are incorporated into all cellular proteins, with rates ranging from 4% to 40% and have been used for MS identification of histone binding proteins. Site-specific incorporation of photoreactive amino acids into specific proteins can be achieved using unnatural amino acid incorporation[86] or split intein technologies[87] to specifically interrogate the interactome of one protein, such as the membrane protein IFITM3 for which traditional affinity purification methods failed to recover true interactors.[88]

3. MS-BASED TECHNIQUES FOR STUDYING ENDOGENOUS COMPLEXES

3.1. Biological MS Delivers Insights Across Multiple Levels of Complex Regulation

A range of biological MS techniques exist which differ in the details of sample preparation and data analysis but have in common the generation of gas-phase ions of proteins or protein fragments, mass separation, and detection (Figure 2). Techniques can be classified as either "bottom-up", in which proteins are digested into peptides before gas-phase ion creation, or "top-down", in which proteins are transferred intact into the mass spectrometer.[89] Middle-down approaches, in which proteins are partially digested into large fragments (>7 kDa[89]) are also emerging but will not be discussed further here. In all forms of MS, specific ions can be selected for gas-phase ion fragmentation in tandem MS experiments.[90,91] By fragmenting ions and analyzing the daughter ions, sequence and structural information about the parent ion can be determined.

We will cover a range of bottom-up approaches, including shotgun proteomics, cross-linking MS, HDX MS, chemical labeling MS, as well as top-down MS and native-MS. After providing a brief overview of each technique, highlighting its pros and cons, we will describe what information it can provide for endogenous complexes. By combining different sample preparation protocols described above with these varied MS modalities, many different types of information can be extracted (Figure 1).

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

For a given complex, it is important to determine both its *composition*, namely the distinct subunits that form the complex, as well as the *stoichiometry* of those proteins, namely how many copies of each subunit is present. The subunit interaction network, describing which proteins specifically interact, should be defined, and subunits classified as core or peripheral. Under the core-attachment model, proteins form core complexes, which have a stable, permanent relationship but are often temporally regulated through the attachment of peripheral proteins.[92−94] Members of the core complex typically share functional annotations, protein localization, and even specific deletion or mutational phenotypes, while attachment members do not necessarily share these properties. The interaction of attachment members with the core complex is highly controlled. Bottom-up proteomics, cross-linking MS, and native-MS can inform on complex composition and stoichiometry for both core and attachment complexes, although the former are easier to preserve through an MS work-flow.

A given genetic sequence can produce multiple proteoforms, a term which refers to a defined sequence of amino acids with localized modifications.[95] Different proteoforms arise due to alternate initiation sites, splicing events, and post-translational modifications (PTMs). PTMs are chemical modifications that occur following protein biosynthesis.[96] As many as 300 PTMs are known to occur physiologically[96] and can range from addition of small acetyl groups to conjugation of ubiquitin and other signaling proteins. PTMs are a key way that the cell temporally regulates the activity of different proteins and protein complexes, with evidence suggesting that proteins at the center of interaction networks are more likely to be post-translationally modified.[97] Analysis of PTMs as a function of cellular state is an important component of understanding protein complex biology. PTMs clearly manifest in the MS spectrum as shifts in protein and peptide mass, and both bottom-up and top-down MS can inform on the PTMs of specific complexes and on large-scale proteome wide screens of PTMs.

MS-based methods, in particular cross-linking MS and ion mobility (IM)-MS, can also determine the topologies of protein complexes, which is especially relevant for endogenous, heterogeneous complexes available in limited quantities that prohibit high-resolution structure determination. For large complexes, if high resolution structures of individual subunits are available, they can be combined with MS-based topologies to create a 3D map of the entire complex. MS-based information can be converted into restraints for a particular protein or protein complex and then used to determine structures and structural ensembles, often in combination with computational modeling tools (see section 3.5).

In addition to determining static topologies and structures, complexes undergo conformational fluctuations as they move between different functional states. Footprinting MS techniques can inform on these conformational fluctuations and the ensemble of different structures present in solution.

Lastly, MS methodologies can be used to study the binding of cofactors and ligands to complexes. It is especially important to identify which small molecules are bound to the complex in its endogenous state and to unravel how these small molecules regulate complex activity. Identifying small-molecule binding is primarily accomplished via native-MS, but information can also be extracted from specialized bottom-up experiments.

3.2. Sample Separation and Fractionation

The power of mass spectrometry as an analytical tool stems from its ability to distinguish between components with different m/z ratios, enabling the detection of the full distribution of coexisting states of a given peptide or protein complex. However, like many analytical techniques, MS has a limited dynamic range,[98,99] which causes intense ions to suppress the ionization, and consequently signal, of weaker, less abundant ions. The signal of protein and peptide ions can also be suppressed by the presence of salts and other matrix components.[100] Spectral complexity is also a challenge and congested spectra containing overlapping peaks are difficult to analyze and assign.

Extensive sample fractionation is used to combat these challenges. By separating samples into different fractions, broader coverage can be achieved. Chromatographic separations can also reduce ion suppression by matrix sample components by enabling effective buffer exchange of the desired analytes. As mentioned above, different types of chromatographic separations are an important component of isolating endogenous complexes; however, here we focus on separation methods that typically constitute part of the MS analytical workflow, acknowledging that they overlap to some extent.

3.2.1. Electrophoresis-Based Methods. MS proteomic analysis of protein mixtures was first coupled to polyacrylamide gel electrophoresis, or PAGE, as an offline chromatographic separation tool. In these experiments, protein spots were excised from gels, either Coomassie or silver-stained, and subjected to in-gel digestion,[101] followed by LC/MS analysis of the digested peptides.[102] Most typically, 2D PAGE was used,[103,104] in which proteins are first separated according to isoelectric point and then by molecular weight. However, its practicality is limited by the fact that it requires extensive sample handling, is time-consuming, and cannot be coupled directly to the mass spectrometer. Moreover, 2D PAGE has limited ability to identify medium-to-low abundant proteins.[105] 2D PAGE also has limited loading power and a resolving power limited by the combination of isoelectric focusing and molecular weight. Therefore, it has been primarily supplanted by liquid chromatography methods for MS analyses, although some groups continue to advance methods for increasing the separation efficiency of 2D PAGE.[106]

A related electrophoretic method is the GELFrEE method, or gel-eluted liquid fraction entrapment electrophoresis, which also separates proteins based on molecular weight.[107,108] In GELFrEE, proteins elute off the end of a polyacrylamide gel column and are trapped in a collection chamber in multiple fractions. While GelFREE is not as robust as other fractionation methods,[109] it has the advantage that it can be performed in denatured or native mode,[24] permitting size-based separation of native proteins for MS applications.

3.2.2. Liquid Chromatography. Liquid chromatography, most commonly high-pressure liquid chromatography (HPLC), is one of the most popular methods for protein and peptide separation prior to MS/MS analysis. It is especially useful because, in combination with electrospray (ESI) methods of ionization, it can be directly coupled to the inlet of the mass spectrometer. Thus, samples can be simultaneously separated and exchanged into MS compatible buffers. Programs can be controlled by software, leading to high reproducibility and easy optimization.

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

The most popular HPLC method to couple to an MS system is reversed-phase HPLC(RP-HPLC).[110] RP-HPLC separates proteins and peptides based on hydrophobicity, using a hydrophobic stationary phase and eluting with organic solvents. These organic solvents are MS-compatible, making RP-HPLC an excellent tool for MS analysis. RP-HPLC also offers good resolution, easy tuning of elution conditions, and is particularly suited to peptide analysis because peptides are typically recovered well from RP columns. However, RP-HPLC requires denaturing and is not applicable to native-MS applications; moreover, recovery of intact proteins is variable.

For complex proteomic samples, 2D LC/LC/MS experiments increase resolution with multiple dimensions. For example, MudPIT, introduced by Yates and co-workers in 2001, combines strong cation exchange chromatography with RP-HPLC to separate and detect a wider range of peptides.[111] Generally, ion exchange columns can also be used for fractionation but suffer in terms of coupling to the mass spectrometry because the salt used for elution will be injected into the MS system, causing severe signal suppression. Directly coupling ion exchange columns to the mass spectrometer requires the use of MS compatible volatile buffers such as ammonium acetate for ion exchange. Therefore, for analysis of peptides, RP-HPLC is almost always used as the final step before the mass spectrometer.

LC-MS is emerging as a technique for buffer exchange[112] and even purification[113] of intact proteins and complexes for native-MS. In these experiments, a small size-exchange column buffer exchanges proteins into MS-compatible buffers directly prior to analysis, and tandem chromatography methods can be used to combine affinity columns with buffer exchange. Other advances in liquid chromatography for MS analysis, as reviewed in ref 114, include the development of smaller columns which enable faster separation and introduction of new resins.

3.2.3. Capillary Zone Electrophoresis.

Capillary zone electrophoresis (CZE) separates ions in solution based on their electrophoretic mobility under applied voltage.[115] The electrophoretic mobility is dependent on charge, shape, and size of proteins or peptides. CZE provides orthogonal, complementary data to liquid chromatography. Advantages of CZE include low sample requirement and low carry-over between separation experiments. It also provides relatively fast separation, which can be an advantage or a disadvantage, because analytes can elute too quickly to be separately fragmented in tandem MS experiments.

CZE has been applied to proteomics workflows in bottom-up mode, detecting patterns of PTMs[116] in endogenous proteins and to rapidly separate peptides in complex proteome digests.[117−119] However, a back-to-back comparison of CZE and RPLC indicates that RPLC may identify more peptides and post-translational modifications,[120] and CZE has limited loading capacity. Both methods can be combined for increased separation capacity.[120]

CZE can also be used to analyze proteins in top-down mode, demonstrating separation of antibody variants differing by single deamidation events.[121] Importantly, CZE is compatible with separation under native conditions[122] in which proteins retain their tertiary and quaternary structure. Recent work by Kelleher and co-workers applied native CZE to analysis of endogenous nucleosomes,[123] using tandem MS to characterize differences in proteoforms between endogenous nucleosomes in different cell lines.

3.2.4. Ion Mobility.

Ion mobility (IM) separates protein and peptide ions within the mass spectrometer after gas phase ionization. In IM experiments, ions traverse a gas-filled tube and are separated according to their collision cross section. The collision cross section is a property of the ion-gas pair and has dependence on the charge, size, and shape of the ion.[124] An extended ion will experience more collisions with the gas and thus travel more slowly than a ion with the same mass but a more compact structure.[125−127] There are multiple IM MS devices, such as drift tube, traveling wave, differential mobility, transversal modulation overtone, field asymmetric, and trapped IM-MS, each with its own advantages and disadvantages, and we refer the readers to detailed reviews on this topic.[128,129]

For analysis of peptides in bottom-up proteomics experiments, including IM greatly increases the detection capacity of the experiment,[130] enhancing selectivity, resolution, and dynamic range by spreading peptides out over an addition dimension.[131,132] Addition of IM to proteomic workflows advances the ability of proteomics to study increasingly smaller quantities of protein[133] toward the goal of single-cell proteomics.

IM is also a powerful tool for native MS analysis of proteins. Particularly, it is used not only to separate sample components but also as a structural biology tool. Extraction of collision cross sections from IM data constitute a restraint that can be used to generate information the 3D shape and conformation of protein complexes[134] (see below sections 3.4 and 3.5).

3.3. Bottom-Up Approaches

3.3.1. Proteomics.

3.3.1.1. Overview of Bottom-Up Protein Identification. The most popular form of biological MS is bottom-up proteomics, described in detail in refs 6, 135, and 136. For bottom-up proteomics experiments, proteins are digested into peptides by proteases, trypsin being by far the most popular. Trypsin's popularity stems from the fact that it is an extremely efficient protease that cleaves after basic arginine or lysine residues, producing peptides that are efficiently ionized and fragmented in tandem MS experiments. However, because of its cleavage specificity, trypsin has a bias toward hydrophilic portions of the protein. Higher sequence coverage and complementary data sets,[137] particularly in terms of PTM identification, can be accomplished by digestion with complementary enzymes, including chymotrypsin, LysC, LysN, AspN, GluC, and ArgC, as reviewed in ref 138. Enzymes can also be sequentially applied to improve protein identification.[139]

As discussed above, peptides are then typically separated on RP-HPLC columns and injected into the mass spectrometer. Tandem-MS techniques[90,91] are then applied for sequence analysis via gas-phase fragmentation patterns. Different strategies for choosing which peptides to analyze via tandem-MS strategies are available, as even with the best liquid chromatography separation strategies, a large number of peptides elute simultaneously. In targeted proteomics, or selected reaction monitoring experiments, specific ions corresponding to proteins of interest are chosen a priori for fragmentation and quantification.[140] These experiments are fast, reproducible, and extremely sensitive to the protein of interest but cannot be used to discover new proteins. For analysis of complex mixtures, researchers rely on shotgun proteomics methods which attempt to sequence all proteins in complex peptide mixtures. In data-dependent acquisition (DDA), the mass spectrometer selects peptides meeting a

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

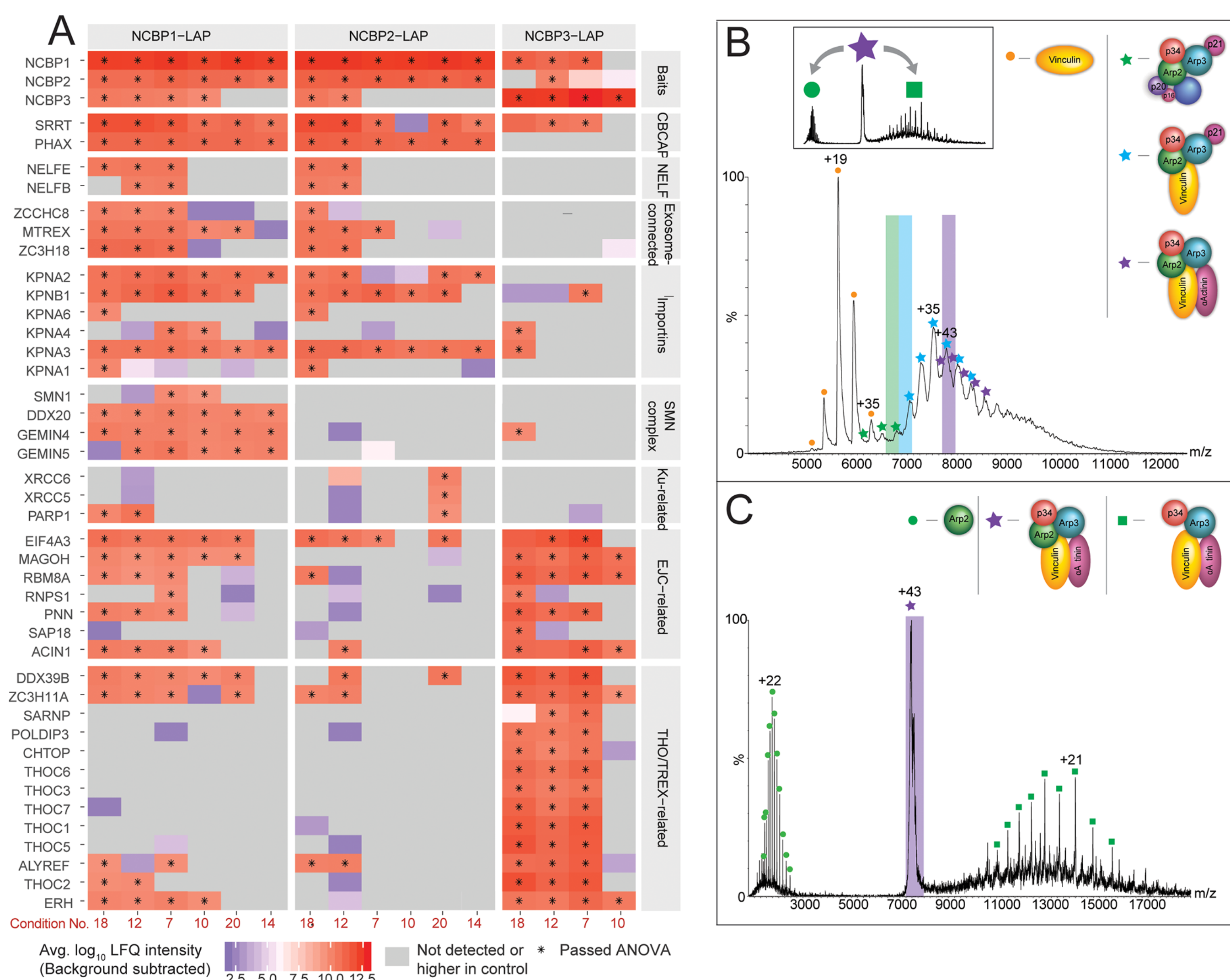


Figure 4. Identifying complex modularity via bottom-up proteomics vs native MS. (A) In bottom-up proteomics experiments, complexes are revealed by identifying copurified proteins. Here, three different tagged proteins were used as "bait" (top row), namely NCPB1, NCBP2, and NCBP3, which are different nuclear cap binding proteins that bind to nascent RNA. The data can be visualized as a heatmap of control-subtracted label free quantification (LFQ) intensity for each "prey" protein identified by MS. This data demonstrates that NCPB3 differs from NCBP1 and NCBP2 in selectivity, coprecipitating with members of the THO/TREX complexes. These interactions were later confirmed as supporting a biological role for NCBP3 in mRNA expression that differs from NCBP1 and NCBP2.[351] (A) Figure reproduced and adapted with permission from ref 352. Copyright 2020 Oxford University Press. (B) Shown is a native mass spectrum of multiple coexisting complexes purified from chicken muscle. While the main charge series corresponds to monomeric vinculin, additional charge states corresponding to three protein complexes are also seen. The masses of these complexes indicate the presence of intact Arp2/3, a vinculin-associated Arp2/3 complex, and a vinculin-α actinin-associated Arp2/3 complex. (C) Tandem-MS is used to validate the complex composition; shown is the MS/MS spectrum of the charge state selected in purple, with circles denoting released subunits and squares corresponding to the remaining stripped complex. Tandem MS reveals that Arp2 is released from a complex composed of vinculin, α-actinin, and the Arp2/3 complex. (B,C) Reproduced and adapted with permissions from ref 18. Copyright 2014 Springer Nature.

certain threshold in the first MS scan and then selects them for fragmentation. DDA is the most widely used method for shotgun proteomic analysis and it aids peptide identification because the precursor mass of the ion is known. However, it is estimated that an extremely large percentage of peptides are missed,[141] prompting the development of data independent acquisition (DIA) strategies. In these methods, first proposed by Yates and co-workers,[142] specific ranges of the MS spectrum are selected sequentially for tandem MS analysis automatically. DIA holds promise for better proteome coverage[143] but the data produced is significantly more complicated. Different variations of DIA are a promising area of research to advance shotgun protoeomics.[144,145]

Protein amounts can also be quantified via quantitative proteomic methods. Quantification methods generally fall into label-free methods, which include peak intensity based methods and spectral counting[146] or through isotope labeling. Peak intensity based methods, such as iBAQ,[147] rely on integrating peak intensities, which were shown to correlate well with protein concentration.[148] In spectral counting,[146] the number of peptide spectra assigned to a protein of interest is used as a proxy for relative protein abundance. Generally, peak intensity based methods are considered to be more accurate than spectral counting, but both have limitations because they rely on accurate identification of peptides across multiple spectra. The gold standard for relative quantification is the

isotope labeling methods. These methods rely on the fact that peptides with the same sequence but different isotope enrichment will coelute but be separated in a single spectrum for quantification. Labeling methods include stable isotope labeling in cell culture (SILAC), in which proteins are grown in isotope enriched media,[149] or the addition of mass tags to samples postisolation.[150,151]

After the MS data is generated, significant software analysis is required to identify and quantify proteins in the MS spectra; choice of data processing strategy is an integral component of experimental design. The tandem-MS experiments must be analyzed to determine the peptides present, the proteins they belong to, and to quantify their percentages. Reference 152 describes some of the different strategies used for analyzing shotgun proteomic data. Typically, MS/MS spectra are collected and compared to databases involving all spectra corresponding to a digest of the parent proteome, either theoretical or experimentally generated. Software with this capability include MASCOT,[153] MaxQuant,[154] Progenesis, and PEAKS.[155] However, de novo sequencing, or forward prediction of peptide sequence based on tandem-MS data, is also an option, with other softwares incorporating this capability.[156] When analyzing targeted proteomic data, Skyline is the most widely used.[157] The ProteoWizard toolkit provides an open source platform for converting data from different mass spectrometer formats and cross-platform analysis.[158] Several groups have compared the different available softwares and shown that the software and data processing pipeline chosen has a significant impact on the results of the LC-MS/MS analysis,[159,160] both finding that with suitable data processing, MaxQuant is the most robust. These researchers also demonstrated, as expected, that for quantification, targeted MS has a higher dynamic range and limit of linearity because specific peptides are targeted and chosen.[160]

For proteomics-based methods, the absolute sensitivity threshold is quite low; a modern mass spectrometer is able to detect as little as 1×10^{-18} mol of peptide.[161] Other sources of error throughout the workflow lead to proteins being missed or misidentified. In the sample preparation step, poor cellular lysis, poor solubilization of proteins, and contamination from abundant proteins, keratin, or dust are all significant sources of error. As mentioned above, another challenge is the dynamic range of the proteome; as discussed in ref 99, the dynamic range of MS is only 4 orders of magnitude while that of the proteins in the cell is 7 orders of magnitude at least. Therefore, many low abundant proteins, which may constitute real interactors, are missed in analyses of protein interactomics, and enhancing proteome coverage is a major challenge for bottom-up proteomics. The limited ability to select and fragment peptides can also cause peptides to be missed, even if they are present in the spectrum, and moreover, the software can misidentify or miss peptides that are present at low signal-to-noise levels. Therefore, when analyzing a particular endogenous protein complex, users should be aware that low-abundant or transient interactions may be missed even in these extremely sensitive experiments.

One area of advance in proteomics is focused on platforms that can analyze proteins from single cells. Termed "single cell proteomics",[162,163] advances towards this effort include development of miniaturized automated nanodroplet processing platforms capable of bottom-up digestion of single cells,[164,165] usage of narrow-bore LC-MS[166] and new methods for cell lysis.[167,168] This increased sensitivity has relevance for studying complex heterogeneity at the single cell level.

In addition to identifying and quantifying proteins, bottom-up proteomics can identify a wide range of PTMs, which manifest as shifts in peptide mass and fragmentation patterns.[169,170] However, a number of pitfalls exist in the use of bottom-up proteomics to identify PTMs, reviewed in ref 171. Information on the coexistence of different modifications in a single polypeptide is lost because a mixture of peptides is generated and analyzed together. Similarly, it is difficult to determine which proteoforms coexist in a complex together because the bottom-up results only report on the presence or absence of the PTM. Therefore, while post-translationally modified proteins are often detected in bottom-up proteomic experiments of protein complexes, including *in cell* cross-linking experiments,[172] it is difficult to functionally characterize how PTMs regulate complex formation from bottom-up data alone.

3.3.1.2. Bottom-Up Proteomics Illuminates Complex Composition. In terms of characterizing endogenous protein complexes, bottom-up proteomics is typically used to provide information on the complex composition when investigating a particular complex or configurational differences that occur under different cellular states by comparative analysis of multiple conditions (Figure 4A). In this experimental format, protein complexes are predicted from observation of which peptides, and therefore proteins, are detected as copurifying. The experimental pipeline typically involves a method of copurifying proteins under native conditions, followed by bottom-up MS analysis of the purified fractions and computational filtering to predict complexes on the basis of which proteins copurify.

One method of generating complexes for copurification identification is intensive biochemical fractionation. This has been used to fractionate lysates from a range of cell types,[21,23,173−175] with cofractionating proteins identified by MS and endogenous complexes predicted by observing which proteins cofractionate. Importantly, multiple studies corroborate that many proteins elute at a MW larger than that predicted by their sequence, implying preservation of native complexes throughout native extraction methods.[23,175] In a recent example, Kastritiis et al. used SEC to characterize endogenous protein supercomplexes, a term that refers to megadalton-sized macromolecular assemblies composed of multiple functionally related protein complexes, from *Chaetomium thermophilum*.[175]

More popularly, co-IP, followed by MS identification, is used to identify protein−protein interactions and thereby infer protein complex composition (see ref 176 for a review). Several high-throughput, broad studies combining IP and MS have been conducted which provide excellent starting resources for understanding cellular interactions and complexes. These include endogenous TAP tag based studies in yeast,[94,177] which identified 491[94] and 547[177] heteromeric complexes. Endogenous human complexes were IP'ed from HeLa cells using over a thousand primary antibodies[178] and previously unknown complexes identified, including a complex suggested to be involved in transcriptional coregulation. Human affinity-tagged lines were also created, with Hein et al.[179] using BAC transgenes while the Gygi lab took advantage of the human ORFeome project[180] to create lentiviral lines coding for a wide range of human proteins in HEK and HELa cells.[181,182] Both of these studies yielded rich data sets that

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

map a wide range of interactions connecting thousands of endogenous proteins.

To determine complex stoichiometries via copurification and bottom-up proteomic experiments, subunit stoichiometry must be inferred from relative abundance of peptides corresponding to different proteins. Label-free quantification of peptides can determine relative abundances of different proteins and was used to determine the stoichiometry of NPC complexes,[183] chromatin-associated complexes,[184] and SET/MLL complexes,[185] to give some examples. Information about total cellular protein abundance can also be incorporated to extract subunit stoichiometries[179] because the amount of protein that copurifies with a particular bait will depend on the cellular abundance of the interacting components as well as on the stoichiometry of the individual complex.

The computational approach used to predict protein complexes is a key component of the experimental workflow. As reviewed in refs 186−188, a variety of strategies exist to generate complexes from protein−protein interaction lists. For example, one algorithm, SAINT-MS,[189] uses a Bayesian model to estimate the probability of a given interaction being true or false, incorporating information from negative controls and spectral intensities, while another, COMPASS, incorporates information about spectral counts and reproducibility to score interactions.[190] However, it should be mentioned that poor reproducibility was found between different algorithms applied to the same data set.[191] Including orthogonal pieces of information, such as gene ontology annotations and gene expression data, with MS-derived interaction data can increase reliability and help remove false positives.

Users should be aware that affinity copurification methods to identify protein complexes can suffer from lack of reproducibility, false positives, and false negatives. Gavin[94] et al. found that duplicate IP-MS measurements from budding yeast had only 69% replicability, although others found that the same protocol conducted by two laboratories had 81% reproducibility.[192] In a meta-analysis of three data sets,[21,179,182] Drew and co-workers observed limited overlap between different high-throughput studies[193] with many false positives. False positives, or contaminants, generally result from nonspecific binding to either the bait protein or the affinity resin used. Specific isotope enrichment strategies, such as I-DIRT, can distinguish contaminants,[194] and the CRAPome is a repository of common contaminants for different affinity purification strategies[195] that should be checked against identified complex components to identify contaminants. False negatives in affinity-purification MS measurements generally result from complex dissociation during purification, from poor tryptic digestion of peptide interacting partners, or from data undersampling in the mass spectrometer.

Besides for high-throughput experiments, affinity-purification proteomic experiments have broad applicability to studying specific proteins and their formed complexes under conditions as close to endogenous as possible. In an early example, Yates and co-workers analyzed proteins present in the ribosome by ribosome purification and bottom-up proteomic analysis of all present components.[196] Examples of proteins studied span the gamut of cellular proteins, from identifying shared targets of an E3 ligase substrate receptor[197] to the RNA exosome[198] and LINE-1 transposon elements.[199] Figure 4A describes a recent study examining interactions of different RNA cap binding proteins that revealed a unique role for one protein in regulating mRNA expression.

An emerging bottom-up proteomics-based method to identify endogenous complexes via MS is thermal proteome profiling, or TPP. In this experiment, whole cells lysates are subjected to a temperature gradient, and then protein abundance as a function of temperature is measured via isotope quantified bottom-up proteomic methods to construct a melting curve for each protein. Although first described as a method for identifying drug-induced protein stabilization,[200] correlation of melting profiles for different proteins, termed thermal protein coaggregation (TPCA),[201] can identify associations based on the assumption that bound proteins have correlated melting curves. In the initial report, 30% of CORUM qualified complexes exhibited correlated TPCA curves.[201] TPCA is less useful for de novo complex prediction and also suffers in the analysis of low-stoichiometry complexes, for whom melting curves may be dominated by the noncomplexed protein. Nevertheless, TPCA analysis is particularly powerful for analysis of differential protein complex association between states such as throughout the cell cycle[202] or at differing time-points following virus infection.[203]

3.3.2. Protein Footprinting. The sample separation and analysis strategies outlined above represent a powerful tool for detecting and analyzing peptides originating from protein complexes. A wide range of strategies, broadly dubbed "protein footprinting",[204,205] use chemical or enzymatic sample treatment before MS analysis to encode information about the protein complex's structure, topology, and interactions in the MS spectrum. The vast majority of the time bottom-up proteomic methods are used to detect these chemical modifications, but top-down methods have be used occasionally.[206,207]

3.3.2.1. Cross-linking MS. 3.3.2.1.1. Overview of Cross-Linking MS. In cross-linking (XL) MS, reviewed in refs 68 and 208, samples in their native state are treated with bifunctional cross-linkers that react with reactive amino groups, acidic side chains, or cysteine groups. These cross-linkers will covalently link together amino acids that are proximal in the three-dimensional structure of the protein or protein complex. After digestion, peptides generated will be composed of two smaller peptides linked by the cross-linker. LC-MS/MS detection and quantification of cross-linked peptides yields information about intra- and intermolecular linkages between residues. Chavez and co-workers estimate the sensitivity of cross-linking MS to be $1-10 \times 10^{-12}$ mol for purified proteins and 5−10 mg of sample for in situ cross-linking,[209] making it an attractive option for endogenous complexes.

The choice of cross-linker is crucial to the XL-MS experiment; as reviewed in ref 73, a wide range of reagents exist with qualities such as distinctive tandem-MS fragmentation patterns[210−212] to enable clear identification of the cross-linked peptide among the background peptides. The next step, and the main challenge in XL-MS experiments, is to accurately analyze the cross-linking reaction mixture via MS and identify and quantify the peptides. Productive cross-linked products consist of two connected peptides, resulting in a quadratic expansion of the search space for the typical bottom-up proteomic algorithm. However, off-target peptides consisting of a single peptide attached to a cross-linker can also be produced. Moreover, cross-linked peptides may not fragment ideally in the MS/MS process, leading to poor signal-to-noise for one of the constituent peptides. Different softwares exist for the analysis of cross-linking MS data; XlinX[80] and Merox[13] are

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

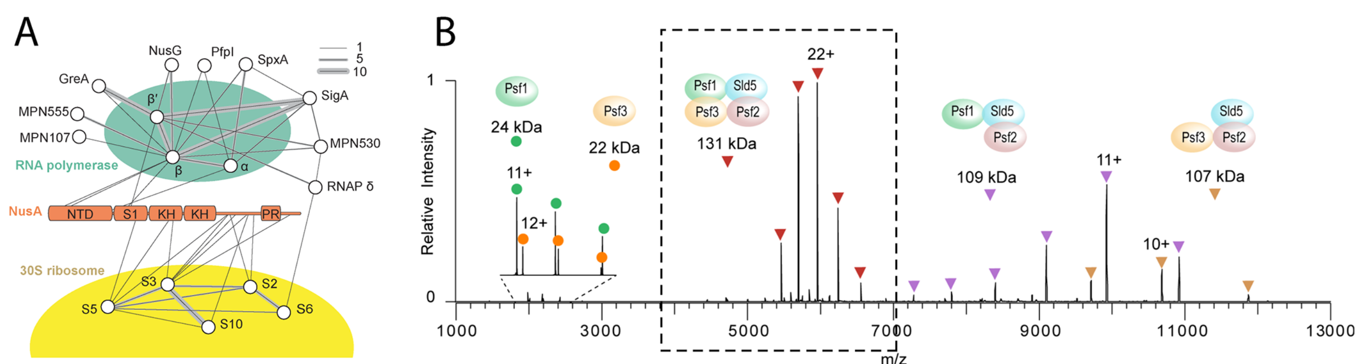


Figure 5. Native-MS and cross-linking MS can identify complex topology. (A) Cross-linking MS identifies the architecture of complexes via cross-links detected between subunits. Here, cross-links were generated in situ for the *Mycoplasma pneumoniae* expressome, a supercomplex composed of RNA polymerase and the ribosome linked via NusA. Cross-links oriented the C-terminal domain of NusA to RNA polymerase and the N-terminal domain to the mRNA entry site of the ribosome. A hybrid structure was later constructed that combined data from cryoEM and cross-linking MS. Reproduced and adapted with permission from ref 238. Copyright 2020 The American Association for the Advancement of Science. (B) Endogenous affinity-isolated GINS complex, a 131 kDa heterotetramer (box), was subjected to HCD activation, causing it to dissociate into subcomplexes, which appear at lower and higher m/z. By identifying the subcomplexes generated, a subunit connectivity map could be generated which was consistent with the known structures of homologous human GINS complexes. Reproduced and adapted with permission from ref 311. Copyright 2016 American Chemical Society.

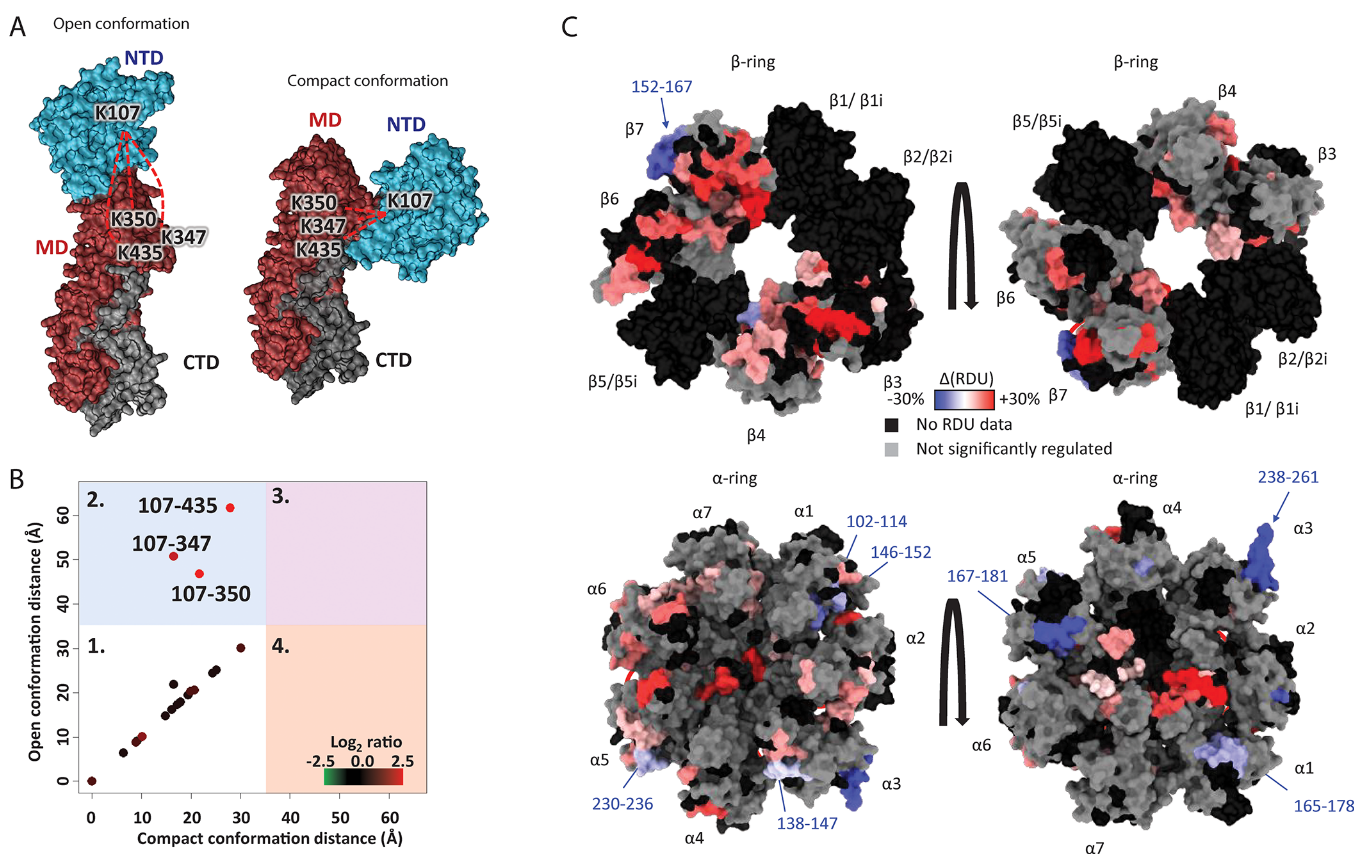


Figure 6. MS illuminates structural dynamics of endogenous complexes. (A,B) Cross-linking MS reveals conformational dynamics of HSP90 in vivo. HeLa cells were cross-linked, and cross-linked peptides corresponding to HSP90 were quantified with and without 17-AAG treatment. (A) Two different conformations of HSP90, with identified cross-links between the N-terminal domain (NTD) and middle domain (MD) shown in red dashed lines. (B) Plot comparing Euclidean Cα−Cα distances calculated for different HS90B experimental cross-links mapped onto the two structures: quadrant 1 is consistent with both models and quadrant 2 with compact model only. The color of the cross-links represents the $\log_2$ of the ratio between the two treatment conditions. This data demonstrates that the compact conformation increases in abundance after 17-AAG treatment. Reproduced with permission from ref 242. Copyright 2016 Cell Chemical Biology. (C) HDX-MS shows differences in structural dynamics between the standard 20S proteasome and the immunoproteasome. Shown is the difference in relative deuterium uptake for specific regions of the α and β rings of the immune-20S proteasome vs the standard 20S proteasome, demonstrating that some regions are more dynamic in the immunoproteasome (red) and some are more dynamic in the standard 20S proteasome (blue). The authors explained these differences as relating to differences in activity between these two proteoforms of the 20S proteasome. Reproduced with permission from ref 255. Copyright 2020 Springer Nature.

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

focused on MS-cleavable linkers, while Xquest,[213] Kojak,[214] StavroX,[215] and pLink[216] can be used for noncleavable linkers. A recent comparison of different softwares using a benchmarked library of peptides demonstrated that false discovery rates vary greatly depending on the software used, with FDR rates ranging from 2.4 to 32%.[217] Reference 218 discusses future initiatives to make XL-MS more reliable and reproducible.

Once cross-linked peptides have been accurately detected and quantified, they can be converted to distance restraints based on assumptions about the maximal distance between cross-linked residues and used to determine complex topology. One challenge is generating sufficient cross-links to properly orient subunits within the complex; Chait and co-workers estimate that at least 4−5 cross-links are needed to determine dimer orientation.[219] Combining orthogonal cross-linkers[220,221] can generate enough cross-links to reliably determine the architectures of affinity purified complexes. As discussed below (section 3.5), cross-linking MS is also commonly combined with computational modeling[222] to derive full structures and architectures of endogenous complexes. For example, the MXNL web server was developed for model validation using restraints from cross-linking MS.[223]

3.3.2.1.2. Application to Endogenous Complex Analysis. Cross-linking has illuminated endogenous protein complex composition by performing the cross-linking reaction before affinity purification to prevent complex dissociation and enrich for transient interactions. This has successfully identified interactors of M-Ras,[224] the proteasome,[79,225] and estrogen receptors,[76,226] among numerous others. Cross-linking can also be used to map protein−protein interactions without affinity purification. If cross-linkers are applied to cells and cross-linked peptides subsequently identified, proteins identified as cross-linked are presumed to interact.[80] However, this approach typically targets the most abundant proteins in the sample;[80,227] in one example, proteins identified as cross-linked in synaptosomes were three times more abundant than the average synaptic protein.[227]

The main power of XL-MS is its ability to illuminate endogenous protein complex topology and architecture. This approach has been used to determine the architecture of diverse endogenous complexes ranging from the coatamer module of the nuclear pore complex (NPC),[219] chromatin−protein complexes,[228] variants of the 26S proteasome,[221,229,230] TAP- purified PolII-TFIIF complexes,[231] ATP-synthase,[232] and substrate-bound TriC,[233] among others.

One strength of XL-MS for determining complex architecture is that cell permeable cross-linkers can be added to intact cells and tissues, yielding inter- and intramolecular cross-links that reflect topologies found in vivo.[234] Workflows for XL-MS in cells and tissues are described in refs 209 and 235. This approach has been used to determine the topology of respirasome supercomplexes directly from mitochondria,[236] sarcomere protein complexes directly from heart tissue,[237] and the expressome composed of RNA polymerase and the ribosome linked by NusA and NusG[238] (Figure 5A). Rappsilber and co-workers recently demonstrated that it is possible to generate de novo cross-linked structures of bovine serum albumin in plasma.[239] Blankenship and co-workers[240] applied XL-MS to intact cyanobacteria, using affinity purification of photosystem II to generate a structure of the megacomplex composed of photosystem I, photosystem II, and the phycobilisome.

Because cross-links report on a given distance cutoff between two residues, cross-links can be a source of information on the dynamics and conformational fluctuations in a complex. Cross-links can be compared between different samples, for example, following a biological treatment. Schmidt and co-workers noted that upon dephosphorylation of the F-type ATPase, intersubunit cross-links in the head, stators, and stalk portions changed in abundance, likely reflective of phosphorylation dependent nucleotide binding.[241] Chavez and co-workers used a quantitative XL approach in live cells to monitor changes in conformations of HSP90 upon treatment with 17-AAG[242] (Figure 6A). Cross-links can also be analyzed to extract information about pools of a protein complex population in a single sample. The coexistence of cross-links that satisfy mutually exclusive conformations can reflect intrinsic heterogeneity and flexibility of the complex. Sinz and co-workers recently used XL-MS to examine the conformational plasticity of ribosomes, mapping a wide range of cross-links onto the model of the *E. coli* 70S ribosome and demonstrating that no single static structure satisfies all cross-links.[243]

3.3.2.2. Hydrogen−Deuterium Exchange MS. *3.3.2.2.1. Overview of the Method.* In HDX-MS, reviewed in ref 244, the "footprint" is produced by exchange of protons in a protein sample with deuterons. Protonated protein samples are incubated in a deuterium-based buffer for variable periods of time to allow exchange of amide backbone protons with deuterium in the buffer. Exchange of amide protons depends on the position of the amide site (backbone or side chain) and its dynamics (see below), with rapidly fluctuating regions exchanging quickly while rigid regions can remain stable for hours.[245] Amide exchange is then quenched by pH and/or temperature[246] and the extent of deuterium exchange measured. Historically, HDX exchange rates were measured by a range of methods, including NMR, but the coupling of HDX with MS readout, which easily distinguishes between deuterated and protonated peptides based on mass, greatly accelerated the application of this technique.[247]

The intrinsic exchange rate for an amide backbone is on the order of milliseconds. The rate and extent of deuterium exchange observed for a particular peptide in a folded protein, specifically reflects the local rate of hydrogen bonding for the amides in question.[248] Generally, it can be assumed that this exchange rate is faster for amides that are solvent exposed and not involved in hydrogen bonds. If a particular region is undergoing dynamic fluctuations, local hydrogen bonds will be disrupted and potentially engage in exchange. Therefore, quantification of peptide deuteration as a function of time is generally assumed to report on solvent exposure and dynamics of a particular region. Moreover, HDX rates can be measured as a function of protein complex or protein state and therefore report on differences in subunit interfaces or dynamics between the states.

Measurements of HDX at the residue level are challenging, and therefore typically exchange rates are reported on the peptide level, meaning that this method does not have residue level information but rather information corresponding to the peptide level. HDX-MS is not be applied in vivo or in cell because following deuterium exchange for a given period of time, exchange must be quenched and the level of deuterium exchange preserved until peptides are analyzed. This prohibits sample purification after labeling and means that HDX-MS is only applicable to purified protein complexes. Much effort has

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

been expended toward optimizing workflows to prevent back-exchange, as back-exchange during sample digestion and LC-MS analysis is a key drawback of the method. Recent developments in HDX-MS are reviewed in ref 249.

Effective HDX-MS requires approximately 100×10^{-12} mol of protein,[250] which is a fairly high sample requirement relative to other MS methods. HDX-MS provides information on complex dynamics[251] and is especially powerful when applied in a comparative format between two different states, such as upon binding of regulatory subunits[252] or ligands.[253] It has also been used to study protein folding and aggregation kinetics.[254]

3.3.2.2.2. Application to Endogenous Protein Complexes Analysis. While most examples of HDX-MS are for recombinant proteins, it has tremendous potential for the dynamic analysis of endogenous protein complexes. In a recent example, the 20S proteasome purified from erythrocytes, as well as the immunoproteasomes purified from spleens, were analyzed and compared via HDX-MS,[255] revealing differences in flexibility between the two isoforms (Figure 6B). The interaction of these complexes with the PA28 proteasome regulators were also studied. Another example using F-type ATPase, probed *E. coli* derived membrane vesicles, in which Konerman and co-workers used HDX-MS to examine mechanical stress in the motor during active and inactive states.[256]

3.3.2.3. Chemical Labeling of MS. 3.3.2.3.1. Overview of the Method. As opposed to HDX-MS, which represents a reversible change (the swapping of a proton for a deuteron), footprinting via irreversible modification of amino acids is also used.[204] This can be accomplished either with chemical modifiers or with radicals, most typically oxygen and hydroxyl radicals. These provide a permanent modification to the protein sequence that will not be lost during bottom-up proteomic analysis. Typically, solvent exposed amino acids will be chemically modified, again provided that the structure allows for reactive accessibility.

A wide range of nonradical chemical modifiers exist (reviewed in ref 204), which react on the second to millisecond time scale. Many specifically target one type of amino acid, such as *n*-ethylmaleimide, which targets cysteine, or glycine ethyl ester (GEE), which targets aspartic and glutamic acids. Diethylpyrocarbonate (DEPC) labels a broad variety of nucleophilic residues,[257] including histidine, tyrosine, threonine, serine, lysine, and cysteine, yielding approximately 30% coverage over the range of surface exposed amino acids in the average protein.

Radicals, and specifically hydroxyl radicals, are a more popular choice. They can modify 19/20 of the amino acids, and thus, similar to HDX, provide broader coverage over the sequence. Radicals are generally produced on the millisecond and microsecond time scales. A variety of methods for generating radicals exist, including Fenton chemistry and radiolysis of water; these methods are reviewed in refs 258 and 259. One broadly applied method of generating hydroxyl radicals, fast photochemical oxidation of proteins, or FPOP,[260] reviewed in ref 261, generates radicals on even faster time scales of the nanosecond to microsecond time scale. This enables FPOP to be used to study protein folding events[262] and other fast dynamic events. In FPOP, samples are incubated with H_2O_2, and laser irradiation of the sample creates OH radicals which broadly label nearby amino acids.

Similar to HDX, chemical labeling methodologies provide information on the solvent accessible surface area of the protein or protein complex and therefore report on the topology of the complex. They are typically applied comparatively between two states to examine changes that occur upon treatment, for example, during heat treatment of an antibody.[263] Fast labeling approaches, such as FPOP, can illuminate fast dynamic events including protein folding.[262]

Drawbacks of these chemical labeling methods include the requirement to work in special buffers compatible with the chemical or radical of choice, as well as separation of peptides postlabeling. Moreover, it can be difficult to determine the exact location of an oxidative modification in the case of radical mediated footprinting, especially if multiple oxidations occur. Quantification of % of labeling based on spectral intensities is also difficult to assess accurately. However, Sharp and co-workers demonstrated that electron transfer dissociation fragmentation tandem MS can efficiently generate quantitative tandem MS product ions[264] for analysis.

3.3.2.3.2. Application to Endogenous Protein Complexes Analysis. Because of the irreversible nature of chemical labeling applied, the method can be applied in vivo to endogenous proteins, unlike HDX experiments. Zhu et al. applied hydroxyl radical footprinting to proteins in the *E. coli* outer membrane and studied the voltage gating of OmpF, observing changes in oxidation inside the porin upon changing the conditions of the cells.[265] Jones and co-workers have advanced the use of FPOP to analyze proteins both in live cells[266,267] and in *Caenorhabditis elegans*,[268] demonstrating its potential for use as a structural technique in a multiorgan system. Chemical labeling was also applied to study human vitamin K epoxide reductase *in cell* using cysteine alkylation to determine the redox status of key catalytic cysteines.[269]

3.3.2.4. Proximity-Induced Labeling. A related approach is proximity induced labeling, in which a protein of interest is genetically tagged with a protein that catalyzes the addition of a chemical tag, usually biotin, onto spatially proximal proteins. Popular proteins include APEX,[270,271] an engineered version of ascorbate peroxidase, which is fast and quite promiscuous,[272] or variants of biotin ligase, including BioID[273,274] and TurboID.[275] This technology provides information on protein−protein interactions, as opposed to topology and dynamics. Proteins can be enriched specifically if interactors are tagged with biotin or other affinity handles. However, these methods can struggle to identify specific interactors because they can promiscuously label all proximal proteins, although approaches have been developed to filter for specific interactions. For example, Krogan and co-workers combined APEX labeling with spatial reference proteins to filter for specific interactors, identifying interactors of GPCRs in response to agonist stimulation.[272]

3.3.2.5. Limited Proteolysis MS. In limited proteolysis (LiP) mass spectrometry, proteins are briefly exposed to a promiscuous protease that cleaves exposed regions before denaturation and digestion for MS quantification.[276,277] The promiscuous protein will produce a cleavage pattern, or footprint, reflective of the structure of the protein or complex. The peptides detected by bottom-up proteomics will reflect these structure-dependent cleavages. By comparing between two samples, the effects of an external condition can be monitored. Picotti and co-workers have advanced LiP MS for the unbiased analysis of crude biological samples[277,278] by combining a promiscuous digestion step with a bottom-up proteomic workflow on a whole crude lysate. This approach has shed light on protein−metabolite interactions. By applying

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

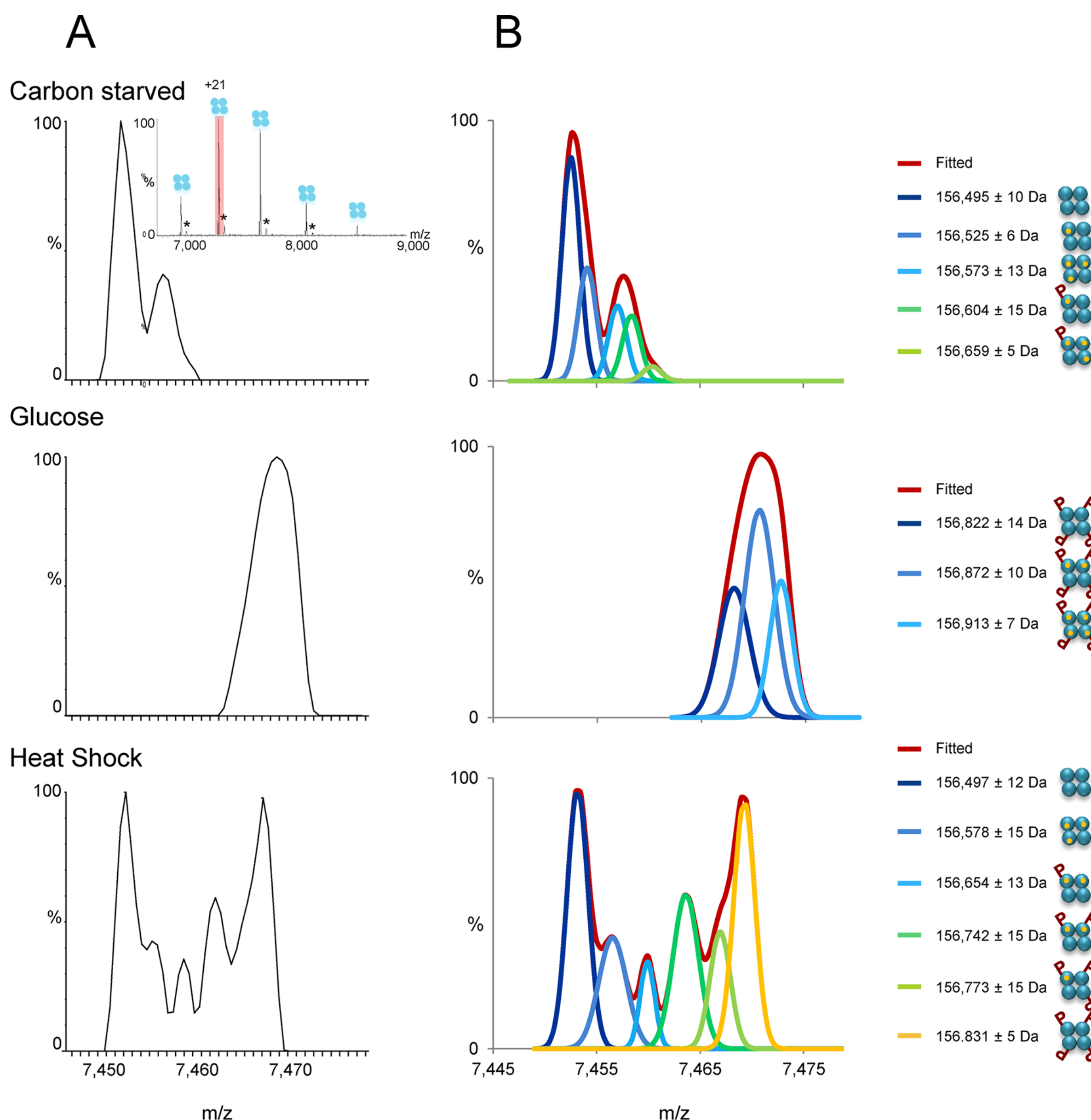


Figure 7. Top-down native MS identifies the proteoform composition of endogenous complexes. (A) Mass spectra of the intact intact fructose-1,6-bisphosphatase 1 (FBP1) homotetramer, purified endogenously from yeast grown under different conditions (carbon starved, glucose and heat shock). The inset shows the full charge series, and each panel displays the 21^+ charge state for each growth condition. Spectral deconvolution (B) reveals that glucose leads to a uniformly tetra-phosphorylated, while heat shock leads to a mixture of tetramers with different numbers of phosphorylations. MS/MS analysis (not shown) revealed that each subunit is monophosphorylated in the tetramer, while MS/MS/MS fragmentation analysis localized the modifications to the ^{12}Ser/^{13}Thr site. Magnesium ions, known binders of FBP1, were also observed bound to the complex. Each FBP1 subunit is graphically depicted as a cyan circle. Mg^{2+} ions are indicated as small orange circles, and phosphorylation is labeled as "P" Reproduced and adapted with permission from ref 291. Copyright 2017 American Chemical Society.

LiP to whole cell lysates[279] fractionated in the presence and absence of metabolites, they detected metabolite-dependent changes in homomeric and heteromeric protein complexes, including the ribosome. These changes could be confirmed via SEC-MS profiles of proteins in the presence and absence of cofactors.

3.4. "Top-Down" Approaches

3.4.1. Top-Down MS for PTM Analysis. In top-down MS, intact proteins or protein complexes are transferred directly into the gas phase.[280] Once in the gas phase, tandem MS can be used to cleave proteins into fragments for sequence analysis. Top-down MS is uniquely suited to identification of proteoforms.[95] This is because the mass of the intact proteoform is known, and thus modifications to individual amino acids can be cross-correlated with each other for an accurate description of the whole proteoform's simultaneous or mutually exclusive PTMs.

However, top-down proteomics methods are much less developed than bottom-up methods and face a range of challenges described in ref 281, including poor solubility of intact proteins, challenges in adequate fractionation of intact proteins, low sensitivity due to dynamic range issues, low-throughput, and the complexity of data analysis. Thus, top-down proteomics often detects only the most-abundant proteins or those present at lower molecular weights.

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

Additionally, there can be a lack of residue level sequence coverage due to poor fragmentation of intact proteins; efforts to increase fragmentation of intact proteins are ongoing.[282,283]

Top-down MS has been used to study PTMs to specific endogenous proteins, often in combination with affinity purification strategies. This strategy has been used to study glycan heterogeneity in plasma proteins,[284] phosphorylation of endogenous mouse cardiac troponin I,[285] post-translational modifications of pili proteins from *Neisseria meningitidis* upon host-cell contact,[286] and examined how endogenously isolated K-Ras proteoforms are correlated with mutational status.[287] Unbiased high throughput top-down proteomics workflows, pioneered by Kelleher and co-workers, have been used to study post-translational modifications in response to induction of DNA damage and treatment with chemical stressors[288] as well as PTMs of cardiac proteins upon myocardial infarctions.[289]

Proteoforms of whole protein complexes, correlating PTMs on one subunit to those on another, can be studied by combining native-MS ionization of whole complexes (see section 3.4.2) with top-down sequencing. Skinner et al. developed a native proteomics workflow to identify 217 proteoforms, including those of complexes such as dimeric creatine kinase M or DJ-1. Hybrid native and proteomic techniques were applied to identify the proteoforms of the C8 complex directly from human plasma.[290] In our lab, a multistage native MS strategy was used to examine how the endogenous yeast FBP1 complex adapts to glucose-rich conditions and heat shock[291] (see Figure 7). This study revealed that the complex responds differently to changes in growth conditions by tuning phosphorylation dynamics, thus illuminating the role of this PTM in FBP1. In another study, we used a top-down native approach to identify a new proteoform of the 20S proteasome PSMA7 subunit, a proteoform which would have eluded detection by bottom-up proteomics due to its proximity to a lysine-rich region subject to tryptic digestion.[292]

3.4.2. Native MS of Protein Complexes.

3.4.2.1. Overview of the Method. Native mass spectrometry takes the top-down approach one step further by retaining aspects of the protein's native tertiary and quaternary structure in the gas phase.[293] In native MS, conditions are optimized to ionize protein complexes while retaining noncovalent interactions, with experimental evidence confirming that aspects of native complex structure are maintained in the gas phase.[294] Complexes can then be broken down in the gas phase by tandem-MS techniques. Native-MS can have increased signal-to-noise relative to denatured top-down methods because folded proteins populate fewer charge state peaks, increasing the intensity of the remaining peaks.[295,296]

The addition of IM (see section 3.2.4) to the native-MS workflow yields information on the 3D shape of protein complexes.[134] IM separates ions according to their collision cross section, which reflects complex conformation and mass, providing a source of structural information.[125−127] This data can be combined with modeling tools that use collision cross sections to refine and develop structural models[297,298] (see section 3.5).

Native MS studies of protein complexes directly measure both composition and stoichiometry of protein complexes because quaternary structure is preserved in the gas phase (see Figure 4B for an example). Moreover, different coexisting complexes are separated in the spectrum based on their masses and can be compared in terms of their relative abundances for a particular sample preparation. Tandem-MS dissociates the complexes into their components to assign and confirm complex composition, to distinguish between core and peripheral subunits and determine subunit network. Contaminants are also easily distinguished because they are not part of the complex charge series.

Analyzing native-MS spectra involves assigning different peaks in the spectra to charge state series for proteins or complexes of interest. This is can be done manually or via computational spectral deconvolution. A variety of softwares exist to enable this, including open source softwares such as UniDec, which deconvolutes spectra using Bayesian inference,[299,300] NaVia, a program to augment manual assignment,[301] as well as vendor-provided software. Software for analyzing native IM-MS experiments presents several challenges, as reviewed in ref 298. Data must be interpreted to assign and extrapolate charge states as well as determine collision cross sections, which can then be used to develop structural models. Determining the collision cross section is typically done via a standard calibration curve,[302] with an open source automated collision cross section software recently introduced by Metz and co-workers.[303] Software has also been introduced to analyze IM experiments coupled to gas-phase activation or collision induced unfolding experiments.[304,305]

The main drawback of native-MS for determining protein complex composition is that relatively large quantities of complexes are needed for each experiment, with at least 10 μL of midnanomolar concentration required for a detailed native-MS experiment. Assignment of native MS peaks to particular complexes also requires some preknowledge of expected molecular weights in the sample. Therefore, it is useful to have preknowledge of the proteins present in solution and the exact molecular weights of the different subunits, information that can be obtained from proteomics and intact MS of denatured samples, respectively (see section 3.4).

Recent advances in native MS have the potential to extend its applicability to endogenous complexes available in limited quantities. One example is the development of individual ion MS via charge detection mass spectrometry,[306] in which the ion charge is measured directly to generate a mass spectrum rather than a spectrum showing mass-to-charge ratio. Implemented on various experimental platforms including Orbitrap[307] and ion trap FT instruments,[308] this method has provided isotope resolved spectra of complexes as large as β-galactosidase, 466 kDa. This can enable very high resolution spectra of endogenous assemblies in complex mixtures because the measurement of an ion's true mass resolves ambiguities in crowded spectra.

3.4.2.2. Application to Endogenous Complexes Analysis. Endogenous affinity purification followed by native-MS has been used by multiple laboratories to characterize the yeast exosome.[309−312] This work has determined both the stoichiometry of various exosome components as well as compositional differences between nuclear and cytoplasmic exosome complexes. Synowsky and Heck showed via native-MS that the Ski complex, which associates with the yeast exosome, is a heterotetramer as opposed to a heterodimer as previously assumed.[313] Novel complexes have also been directly identified via native-MS; using endogenous purification from muscle tissue followed by native-MS, we identified novel complexes consisting of Arp2/3 subunits and vinculin, likely involved in the formation of focal adhesion complexes[18] (Figure 4B). New subunits for existing complexes can also be

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

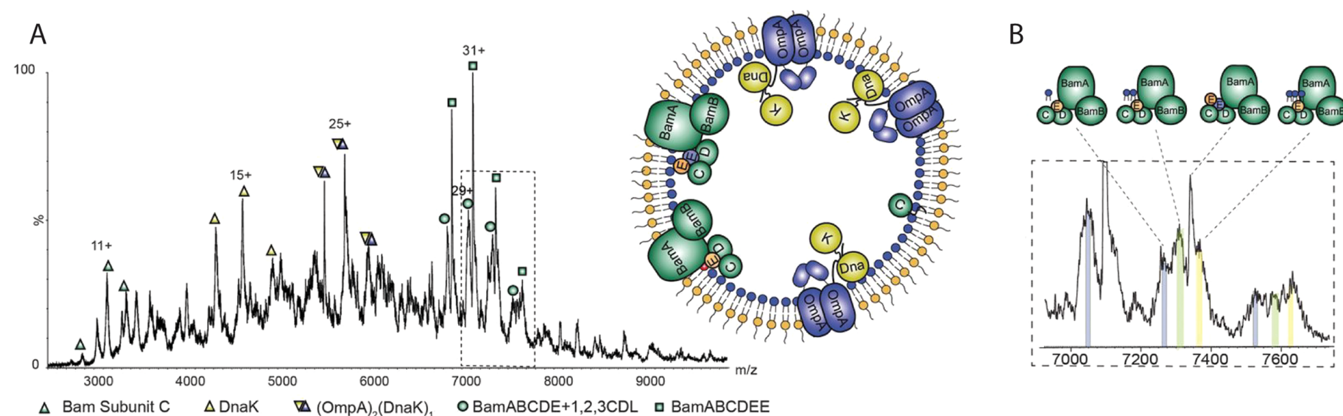


Figure 8. Native-MS spectra show lipid binding to endogenous membrane proteins. (A) Full spectrum of protein complexes observed from the *E. coli* outer membrane. The inset depicts observed complexes of an outer membrane vesicle. (B) Expansion of the mass spectrum assigned to the Bam complex (boxed region in (A)), with monomeric BamE binding to one, two, and three cardiolipins (gray, green, and yellow, respectively). Native-MS thus uncovers the range of lipid bound complexes present in the sample. Reproduced with permission from ref 315. Copyright 2018 The American Association for the Advancement of Science.

seen in native-MS spectra. For example, COP9 signalosome (CSN) complexes purified directly from erythrocytes harbored a novel subunit, named CSNAP, that had previously evaded detection.[17] Semi-high throughput native proteomics workflows, analogous to the fractionation-MS bottom-up methods described above, have also been developed.[19,25] Intriguingly, Skinner et al.[19] detected novel homomers, which often elude bottom-up proteomic approaches that rely on indirect measurement of protein association.

In addition to soluble complexes, native MS can be applied to membrane proteins, work pioneered by Robinson and coworkers.[314] By ejecting endogenous membrane complexes directly from native membranes,[315] complexes of the chaperone DnaK and the outer membrane protein OmpA were observed, confirming the involvement of DnaK in membrane porin assembly. Endogenous membrane proteins can also be purified, and Gross and co-workers studied endogenous photosynthetic reaction centers via native-MS in detergent micelles.[316]

Moreover, the method can be applied to very large protein complexes such as the ribosome (2.30 mDa) and its subcomplexes. Native and tandem-MS of the ribosome stalk complexes from a wide range of organisms revealed differences in stalk composition and stoichiometry.[317] Intact protein MS of ribosomes examined the stoichiometry of different ribosomal proteins, revealing heterogeneity and substoichiometric binding of ribosome associated proteins.[318,319] Affinity captured NPC endogenous complexes, which are upward of 80 megadaltons, were analyzed by native MS, confirming the overall stoichiometry and membrane components of the NPC complex.[183]

Because protein complexes retain their solution noncovalent interactions, complex topology can be determined by analyzing the pattern of subunit dissociation in tandem native MS experiments (Figure 5B). Generation of subcomplexes can be aided by addition of small amounts of methanol or other solvents that disrupt hydrophobic interactions, thus encouraging dissociation across weaker quaternary interfaces. Subunit connectivity is then determined by assuming that when large complexes dissociate, they generate subcomplexes from proteins already proximal in the intact structure. This method was used to determine the topologies of a range of endogenous yeast complexes,[310−313,320] including structural models of the yeast exosome that were later confirmed by crystallographic studies.[310,320]

As discussed above, IM can also be used to derive structural information on endogenous complexes. While IM has not been broadly applied to purified endogenous complexes, our lab recently characterized orthologous endogenous 20S proteasome complexes from a range of species using IM, demonstrating an enlargement of the complex size from the archaeal prokaryotic complex to the eukaryotic 20S proteasomes in yeast and mammals.[292]

Protein complexes are often modulated by cofactor binding. Native mass spectrometry is uniquely suited to the analysis of cofactor/complex binding because cofactors will appear as mass shifts in the MS spectrum of the complex. Moreover, a single native-MS experiment can reveal the full range of coexisting complexes, allowing relative ratios to be quantified in a single experiment.[321] Skinner et al. in their top-down proteomic study[19] demonstrated that the holoenzyme of α-enolase was only present in the dimer, as opposed to the monomer, indicating that Mg stabilizes the dimer, and that GDP and magnesium supported formation of the RHOA heterodimers with RHOGDI1. In a similar study from *E. coli*, Shen and co-workers[25] identified protein complexes bound to metals and glutamine.

The impact of biomolecule binding can be also studied in the context of membrane protein complexes. Robinson and coworkers have made seminal contributions to the use of mass spectrometry to study the binding of lipids to protein complexes, many of them endogenously isolated from host membranes[241,315] (Figure 8). Membrane proteins display preferences for which endogenous lipids they bind,[322] and special tandem MS methodologies enable the identification of lipids directly from samples of purified membrane proteins.[323] These experiments have revealed the stoichiometry of lipid-membrane protein binding[241] and investigated the effect of lipids on membrane protein stability.[324] For example, in complexes ejected directly from endogenous membranes, the BAM complex from *E. coli* membranes was observed bound specifically to cardiolipin (Figure 8), while ANT-1 from the mitochondrial membrane bound fatty acids with palmitate headgroups, among others.[315] These studies are often

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

complementary to high-resolution cryo-EM investigations of membrane proteins, where the identity of the bound lipid is unknown.

Native MS can also be accomplished directly from crude samples, detecting abundant proteins without the need for purification.[325] This approach uses the limited dynamic range of MS as an advantage, and clear spectra can be acquired of abundant proteins. In an extension of the direct-MS approach, native MS was recently applied to individual erythrocytes yielding spectra of hemoglobin, which constitutes the vast majority of protein in erthrocytes.[326] Individual cells were selected and a microinjector used to transfer the cells into an MS emitter. However, these methods are limited to very abundant proteins; to apply this technique to a broader range of proteins, online separation and increased signal-to-noise must be applied.

3.5. Integrative Approaches and Computational Modeling

As described above, each MS method provides specific information about endogenous complexes, and each has their own pros and cons. The information provided from different MS experiments is complementary; for example, bottom-up proteomics can be exquisitely sensitive to complex components while native-MS directly detects whole complexes. Therefore, it is often advantageous to combine multiple MS modalities to derive a more complete functional and structural model of a particular endogenous complex. Termed "hybrid" mass spectrometric approaches,[229,319] in this method, information from different experimental modalities is combined to characterize complexes. For example, Robinson, Aebersold, and co-workers combined information from native-MS, bottom-up proteomics, chemical cross-linking, and IM to derive integrative models of the 26S proteasome lid complex and its submodules.[229] Similarly, Heck and co-workers used hybrid MS approaches to analyze a ribosomes[319] from multiple kingdoms of life. For native MS experiments that determine complex stoichiometry, bottom-up proteomics are often important to identify the composing subunits,[310,313] with complex stoichiometry determined by native-MS measurement of the intact complex.

Moreover, information from MS can be used in combination with computational molecular modeling tools, with MS data serving as a restraint for the generation of complete structural models. Cross-linking MS is a natural companion to computational modeling, as described in ref 222. Chemical labeling methodologies can also be combined with computational methods to determine protein structures and topographies. Huang et al. advanced the field by converting hydroxyl radical footprinting data into a protection factor that can be used as a restraint in molecular modeling.[327] Similarly, hydroxyl radical footprinting data has been used to differentiate between molecular models[328] and combined with molecular dynamics data for model determination data.[329] IM can also be combined with modeling tools that use the collision cross section as a restraint to refine and develop structural models.[297,298,330] Because of the sensitivity and robustness of these methods, they hold great promise for the study of endogenous complexes available in limited quantity.

Moreover, modeling tools form the basis of integrative structural biology,[332] of which MS is often an important component. As described in refs 332 and 333, this experimental approach combines data from multiple experimental and theoretical methods to build a model of a large biological system for which no single method is sufficient. A wide range of experimental data can be used, including different MS modalities, cryo-EM, SAXS, FRET, and protein–protein interaction data. Different types of experimental and theoretical data are phrased as restraints, and an algorithm driven by a scoring function determines the structural and dynamic model that satisfies most data. Central to this experimental approach is flexible software that can generate the complete range of structures that satisfy the available data. Popular available programs include Rosetta[334] and the Integrative Modeling Platform (IMP),[335] although there are other options available that may be tailored to a particular experimental tool, as described in refs 333 and 336.

Several important and extremely large structural models have been built using integrative tools that use MS data as a central modality. For example, MS data has played an integral role in the structural biology of the nuclear pore complex. One of the first integrative models was reported in 2007 by Alber and co-workers,[337] later updated in 2018,[183] combining information from proteomic affinity experiments, charge detection native-MS, cross-linking MS, cryo-EM, among other information, to develop a full structural model of the complete NPC, which is greater than 40 mDa. Beck and co-workers also used integrative modeling, relying again on cross-linking MS and cryo-EM tomography, to analyze the full nuclear pore complex[338] and build a model of the inner scaffold ring.[339] Cross-linking MS along with cryo-EM data was used to describe the architecture of the 10 mega Da pyruvate dehydrogenase supercomplex.[331] Other examples of structural models relying on MS data include the 26S proteasome,[230] constructed by combining cross-linking MS, proteomics interaction data, and cryo-EM, and the human Polycomb repressive complex 2 complex, generated primarily by cryo-EM but also with cross-linking MS data.[340]

Another related area is the increasing use of MS for structural proteomics efforts that combine MS and cryo-EM for the analysis of cryo-EM grids prepared from crude fractions as reviewed in ref 341. MS identification of components in SEC or sucrose gradient fractions is key to enabling the fitting of electron density maps to specific proteins.[342−344] Integrating the MS methods described here with these and other emerging structural biology techniques will only further the study of endogenous complexes.

4. CONCLUSION AND FUTURE DIRECTIONS

Biological MS has made seminal contributions to the analysis of endogenous protein complexes. Depending on the technology chosen, insights range from characterizing the interaction partners of a specific protein to determining the topology of endogenous complexes to analyzing conformational changes between different states. Importantly, by determining these properties for endogenously isolated protein complexes, results are generated that report on biologically relevant interactions and conformations.

Over the course of this review, it has been repeatedly highlighted that one of the main challenges in characterizing endogenous complexes is their low abundance, which is in part what makes MS a useful tool for characterizing such complexes. One of the main solutions to this issue is to process larger amounts of starting material. While it is relatively straightforward to scale up yeast and bacterial growth protocols, even in an academic setting, it is more difficult and costly for eukaryotic cells. HEK293 cells and other strains have

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

been adapted to growth in high density suspension culture but may not be the most relevant physiologically due to altered gene expression.[345] Importantly, for native-MS, emerging research indicates that enrichment, rather than complete purification, can yield well-resolved spectra of proteins.[325] Thus, for native MS efforts where intact complexes are detected, proteins may not need to be fully purified for complex identification.

Advances in mass spectrometer technology are also increasing the sensitivity of measurements and enabling application of biological MS to less abundant proteins. Single-molecule MS,[346] in which both charge and m/z ratio are measured to directly detect the mass of the ion, was recently applied to native macromolecular protein assemblies on Orbitrap instruments[306,347] by two different groups. Single-particle charge detection not only increases the sensitivity of native-MS but also permits the resolution of overlapped charge states in heterogeneous samples. Very small amounts of sample can be analyzed with novel microfluidic devices.[348] For bottom-up proteomics, new instruments that includes variants of IM devices can greatly increase sensitivity and sequencing speed[145] and were recently shown to identify proteins from single cells.[349,350] Applying these technologies to the analysis of protein complexes will require miniaturization of purification techniques but can potentially enable comparison of complexes across different tissues available in limited amounts.

This latter goal, namely understanding how complexes vary between tissues and physiological states, is a crucial frontier for biological MS. Thus far, however, characterization of endogenous complexes has largely been limited to model organisms such as yeast and model cell lines such as HEK and HeLa. Extending this approach to the analysis of protein complexes and proteoforms from a range of tissues can be enabled by CRISPR genome tagging approaches, which permit affinity labeling of proteins from a broader range of germ lines. It is also conceivable to construct affinity tagged mice and purify and compare complexes from a broad range of tissues. We expect that in the future biological mass spectrometry will be applied to a broader range of organisms and under healthy and disease states.

It is our hope that this review, while broad and general in its outlook, will serve as a guide to scientists interested in studying endogenous protein complexes and will stimulate the application of biological MS to a broader range of problems.

AUTHOR INFORMATION

Corresponding Author

Michal Sharon − *Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot 7610001, Israel*; orcid.org/0000-0003-3933-0595; Phone: 972-8-9343947; Email: michal.sharon@weizmann.ac.il

Author

Rivkah Rogawski − *Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot 7610001, Israel*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.chemrev.1c00217

Notes

The authors declare no competing financial interest.

Biographies

Rivkah Rogawski received her B.A. in Chemistry from Yeshiva University in 2011. She then completed her Ph.D. at Columbia University under the supervision of Professor Ann McDermott, focusing on dynamic nuclear polarization solid-state NMR and its application to protein characterization in complex mixtures. She then received a Fulbright Postdoctoral Fellowship to join the lab of Michal Sharon at the Weizmann Institute of Science, where her research focuses on applications of native mass spectrometry to protein characterization.

Michal Sharon received a B.Sc. in Chemistry from the Hebrew University of Jerusalem. Her Ph.D. studies at the Weizmann Institute of Science focused on studying the three-dimensional structure of proteins by NMR. In 2003, she became a postdoctoral fellow at the University of Cambridge, UK, in the lab of Prof. Carol Robinson. During this period, she developed approaches to analyze large protein complexes by mass spectrometry. The Prof. Michal Sharon research lab focuses on investigating the mechanisms that control and coordinate the activity of molecular machines involved in the protein degradation pathway. The group applies novel native mass spectrometry approaches, in conjunction with fluorescence microscopy, biochemical, and cell biology methods, generating an integrative mode of analysis combining in vitro and in vivo findings.

ACKNOWLEDGMENTS

M.S. is grateful for the support of an (300/17), a Sagol Institute for Longevity Research grant and a Moross Proof-of-Concept grant. M.S. is the incumbent of the Aharon and Ephraim Katzir Memorial Professorial Chair. R.R. was a Fulbright postdoctoral fellow for the years 2019−2020.

REFERENCES

(1) Alberts, B. The Cell as a Collection of Protein Machines: Preparing the Next Generation of Molecular Biologists. *Cell* **1998**, *92*, 291−294.

(2) Giurgiu, M.; Reinhard, J.; Brauner, B.; Dunger-Kaltenbach, I.; Fobo, G.; Frishman, G.; Montrone, C.; Ruepp, A. CORUM: The Comprehensive Resource of Mammalian Protein Complexes - 2019. *Nucleic Acids Res.* **2019**, *47*, D559−D563.

(3) Stumpf, M. P. H.; Thorne, T.; De Silva, E.; Stewart, R.; An, H. J.; Lappe, M.; Wiuf, C. Estimating the Size of the Human Interactome. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 6959−6964.

(4) Chen, Z. A.; Rappsilber, J. Quantitative Cross-Linking/Mass Spectrometry to Elucidate Structural Changes in Proteins and Their Complexes. *Nat. Protoc.* **2019**, *14*, 171−201.

(5) Benesch, J. L. P.; Ruotolo, B. T.; Simmons, D. A.; Robinson, C. V. Protein Complexes in the Gas Phase: Technology for Structural Genomics and Proteomics. *Chem. Rev.* **2007**, *107*, 3544−3567.

(6) Gillet, L. C.; Leitner, A.; Aebersold, R. Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing. *Annu. Rev. Anal. Chem.* **2016**, *9*, 449−472.

(7) Hakhverdyan, Z.; Domanski, M.; Hough, L. E.; Oroskar, A. A.; Oroskar, A. R.; Keegan, S.; Dilworth, D. J.; Molloy, K. R.; Sherman, V.; Aitchison, J. D.; et al. Rapid, Optimized Interactomic Screening. *Nat. Methods* **2015**, *12*, 553−560.

(8) Peth, A.; Berndt, C.; Henke, W.; Dubiel, W. Downregulation of COP9 Signalosome Subunits Differentially Affects the CSN Complex and Target Protein Stability. *BMC Biochem.* **2007**, *8*, 27.

(9) Cox, B.; Emili, A. Tissue Subcellular Fractionation and Protein Extraction for Use in Mass-Spectrometry-Based Proteomics. *Nat. Protoc.* **2006**, *1*, 1872−1878.

(10) Masuda, T.; Sugiyama, N.; Tomita, M.; Ohtsuki, S.; Ishihama, Y. Mass Spectrometry-Compatible Subcellular Fractionation for Proteomics. *J. Proteome Res.* **2020**, *19*, 75−84.

(11) Liu, C. W.; Li, X.; Thompson, D.; Wooding, K.; Chang, T.; Tang, Z.; Yu, H.; Thomas, P. J.; DeMartino, G. N. ATP Binding and ATP Hydrolysis Play Distinct Roles in the Function of 26S Proteasome. *Mol. Cell* **2006**, *24*, 39−50.
(12) Ben-Nissan, G.; Vimer, S.; Tarnavsky, M.; Sharon, M. Structural Mass Spectrometry Approaches to Study the 20S Proteasome. *Methods Enzymol.* **2019**, *619*, 179−223.
(13) Iacobucci, C.; Götze, M.; Ihling, C. H.; Piotrowski, C.; Arlt, C.; Schäfer, M.; Hage, C.; Schmidt, R.; Sinz, A. A Cross-Linking/Mass Spectrometry Workflow Based on MS-Cleavable Cross-Linkers and the MeroX Software for Studying Protein Structures and Protein−Protein Interactions. *Nat. Protoc.* **2018**, *13*, 2864−2889.
(14) Donnelly, D. P.; Rawlins, C. M.; DeHart, C. J.; Fornelli, L.; Schachner, L. F.; Lin, Z.; Lippens, J. L.; Aluri, K. C.; Sarin, R.; Chen, B.; et al. Best Practices and Benchmarks for Intact Protein Analysis for Top-down Mass Spectrometry. *Nat. Methods* **2019**, *16*, 587−594.
(15) Burgess, R. R.; Deutscher, M. P. Guide to Protein Purification. *Methods Enzymol.* **2009**, *463*, xxv−xxvi.
(16) Greber, B. J.; Bieri, P.; Leibundgut, M.; Leitner, A.; Aebersold, R.; Boehringer, D.; Ban, N. The Complete Structure of the 55S Mammalian Mitochondrial Ribosome. *Science* **2015**, *348*, 303−308.
(17) Rozen, S.; Füzesi-Levi, M. G.; Ben-Nissan, G.; Mizrachi, L.; Gabashvili, A.; Levin, Y.; Ben-Dor, S.; Eisenstein, M.; Sharon, M. CSNAP Is a Stoichiometric Subunit of the COP9 Signalosome. *Cell Rep.* **2015**, *13*, 585−598.
(18) Chorev, D. S.; Moscovitz, O.; Geiger, B.; Sharon, M. Regulation of Focal Adhesion Formation by a Vinculin-Arp2/3 Hybrid Complex. *Nat. Commun.* **2014**, *5*, 3758.
(19) Skinner, O. S.; Haverland, N. A.; Fornelli, L.; Melani, R. D.; Do Vale, L. H. F.; Seckler, H. S.; Doubleday, P. F.; Schachner, L. F.; Srzentić, K.; Kelleher, N. L.; et al. Top-down Characterization of Endogenous Protein Complexes with Native Proteomics. *Nat. Chem. Biol.* **2018**, *14*, 36−41.
(20) 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. A Census of Human Soluble Protein Complexes. *Cell* **2012**, *150*, 1068−1081.
(21) Wan, C.; Borgeson, B.; Phanse, S.; Tu, F.; Drew, K.; Clark, G.; Xiong, X.; Kagan, O.; Kwan, J.; Bezginov, A.; et al. Panorama of Ancient Metazoan Macromolecular Complexes. *Nature* **2015**, *525*, 339−334.
(22) McBride, Z.; Chen, D.; Lee, Y.; Aryal, U. K.; Xie, J.; Szymanski, D. B. A Label-Free Mass Spectrometry Method to Predict Endogenous Protein Complex Composition. *Mol. Cell. Proteomics* **2019**, *18*, 1588−1606.
(23) Kirkwood, K. J.; Ahmad, Y.; Larance, M.; Lamond, A. I. Characterization of Native Protein Complexes and Protein Isoform Variation Using Size- Fractionation-Based Quantitative Proteomics. *Mol. Cell. Proteomics* **2013**, *12*, 3851−3873.
(24) Skinner, O. S.; Do Vale, L. H. F.; Catherman, A. D.; Havugimana, P. C.; Sousa, M. V. De; Compton, P. D.; Kelleher, N. L. Native GELFrEE: A New Separation Technique for Biomolecular Assemblies. *Anal. Chem.* **2015**, *87*, 3032−3038.
(25) Shen, X.; Kou, Q.; Guo, R.; Yang, Z.; Chen, D.; Liu, X.; Hong, H.; Sun, L. Native Proteomics in Discovery Mode Using Size-Exclusion Chromatography-Capillary Zone Electrophoresis-Tandem Mass Spectrometry. *Anal. Chem.* **2018**, *90*, 10095−10099.
(26) Federspiel, J. D.; Cristea, I. M. Considerations for Identifying Endogenous Protein Complexes from Tissue via Immunoaffinity Purification and Quantitative Mass Spectrometry. *Methods Mol. Biol.* **2019**, *1977*, 115−143.
(27) Bousquet-Dubouch, M. P.; Baudelet, E.; Guérin, F.; Matondo, M.; Uttenweiler-Joseph, S.; Burlet-Schiltz, O.; Monsarrat, B. Affinity Purification Strategy to Capture Human Endogenous Proteasome Complexes Diversity and to Identify Proteasome-Interacting Proteins. *Mol. Cell. Proteomics* **2009**, *8*, 1150−1164.
(28) Jarvik, J. W.; Telmer, C. A. Epitope Tagging. *Annu. Rev. Genet.* **1998**, *32*, 601−618.
(29) Kimple, M. E.; Brill, A. L.; Pasker, R. L. Overview of Affinity Tags for Protein Purification. *Curr. Protoc. Protein Sci.* **2013**, *73*, 9.9.1.
(30) Young, C. L.; Britton, Z. T.; Robinson, A. S. Recombinant Protein Expression and Purification: A Comprehensive Review of Affinity Tags and Microbial Applications. *Biotechnol. J.* **2012**, *7*, 620−634.
(31) Rigaut, G.; Shevchenko, A.; Rutz, B.; Wilm, M.; Mann, M.; Seraphin, B. A Generic Protein Purification Method for Protein Complex Characterization and Proteome Exploration. *Nat. Biotechnol.* **1999**, *17*, 1030−1032.
(32) Puig, O.; Caspary, F.; Rigaut, G.; Rutz, B.; Bouveret, E.; Bragado-Nilsson, E.; Wilm, M.; Séraphin, B. The Tandem Affinity Purification (TAP) Method: A General Procedure of Protein Complex Purification. *Methods* **2001**, *24*, 218−229.
(33) Babu, M.; Butl, G.; Pogoutse, O.; Li, J.; Greenblatt, J. F.; Emili, A. Sequential Peptide Affinity Purification System for Systematic Isolation and Identification of Protein Complexes from Escherichia Coli. *Methods Mol. Biol.* **2009**, *564*, 373−400.
(34) Nakatani, Y.; Ogryzko, V. Immunoaffinity Purification of Mammalian Protein Complexes. *Methods Enzymol.* **2003**, *370*, 430−444.
(35) Pappenberger, G.; McCormack, E. A.; Willison, K. R. Quantitative Actin Folding Reactions Using Yeast CCT Purified via an Internal Tag in the CCT3/γ Subunit. *J. Mol. Biol.* **2006**, *360*, 484−496.
(36) Field, J.; Nikawa, J.; Broek, D.; MacDonald, B.; Rodgers, L.; Wilson, I. A.; Lerner, R. A.; Wigler, M. Purification of a RAS-Responsive Adenylyl Cyclase Complex from Saccharomyces Cerevisiae by Use of an Epitope Addition Method. *Mol. Cell. Biol.* **1988**, *8*, 2159−2165.
(37) Nilsson, B.; Moks, T.; Jansson, B.; Abrahmsén, L.; Elmblad, A.; Holmgren, E.; Henrichson, C.; Jones, T. A.; Uhlén, M. A Synthetic IgG-Binding Domain Based on Staphylococcal Protein A. *Protein Eng., Des. Sel.* **1987**, *1*, 107−113.
(38) Hopp, T. P.; Prickett, K. S.; Price, V. L.; Libby, R. T.; March, C. J.; Cerretti, D. P.; Urdal, D. L.; Conlon, P. J. A Short Polypeptide Marker Sequence Useful for Recombinant Protein Identification and Purification. *Bio/Technology* **1988**, *6*, 1204−1210.
(39) Schmidt, T. G. M.; Skerra, A. The Strep-Tag System for One-Step Purification and High-Affinity Detection or Capturing of Proteins. *Nat. Protoc.* **2007**, *2*, 1528−1535.
(40) Hochuli, E.; Bannwarth, W.; Dobeli, H.; Gentz, R.; Stuber, D. Genetic Approach To Facilitate Purification of Recombinant Proteins With a Novel Metal Chelate Adsorbent. *Nat. Biotechnol.* **1988**, *6*, 1321−1325.
(41) Court, D. L.; Sawitzke, J. A.; Thomason, L. C. Genetic Engineering Using Homologous Recombination. *Annu. Rev. Genet.* **2002**, *36*, 361−388.
(42) Gelot, C.; Le-Guen, T.; Ragu, S.; Lopez, B. S. Chapter 20-Double-Strand Break Repair: Homologous Recombination in Mammalian Cells. *Genome Stab.* **2016**, 337−351.
(43) Hsu, P. D.; Lander, E. S.; Zhang, F. Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* **2014**, *157*, 1262−1278.
(44) Jiang, W.; Bikard, D.; Cox, D.; Zhang, F.; Marraffini, L. A. RNA-Guided Editing of Bacterial Genomes Using CRISPR-Cas Systems. *Nat. Biotechnol.* **2013**, *31*, 233−239.
(45) Pyne, M. E.; Moo-Young, M.; Chung, D. A.; Chou, C. P. Coupling the CRISPR/Cas9 System with Lambda Red Recombineering Enables Simplified Chromosomal Gene Replacement in Escherichia Coli. *Appl. Environ. Microbiol.* **2015**, *81*, 5103−5114.
(46) Buchmuller, B. C.; Herbst, K.; Meurer, M.; Kirrmaier, D.; Sass, E.; Levy, E. D.; Knop, M. Pooled Clone Collections by Multiplexed CRISPR-Cas12a-Assisted Gene Tagging in Yeast. *Nat. Commun.* **2019**, *10*, 2960.
(47) Dewari, P. S.; Southgate, B.; Mccarten, K.; Monogarov, G.; O'Duibhir, E.; Quinn, N.; Tyrer, A.; Leitner, M.-C.; Plumb, C.; Kalantzaki, M.; et al. An Efficient and Scalable Pipeline for Epitope

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

Tagging in Mammalian Stem Cells Using Cas9 Ribonucleoprotein. *eLife* **2018**, *7*, e35069.
(48) Dalvai, M.; Loehr, J.; Jacquet, K.; Huard, C. C.; Roques, C.; Herst, P.; Côté, J.; Doyon, Y. A Scalable Genome-Editing-Based Approach for Mapping Multiprotein Complexes in Human Cells. *Cell Rep.* **2015**, *13*, 621−633.
(49) Savic, D.; Partridge, E. C.; Newberry, K. M.; Smith, S. B.; Meadows, S. K.; Roberts, B. S.; Mackiewicz, M.; Mendenhall, E. M.; Myers, R. M. CETCh-Seq: CRISPR Epitope Tagging ChIP-Seq of DNA-Binding Proteins. *Genome Res.* **2015**, *25* (10), 1581−1589.
(50) Lin, D. W.; Chung, B. P.; Huang, J. W.; Wang, X.; Huang, L.; Kaiser, P. Microhomology-Based CRISPR Tagging Tools for Protein Tracking, Purification, and Depletion. *J. Biol. Chem.* **2019**, *294*, 10877−10885.
(51) Lyu, Q.; Dhagia, V.; Han, Y.; Guo, B.; Wines-Samuelson, M. E.; Christie, C. K.; Yin, Q.; Slivano, O. J.; Herring, P.; Long, X.; et al. CRISPR-Cas9-Mediated Epitope Tagging Provides Accurate and Versatile Assessment of Myocardin-Brief Report. *Arterioscler., Thromb., Vasc. Biol.* **2018**, *38*, 2184−2190.
(52) Zhang, X. H.; Tee, L. Y.; Wang, X. G.; Huang, Q. S.; Yang, S. H. Off-Target Effects in CRISPR/Cas9-Mediated Genome Engineering. *Mol. Ther.–Nucleic Acids* **2015**, *4*, No. e264.
(53) Yu, D.; Ellis, H. M.; Lee, E. C.; Jenkins, N. A.; Copeland, N. G.; Court, D. L. An Efficient Recombination System for Chromosome Engineering in Escherichia Coli. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 5978−5983.
(54) Sharan, S. K.; Thomason, L. C.; Kuznetsov, S. G.; Court, D. L. Recombineering: A Homologous Recombination-Based Method of Genetic Engineering. *Nat. Protoc.* **2009**, *4*, 206−223.
(55) Schneider, B. L.; Seufert, W.; Steiner, B.; Yang, Q. H.; Futcher, A. B. Use of Polymerase Chain Reaction Epitope Tagging for Protein Tagging in Saccharomyces Cerevisiae. *Yeast* **1995**, *11*, 1265−1274.
(56) Bähler, J.; Wu, J. Q.; Longtine, M. S.; Shah, N. G.; McKenzie, A.; Steever, A. B.; Wach, A.; Philippsen, P.; Pringle, J. R. Heterologous Modules for Efficient and Versatile PCR-Based Gene Targeting in Schizosaccharomyces Pombe. *Yeast* **1998**, *14*, 943−951.
(57) Booher, K. R.; Kaiser, P. A PCR-Based Strategy to Generate Yeast Strains Expressing Endogenous Levels of Amino-Terminal Epitope-Tagged Proteins. *Biotechnol. J.* **2008**, *3*, 524−529.
(58) Knop, M.; Siegers, K.; Pereira, G.; Zachariae, W.; Winsor, B.; Nasmyth, K.; Schiebel, E. Epitope Tagging of Yeast Genes Using a PCR-Based Strategy: More Tags and Improved Practical Routines. *Yeast* **1999**, *15*, 963−972.
(59) Zhang, Y.; Serratore, N. D.; Briggs, S. D. N-ICE Plasmids for Generating N-Terminal 3 × FLAG Tagged Genes That Allow Inducible, Constitutive or Endogenous Expression in Saccharomyces Cerevisiae. *Yeast* **2017**, *34*, 223−235.
(60) Yofe, I.; Weill, U.; Meurer, M.; Chuartzman, S.; Zalckvar, E.; Goldman, O.; Ben-Dor, S.; Schütze, C.; Wiedemann, N.; Knop, M.; et al. One Library to Make Them All: Streamlining the Creation of Yeast Libraries via a SWAp-Tag Strategy. *Nat. Methods* **2016**, *13*, 371−378.
(61) Meurer, M.; Duan, Y.; Sass, E.; Kats, I.; Herbst, K.; Buchmuller, B. C.; Dederer, V.; Huber, F.; Kirrmaier, D.; Štefl, M.; et al. Genome-Wide C-SWAT Library for High-Throughput Yeast Genome Tagging. *Nat. Methods* **2018**, *15*, 598−600.
(62) Poser, I.; Sarov, M.; Hutchins, J. R A; Heriche, J.-K.; Toyoda, Y.; Pozniakovsky, A.; Weigl, D.; Nitzsche, A.; Hegemann, B.; Bird, A. W BAC TransgeneOmics: A High-Throughput Method for Exploration of Protein Function in Mammals. *Nat. Methods* **2008**, *5*, 409−415.
(63) Andersen, P. R.; Domanski, M.; Kristiansen, M. S.; Storvall, H.; Ntini, E.; Verheggen, C.; Schein, A.; Bunkenborg, J.; Poser, I.; Hallais, M.; et al. The Human Cap-Binding Complex Is Functionally Connected to the Nuclear RNA Exosome. *Nat. Struct. Mol. Biol.* **2013**, *20*, 1367−1376.
(64) Van Craenenbroeck, K.; Vanhoenacker, P.; Haegeman, G. Episomal Vectors for Gene Expression in Mammalian Cells. *Eur. J. Biochem.* **2000**, *267*, 5665−5678.
(65) Sadelain, M.; Papapetrou, E. P.; Bushman, F. D. Safe Harbours for the Integration of New DNA in the Human Genome. *Nat. Rev. Cancer* **2012**, *12*, 51−58.
(66) DeKelver, R. C.; Choi, V. M.; Moehle, E. A.; Paschon, D. E.; Hockemeyer, D.; Meijsing, S. H.; Sancak, Y.; Cui, X.; Steine, E. J.; Miller, J. C.; et al. Functional Genomics, Proteomics, and Regulatory DNA Analysis in Isogenic Settings Using Zinc Finger Nuclease-Driven Transgenesis into a Safe Harbor Locus in the Human Genome. *Genome Res.* **2010**, *20*, 1133−1142.
(67) Elegheert, J.; Behiels, E.; Bishop, B.; Scott, S.; Woolley, R. E.; Griffiths, S. C.; Byrne, E. F. X.; Chang, V. T.; Stuart, D. I.; Jones, E. Y.; et al. Lentiviral Transduction of Mammalian Cells for Fast, Scalable and High-Level Production of Soluble and Membrane Proteins. *Nat. Protoc.* **2018**, *13*, 2991−3017.
(68) Yu, C.; Huang, L. Cross-Linking Mass Spectrometry: An Emerging Technology for Interactomics and Structural Biology. *Anal. Chem.* **2018**, *90*, 144−165.
(69) Chen, Z. A.; Rappsilber, J. Protein Dynamics in Solution by Quantitative Crosslinking/Mass Spectrometry. *Trends Biochem. Sci.* **2018**, *43*, 908−920.
(70) Iacobucci, C.; Götze, M.; Sinz, A. Cross-Linking/Mass Spectrometry to Get a Closer View on Protein Interaction Networks. *Curr. Opin. Biotechnol.* **2020**, *63*, 48−53.
(71) Zhang, H.; Tang, X.; Munske, G. R.; Zakharova, N.; Yang, L.; Zheng, C.; Wolff, M. A.; Tolic, N.; Anderson, G. A.; Shi, L.; et al. In Vivo Identification of the Outer Membrane Protein OmcA-MtrC Interaction Network in Shewanella Oneidensis MR-1 Cells Using Novel Hydrophobic Chemical Cross-Linkers. *J. Proteome Res.* **2008**, *7*, 1712−1720.
(72) Bartolec, T. K.; Smith, D. L.; Pang, C. N. I.; Xu, Y. D.; Hamey, J. J.; Wilkins, M. R. Cross-Linking Mass Spectrometry Analysis of the Yeast Nucleus Reveals Extensive Protein-Protein Interactions Not Detected by Systematic Two-Hybrid or Affinity Purification-Mass Spectrometry. *Anal. Chem.* **2020**, *92*, 1874−1882.
(73) Belsom, A.; Rappsilber, J. Anatomy of a Crosslinker. *Curr. Opin. Chem. Biol.* **2021**, *60*, 39−46.
(74) Subbotin, R. I.; Chait, B. T. A Pipeline for Determining Protein-Protein Interactions and Proximities in the Cellular Milieu. *Mol. Cell. Proteomics* **2014**, *13*, 2824−2835.
(75) Kaake, R. M.; Wang, X.; Burke, A.; Yu, C.; Kandur, W.; Yang, Y.; Novtisky, E. J.; Second, T.; Duan, J.; Kao, A.; et al. A New in Vivo Cross-Linking Mass Spectrometry Platform to Define Protein-Protein Interactions in Living Cells. *Mol. Cell. Proteomics* **2014**, *13*, 3533−3543.
(76) Mohammed, H.; D'Santos, C.; Serandour, A. A.; Ali, H. R.; Brown, G. D.; Atkins, A.; Rueda, O. M.; Holmes, K. A.; Theodorou, V.; Robinson, J. L. L.; et al. Endogenous Purification Reveals GREB1 as a Key Estrogen Receptor Regulatory Factor. *Cell Rep.* **2013**, *3*, 342−349.
(77) Mohammed, H.; Taylor, C.; Brown, G. D.; Papachristou, E. K.; Carroll, J. S.; D'Santos, C. S. Rapid Immunoprecipitation Mass Spectrometry of Endogenous Proteins (RIME) for Analysis of Chromatin Complexes. *Nat. Protoc.* **2016**, *11*, 316−326.
(78) Papachristou, E. K.; Kishore, K.; Holding, A. N.; Harvey, K.; Roumeliotis, T. I.; Chilamakuri, C. S. R.; Omarjee, S.; Chia, K. M.; Swarbrick, A.; Lim, E.; et al. A Quantitative Mass Spectrometry-Based Approach to Monitor the Dynamics of Endogenous Chromatin-Associated Protein Complexes. *Nat. Commun.* **2018**, *9*, 2311.
(79) Fabre, B.; Lambour, T.; Delobel, J.; Amalric, F.; Monsarrat, B.; Burlet-Schiltz, O.; Bousquet-Dubouch, M.-P. Subcellular Distribution and Dynamics of Active Proteasome Complexes Unraveled by a Workflow Combining *in Vivo* Complex Cross-Linking and Quantitative Proteomics. *Mol. Cell. Proteomics* **2013**, *12*, 687−699.
(80) Liu, F.; Rijkers, D. T. S.; Post, H.; Heck, A. J. R. Proteome-Wide Profiling of Protein Assemblies by Cross-Linking Mass Spectrometry. *Nat. Methods* **2015**, *12*, 1179−1184.
(81) Shi, Y.; Pellarin, R.; Fridy, P. C.; Fernandez-Martinez, J.; Thompson, M. K.; Li, Y.; Wang, Q. J.; Sali, A.; Rout, M. P.; Chait, B.

T. A Strategy for Dissecting the Architectures of Native Macromolecular Assemblies. *Nat. Methods* **2015**, *12*, 1135−1138.
(82) Hevler, J.; Lukassen, M.; Cabrera-Orefice, A.; Arnold, S.; Pronker, M.; Franc, V.; Heck, A. Selective Cross-Linking of Coinciding Protein Assemblies by in-Gel Cross-Linking Mass-Spectrometry. *EMBO J.* **2021**, *40*, No. e106174.
(83) Mishra, P. K.; Yoo, C. M.; Hong, E.; Rhee, H. W. Photo-Crosslinking: An Emerging Chemical Tool for Investigating Molecular Networks in Live Cells. *ChemBioChem* **2020**, *21*, 924−932.
(84) Suchanek, M.; Radzikowska, A.; Thiele, C. Photo-Leucine and Photo-Methionine Allow Identification of Protein-Protein Interactions in Living Cells. *Nat. Methods* **2005**, *2*, 261−267.
(85) Yang, T.; Li, X. M.; Bao, X.; Fung, Y. M. E.; Li, X. D. Photo-Lysine Captures Proteins That Bind Lysine Post-Translational Modifications. *Nat. Chem. Biol.* **2016**, *12*, 70−72.
(86) Kleiner, R. E.; Hang, L. E.; Molloy, K. R.; Chait, B. T.; Kapoor, T. M. A Chemical Proteomics Approach to Reveal Direct Protein-Protein Interactions in Living Cells. *Cell Chem. Biol.* **2018**, *25*, 110.
(87) Burton, A. J.; Haugbro, M.; Gates, L. A.; Bagert, J. D.; Allis, C. D.; Muir, T. W. In Situ Chromatin Interactomics Using a Chemical Bait and Trap Approach. *Nat. Chem.* **2020**, *12*, 520−527.
(88) Wu, X.; Spence, J. S.; Das, T.; Yuan, X.; Chen, C.; Zhang, Y.; Li, Y.; Sun, Y.; Chandran, K.; Hang, H. C.; et al. Site-Specific Photo-Crosslinking Proteomics Reveal Regulation of IFITM3 Trafficking and Turnover by VCP/P97 ATPase. *Cell Chem. Biol.* **2020**, *27*, 571−585.
(89) Lermyte, F.; Tsybin, Y. O.; O'Connor, P. B.; Loo, J. A. Top or Middle? Up or Down? Toward a Standard Lexicon for Protein Top-Down and Allied Mass Spectrometry Approaches. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 1149−1157.
(90) Coon, J. J.; Syka, J. E. P.; Shabanowitz, J.; Hunt, D. F. Tandem Mass Spectrometry for Peptide and Protein Sequence Analysis. *BioTechniques* **2005**, *38*, 519−523.
(91) McLafferty, F. W. Tandem Mass Spectrometry (MS/MS): A Promising New Analytical Technique for Specific Component Determination in Complex Mixtures. *Acc. Chem. Res.* **1980**, *13*, 33.
(92) Wu, M.; Li, X.; Kwoh, C. K.; Ng, S. K. A Core-Attachment Based Method to Detect Protein Complexes in PPI Networks. *BMC Bioinf.* **2009**, *10*, 169.
(93) Dezso, Z.; Oltvai, Z. N.; Barabási, A. L. Bioinformatics Analysis of Experimentally Determined Protein Complexes in the Yeast Saccharomyces Cerevisiae. *Genome Res.* **2003**, *13*, 2450−2454.
(94) Gavin, A. C.; Aloy, P.; Grandi, P.; Krause, R.; Boesche, M.; Marzioch, M.; Rau, C.; Jensen, L. J.; Bastuck, S.; Dümpelfeld, B.; et al. Proteome Survey Reveals Modularity of the Yeast Cell Machinery. *Nature* **2006**, *440*, 631−636.
(95) Schaffer, L. V.; Millikin, R. J.; Miller, R. M.; Anderson, L. C.; Fellers, R. T.; Ge, Y.; Kelleher, N. L.; LeDuc, R. D.; Liu, X.; Payne, S. H.; et al. Identification and Quantification of Proteoforms by Mass Spectrometry. *Proteomics* **2019**, *19*, 1800361.
(96) Uversky, V. N. *Posttranslational Modification*; Elsevier, 2013; Vol. 5.
(97) Duan, G.; Walther, D. The Roles of Post-Translational Modifications in the Context of Protein Interaction Networks. *PLoS Comput. Biol.* **2015**, *11*, e1004049.
(98) Wu, L.; Han, D. K. Overcoming the Dynamic Range Problem in Mass Spectrometry-Based Shotgun Proteomics. *Expert Rev. Proteomics* **2006**, *3*, 611−619.
(99) Zubarev, R. A. The Challenge of the Proteome Dynamic Range and Its Implications for In-Depth Proteomics. *Proteomics* **2013**, *13*, 723−726.
(100) Furey, A.; Moriarty, M.; Bane, V.; Kinsella, B.; Lehane, M. Ion Suppression; A Critical Review on Causes, Evaluation, Prevention and Applications. *Talanta* **2013**, *115*, 104−122.
(101) Rosenfeld, J.; Capdevielle, J.; Guillemot, J. C.; Ferrara, P. In-Gel Digestion of Proteins for Internal Sequence Analysis after One- or Two-Dimensional Gel Electrophoresis. *Anal. Biochem.* **1992**, *203*, 173−179.
(102) Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M. Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels. *Anal. Chem.* **1996**, *68*, 850−858.
(103) O'Farrell, P. H. High Resolution Two Dimensional Electrophoresis of Proteins. *J. Biol. Chem.* **1975**, *250*, 4007−4021.
(104) Issaq, H. J.; Veenstra, T. D. Two-Dimensional Polyacrylamide Gel Electrophoresis (2D-PAGE): Advances and Perspectives. *BioTechniques* **2008**, *44*, 697−700.
(105) Gygi, S. P.; Corthals, G. L.; Zhang, Y.; Rochon, Y.; Aebersold, R. Evaluation of Two-Dimensional Gel Electrophoresis-Based Proteome Analysis Technology. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 9390−9395.
(106) Wright, E. P.; Partridge, M. A.; Padula, M. P.; Gauci, V. J.; Malladi, C. S.; Coorssen, J. R. Top-down Proteomics: Enhancing 2D Gel Electrophoresis from Tissue Processing to High-Sensitivity Protein Detection. *Proteomics* **2014**, *14*, 872−889.
(107) Tran, J. C.; Doucette, A. A. Multiplexed Size Separation of Intact Proteins in Solution Phase for Mass Spectrometry. *Anal. Chem.* **2009**, *81*, 6201−6209.
(108) Tran, J. C.; Doucette, A. A. Gel-Eluted Liquid Fraction Entrapment Electrophoresis: An Electrophoretic Method for Broad Molecular Weight Range Proteome Separation. *Anal. Chem.* **2008**, *80*, 1568−1573.
(109) Kuljanin, M.; Dieters-Castator, D. Z.; Hess, D. A.; Postovit, L. M.; Lajoie, G. A. Comparison of Sample Preparation Techniques for Large-Scale Proteomics. *Proteomics* **2017**, *17*, 1600337.
(110) Aguilar, M. Reversed-Phase High-Performance Liquid Chromatography. *Methods Mol. Biol.* **2004**, *251*, 9−22.
(111) Wolters, D. A.; Washburn, M. P.; Yates, J. R. An Automated Multidimensional Protein Identification Technology for Shotgun Proteomics. *Anal. Chem.* **2001**, *73*, 5683−5690.
(112) VanAernum, Z. L.; Busch, F.; Jones, B. J.; Jia, M.; Chen, Z.; Boyken, S. E.; Sahasrabuddhe, A.; Baker, D.; Wysocki, V. H. Rapid Online Buffer Exchange for Screening of Proteins, Protein Complexes and Cell Lysates by Native Mass Spectrometry. *Nat. Protoc.* **2020**, *15*, 1132−1157.
(113) Busch, F.; VanAernum, Z. L.; Lai, S. M.; Gopalan, V.; Wysocki, V. H. Analysis of Tagged Proteins Using Tandem Affinity-Buffer Exchange Chromatography Online with Native Mass Spectrometry. *Biochemistry* **2021**, *60*, 1876−1884.
(114) Mitulović, G. New HPLC Techniques for Proteomics Analysis: A Short Overview of Latest Developments. *J. Liq. Chromatogr. Relat. Technol.* **2015**, *38*, 390−403.
(115) Dolník, V. Capillary Zone Electrophoresis of Proteins. *Electrophoresis* **1997**, *18*, 2353−2361.
(116) Boss, H. J.; Watson, D. B.; Rush, R. S. Peptide Capillary Zone Electrophoresis Mass Spectrometry of Recombinant Human Erythropoietin: An Evaluation of the Analytical Method. *Electrophoresis* **1998**, *19*, 2654−2664.
(117) Yan, X.; Essaka, D. C.; Sun, L.; Zhu, G.; Dovichi, N. J. Bottom-up Proteome Analysis of E. Coli Using Capillary Zone Electrophoresis-Tandem Mass Spectrometry with an Electrokinetic Sheath-Flow Electrospray Interface. *Proteomics* **2013**, *13*, 2546−2551.
(118) Sun, L.; Zhu, G.; Zhao, Y.; Yan, X.; Mou, S.; Dovichi, N. J. Ultrasensitive and Fast Bottom-up Analysis of Femtogram Amounts of Complex Proteome Digests. *Angew. Chem., Int. Ed.* **2013**, *52*, 13661−13664.
(119) Zhu, G.; Sun, L.; Yan, X.; Dovichi, N. J. Single-Shot Proteomics Using Capillary Zone Electrophoresis-Electrospray Ionization-Tandem Mass Spectrometry with Production of More than 1 250 Escherichia Coli Peptide Identifications in a 50 min Separation. *Anal. Chem.* **2013**, *85*, 2569−2573.
(120) Kumar, R.; Shah, R. L.; Rathore, A. S. Harnessing the Power of Electrophoresis and Chromatography: Offline Coupling of Reverse Phase Liquid Chromatography-Capillary Zone Electrophoresis-Tandem Mass Spectrometry for Peptide Mapping for Monoclonal Antibodies. *J. Chromatogr. A* **2020**, *1620*, 460954.
(121) Jooß, K.; Hühner, J.; Kiessig, S.; Moritz, B.; Neusüß, C. Two-Dimensional Capillary Zone Electrophoresis−Mass Spectrometry for

the Characterization of Intact Monoclonal Antibody Charge Variants. *Anal. Bioanal. Chem.* **2017**, *409*, 6057−6067.
(122) Jooß, K.; McGee, J. P.; Melani, R. D.; Kelleher, N. L. Standard Procedures for Native CZE-MS of Proteins and Protein Complexes up to 800 KDa. *Electrophoresis* **2021**, *42*, 1050−1059.
(123) Jooß, K.; Schachner, L. F.; Watson, R.; Gillespie, Z. B.; Howard, S. A.; Cheek, M. A.; Meiners, M. J.; Sobh, A.; Licht, J. D.; Keogh, M. C.; et al. Separation and Characterization of Endogenous Nucleosomes by Native Capillary Zone Electrophoresis-Top-Down Mass Spectrometry. *Anal. Chem.* **2021**, *93*, 5151−5160.
(124) Mack, E. Average Cross-Sectional Areas of Molecules by Gasesous Diffusion Methods. *J. Am. Chem. Soc.* **1925**, *47*, 2468−2482.
(125) Verbeck, G. F.; Ruotolo, B. T.; Sawyer, H. A.; Gillig, K. J.; Russell, D. H. A Fundamental Introduction to Ion Mobility Mass Spectrometry Applied to the Analysis of Biomolecules. *J. Biomol. Technol.* **2002**, *13*, 56−61.
(126) Ben-Nissan, G.; Sharon, M. The Application of Ion-Mobility Mass Spectrometry for Structure/Function Investigation of Protein Complexes. *Curr. Opin. Chem. Biol.* **2018**, *42*, 25−33.
(127) Pukala, T. Importance of Collision Cross Section Measurements by Ion Mobility Mass Spectrometry in Structural Biology. *Rapid Commun. Mass Spectrom.* **2019**, *33*, 72−82.
(128) Konijnenberg, A.; Butterer, A.; Sobott, F. Native Ion Mobility-Mass Spectrometry and Related Methods in Structural Biology. *Biochim. Biophys. Acta, Proteins Proteomics* **2013**, *1834*, 1239−1256.
(129) Lanucara, F.; Holman, S. W.; Gray, C. J.; Eyers, C. E. The Power of Ion Mobility-Mass Spectrometry for Structural Characterization and the Study of Conformational Dynamics. *Nat. Chem.* **2014**, *6*, 281−294.
(130) Causon, T. J.; Hann, S. Theoretical Evaluation of Peak Capacity Improvements by Use of Liquid Chromatography Combined with Drift Tube Ion Mobility-Mass Spectrometry. *J. Chromatogr. A* **2015**, *1416*, 47−56.
(131) Baker, E. S.; Burnum-Johnson, K. E.; Ibrahim, Y. M.; Orton, D. J.; Monroe, M. E.; Kelly, R. T.; Moore, R. J.; Zhang, X.; Théberge, R.; Costello, C. E.; et al. Enhancing Bottom-up and Top-down Proteomic Measurements with Ion Mobility Separations. *Proteomics* **2015**, *15*, 2766−2776.
(132) Baker, E. S.; Livesay, E. A.; Orton, D. J.; Moore, R. J.; Danielson, W. F.; Prior, D. C.; Ibrahim, Y. M.; LaMarche, B. L.; Mayampurath, A. M.; Schepmoes, A. A.; et al. An LC-IMS-MS Platform Providing Increased Dynamic Range for High-Throughput Proteomic Studies. *J. Proteome Res.* **2010**, *9*, 997−1006.
(133) Meier, F.; Brunner, A. D.; Koch, S.; Koch, H.; Lubeck, M.; Krause, M.; Goedecke, N.; Decker, J.; Kosinski, T.; Park, M. A.; et al. Online Parallel Accumulation−Serial Fragmentation (PASEF) with a Novel Trapped Ion Mobility Mass Spectrometer. *Mol. Cell. Proteomics* **2018**, *17*, 2534−2545.
(134) Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.; Robinson, C. V. Ion Mobility − Mass Spectrometry Analysis of Large Protein Complexes. *Nat. Protoc.* **2008**, *3*, 1139−1152.
(135) Han, X.; Aslanian, A.; Yates, J. R. Mass Spectrometry for Proteomics. *Curr. Opin. Chem. Biol.* **2008**, *12*, 483−490.
(136) Bantscheff, M.; Lemeer, S.; Savitski, M. M.; Kuster, B. Quantitative Mass Spectrometry in Proteomics: Critical Review Update from 2007 to the Present. *Anal. Bioanal. Chem.* **2012**, *404*, 939−965.
(137) Swaney, D. L.; Wenger, C. D.; Coon, J. J. Value of Using Multiple Proteases for Large-Scale Mass Spectrometry-Based Proteomics. *J. Proteome Res.* **2010**, *9*, 1323−1329.
(138) Giansanti, P.; Tsiatsiani, L.; Low, T. Y.; Heck, A. J. R. Six Alternative Proteases for Mass Spectrometry-Based Proteomics beyond Trypsin. *Nat. Protoc.* **2016**, *11*, 993−1006.
(139) Dau, T.; Bartolomucci, G.; Rappsilber, J. Proteomics Using Protease Alternatives to Trypsin Benefits from Sequential Digestion with Trypsin. *Anal. Chem.* **2020**, *92*, 9523−9527.
(140) Marx, V. Targeted Proteomics. *Nat. Methods* **2013**, *10*, 19−22.
(141) Michalski, A.; Cox, J.; Mann, M. More than 100,000 Detectable Peptide Species Elute in Single Shotgun Proteomics Runs but the Majority Is Inaccessible to Data-Dependent LC-MS/MS. *J. Proteome Res.* **2011**, *10*, 1785−1793.
(142) Venable, J. D.; Dong, M. Q.; Wohlschlegel, J.; Dillin, A.; Yates, J. R. Automated Approach for Quantitative Analysis of Complex Peptide Mixtures from Tandem Mass Spectra. *Nat. Methods* **2004**, *1*, 39−45.
(143) Law, K. P.; Lim, Y. P. Recent Advances in Mass Spectrometry: Data Independent Analysis and Hyper Reaction Monitoring. *Expert Rev. Proteomics* **2013**, *10*, 551−566.
(144) Messner, C. B.; Demichev, V.; Bloomfield, N.; Yu, J. S. L.; White, M.; Kreidl, M.; Egger, A. S.; Freiwald, A.; Ivosev, G.; Wasim, F. Ultra-Fast Proteomics with Scanning SWATH. *Nat. Biotechnol.* **2021**, *39*, 846−854.
(145) Meier, F.; Beck, S.; Grassl, N.; Lubeck, M.; Park, M. A.; Raether, O.; Mann, M. Parallel Accumulation-Serial Fragmentation (PASEF): Multiplying Sequencing Speed and Sensitivity by Synchronized Scans in a Trapped Ion Mobility Device. *J. Proteome Res.* **2015**, *14*, 5378−5387.
(146) Liu, H.; Sadygov, R. G.; Yates, J. R. A Model for Random Sampling and Estimation of Relative Protein Abundance in Shotgun Proteomics. *Anal. Chem.* **2004**, *76*, 4193−4201.
(147) Schwanhausser, B.; Busse, D.; Li, N.; Dittmar, G.; Schuchhardt, J.; Wolf, J.; Chen, W.; Selbach, M. Global Quantification of Mammalian Gene Expression Control. *Nature* **2011**, *473*, 337−342.
(148) Chelius, D.; Bondarenko, P. V. Quantitative Profiling of Proteins in Complex Mixtures Using Liquid Chromatography and Mass Spectrometry. *J. Proteome Res.* **2002**, *1*, 317−323.
(149) Hoedt, E.; Zhang, G.; Neubert, T. A. Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC) for Quantitative Proteomics. *Adv. Exp. Med. Biol.* **2019**, *1140*, 531−539.
(150) Thompson, A.; Schäfer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Hamon, C. Tandem Mass Tags: A Novel Quantification Strategy for Comparative Analysis of Complex Protein Mixtures by MS/MS. *Anal. Chem.* **2003**, *75*, 1895−1904.
(151) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. Quantitative Analysis of Complex Protein Mixtures Using Isotope-Coded Affinity Tags. *Nat. Biotechnol.* **1999**, *17*, 994−999.
(152) Nesvizhskii, A. I.; Vitek, O.; Aebersold, R. Analysis and Validation of Proteomic Data Generated by Tandem Mass Spectrometry. *Nat. Methods* **2007**, *4*, 787−797.
(153) Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S. Probability-Based Protein Identification by Searching Sequence Databases Using Mass Spectrometry Data. *Electrophoresis* **1999**, *20*, 3551−3567.
(154) Tyanova, S.; Temu, T.; Cox, J. The MaxQuant Computational Platform for Mass Spectrometry-Based Shotgun Proteomics. *Nat. Protoc.* **2016**, *11*, 2301−2319.
(155) Zhang, J.; Xin, L.; Shan, B.; Chen, W.; Xie, M.; Yuen, D.; Zhang, W.; Zhang, Z.; Lajoie, G. A.; Ma, B. PEAKS DB: De Novo Sequencing Assisted Database Search for Sensitive and Accurate Peptide Identification. *Mol. Cell. Proteomics* **2012**, *11*, M111.010587.
(156) Tran, N. H.; Qiao, R.; Xin, L.; Chen, X.; Liu, C.; Zhang, X.; Shan, B.; Ghodsi, A.; Li, M. Deep Learning Enables de Novo Peptide Sequencing from Data-Independent-Acquisition Mass Spectrometry. *Nat. Methods* **2019**, *16*, 63−66.
(157) Pino, L. K.; Searle, B. C.; Bollinger, J. G.; Nunn, B.; MacLean, B.; MacCoss, M. J. The Skyline Ecosystem: Informatics for Quantitative Mass Spectrometry Proteomics. *Mass Spectrom. Rev.* **2020**, *39*, 229−244.
(158) Chambers, M. C.; MacLean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; et al. A Cross-Platform Toolkit for Mass Spectrometry and Proteomics. *Nat. Biotechnol.* **2012**, *30*, 918−920.
(159) Välikangas, T.; Suomi, T.; Elo, L. L. A Comprehensive Evaluation of Popular Proteomics Software Workflows for Label-Free Proteome Quantification and Imputation. *Briefings Bioinf.* **2017**, *19*, 1344−1355.

(160) Chawade, A.; Sandin, M.; Teleman, J.; Malmström, J.; Levander, F. Data Processing Has Major Impact on the Outcome of Quantitative Label-Free LC-MS Analysis. *J. Proteome Res.* **2015**, *14*, 676−687.
(161) Timp, W.; Timp, G. Beyond Mass Spectrometry, the next Step in Proteomics. *Sci. Adv.* **2020**, *6*, eaax8978.
(162) Kelly, R. T. Single-Cell Proteomics: Progress and Prospects. *Mol. Cell. Proteomics* **2020**, *19*, 1739−1748.
(163) Slavov, N. Single-Cell Protein Analysis by Mass Spectrometry. *Curr. Opin. Chem. Biol.* **2021**, *60*, 1−9.
(164) Zhu, Y.; Piehowski, P. D.; Zhao, R.; Chen, J.; Shen, Y.; Moore, R. J.; Shukla, A. K.; Petyuk, V. A.; Campbell-Thompson, M.; Mathews, C. E.; et al. Nanodroplet Processing Platform for Deep and Quantitative Proteome Profiling of 10−100 Mammalian Cells. *Nat. Commun.* **2018**, *9*, 882.
(165) Zhu, Y.; Clair, G.; Chrisler, W. B.; Shen, Y.; Zhao, R.; Shukla, A. K.; Moore, R. J.; Misra, R. S.; Pryhuber, G. S.; Smith, R. D.; et al. Proteomic Analysis of Single Mammalian Cells Enabled by Microfluidic Nanodroplet Sample Preparation and Ultrasensitive NanoLC-MS. *Angew. Chem.* **2018**, *130*, 12550−12554.
(166) Cong, Y.; Liang, Y.; Motamedchaboki, K.; Huguet, R.; Truong, T.; Zhao, R.; Shen, Y.; Lopez-Ferrer, D.; Zhu, Y.; Kelly, R. T. Improved Single-Cell Proteome Coverage Using Narrow-Bore Packed NanoLC Columns and Ultrasensitive Mass Spectrometry. *Anal. Chem.* **2020**, *92* (3), 2665−2671.
(167) Specht, H.; Harmange, G.; Perlman, D. H.; Emmott, E.; Niziolek, Z.; Budnik, B.; Slavov, N. Automated Sample Preparation for High-Throughput Single-Cell Proteomics. *bioRxiv* **2018**. DOI: 10.1101/399774
(168) Budnik, B.; Levy, E.; Harmange, G.; Slavov, N. SCoPE-MS: Mass Spectrometry of Single Mammalian Cells Quantifies Proteome Heterogeneity during Cell Differentiation. *Genome Biol.* **2018**, *19*, 161.
(169) Virág, D.; Dalmadi-Kiss, B.; Vékey, K.; Drahos, L.; Klebovich, I.; Antal, I.; Ludányi, K. Current Trends in the Analysis of Post-Translational Modifications. *Chromatographia* **2020**, *83*, 1−10.
(170) Dephoure, N.; Gould, K. L.; Gygi, S. P.; Kellogg, D. R. Mapping and Analysis of Phosphorylation Sites: A Quick Guide for Cell Biologists. *Mol. Biol. Cell* **2013**, *24*, 535−542.
(171) Kim, M. S.; Zhong, J.; Pandey, A. Common Errors in Mass Spectrometry-Based Analysis of Post-Translational Modifications. *Proteomics* **2016**, *16*, 700−714.
(172) Chavez, J. D.; Weisbrod, C. R.; Zheng, C.; Eng, J. K.; Bruce, J. E. Protein Interactions, Post-Translational Modifications and Topologies in Human Cells. *Mol. Cell. Proteomics* **2013**, *12*, 1451−1467.
(173) 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. A Census of Human Soluble Protein Complexes. *Cell* **2012**, *150*, 1068−1081.
(174) Olinares, P. D. B.; Ponnala, L.; Van Wijk, K. J. Megadalton Complexes in the Chloroplast Stroma of Arabidopsis Thaliana Characterized by Size Exclusion Chromatography, Mass Spectrometry, and Hierarchical Clustering. *Mol. Cell. Proteomics* **2010**, *9*, 1594−1615.
(175) 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. Capturing Protein Communities by Structural Proteomics in a Thermophilic Eukaryote. *Mol. Syst. Biol.* **2017**, *13*, 936.
(176) Miteva, Y. V.; Budayeva, H. G.; Cristea, I. M. Proteomics-Based Methods for Discovery, Quantification, and Validation of Protein − Protein Interactions. *Anal. Chem.* **2013**, *85*, 749−768.
(177) Krogan, N. J.; Cagney, G.; Yu, H.; Zhong, G.; Guo, X.; Ignatchenko, A.; Li, J.; Pu, S.; Datta, N.; Tikuisis, A. P.; et al. Global Landscape of Protein Complexes in the Yeast Saccharomyces Cerevisiae. *Nature* **2006**, *440*, 637−643.
(178) Malovannaya, A.; Lanz, R. B.; Jung, S. Y.; Bulynko, Y.; Le, N. T.; Chan, D. W.; Ding, C.; Shi, Y.; Yucer, N.; Krenciute, G.; et al. Analysis of the Human Endogenous Coregulator Complexome. *Cell* **2011**, *145*, 787−799.
(179) 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. A Human Interactome in Three Quantitative Dimensions Organized by Stoichiometries and Abundances. *Cell* **2015**, *163*, 712−723.
(180) Wiemann, S.; Pennacchio, C.; Hu, Y.; Hunter, P.; Harbers, M.; Amiet, A.; Bethel, G.; Busse, M.; Carninci, P.; Dunham, I.; et al. The ORFeome Collaboration: A Genome-Scale Human ORF-Clone Resource. *Nat. Methods* **2016**, *13*, 191−192.
(181) Huttlin, E. L.; Bruckner, R. J.; Paulo, J. A.; Cannon, J. R.; Ting, L.; Baltier, K.; Colby, G.; Gebreab, F.; Gygi, M. P.; Parzen, H.; et al. Architecture of the Human Interactome Defines Protein Communities and Disease Networks. *Nature* **2017**, *545*, 505−509.
(182) Huttlin, E. L.; Ting, L.; Bruckner, R. J.; Gebreab, F.; Gygi, M. P.; Szpyt, J.; Tam, S.; Zarraga, G.; Colby, G.; Baltier, K.; et al. The BioPlex Network: A Systematic Exploration of the Human Interactome. *Cell* **2015**, *162*, 425−440.
(183) 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. Integrative Structure and Functional Anatomy of a Nuclear Pore Complex. *Nature* **2018**, *555*, 475−482.
(184) Smits, A. H.; Jansen, P. W. T. C.; Poser, I.; Hyman, A. A.; Vermeulen, M. Stoichiometry of Chromatin-Associated Protein Complexes Revealed by Label-Free Quantitative Mass Spectrometry-Based Proteomics. *Nucleic Acids Res.* **2013**, *41*, e28.
(185) van Nuland, R.; Smits, A. H.; Pallaki, P.; Jansen, P. W. T. C.; Vermeulen, M.; Timmers, M. H. T. Quantitative Dissection and Stoichiometry Determination of the Human Set1/MLL Histone Methyltransferase Complexes. *Mol. Cell. Biol.* **2013**, *33*, 2067−2077.
(186) Meysman, P.; Titeca, K.; Eyckerman, S.; Tavernier, J.; Goethals, B.; Martens, L.; Valkenborg, D.; Laukens, K. Protein Complex Analysis: From Raw Protein Lists to Protein Interaction Networks. *Mass Spectrom. Rev.* **2017**, *36*, 600−614.
(187) Choi, H. Computational Detection of Protein Complexes in AP-MS Experiments. *Proteomics* **2012**, *12*, 1663−1668.
(188) Wu, Z.; Liao, Q.; Liu, B. A Comprehensive Review and Evaluation of Computational Methods for Identifying Protein Complexes from Protein-Protein Interaction Networks. *Briefings Bioinf.* **2020**, *21*, 1531−1548.
(189) Choi, H.; Larsen, B.; Lin, Z. Y.; Breitkreutz, A.; Mellacheruvu, D.; Fermin, D.; Qin, Z. S.; Tyers, M.; Gingras, A. C.; Nesvizhskii, A. I. SAINT: Probabilistic Scoring of Affinity Purification-Mass Spectrometry Data. *Nat. Methods* **2011**, *8*, 70−73.
(190) Sowa, M. E.; Bennett, E. J.; Gygi, S. P.; Harper, J. W. Defining the Human Deubiquitinating Enzyme Interaction Landscape. *Cell* **2009**, *138*, 389−403.
(191) Pu, S.; Vlasblom, J.; Turinsky, A.; Marcon, E.; Phanse, S.; Trimble, S. S.; Olsen, J.; Greenblatt, J.; Emili, A.; Wodak, S. J. Extracting High Confidence Protein Interactions from Affinity Purification Data: At the Crossroads. *J. Proteomics* **2015**, *118*, 63−80.
(192) Varjosalo, M.; Sacco, R.; Stukalov, A.; Van Drogen, A.; Planyavsky, M.; Hauri, S.; Aebersold, R.; Bennett, K. L.; Colinge, J.; Gstaiger, M.; et al. Interlaboratory Reproducibility of Large-Scale Human Protein-Complex Analysis by Standardized AP-MS. *Nat. Methods* **2013**, *10*, 307−314.
(193) Drew, K.; Lee, C.; Huizar, R. L.; Tu, F.; Borgeson, B.; Mcwhite, C. D.; Ma, Y.; Wallingford, J. B.; Marcotte, E. M. Integration of over 9, 000 Mass Spectrometry Experiments Builds a Global Map of Human Protein Complexes. *Mol. Syst. Biol.* **2017**, *13*, 932.
(194) Tackett, A. J.; DeGrasse, J. A.; Sekedat, M. D.; Oeffinger, M.; Rout, M. P.; Chait, B. T. I-DIRT, a General Method for Distinguishing between Specific and Nonspecific Protein Interactions. *J. Proteome Res.* **2005**, *4*, 1752−1756.
(195) 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. The CRAPome: A Contaminant Repository for Affinity Purification-Mass Spectrometry Data. *Nat. Methods* **2013**, *10*, 730−736.

(196) Link, A. J.; Eng, J. K.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R. Direct Analysis of Protein Complexes Using Mass Spectrometry. *Nat. Biotechnol.* **1999**, *17*, 676−682.
(197) Mena, E. L.; Kjolby, R. A. S.; Saxton, R. A.; Werner, A.; Lew, B. G.; Boyle, J. M.; Harland, R.; Rape, M. Dimerization Quality Control Ensures Neuronal Development and Survival. *Science* **2018**, *362*, No. eaap8236.
(198) Lubas, M.; Christensen, M. S.; Kristiansen, M. S.; Domanski, M.; Falkenby, L. G.; Lykke-Andersen, S.; Andersen, J. S.; Dziembowski, A.; Jensen, T. H. Interaction Profiling Identifies the Human Nuclear Exosome Targeting Complex. *Mol. Cell* **2011**, *43*, 624−637.
(199) Taylor, M. S.; Altukhov, I.; Molloy, K. R.; Mita, P.; Jiang, H.; Adney, E. M.; Wudzinska, A.; Badri, S.; Ischenko, D.; Eng, G.; et al. Dissection of Affinity Captured LINE-1 Macromolecular Complexes. *eLife* **2018**, *7*, e30094.
(200) Savitski, M. M.; Reinhard, F. B. M.; Franken, H.; Werner, T.; Savitski, M. F.; Eberhard, D.; Molina, D. M.; Jafari, R.; Dovega, R. B.; Klaeger, S.; et al. Tracking Cancer Drugs in Living Cells by Thermal Profiling of the Proteome. *Science* **2014**, *346*, 1255784.
(201) Tan, C. S. H.; Go, K. D.; Bisteau, X.; Dai, L.; Yong, C. H.; Prabhu, N.; Ozturk, M. B.; Lim, Y. T.; Sreekumar, L.; Lengqvist, J.; et al. Thermal Proximity Coaggregation for System-Wide Profiling of Protein Complex Dynamics in Cells. *Science* **2018**, *359*, 1170−1177.
(202) Dai, L.; Zhao, T.; Bisteau, X.; Sun, W.; Prabhu, N.; Lim, Y. T.; Sobota, R. M.; Kaldis, P.; Nordlund, P. Modulation of Protein-Interaction States through the Cell Cycle. *Cell* **2018**, *173*, 1481−1494.
(203) Hashimoto, Y.; Sheng, X.; Murray-Nerger, L. A.; Cristea, I. M. Temporal Dynamics of Protein Complex Formation and Dissociation during Human Cytomegalovirus Infection. *Nat. Commun.* **2020**, *11*, 806.
(204) Liu, X. R.; Zhang, M. M.; Gross, M. L. Mass Spectrometry-Based Protein Footprinting for Higher-Order Structure Analysis: Fundamentals and Applications. *Chem. Rev.* **2020**, *120*, 4355−4454.
(205) Sheshberadaran, H.; Payne, L. G. Protein Antigen-Monoclonal Antibody Contact Sites Investigated by Limited Proteolysis of Monoclonal Antibody-Bound Antigen: Protein "Footprinting. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85*, 1−5.
(206) Chen, J.; Cui, W.; Giblin, D.; Gross, M. L. New Protein Footprinting: Fast Photochemical Iodination Combined with Top-down and Bottom-up Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 1306−1318.
(207) Novak, P.; Young, M. M.; Schoeniger, J. S.; Kruppa, G. H. A Top-down Approach to Protein Structure Studies Using Chemical Cross-Linking and Fourier Transform Mass Spectrometry. *Eur. J. Mass Spectrom.* **2003**, *9*, 623−631.
(208) O'Reilly, F. J.; Rappsilber, J. Cross-Linking Mass Spectrometry: Methods and Applications in Structural, Molecular and Systems Biology. *Nat. Struct. Mol. Biol.* **2018**, *25*, 1000−1008.
(209) Chavez, J. D.; Mohr, J. P.; Mathay, M.; Zhong, X.; Keller, A.; Bruce, J. E. Systems Structural Biology Measurements by in Vivo Cross-Linking with Mass Spectrometry. *Nat. Protoc.* **2019**, *14*, 2318−2343.
(210) Gutierrez, C. B.; Yu, C.; Novitsky, E. J.; Huszagh, A. S.; Rychnovsky, S. D.; Huang, L. Developing an Acidic Residue Reactive and Sulfoxide-Containing MS-Cleavable Homobifunctional Cross-Linker for Probing Protein-Protein Interactions. *Anal. Chem.* **2016**, *88*, 8315−8322.
(211) Kao, A.; Chiu, C.; Vellucci, D.; Yang, Y.; Patel, V. R.; Guan, S.; Randall, A.; Baldi, P.; Rychnovsky, S. D.; Huang, L. Development of a Novel Cross-Linking Strategy for Fast and Accurate Identification of Cross-Linked Peptides of Protein Complexes. *Mol. Cell. Proteomics* **2011**, *10*, M110.002170.
(212) Müller, M. Q.; Dreiocker, F.; Ihling, C. H.; Schäfer, M.; Sinz, A. Cleavable Cross-Linker for Protein Structure Analysis: Reliable Identification of Cross-Linking Products by Tandem MS. *Anal. Chem.* **2010**, *82*, 6958−6968.
(213) Leitner, A.; Walzthoeni, T.; Aebersold, R. Lysine-Specific Chemical Cross-Linking of Protein Complexes and Identification of Cross-Linking Sites Using LC-MS/MS and the XQuest/XProphet Software Pipeline. *Nat. Protoc.* **2014**, *9*, 120−137.
(214) Hoopmann, M. R.; Zelter, A.; Johnson, R. S.; Riffle, M.; Maccoss, M. J.; Davis, T. N.; Moritz, R. L. Kojak: Efficient Analysis of Chemically Cross-Linked Protein Complexes. *J. Proteome Res.* **2015**, *14*, 2190−2198.
(215) Götze, M.; Pettelkau, J.; Schaks, S.; Bosse, K.; Ihling, C. H.; Krauth, F.; Fritzsche, R.; Kühn, U.; Sinz, A. StavroX-A Software for Analyzing Crosslinked Products in Protein Interaction Studies. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 76−87.
(216) Chen, Z. L.; Meng, J. M.; Cao, Y.; Yin, J. L.; Fang, R. Q.; Fan, S. B.; Liu, C.; Zeng, W. F.; Ding, Y. H.; Tan, D.; et al. A High-Speed Search Engine PLink 2 with Systematic Evaluation for Proteome-Scale Identification of Cross-Linked Peptides. *Nat. Commun.* **2019**, *10*, 3404.
(217) Beveridge, R.; Stadlmann, J.; Penninger, J. M.; Mechtler, K. A Synthetic Peptide Library for Benchmarking Crosslinking-Mass Spectrometry Search Engines for Proteins and Protein Complexes. *Nat. Commun.* **2020**, *11*, 742.
(218) Leitner, A.; Bonvin, A. M. J. J.; Borchers, C. H.; Chalkley, R. J.; Chamot-Rooke, J.; Combe, C. W.; Cox, J.; Dong, M. Q.; Fischer, L.; Götze, M.; et al. Toward Increased Reliability, Transparency, and Accessibility in Cross-Linking Mass Spectrometry. *Structure* **2020**, *28*, 1259−1268.
(219) Shi, Y.; Fernandez-Martinez, J.; Tjioe, E.; Pellarin, R.; Kim, S. J.; Williams, R.; Schneidman-Duhovny, D.; Sali, A.; Rout, M. P.; Chait, B. T. Structural Characterization by Cross-Linking Reveals the Detailed Architecture of a Coatomer-Related Heptameric Module from the Nuclear Pore Complex. *Mol. Cell. Proteomics* **2014**, *13*, 2927−2943.
(220) Gutierrez, C.; Chemmama, I. E.; Mao, H.; Yu, C.; Echeverria, I.; Block, S. A.; Rychnovsky, S. D.; Zheng, N.; Sali, A.; Huang, L. Structural Dynamics of the Human COP9 Signalosome Revealed by Cross-Linking Mass Spectrometry and Integrative Modeling. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 4088−4098.
(221) Mintseris, J.; Gygi, S. P. High-Density Chemical Cross-Linking for Modeling Protein Interactions. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 93−102.
(222) Rappsilber, J. The Beginning of a Beautiful Friendship: Cross-Linking/Mass Spectrometry and Modelling of Proteins and Multi-Protein Complexes. *J. Struct. Biol.* **2011**, *173*, 530−540.
(223) Bullock, J. M. A.; Thalassinos, K.; Topf, M. Jwalk and MNXL Web Server: Model Validation Using Restraints from Crosslinking Mass Spectrometry. *Bioinformatics* **2018**, *34*, 3584−3585.
(224) Vasilescu, J.; Guo, X.; Kast, J. Identification of Protein-Protein Interactions Using in Vivo Cross-Linking and Mass Spectrometry. *Proteomics* **2004**, *4*, 3845−3854.
(225) Guerrero, C.; Tagwerker, C.; Kaiser, P.; Huang, L. An Integrated Mass Spectrometry-Based Proteomic Aproach. *Mol. Cell. Proteomics* **2006**, *5*, 366−378.
(226) Mohammed, H.; Russell, I. A.; Stark, R.; Rueda, O. M.; Hickey, T. E.; Tarulli, G. A.; Serandour, A. A. A.; Birrell, S. N.; Bruna, A.; Saadi, A.; et al. Progesterone Receptor Modulates ERα Action in Breast Cancer. *Nature* **2015**, *523*, 313−317.
(227) Gonzalez-Lozano, M. A.; Koopmans, F.; Sullivan, P. F.; Protze, J.; Krause, G.; Verhage, M.; Li, K. W.; Liu, F.; Smit, A. B. Stitching the Synapse: Cross-Linking Mass Spectrometry into Resolving Synaptic Protein Interactions. *Sci. Adv.* **2020**, *6*, eaax5783.
(228) Makowski, M. M.; Willems, E.; Jansen, P. W. T. C.; Vermeulen, M. Cross-Linking Immunoprecipitation-MS (XIP-MS): Topological Analysis of Chromatin-Associated Protein Complexes Using Single Affinity Purification. *Mol. Cell. Proteomics* **2016**, *15*, 854−865.
(229) Politis, A.; Stengel, F.; Hall, Z.; Hernández, H.; Leitner, A.; Walzthoeni, T.; Robinson, C. V.; Aebersold, R. A Mass Spectrometry-Based Hybrid Method for Structural Modeling of Protein Complexes. *Nat. Methods* **2014**, *11*, 403−406.

https://doi.org/10.1021/acs.chemrev.1c00217
Chem. Rev. 2022, 122, 7386−7414

(230) Lasker, K.; Förster, F.; Bohn, S.; Walzthoeni, T.; Villa, E.; Unverdorben, P.; Beck, F.; Aebersold, R.; Sali, A.; Baumeister, W. Molecular Architecture of the 26S Proteasome Holocomplex Determined by an Integrative Approach. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 1380−1387.
(231) Chen, Z. A.; Jawhari, A.; Fischer, L.; Buchen, C.; Tahir, S.; Kamenski, T.; Rasmussen, M.; Lariviere, L.; Bukowski-Wills, J. C.; Nilges, M.; et al. Architecture of the RNA Polymerase II-TFIIF Complex Revealed by Cross-Linking and Mass Spectrometry. *EMBO J.* **2010**, *29*, 717−726.
(232) Lee, J.; Ding, S. J.; Walpole, T. B.; Holding, A. N.; Montgomery, M. G.; Fearnley, I. M.; Walker, J. E. Organization of Subunits in the Membrane Domain of the Bovine F-ATPase Revealed by Covalent Cross-Linking. *J. Biol. Chem.* **2015**, *290*, 13308−13320.
(233) Herzog, F.; Kahraman, A.; Boehringer, D.; Mak, R.; Bracher, A.; Walzthoeni, T.; Leitner, A.; Beck, M.; Hartl, F.-U.; Ban, N.; Malmstrom, L.; Aebersold, R. Structural Probing of a Protein Phosphatase 2A Network by Chemical Cross-Linking and Mass Spectrometry. *Science* **2012**, *337*, 1348−1353.
(234) Barysz, H. M.; Malmström, J. Development of Large-Scale Cross-Linking Mass Spectrometry. *Mol. Cell. Proteomics* **2018**, *17*, 1055−1066.
(235) Götze, M.; Iacobucci, C.; Ihling, C. H.; Sinz, A. A Simple Cross-Linking/Mass Spectrometry Workflow for Studying System-Wide Protein Interactions. *Anal. Chem.* **2019**, *91*, 10236−10244.
(236) Schweppe, D. K.; Chavez, J. D.; Lee, C. F.; Caudal, A.; Kruse, S. E.; Stuppard, R.; Marcinek, D. J.; Shadel, G. S.; Tian, R.; Bruce, J. E. Mitochondrial Protein Interactome Elucidated by Chemical Cross-Linking Mass Spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 1732−1737.
(237) Chavez, J. D.; Lee, C. F.; Caudal, A.; Keller, A.; Tian, R.; Bruce, J. E. Chemical Crosslinking Mass Spectrometry Analysis of Protein Conformations and Supercomplexes in Heart Tissue. *Cell Syst.* **2018**, *6*, 136−141.
(238) O'Reilly, F. J.; Xue, L.; Graziadei, A.; Sinn, L.; Lenz, S.; Tegunov, D.; Blötz, C.; Singh, N.; Hagen, W. J. H.; Cramer, P.; et al. In-Cell Architecture of an Actively Transcribing-Translating Expressome. *Science* **2020**, *369*, 554−557.
(239) Belsom, A.; Schneider, M.; Fischer, L.; Brock, O.; Rappsilber, J. Serum Albumin Domain Structures in Human Blood Serum by Mass Spectrometry and Computational Biology. *Mol. Cell. Proteomics* **2016**, *15*, 1105−1116.
(240) Liu, H.; Zhang, H.; Niedzwiedzki, D. M.; Prado, M.; He, G.; Gross, M. L.; Blankenship, R. E. Phycobilisomes Supply Excitations to Both Photosystems in a Megacomplex in Cyanobacteria. *Science* **2013**, *342*, 1104−1107.
(241) Schmidt, C.; Zhou, M.; Marriott, H.; Morgner, N.; Politis, A.; Robinson, C. V. Comparative Cross-Linking and Mass Spectrometry of an Intact F-Type ATPase Suggest a Role for Phosphorylation. *Nat. Commun.* **2013**, *4*, 1985.
(242) Chavez, J. D.; Schweppe, D. K.; Eng, J. K.; Bruce, J. E. In Vivo Conformational Dynamics of Hsp90 and Its Interactors. *Cell Chem. Biol.* **2016**, *23*, 716−726.
(243) Tüting, C.; Iacobucci, C.; Ihling, C. H.; Kastritis, P. L.; Sinz, A. Structural Analysis of 70S Ribosomes by Cross-Linking/Mass Spectrometry Reveals Conformational Plasticity. *Sci. Rep.* **2020**, *10*, 12618.
(244) Konermann, L.; Pan, J.; Liu, Y. H. Hydrogen Exchange Mass Spectrometry for Studying Protein Structure and Dynamics. *Chem. Soc. Rev.* **2011**, *40*, 1224−1234.
(245) Englander, S. W.; Kallenbach, N. R. Hydrogen Exchange and Structural Dynamics of Proteins and Nucleic Acids. *Q. Rev. Biophys.* **1983**, *16*, 521−655.
(246) Walters, B. T.; Ricciuti, A.; Mayne, L.; Englander, S. W. Minimizing Back Exchange in the Hydrogen Exchange-Mass Spectrometry Experiment. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 2132−2139.
(247) Zhang, Z.; Smith, D. L. Determination of Amide Hydrogen Exchange by Mass Spectrometry: A New Tool for Protein Structure Elucidation. *Protein Sci.* **1993**, *2*, 522−531.
(248) Marcsisin, S. R.; Engen, J. R. Hydrogen Exchange Mass Spectrometry: What Is It and What Can It Tell Us? *Anal. Bioanal. Chem.* **2010**, *397*, 967−972.
(249) Engen, J. R.; Botzanowski, T.; Peterle, D.; Georgescauld, F.; Wales, T. E. Developments in Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal. Chem.* **2021**, *93*, 567−582.
(250) Guttman, M.; Lee, K. K. Isotope Labeling of Biomolecules: Structural Analysis of Viruses by HDX-MS, 1st ed.; Elsevier, 2016; Vol. *566*.
(251) Trabjerg, E.; Nazari, Z. E.; Rand, K. D. Conformational Analysis of Complex Protein States by Hydrogen/Deuterium Exchange Mass Spectrometry (HDX-MS): Challenges and Emerging Solutions. *TrAC, Trends Anal. Chem.* **2018**, *106*, 125−138.
(252) Hamuro, Y.; Anand, G. S.; Kim, J. S.; Juliano, C.; Stranz, D. D.; Taylor, S. S.; Woods, V. L. Mapping Intersubunit Interactions of the Regulatory Subunit (RIα) in the Type I Holoenzyme of Protein Kinase A by Amide Hydrogen/Deuterium Exchange Mass Spectrometry (DXMS). *J. Mol. Biol.* **2004**, *340*, 1185−1196.
(253) Marciano, D. P.; Dharmarajan, V.; Griffin, P. R. HDX-MS Guided Drug Discovery: Small Molecules and Biopharmaceuticals. *Curr. Opin. Struct. Biol.* **2014**, *28*, 105−111.
(254) Illes-Toth, E.; Meisl, G.; Rempel, D. L.; Knowles, T. P. J.; Gross, M. L. Pulsed Hydrogen−Deuterium Exchange Reveals Altered Structures and Mechanisms in the Aggregation of Familial Alzheimer's Disease Mutants. *ACS Chem. Neurosci.* **2021**, *12*, 1972−1982.
(255) Lesne, J.; Locard-Paulet, M.; Parra, J.; Zivković, D.; Menneteau, T.; Bousquet, M. P.; Burlet-Schiltz, O.; Marcoux, J. Conformational Maps of Human 20S Proteasomes Reveal PA28- and Immuno-Dependent Inter-Ring Crosstalks. *Nat. Commun.* **2020**, *11*, 6140.
(256) Vahidi, S.; Bi, Y.; Dunn, S. D.; Konermann, L. Load-Dependent Destabilization of the γ-Rotor Shaft in FOF1 ATP Synthase Revealed by Hydrogen/Deuterium-Exchange Mass Spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 2412−2417.
(257) Mendoza, V. L.; Vachet, R. W. Protein Surface Mapping Using Diethylpyrocarbonate with Mass Spectrometric Detection. *Anal. Chem.* **2008**, *80*, 2895−2904.
(258) Xu, G.; Chance, M. R. Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics. *Chem. Rev.* **2007**, *107*, 3514−3543.
(259) Maleknia, S. D.; Downard, K. M. Advances in Radical Probe Mass Spectrometry for Protein Footprinting in Chemical Biology Applications. *Chem. Soc. Rev.* **2014**, *43*, 3244−3258.
(260) Hambly, D. M.; Gross, M. L. Laser Flash Photolysis of Hydrogen Peroxide to Oxidize Protein Solvent-Accessible Residues on the Microsecond Timescale. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 2057−2063.
(261) Johnson, D. T.; Di Stefano, L. H.; Jones, L. M. Fast Photochemical Oxidation of Proteins (FPOP): A Powerful Mass Spectrometry−Based Structural Proteomics Tool. *J. Biol. Chem.* **2019**, *294*, 11969−11979.
(262) Wu, L.; Lapidus, L. J. Combining Ultrarapid Mixing with Photochemical Oxidation to Probe Protein Folding. *Anal. Chem.* **2013**, *85*, 4920−4924.
(263) Tremblay, C. Y.; Limpikirati, P.; Vachet, R. W. Complementary Structural Information for Stressed Antibodies from Hydrogen−Deuterium Exchange and Covalent Labeling Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2021**, *32*, 1237−1248.
(264) Li, X.; Li, Z.; Xie, B.; Sharp, J. S. Improved Identification and Relative Quantification of Sites of Peptide and Protein Oxidation for Hydroxyl Radical Footprinting. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 1767−1776.
(265) Zhu, Y.; Guo, T.; Park, J. E.; Li, X.; Meng, W.; Datta, A.; Bern, M.; Lim, S. K.; Sze, S. K. Elucidating in Vivo Structural Dynamics in Integral Membrane Protein by Hydroxyl Radical Footprinting. *Mol. Cell. Proteomics* **2009**, *8*, 1999−2010.

(266) Espino, J. A.; Mali, V. S.; Jones, L. M. In Cell Footprinting Coupled with Mass Spectrometry for the Structural Analysis of Proteins in Live Cells. *Anal. Chem.* **2015**, *87*, 7971−7978.
(267) Rinas, A.; Mali, V. S.; Espino, J. A.; Jones, L. M. Development of a Microflow System for In-Cell Footprinting Coupled with Mass Spectrometry. *Anal. Chem.* **2016**, *88*, 10052−10058.
(268) Espino, J. A.; Jones, L. M. Illuminating Biological Interactions with in Vivo Protein Footprinting. *Anal. Chem.* **2019**, *91*, 6577−6584.
(269) Shen, G.; Cui, W.; Zhang, H.; Zhou, F.; Huang, W.; Liu, Q.; Yang, Y.; Li, S.; Bowman, G. R.; Sadler, J. E.; et al. Warfarin Traps Human Vitamin K Epoxide Reductase in an Intermediate State during Electron Transfer. *Nat. Struct. Mol. Biol.* **2017**, *24*, 69−76.
(270) Rhee, H. W.; Zou, P.; Udeshi, N. D.; Martell, J. D.; Mootha, V. K.; Carr, S. A.; Ting, A. Y. Proteomic Mapping of Mitochondria in Living Cells via Spatially Restricted Enzymatic Tagging. *Science* **2013**, *339*, 1328−1331.
(271) Joeh, E.; Reeves, A. E.; Parker, C. G.; Huang, M. L. Mapping Interactions between Glycans and Glycan-Binding Proteins by Live Cell Proximity Tagging. *Curr. Protoc.* **2021**, *1*, e104.
(272) Lobingier, B. T.; Hüttenhain, R.; Eichel, K.; Miller, K. B.; Ting, A. Y.; von Zastrow, M.; Krogan, N. J. An Approach to Spatiotemporally Resolve Protein Interaction Networks in Living Cells. *Cell* **2017**, *169*, 350−360.
(273) Roux, K. J.; Kim, D. I.; Raida, M.; Burke, B. A Promiscuous Biotin Ligase Fusion Protein Identifies Proximal and Interacting Proteins in Mammalian Cells. *J. Cell Biol.* **2012**, *196* (6), 801−810.
(274) Kim, D. I.; Jensen, S. C.; Noble, K. A.; Kc, B.; Roux, K. H.; Motamedchaboki, K.; Roux, K. J. An Improved Smaller Biotin Ligase for BioID Proximity Labeling. *Mol. Biol. Cell* **2016**, *27*, 1188−1196.
(275) Branon, T. C.; Bosch, J. A.; Sanchez, A. D.; Udeshi, N. D.; Svinkina, T.; Carr, S. A.; Feldman, J. L.; Perrimon, N.; Ting, A. Y. Efficient Proximity Labeling in Living Cells and Organisms with TurboID. *Nat. Biotechnol.* **2018**, *36*, 880−898.
(276) Monti, M.; Pucci, P. Limited Proteolysis Mass Spectrometry of Protein Complexes. *Mass Spectrom. Protein Interact.* **2007**, 63−82.
(277) Feng, Y.; De Franceschi, G.; Kahraman, A.; Soste, M.; Melnik, A.; Boersema, P. J.; De Laureto, P. P.; Nikolaev, Y.; Oliveira, A. P.; Picotti, P. Global Analysis of Protein Structural Changes in Complex Proteomes. *Nat. Biotechnol.* **2014**, *32*, 1036−1044.
(278) Schopper, S.; Kahraman, A.; Leuenberger, P.; Feng, Y.; Piazza, I.; Müller, O.; Boersema, P. J.; Picotti, P. Measuring Protein Structural Changes on a Proteome-Wide Scale Using Limited Proteolysis-Coupled Mass Spectrometry. *Nat. Protoc.* **2017**, *12*, 2391−2410.
(279) Piazza, I.; Kochanowski, K.; Cappelletti, V.; Fuhrer, T.; Noor, E.; Sauer, U.; Picotti, P. A Map of Protein-Metabolite Interactions Reveals Principles of Chemical Communication. *Cell* **2018**, *172*, 358−372.
(280) Toby, T. K.; Fornelli, L.; Kelleher, N. L. Progress in Top-Down Proteomics and the Analysis of Proteoforms. *Annu. Rev. Anal. Chem.* **2016**, *9*, 499−519.
(281) Melby, J. A.; Roberts, D. S.; Larson, E. J.; Brown, K. A.; Bayne, E. F.; Jin, S.; Ge, Y. Novel Strategies to Address the Challenges in Top-Down Proteomics. *J. Am. Soc. Mass Spectrom.* **2021**, *32*, 1278−1294.
(282) Polasky, D. A.; Dixit, S. M.; Keating, M. F.; Gadkari, V. V.; Andrews, P. C.; Ruotolo, B. T. Pervasive Charge Solvation Permeates Native-like Protein Ions and Dramatically Influences Top-down Sequencing Data. *J. Am. Chem. Soc.* **2020**, *142*, 6750−6760.
(283) Shaw, J. B.; Malhan, N.; Vasil'Ev, Y. V.; Lopez, N. I.; Makarov, A.; Beckman, J. S.; Voinov, V. G. Sequencing Grade Tandem Mass Spectrometry for Top-Down Proteomics Using Hybrid Electron Capture Dissociation Methods in a Benchtop Orbitrap Mass Spectrometer. *Anal. Chem.* **2018**, *90*, 10819−10827.
(284) Wu, D.; Struwe, W. B.; Harvey, D. J.; Ferguson, M. A. J.; Robinson, C. V. N-Glycan Microheterogeneity Regulates Interactions of Plasma Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 8763−8768.
(285) Ayaz-Guner, S.; Zhang, J.; Li, L.; Walker, J. W.; Ge, Y. In Vivo Phosphorylation Site Mapping in Mouse Cardiac Troponin I by High Resolution Top-down Electron Capture Dissociation Mass Spectrometry: Ser22/23 Are the Only Sites Basally Phosphorylated. *Biochemistry* **2009**, *48*, 8161−8170.
(286) Chamot-Rooke, J.; Mikaty, G.; Malosse, C.; Soyer, M.; Dumont, A.; Gault, J.; Imhaus, A.-F.; Martin, P.; Trellet, M.; Clary, G.; et al. Posttranslational Modification of Pili upon Cell Contact Triggers N. Meningitidis Dissemination. *Science* **2011**, *331*, 778−782.
(287) Ntai, I.; Fornelli, L.; DeHart, C. J.; Hutton, J. E.; Doubleday, P. F.; LeDuc, R. D.; van Nispen, A. J.; Fellers, R. T.; Whiteley, G.; Boja, E. S.; et al. Precise Characterization of KRAS4b Proteoforms in Human Colorectal Cells and Tumors Reveals Mutation/Modification Cross-Talk. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 4140−4145.
(288) Tran, J. C.; Zamdborg, L.; Ahlf, D. R.; Lee, J. E.; Catherman, A. D.; Durbin, K. R.; Tipton, J. D.; Vellaichamy, A.; Kellie, J. F.; Li, M.; et al. Mapping Intact Protein Isoforms in Discovery Mode Using Top-down Proteomics. *Nature* **2011**, *480*, 254−258.
(289) Peng, Y.; Gregorich, Z. R.; Valeja, S. G.; Zhang, H.; Cai, W.; Chen, Y. C.; Guner, H.; Chen, A. J.; Schwahn, D. J.; Hacker, T. A.; et al. Top-down Proteomics Reveals Concerted Reductions in Myofilament and Z-Disc Protein Phosphorylation after Acute Myocardial Infarction. *Mol. Cell. Proteomics* **2014**, *13*, 2752−2764.
(290) Franc, V.; Zhu, J.; Heck, A. J. R. Comprehensive Proteoform Characterization of Plasma Complement Component C8$\alpha\beta\gamma$ by Hybrid Mass Spectrometry Approaches. *J. Am. Soc. Mass Spectrom.* **2018**, *29*, 1099−1110.
(291) Ben-Nissan, G.; Belov, M. E.; Morgenstern, D.; Levin, Y.; Dym, O.; Arkind, G.; Lipson, C.; Makarov, A. A.; Sharon, M. Triple-Stage Mass Spectrometry Unravels the Heterogeneity of an Endogenous Protein Complex. *Anal. Chem.* **2017**, *89*, 4708−4715.
(292) Vimer, S.; Ben-Nissan, G.; Morgenstern, D.; Kumar-Deshmukh, F.; Polkinghorn, C.; Quintyn, R. S.; Vasil'Ev, Y. V.; Beckman, J. S.; Elad, N.; Wysocki, V. H.; et al. Comparative Structural Analysis of 20S Proteasome Ortholog Protein Complexes by Native Mass Spectrometry. *ACS Cent. Sci.* **2020**, *6*, 573−588.
(293) Leney, A. C.; Heck, A. J. R. Native Mass Spectrometry: What Is in the Name? *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 5−13.
(294) Ruotolo, B. T.; Giles, K.; Campuzano, I.; Sandercock, A. M.; Bateman, R. H.; Robinson, C. V. Evidence for Macromolecular Protein Rings in the Absence of Bulk Water. *Science* **2005**, *310*, 1658−1661.
(295) Hamdy, O. M.; Julian, R. R. Reflections on Charge State Distributions, Protein Structure, and the Mystical Mechanism of Electrospray Ionization. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 1−6.
(296) Kafader, J. O.; Melani, R. D.; Schachner, L. F.; Ives, A. N.; Patrie, S. M.; Kelleher, N. L.; Compton, P. D. Native vs Denatured: An in Depth Investigation of Charge State and Isotope Distributions. *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 574−581.
(297) Marklund, E. G.; Degiacomi, M. T.; Robinson, C. V.; Baldwin, A. J.; Benesch, J. L. P. Collision Cross Sections for Structural Proteomics. *Structure* **2015**, *23*, 791−799.
(298) Allison, T. M.; Barran, P.; Cianférani, S.; Degiacomi, M. T.; Gabelica, V.; Grandori, R.; Marklund, E. G.; Menneteau, T.; Migas, L. G.; Politis, A.; et al. Computational Strategies and Challenges for Using Native Ion Mobility Mass Spectrometry in Biophysics and Structural Biology. *Anal. Chem.* **2020**, *92*, 10872−10880.
(299) Marty, M. T.; Baldwin, A. J.; Marklund, E. G.; Hochberg, G. K. A.; Benesch, J. L. P.; Robinson, C. V. Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. *Anal. Chem.* **2015**, *87*, 4370−4376.
(300) Reid, D. J.; Diesing, J. M.; Miller, M. A.; Perry, S. M.; Wales, J. A.; Montfort, W. R.; Marty, M. T. MetaUniDec: High-Throughput Deconvolution of Native Mass Spectra. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 118−127.
(301) Quetschlich, D.; Esser, T. K; Newport, T. D; Fiorentino, F.; Shutin, D.; Chen, S.; Davis, R.; Lovera, S.; Liko, I.; Stansfeld, P. J; Robinson, C. V NaViA: A Program for the Visual Analysis of Complex Mass Spectra. *Bioinformatics* **2021**, btab436.
(302) Bush, M. F.; Hall, Z.; Giles, K.; Hoyes, J.; Robinson, C. V.; Ruotolo, B. T. Collision Cross Sections of Proteins and Their

Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology. *Anal. Chem.* **2010**, *82*, 9557−9565.
(303) Lee, J.-Y.; Bilbao, A.; Conant, C. R.; Bloodsworth, K. J.; Orton, D. J.; Zhou, M.; Wilson, J. W.; Zheng, X.; Webb, I. K.; Li, A.; et al. AutoCCS: Automated Collision Cross-Section Calculation Software for Ion Mobility Spectrometry−Mass Spectrometry. *Bioinformatics* **2021**, btab429.
(304) Polasky, D. A.; Dixit, S. M.; Fantin, S. M.; Ruotolo, B. T. CIUSuite 2: Next-Generation Software for the Analysis of Gas-Phase Protein Unfolding Data. *Anal. Chem.* **2019**, *91*, 3147−3155.
(305) Migas, L. G.; France, A. P.; Bellina, B.; Barran, P. E. ORIGAMI: A Software Suite for Activated Ion Mobility Mass Spectrometry (AIM-MS) Applied to Multimeric Protein Assemblies. *Int. J. Mass Spectrom.* **2018**, *427*, 20−28.
(306) Kafader, J. O.; Melani, R. D.; Durbin, K. R.; Ikwuagwu, B.; Early, B. P.; Fellers, R. T.; Beu, S. C.; Zabrouskov, V.; Makarov, A. A.; Maze, J. T.; et al. Multiplexed Mass Spectrometry of Individual Ions Improves Measurement of Proteoforms and Their Complexes. *Nat. Methods* **2020**, *17*, 391−394.
(307) McGee, J. P.; Melani, R. D.; Yip, P. F.; Senko, M. W.; Compton, P. D.; Kafader, J. O.; Kelleher, N. L. Isotopic Resolution of Protein Complexes up to 466 KDa Using Individual Ion Mass Spectrometry. *Anal. Chem.* **2021**, *93*, 2723−2727.
(308) Todd, A. R.; Jarrold, M. F. Dynamic Calibration Enables High-Accuracy Charge Measurements on Individual Ions for Charge Detection Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 1241−1248.
(309) Synowsky, S. A.; van den Heuvel, R. H. H.; Mohammed, S.; Pim Pijnappel, W. W. M.; Heck, A. J. R. Probing Genuine Strong Interactions and Post-Translational Modifications in the Heterogeneous Yeast Exosome Protein Complex. *Mol. Cell. Proteomics* **2006**, *5*, 1581−1592.
(310) Synowsky, S. A.; van Wijk, M.; Raijmakers, R.; Heck, A. J. R. Comparative Multiplexed Mass Spectrometric Analyses of Endogenously Expressed Yeast Nuclear and Cytoplasmic Exosomes. *J. Mol. Biol.* **2009**, *385*, 1300−1313.
(311) Olinares, P. D. B.; Dunn, A. D.; Padovan, J. C.; Fernandez-Martinez, J.; Rout, M. P.; Chait, B. T. A Robust Workflow for Native Mass Spectrometric Analysis of Affinity-Isolated Endogenous Protein Assemblies. *Anal. Chem.* **2016**, *88*, 2799−2807.
(312) Olinares, P. D. B.; Chait, B. T. Native Mass Spectrometry Analysis of Affinity-Captured Endogenous Yeast RNA Exosome Complexes. *Methods Mol. Biol.* **2020**, *2062*, 357−382.
(313) Synowsky, S. A.; Heck, A. J. R. The Yeast Ski Complex Is a Hetero-Tetramer. *Protein Sci.* **2007**, *17*, 119−125.
(314) Barrera, N. P.; Isaacson, S. C.; Zhou, M.; Bavro, V. N.; Welch, A.; Schaedler, T. A.; Seeger, M. A.; Miguel, R. N.; Korkhov, V. M.; van Veen, H. W.; et al. Mass Spectrometry of Membrane Transporters Reveals Subunit Stoichiometry and Interactions. *Nat. Methods* **2009**, *6*, 585−587.
(315) 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.; Stewart, A. G.; Matthews, S. J.; Grunewald, K.; Robinson, C. V. Protein Assemblies Ejected Directly from Native Membranes Yield Complexes for Mass Spectrometry. *Science* **2018**, *362*, 829−834.
(316) Zhang, H.; Harrington, L. B.; Lu, Y.; Prado, M.; Saer, R.; Rempel, D.; Blankenship, R. E.; Gross, M. L. Native Mass Spectrometry Characterizes the Photosynthetic Reaction Center Complex from the Purple Bacterium Rhodobacter Sphaeroides. *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 87−95.
(317) Gordiyenko, Y.; Videler, H.; Zhou, M.; McKay, A. R.; Fucini, P.; Biegel, E.; Muller, V.; Robinson, C. V. Mass Spectrometry Defines the Stoichiometry of Ribosomal Stalk Complexes across the Phylogenetic Tree. *Mol. Cell. Proteomics* **2010**, *9*, 1774−1783.
(318) Van De Waterbeemd, M.; Fort, K. L.; Boll, D.; Reinhardt-Szyba, M.; Routh, A.; Makarov, A.; Heck, A. J. R. High-Fidelity Mass Analysis Unveils Heterogeneity in Intact Ribosomal Particles. *Nat. Methods* **2017**, *14*, 283−286.
(319) van de Waterbeemd, M.; Tamara, S.; Fort, K. L.; Damoc, E.; Franc, V.; Bieri, P.; Itten, M.; Makarov, A.; Ban, N.; Heck, A. J. R. Dissecting Ribosomal Particles throughout the Kingdoms of Life Using Advanced Hybrid Mass Spectrometry Methods. *Nat. Commun.* **2018**, *9*, 2493.
(320) Hernández, H.; Dziembowski, A.; Taverner, T.; Séraphin, B.; Robinson, C. V. Subunit Architecture of Multimeric Complexes Isolated Directly from Cells. *EMBO Rep.* **2006**, *7*, 605−610.
(321) Sharon, M.; Horovitz, A. Probing Allosteric Mechanisms Using Native Mass Spectrometry. *Curr. Opin. Struct. Biol.* **2015**, *34*, 7−16.
(322) Gault, J.; Donlan, J. A. C.; Liko, I.; Hopper, J. T. S.; Gupta, K.; Housden, N. G.; Struwe, W. B.; Marty, M. T.; Mize, T.; Bechara, C.; et al. High-Resolution Mass Spectrometry of Small Molecules Bound to Membrane Proteins. *Nat. Methods* **2016**, *13*, 333−336.
(323) Chorev, D. S.; Tang, H.; Rouse, S. L.; Bolla, J. R.; von Kügelgen, A.; Baker, L. A.; Wu, D.; Gault, J.; Grünewald, K.; Bharat, T. A. M.; et al. The Use of Sonicated Lipid Vesicles for Mass Spectrometry of Membrane Protein Complexes. *Nat. Protoc.* **2020**, *15*, 1690−1706.
(324) Gupta, K.; Donlan, J. A. C.; Hopper, J. T. S.; Uzdavinys, P.; Landreh, M.; Struwe, W. B.; Drew, D.; Baldwin, A. J.; Stansfeld, P. J.; Robinson, C. V. The Role of Interfacial Lipids in Stabilizing Membrane Protein Oligomers. *Nature* **2017**, *541*, 421−424.
(325) Vimer, S.; Ben-Nissan, G.; Sharon, M. Direct Characterization of Overproduced Proteins by Native Mass Spectrometry. *Nat. Protoc.* **2020**, *15*, 236−265.
(326) Sakamoto, W.; Azegami, N.; Konuma, T.; Akashi, S. Single-Cell Native Mass Spectrometry of Human Erythrocytes. *Anal. Chem.* **2021**, *93*, 6583−6588.
(327) Huang, W.; Ravikumar, K. M.; Chance, M. R.; Yang, S. Quantitative Mapping of Protein Structure by Hydroxyl Radical Footprinting-Mediated Structural Mass Spectrometry: A Protection Factor Analysis. *Biophys. J.* **2015**, *108*, 107−115.
(328) Xie, B.; Sood, A.; Woods, R. J.; Sharp, J. S. Quantitative Protein Topography Measurements by High Resolution Hydroxyl Radical Protein Footprinting Enable Accurate Molecular Model Selection. *Sci. Rep.* **2017**, *7*, 4552.
(329) Biehn, S. E.; Lindert, S. Accurate Protein Structure Prediction with Hydroxyl Radical Protein Footprinting Data. *Nat. Commun.* **2021**, *12*, 341.
(330) Politis, A.; Park, A. Y.; Hyung, S. J.; Barsky, D.; Ruotolo, B. T.; Robinson, C. V. Integrating Ion Mobility Mass Spectrometry with Molecular Modelling to Determine the Architecture of Multiprotein Complexes. *PLoS One* **2010**, *5*, No. e12080.
(331) Kyrilis, F. L.; Semchonok, D. A.; Skalidis, I.; Tuting, C.; Hamdi, F.; O'Reilly, F. J.; Rappsilber, J.; Kastritis, P. L. Integrative Structure of a 10-Megadalton Eukaryotic Pyruvate Dehydrogenase Complex from Native Cell Extracts. *Cell Rep.* **2021**, *34*, 108727.
(332) Ward, A. B.; Sali, A.; Wilson, I. A. Integrative Structural Biology. *Science* **2013**, *339*, 913−916.
(333) Rout, M. P.; Sali, A. Principles for Integrative Structural Biology Studies. *Cell* **2019**, *177*, 1384−1403.
(334) Das, R.; Baker, D. Macromolecular Modeling with Rosetta. *Annu. Rev. Biochem.* **2008**, *77*, 363−382.
(335) Russel, D.; Lasker, K.; Webb, B.; Velázquez-Muriel, J.; Tjioe, E.; Schneidman-Duhovny, D.; Peterson, B.; Sali, A. Putting the Pieces Together: Integrative Modeling Platform Software for Structure Determination of Macromolecular Assemblies. *PLoS Biol.* **2012**, *10*, 1001244.
(336) Braitbard, M.; Schneidman-Duhovny, D.; Kalisman, N. Integrative Structure Modeling: Overview and Assessment. *Annu. Rev. Biochem.* **2019**, *88*, 113−135.
(337) Alber, F.; Dokudovskaya, S.; Veenhoff, L. M.; Zhang, W.; Kipper, J.; Devos, D.; Suprapto, A.; Karni-Schmidt, O.; Williams, R.; Chait, B. T.; et al. Determining the Architectures of Macromolecular Assemblies. *Nature* **2007**, *450*, 683−694.
(338) Von Appen, A.; Kosinski, J.; Sparks, L.; Ori, A.; DiGuilio, A. L.; Vollmer, B.; Mackmull, M. T.; Banterle, N.; Parca, L.; Kastritis, P.;

et al. In Situ Structural Analysis of the Human Nuclear Pore Complex. *Nature* **2015**, *526*, 140−143.
(339) Kosinski, J.; Mosalaganti, S.; von Appen, A.; Teimer, R.; Diguilio, A. L.; Wan, W.; Bui, K. H.; Hagen, W. J. H.; Briggs, J. a G.; Glavy, J. S.; et al. Molecular Architecture of the Inner Ring Scaffold of the Human Nuclear Pore Complex. *Science* **2016**, *352*, 363−365.
(340) Kasinath, V.; Faini, M.; Poepsel, S.; Reif, D.; Feng, X. A.; Stjepanovic, G.; Aebersold, R.; Nogales, E. Structures of Human PRC2 with Its Cofactors AEBP2 and JARID2. *Science* **2018**, *359*, 940−944.
(341) McCafferty, C. L.; Verbeke, E. J.; Marcotte, E. M.; Taylor, D. W. Structural Biology in the Multi-Omics Era. *J. Chem. Inf. Model.* **2020**, *60*, 2424−2429.
(342) Ho, C.; Li, X.; Lai, M.; Terwilliger, T. C.; Beck, J. R.; Wohlschlegel, J.; Goldberg, D. E.; Fitzpatrick, A. W. P.; Zhou, Z. H. Bottom-up Structural Proteomics: CryoEM of Complexes Enriched from the Cellular Milieu. *Nat. Methods* **2020**, *17*, 79−85.
(343) Verbeke, E. J.; Mallam, A. L.; Drew, K.; Marcotte, E. M.; Taylor, D. W. Classification of Single Particles from Human Cell Extract Reveals Distinct Structures. *Cell Rep.* **2018**, *24*, 259.
(344) Su, C. C.; Lyu, M.; Morgan, C. E.; Bolla, J. R.; Robinson, C. V.; Yu, E. W. A 'Build and Retrieve' Methodology to Simultaneously Solve Cryo-EM Structures of Membrane Proteins. *Nat. Methods* **2021**, *18*, 69−75.
(345) Malm, M.; Saghaleyni, R.; Lundqvist, M.; Giudici, M.; Chotteau, V.; Field, R.; Varley, P. G.; Hatton, D.; Grassi, L.; Svensson, T.; et al. Evolution from Adherent to Suspension: Systems Biology of HEK293 Cell Line Development. *Sci. Rep.* **2020**, *10*, 18966.
(346) Keifer, D. Z.; Jarrold, M. F. Single-Molecule Mass Spectrometry. *Mass Spectrom. Rev.* **2017**, *36*, 715−733.
(347) Wörner, T. P.; Snijder, J.; Bennett, A.; Agbandje-McKenna, M.; Makarov, A. A.; Heck, A. J. R. Resolving Heterogeneous Macromolecular Assemblies by Orbitrap-Based Single-Particle Charge Detection Mass Spectrometry. *Nat. Methods* **2020**, *17*, 395−398.
(348) Kempa, E. E.; Smith, C. A.; Li, X.; Bellina, B.; Richardson, K.; Pringle, S.; Galman, J. L.; Turner, N. J.; Barran, P. E. Coupling Droplet Microfluidics with Mass Spectrometry for Ultrahigh-Throughput Analysis of Complex Mixtures up to and above 30 Hz. *Anal. Chem.* **2020**, *92*, 12605−12612.
(349) Cong, Y.; Motamedchaboki, K.; Misal, S. A.; Liang, Y.; Guise, A. J.; Truong, T.; Huguet, R.; Plowey, E. D.; Zhu, Y.; Lopez-Ferrer, D.; et al. Ultrasensitive Single-Cell Proteomics Workflow Identifies > 1000 Protein Groups per Mammalian Cell. *Chem. Sci.* **2021**, *12*, 1001−1006.
(350) Brunner, A. D.; Thielert, M.; Vasilopoulou, C.; Ammar, C.; Coscia, F.; Mund, A.; Horning, O. B.; Bache, N.; Apalategui, A.; Lubeck, M.; et al. Ultra-High Sensitivity Mass Spectrometry Quantifies Single-Cell Proteome Changes upon Perturbation. *bioRxiv* **2020**, DOI: 10.1101/2020.12.22.423933.
(351) Dou, Y.; Barbosa, I.; Jiang, H.; Iasillo, C.; Molloy, K. R; Schulze, W. M.; Cusack, S.; Schmid, M.; Le Hir, H.; LaCava, J.; Jensen, T. H.; et al. NCBP3 Positively Impacts MRNA Biogenesis. *Nucleic Acids Res.* **2020**, *48* (18), 10413−10427.
(352) Dou, Y.; Kalmykova, S.; Pashkova, M.; Oghbaie, M.; Jiang, H.; Molloy, K. R.; Chait, B. T.; Rout, M. P.; Fenyö, D.; Jensen, T. H.; Altukhov, I.; LaCava, J. Affinity proteomic dissection of the human nuclear cap-binding complex interactome. *Nucleic Acids Res.* **2020**, *48*, 10456.