

Native MS for structural biology research

Keywords: Orbitrap mass analyzer, mass spectrometry, Q Exactive UHMR MS, m/z , protein, complex, intact, liquid chromatography, biomolecule

Executive summary

The characterization of protein and protein complex structure is crucial to understanding protein function and the roles proteins play in biological systems. Traditional techniques used to examine protein structures have a number of limitations, including the sample amount needed for the experiment, the size of the structure and the availability of the sample in the condition needed for analysis. Historically, mass spectrometry has been an alternative, albeit a minor player in protein structural analysis. However, recent innovations in hardware, software and workflows have pushed mass spectrometry to the forefront of structural analysis. Native mass spectrometry is a powerful tool in the mass spectrometry toolbox. This white paper will focus on the information provided by native mass spectrometry, and how the development of the latest tools, from sample preparation through to data analysis, have made this workflow more informative and streamlined.

Introduction

It is essential to characterize protein complex assembly and structure to understand protein function and mechanism of action in biological systems. It is these protein complexes, the networks they form and the various interactions that occur during complex formation that govern biological

activities. Protein structure elucidation is also important in the area of drug design. Researchers synthesizing drugs require structural information to design molecules that could fit into the pockets of proteins which in turn can aid in the treatment of diseases.

Current techniques used to study protein complex structures and protein-protein interactions include X-ray crystallography, cryo-electron microscopy (cryo-EM) and nuclear magnetic resonance (NMR). However, these approaches require significant amounts of highly purified proteins and may not allow for the analysis of proteins in their native conditions. They also struggle with intrinsically disordered proteins and protein flexibility. These techniques can also be limited by the size of the biomolecule that is being examined. Furthermore, many proteins are simply not amenable to these types of analysis, thereby limiting the accessibility of these techniques. For example, in X-ray crystallography it is often very difficult to produce well diffracting crystals needed for analysis. This is further exacerbated in complexes where more than one protein needs to be crystalized. Mass spectrometry (MS) techniques have been used for many years to study protein complexes, protein structures and protein-protein interactions, but limited to highly specialized and instrument savvy research laboratories. Additionally, due to their complexity and requirement for specialized sample preparation, advanced MS feature requirements, as well as fit-for-purpose data analysis, they have lagged behind general proteomics analyses.

The white paper presented here is intended to introduce native mass spectrometry (native MS) as a powerful solution for studying protein structures, complexes and interactions. This paper will address the advantage of native MS, the information it provides and the complimentary role it plays in concert with traditional techniques. The latest tools and workflows that have been developed to democratize native MS will be discussed.

Protein and protein complex studies

Proteins are pivotal to virtually every biological process that occurs within a cell, from gene expression to cell growth and proliferation, intercellular communication and apoptosis. Examining the roles proteins play in biological processes can be challenging due to their dynamic nature. Proteins within a cell are continuously stimulated by external factors that change their dynamics and properties. Since proteins are expressed in a cell-dependent fashion, they can vary in type and in structure from one cell to another. These protein characteristics suggest a complexity that can be difficult to investigate when trying to understand protein function in a biological context. The complexity is further compounded by the fact that the majority of proteins interact with one another forming larger complexes via non-covalent interactions for biological activity.

The specific structure of a protein influences the interaction with other proteins, as well as lipids, glycans, small molecules and nucleic acids. This in turn can affect function. The characterization of the three-dimensional (3-D) structure of a protein or protein complex helps explain the role of proteins in biological functions. The study of characterizing protein and protein complex structures is called structural biology.

Native mass spectrometry

There are number of MS tools available for elucidating protein or protein complex structures (Figure 1). Each approach provides complementary information about the structures of proteins and protein complexes. Native MS is one such tool in the MS toolbox for structural biology. In native MS, protein structures are kept intact and introduced

into the mass spectrometer in the similar structural configuration that they exist in biological conditions. Researchers try to refrain from using any conditions that can alter the structure of the protein in solution, such as use of acids or denaturing solvents. If the focus is on protein complexes then the analysis is attractive because the approach can maintain the non-covalent interactions between the protein and its binding partners (other proteins, biomolecules or small molecules). Hence, why the technique is called native MS.

The primary information obtained by native MS is listed below:

- **Mass of the intact protein complex or the intact protein** – measures the overall mass of a protein or protein complex. The initial mass measurement is the starting point for obtaining additional information such as the masses of the subunits in a protein complex.
- **Subunit stoichiometry** – provides information on the composition of the subunits that make up the protein complex. Specifically, the number of each unique subunits that are present.
- **Subunit identification** – reveals information on whether the subunit is bound to the complex of interest and not simply copurified.
- **Biomolecule binding** – provides information on the interaction between the different subunits or ligands. Whether the interaction is happening between a protein and another protein, or with nucleic acids, lipids, glycans and small molecules.
- **Protein complex topology** – describes the interconnectivity of the subunits. For example, the location of the subunits within the protein complex, and whether these subunits present within the inner or outer portion of the complex.
- **Protein dynamics** – protein or protein complexes can undergo conformational changes. The changes that occur between the static state is defined as protein dynamics.

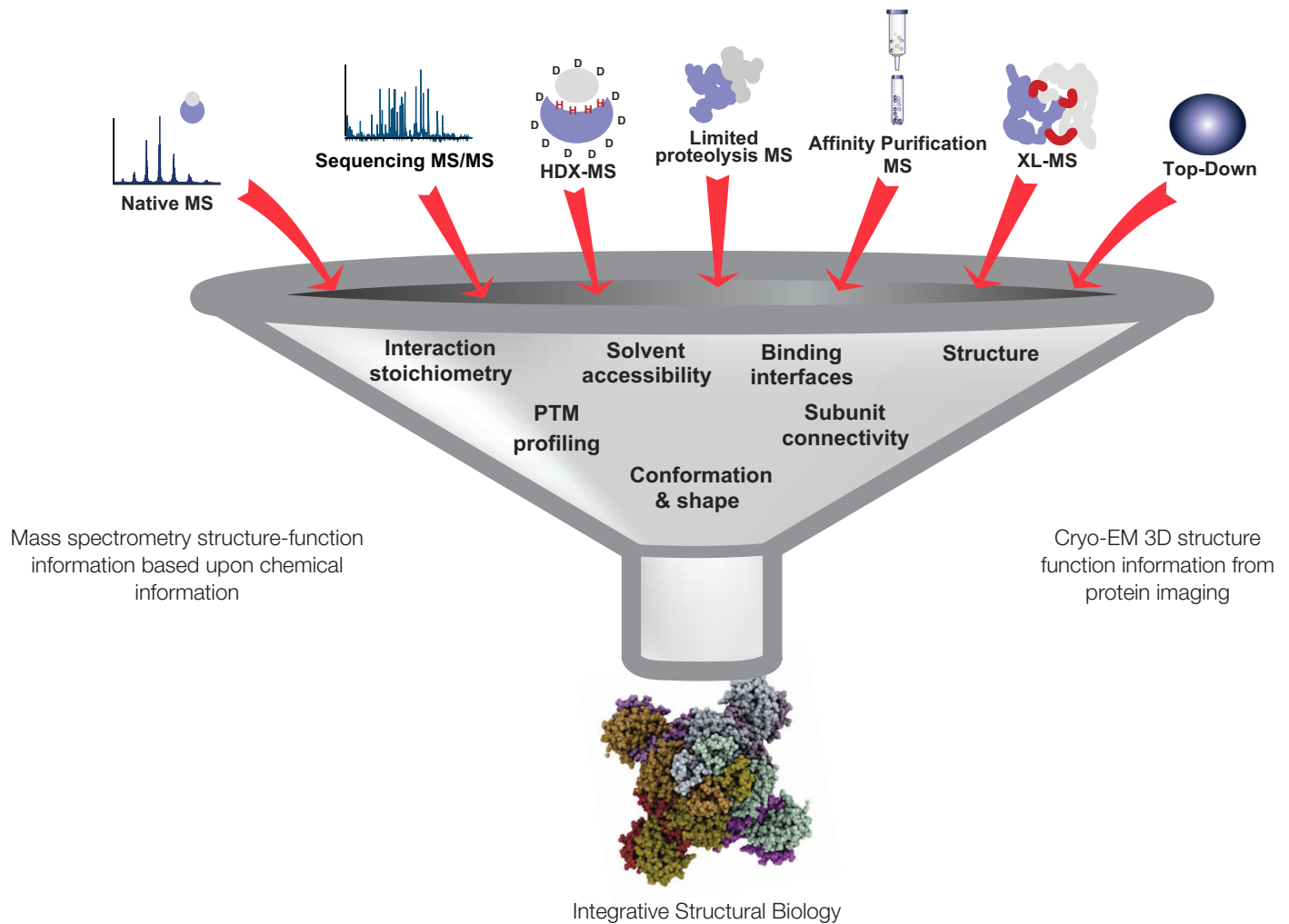


Figure 1. MS tools for performing integrative structural biology.

Intact mass

The simplest experiment performed in native MS is the measurement of the intact protein or protein complex mass. The mass can inform whether the protein of interest is being examined or if it's another protein. If the analysis involves a single protein, the mass measurement can provide information on the presence of posttranslational modifications (PTMs) such as glycosylation (Figure 2). The number and the type of glycoform compositions that are present can be obtained using native MS. Phosphorylation is another PTM that can be elucidated using native MS in protein or protein complex. From the mass measurement, native MS can also reveal information on sample heterogeneity, whether there is a single protein or multiple proteins present in the sample. Such information is useful for obtaining insights into the dynamics of protein complexes. This information can also be used as a screening tool for analysis with techniques such as cryo-EM.

Subunit stoichiometry

If dealing with protein complexes, native MS mass measurement can provide information on the stoichiometry of the subunits, the exact ratios of the individual components of the protein complex. As a first step, intact mass measurements are taken of the protein complex in denatured condition, providing masses of the individual subunits. This is followed by native MS of the protein complex, providing the mass of the intact complex. The combined information from the masses of subunits and the protein complex allow researchers to calculate the stoichiometry of the subunits, composition of the complex, and aids in deducing the presence of tightly binding ligands or homogenous PTMs. Stoichiometry information is important because many cellular processes are regulated by association or dissociation of specific protein complexes, ligands, or enzymes. Protein stoichiometry aims to measure the exact ratios of the individual components of these protein complexes, which is a requirement for fully understanding their overall function.¹

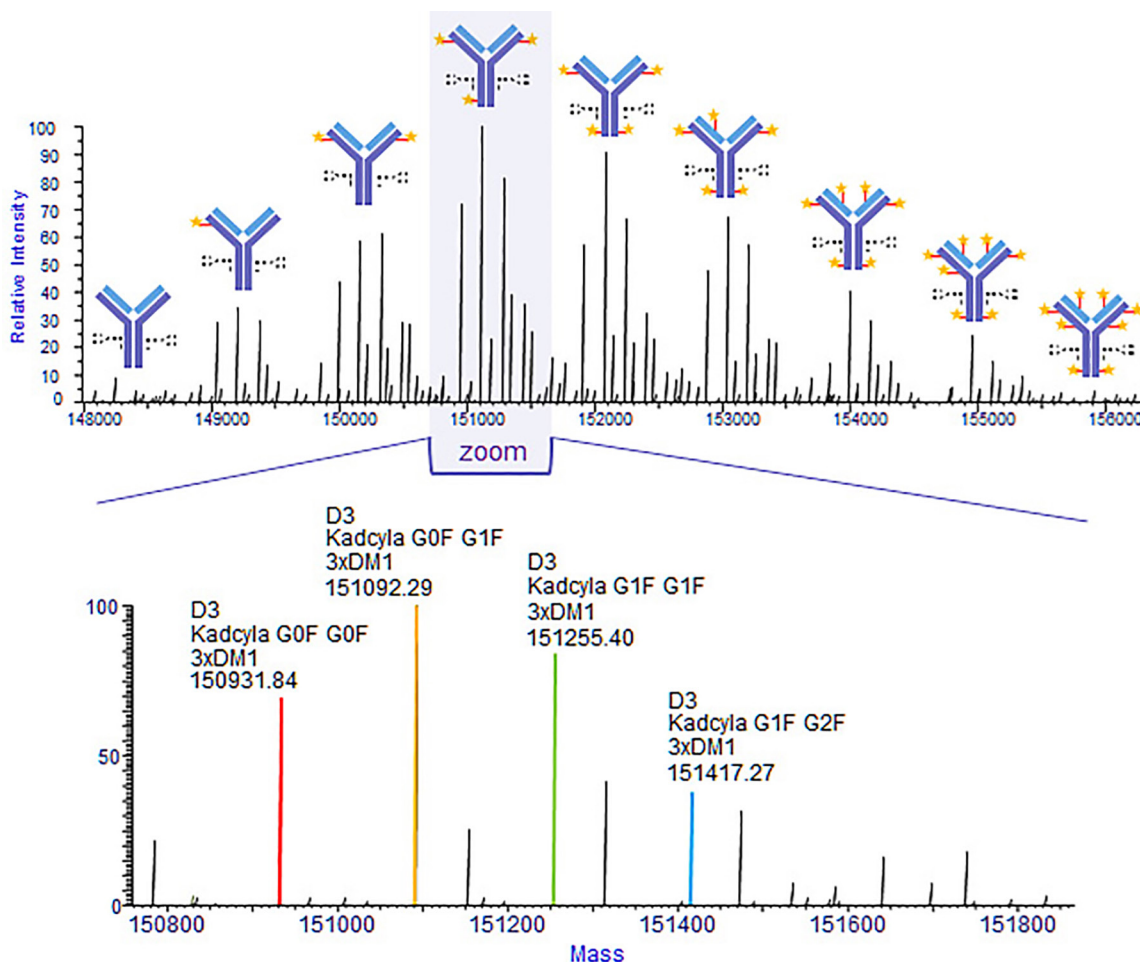


Figure 2. Deconvoluted spectrum of native MS of Trastuzumab Emtansine Lys-linked ADC. Multiple glycoforms could be detected and their composition assigned.²

Biomolecule binding

Native MS can also provide information on binding partners (Figure 3). For example, whether the interactions are in fact occurring and what the binding partners might be (proteins, nucleic acids, lipids, glycans or small molecules). By further fragmenting the structure within the mass spectrometer and combining it with native MS profile, it is possible to get the identity and the stoichiometry of the binding partner.

Protein complex topology

Native MS can also provide information on protein complex topology – the organization of the subunits that make up the larger complex.^{3,4} The intricacy of the protein complexes can vary, from being simple structures made up of symmetrical homodimers to very complex structures with no symmetry containing heterogeneous subunits. There are a number of approaches currently available for determining topology relying on the use of MS/MS. Fragmentation techniques such as collision-induced

dissociation (CID), surface-induced dissociation (SID) and ultraviolet photodissociation (UVPD) have been used to characterize protein topology.

The key factor in all these approaches is to maintain the arrangement of the subunits as they are present in the much larger complex during dissociation. Thus, starting with native MS acquisition ensures that the subunit arrangement is maintained, further fragmentation of the complex into smaller subunits provides information regarding which subunits are involved in the interaction. Although CID can dissociate subunits from the larger protein complex, its use here is limited as it always ejects the weakest interacting subunits, which in many cases means that it will eject the smallest subunit. Furthermore, the ejected monomer is also unfolded in the process. Additionally, CID removes large amounts of charge away from the complex, making any further fragmentation attempts difficult.

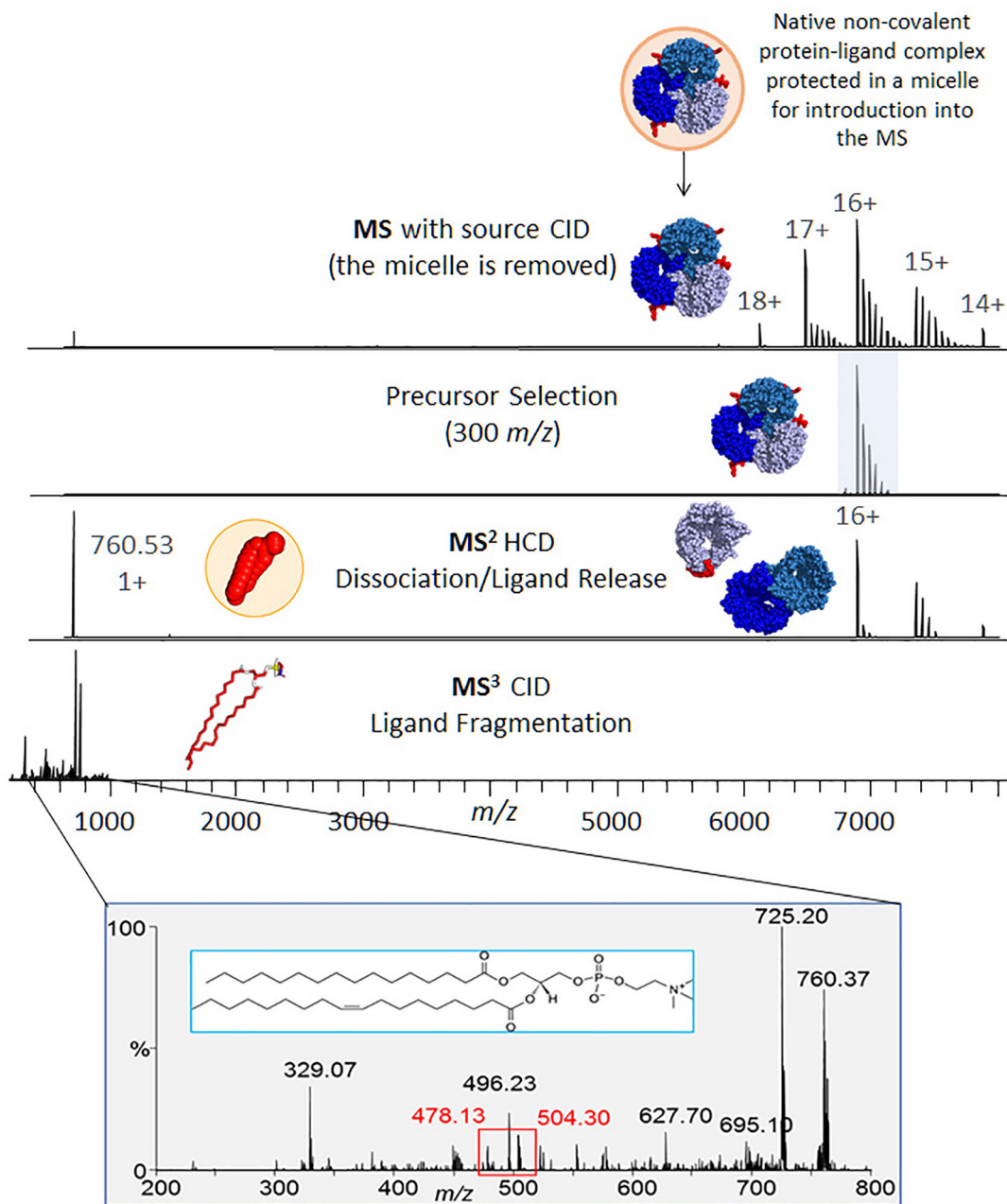


Figure 3. Native MS and multistage fragmentation (MSⁿ) of 111 kDa membrane protein complex. Detergent micelle is removed using in source activation. The molecular ion had several non-covalent ligand molecules attached to it, measuring 760 Da. MS⁴ analysis identified the ligand as phosphatidylcholine 16:0/18:1.

Alternative fragmentations such as SID and UVPD have emerged, which provide a deeper insight into the structural organization of the complex. SID has been a promising fragmentation for topology determination because the fragmentation of the complex reflects its structural organization.⁴ For example, a complex that is organized as a dimer of trimers will yield two trimers upon dissociation.

Topology information can provide a number of insights from aiding in identifying functions, which subunits are involved in protein interactions, where the binding takes place simplifying the process of predicting conformations to guide molecular modelling for electron microscopy or hybrid structural methods.

An alternative approach for large heteromeric protein complexes is to use solution disruption to generate smaller subcomplexes.⁵ Typically low concentrations (10-30%) of chaotropic solvents such as methanol or isopropanol are used to dissociate the weakest interacting subunits, thereby producing subcomplexes that retain their structural organization, which can then be selected for further investigations by CID or higher-energy collisional dissociation (HCD). Combining these results with the stoichiometry obtained through native MS, means the organization of the complex can be determined.

Protein dynamics

One of the biggest advantages of native MS is its ability to examine the dynamics of protein complexes.⁶ For example, the oligomeric state of a protein complex can be influenced by many factors including the composition of the proteins and the environment in which the complex exists. This translates to many different compositions/stoichiometries existing often at the same time. Native MS is a powerful tool to study these changes. In a single analysis, various oligomeric states of a protein complex can be recorded.⁶ Each oligomeric state produces distinct masses that can be detected with high-resolution accurate-mass (HRAM) mass spectrometers. This ability enables native MS to examine the dynamic state of protein complexes all at the same time, but more importantly the information can be obtained instantly. This approach has been used to monitor protein assembly and disassembly.

Native mass spectrometry workflow

Since native MS examines the structure that proteins have in biological solution prior to introduction into the mass spectrometer, the experiment must preserve the native structure during sample introduction. The pH, ionic strength and the composition of the buffer can influence the structure of the proteins. Additionally, salts and background matrices present in the sample can interfere with analysis by suppressing MS signals from the proteins of interest.

Typical native MS experiments involve a buffer exchange of the sample to remove the salts that might interfere with analysis (Figure 4). Most buffers used in biological experiments tend not to be suitable for native MS due to the presence of non-volatile salts such as sodium/potassium ions. Aqueous (volatile) buffers are better suited for native MS experiments. The most commonly used buffers are ammonium acetate, ethylenediammonium

diacetate, or triethylammonium acetate. Thus, after desalting, a buffer exchange is performed to remove the original buffer and replace it with the aqueous buffer such as ammonium acetate.



Figure 4. Schematic representation of a typical native MS workflow

The native MS workflow is different from the traditional intact/top-down or bottom-up proteomics workflow because large intact proteins or protein complexes are involved rather than individual proteins or digested peptides. In native MS, samples can be introduced into the mass spectrometer by direct infusion, liquid chromatography (LC) or capillary electrophoresis (CE) depending on the experiment and the question at hand.

Direct infusion or static spray is the simplest way of introducing sample into the mass spectrometer. Samples are introduced under nanospray conditions to minimize sample consumption and to minimize the need for desolvation. It is also ideal for examining very weak non-covalent interactions.

LC is used to increase throughput, analysis of complex mixtures, heterogeneous species or for on-line desalting. There are two types of separation modes used in native MS, size-exclusion chromatography (SEC) and ion-exchange chromatography (IEC) with pH gradients or salt gradients.

CE can also be used up front as separation prior to MS analysis. CE can be an attractive separation mode due to charge-based separation of analytes in short period of time.

Once the samples are introduced into the mass spectrometer the data is acquired. Specifically designed software is used for data interpretation. The following sections will discuss these in more detail.

Technical advances in mass spectrometry

MS has made it possible for the elucidation of proteins and protein complex structures. It has democratized protein structural studies, making it accessible to a wider community of researchers. At the present time, mass spectrometers are available in most research institutes providing services and resources to researchers engaged in proteomics experiments. These same instruments can be used for protein structural studies, making it possible for the elucidation of proteins and protein complex structures. In the majority of these studies, the researchers engaged in structural work do need to be the operator or the expert in mass spectrometry. By tapping into these resource laboratories, structural biologists can obtain a much more complete picture of structures.

However, there are challenges associated with native MS. Primarily, researchers are dealing with very large biomolecules such as ribosomes, membrane ion channels, and virus capsids. These biomolecules range in mass from kilodalton (kDa) to millidalton (mDa), producing mass spectra that are quite complex containing multiple charge states for each protein. These charge state envelopes can overlap in the mass spectrum and be close in mass to each other. In native MS, the mass spectrum acquisition requires mass spectrometers with ultra-high resolution to resolve and to assign accurate charge states and masses to the spectrum. The resolution needs to be high enough to differentiate small mass differences that can aid in resolving near mass proteoforms, reveal key ligands, PTMs and interactions. For example, in native MS analysis of viruses, high resolution mass spectrometers are needed to resolve heterogeneity associated with different proteoforms.

Native MS also requires mass spectrometers that can transmit and detect high mass-to-charge (m/z) ions associated with these large biomolecules. Here, when dealing with large protein complexes like protein-nucleic acid complexes that contain few charges, the charge envelope will appear in high m/z area.⁷ Mass spectrometers that can transmit and detect such a high m/z range would be desirable. The ability to do this also increases the sensitivity of detection at the high m/z range, needed in experiments where sample is limited.

Historically, native MS has been performed on mass spectrometers with time-of-flight (TOF) detectors. However, in recent years there has been paradigm shift in native MS with the emergence of Thermo Scientific™ Orbitrap™ mass analyzer-based mass spectrometers.⁷⁻¹³ These instruments provide ultra-high resolution, high mass accuracy and high sensitivity which are all mandatory requirements for performing native MS.

Mass spectrometers with high resolution can aid in mitigating overlapping charge states and minimizing mass interferences for proteins, making it easier to assign correct charge states and masses (Figure 5). An example of how high-resolution aids in the differentiation and identification of PTMs is shown in Figure 2.

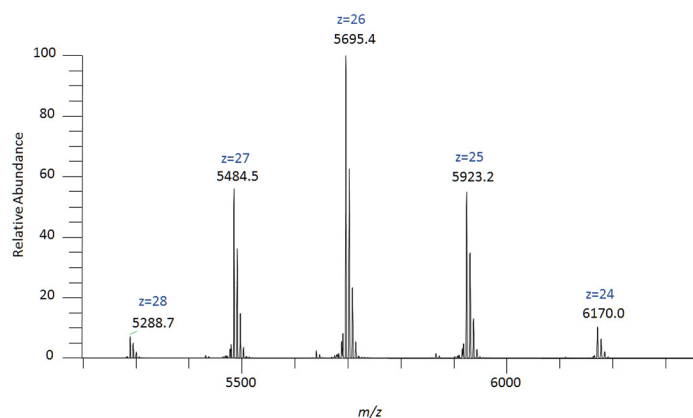


Figure 5. Native MS analysis of Trastuzumab. The high resolution provided by the mass spectrometer resolves the overlapping charge states. It also reduces mass interferences enabling clear charge state assignment to the spectrum.

The term native MS might imply just acquiring intact mass measurement, however, the ability to fragment and to characterize proteins or protein complexes are all a part of this type of experiment. Mass spectrometers that can

isolate and fragment high mass species would be attractive for native MS. Recent innovations in MS have produced instruments that have high mass isolating quadrupoles that can isolate up to 25,000 m/z , thereby, enabling fragmentation of very large intact proteins (Figure 6). Furthermore, these instruments can also detect up to high m/z range.^{3,12}

Another recent innovation is the ability to perform pseudo-MS³ on complexes in a native MS experiment.^{3,13} As part of this strategy, in-source trapping has been implemented within the current generation of commercial mass spectrometers. This allows for efficient desolvation of large protein ions which can help in the removal of solvent adducts and detergent micelles thereby improving transmission and resolution for MS spectrum. Furthermore, by applying more in-source energy during in-source trapping, protein complexes can be dissociated into their subunits which can then be transmitted into the collision cell and further fragmented for sequence analysis (pseudo-MS³, Figure 7).

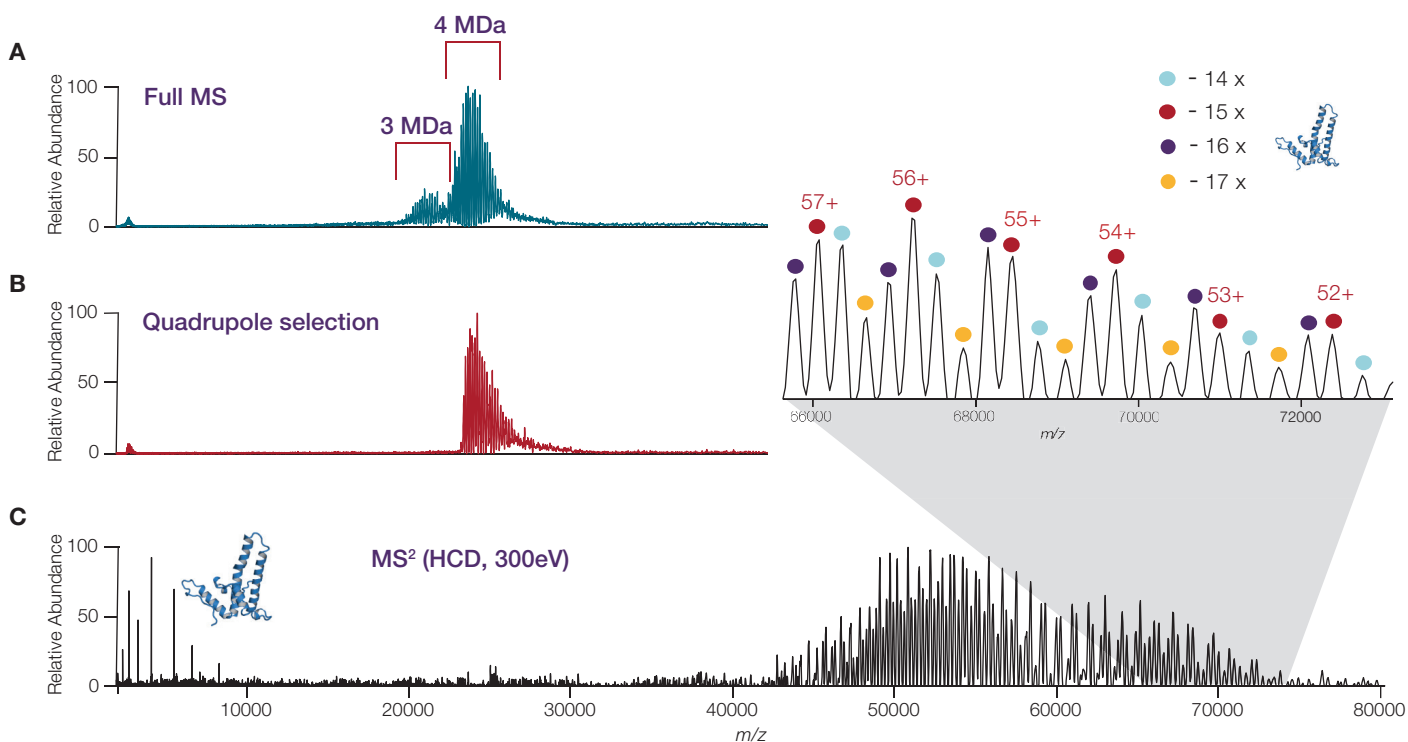


Figure 6. (A) Native mass spectrum of Hepatitis B Virus capsids. (B) Shows high mass quadrupole isolation of the 4 MDa particle (C) Fragmentation of the 4 MDa particle in the HCD cell and the resultant MS² spectra detected in the Orbitrap mass analyzer up to 80,000 m/z . Also shown, sequential ejection of up to 17 out of the 240 copies of the capsid proteins. The ejected monomers appear at low m/z while the product ions appear at increasingly higher m/z . The inset shows an enlargement of the spectrum at 70,000 m/z , which contains baseline resolved ions of HBV capsids that have lost between 14 and 17 capsid proteins (6.5% of the original mass) and 68% of the original charge.¹⁴

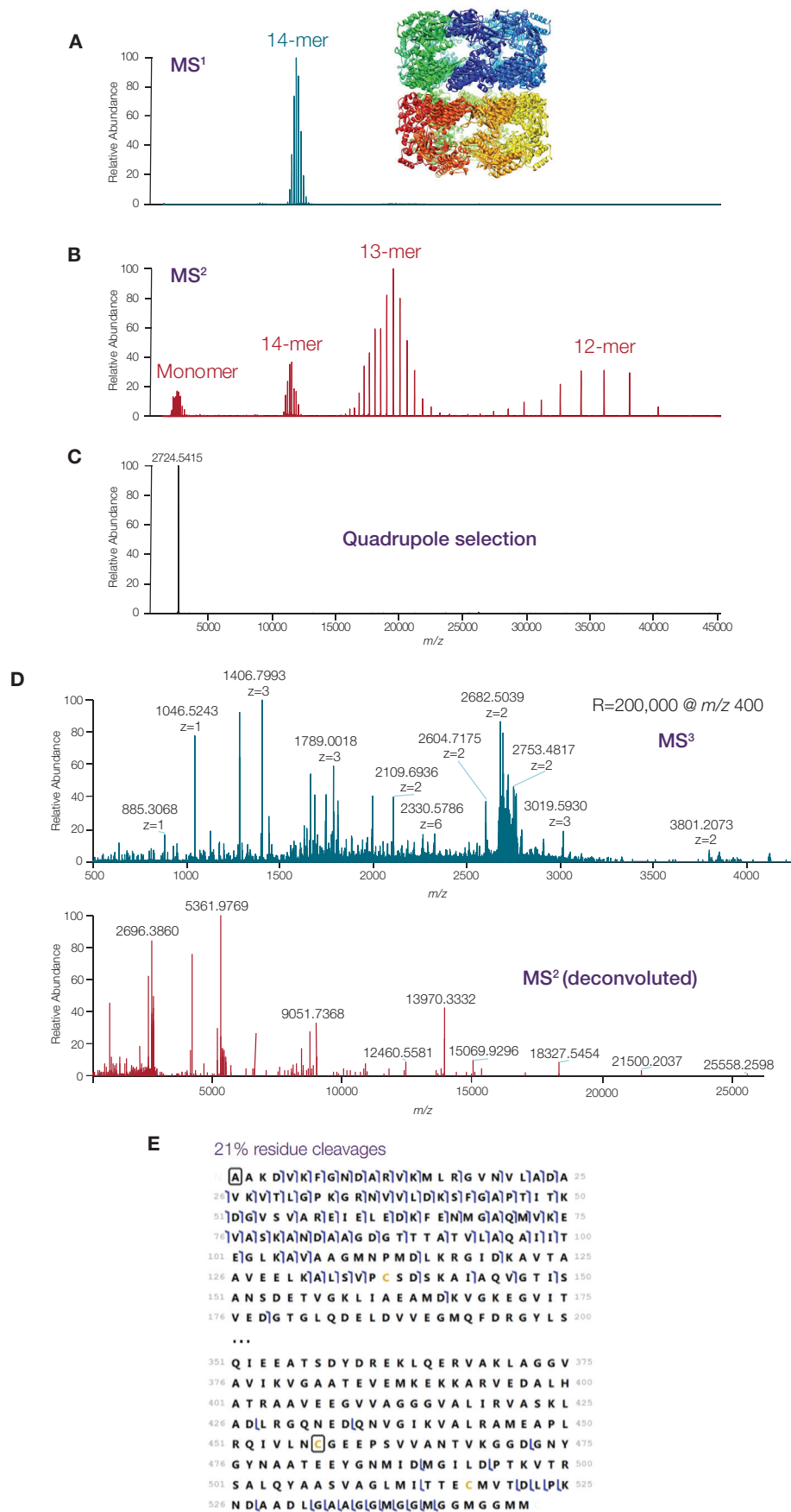


Figure 7. (A) Native MS GroEL protein complex (B) Using “in-source trapping”, desolvation and fragmentation of the GroEL 14-mer complex into monomer and stripped complexes. (C) High mass quadrupole isolation of the monomer (D) Fragmentation within the HCD cell and (E) Resultant MS² spectra shown as well as the sequence coverage observed.³

Native MS has benefitted immensely from advancements in fragmentation in mass spectrometers. The availability of multiple fragmentation techniques (CID, HCD, ETD, EThcD), MSⁿ and the ability to perform these fragmentations at any stage of MSⁿ has enabled researchers to increase sequence coverage and thoroughly characterize protein structures. Implementation of proton transfer charge reduction (PTCR) reaction capabilities in mass spectrometers has also helped simplify MS and MSⁿ spectra making them easier to interpret (Figure 8).

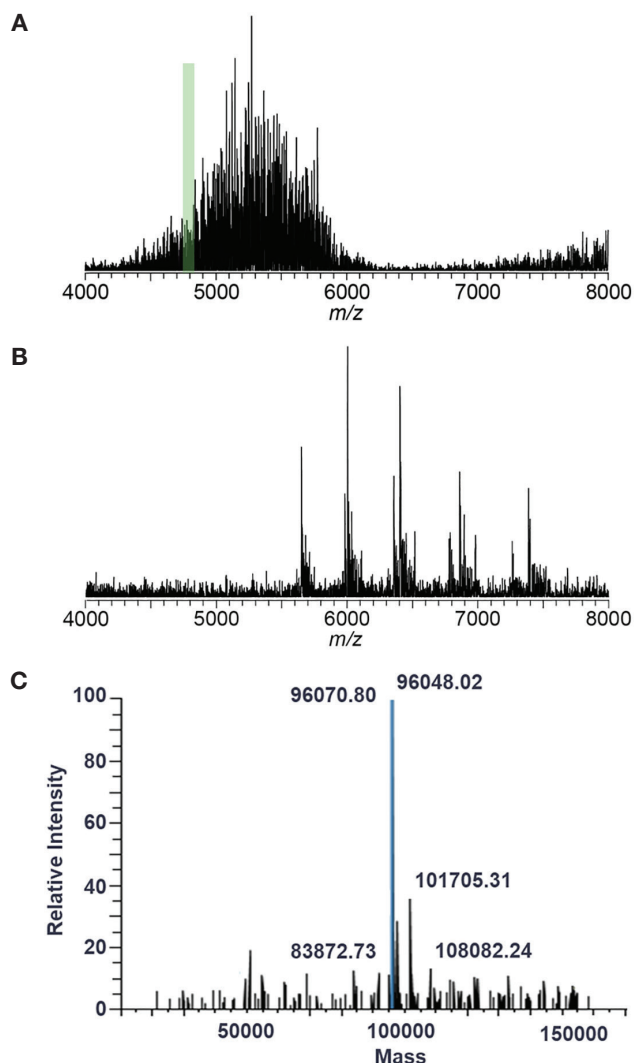


Figure 8. Native MS analysis of Cytokine-Fc fusion protein. (A) The native MS produces a very complex spectrum that makes it challenging to interpret. (B) PTCR of 80 amu window, m/z 5677-5757 results in interpretable spectrum. (C) Deconvolution of spectrum produces expected masses.

Data analysis

Mass spectra produced by intact proteins and protein complexes are far more complex than spectra from peptides. A single protein mass spectrum contains multiple charge states each with its own m/z value. These charge state envelopes can be overlapping in the spectrum and close in mass to each other. Without the use of software, the spectra are very hard to interpret manually.

Currently, software uses a process called deconvolution to transform the series of charge states of a protein into a molecular mass for the protein. This is performed by identifying multiple peaks in the mass spectrum associated with different charge states of the same component and displaying the information about the masses and the abundance of that component. For example, peaks at m/z 1000, 1111, and 1250 might be the charge states 10, 9, and 8 for a protein with a mass of 10,000 Da. The challenge of this process comes from whether the acquired spectrum is isotopically resolved or unresolved. This is dependent upon the resolution that the mass spectrometer used, the mass of the protein and the condition of the experiment. Researchers have developed specific software to address these issues. One such software is Thermo Scientific™ BioPharma Finder™ software that contains two specific deconvolution algorithms, Xtract and ReSpect™, that take advantage of the high-quality HRAM intact protein data produced by current generation of mass spectrometers. The Xtract algorithm handles isotopically resolved spectra data while the ReSpect algorithm deals with unresolved data. These algorithms produce highly accurate results, even for low-abundance proteins, and enable detection of extremely small protein modifications with mass shifts of just a few Daltons.

Since current experiments tap into modern instruments' ability to perform top-down on native MS, software is needed to interpret data. There are challenges that need to be addressed. Namely, the fragmentation of intact proteins creates complex spectra due to the number of fragment ions produced, each often present at multiple charge states. Presence of PTMs can complicate the fragment ion spectra.

Thermo Scientific™ BioPharma Finder™ software has been designed to meet these challenges. It supports both native intact and top down analysis.

Conclusion

Native MS enables the study of proteins and protein complex structures, which promotes a better understanding of protein function and mechanism of action in biological systems. It also is complementary to traditional techniques such as cryo-EM, X-ray crystallography, NMR and MS structural biology techniques – such as hydrogen deuterium exchange (HDX) and crosslinking mass spectrometry (XL-MS). Since experiments using native MS involve looking at samples comparable to the physiological state of an organism, it can generate biologically relevant information on structure. In recent years with innovations in hardware – such as improvements in mass resolution, increase in transmission and detection of large ions, and the ability to perform multistage fragmentation, has allowed researchers to study structures that were previously inaccessible. Advancements in software and overall workflow for native MS has enabled the application to transition from specialized research groups to the mainstream.

Acknowledgment

Figure 3 was provided by Dr. Joseph Gault, University of Oxford.

References

1. Snijder, J.; Schuller, J.M.; Wiegand, A.; Lössl, P.; Schmelling, N.; Axmann, I.M.; Plitzko, J.M.; Förster, F.; Heck, A.J.R., Structures of the cyanobacterial circadian oscillator frozen in a fully assembled state. *Science* **2017**, *355*(6330), 1181–1184.
2. Bailey, A.O.; Houel, S.; Scheffler, K.; Damoc, E.; Sutton, J.; Josephs, J.L., Complete characterization of a lysine-linked antibody drug conjugate by native LC/MS intact mass analysis and peptide mapping. *Application Note* 72511.
3. Damoc, E.; Fort, K.; Reinhardt-Szyba, M.; Belov, M.; Makarov, A.; Viner, R.; Konijnenberg, A., Advancing native top-down MS analysis of non-covalent protein complexes: The Thermo Scientific Q Exactive UHMR mass spectrometer. *Technical Note* 65379.
4. VanAernum, Z.L.; Gilbert, J.D.; Belov, M.E.; Makarov, A.A.; Horning, S.R.; Wysocki, V.H., Surface-Induced Dissociation of Ion Mobility-Separated Noncovalent Complexes in a Quadrupole/Time-of-Flight Mass Spectrometer. *Anal Chem* **2012**, *84*(14), 6016–6023.
5. Politis, A.; Schmidt, C.; Tjioe, E.; Sandercock, A.M.; Lasker, K.; Gordiyenko, Y.; Russel, D.; Sali, A.; Robinson, C.V., Topological models of heteromeric protein assemblies from mass spectrometry: application to the yeast eIF3:eIF5 complex. *Chem Biol.* **2015**, *22*(1), 117–28.
6. Erba, E.B.; Petosa, C., The emerging role of native mass spectrometry in characterizing the structure and dynamics of macromolecular complexes. *Protein Science* **2015**, *24*, 1176–1192.
7. 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. *Nature* **2017**, *14*(3), 283–6.
8. Rose, R.J.; Damoc, E.; Denisov, E.; Makarov, A.; Heck, A.J., High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. *Nat Methods* **2012**, *9*(11), 1084–6.
9. Rosati, S.; Rose, R.J.; Thompson, N.J.; van Duijn, E.; Damoc, E.; Denisov, E.; Makarov, A.; Heck, A.J., Exploring an orbitrap analyzer for the characterization of intact antibodies by native mass spectrometry. *Angew Chem Int Ed Engl* **2012**, *51*(52):12992–6.
10. Gault, J.; Donlan, J.A.; Liko, I.; Hopper, J.T.; Gupta, K.; Housden, N.G.; Struwe, W.B.; Marty, M.T.; Mize, T.; Bechara, C.; Zhu, Y.; Wu, B.; Kleanthous, C.; Belov, M.; Damoc, E.; Makarov, A.; Robinson, C.V., High-resolution mass spectrometry of small molecules bound to membrane proteins. *Nat Methods* **2016**, *13*(4), 333–6.
11. 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*(1), 2493.
12. 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.; Grünwald, K.; Robinson, C.V., Protein assemblies ejected directly from native membranes yield complexes for mass spectrometry. *Science* **2018**, *362*(6416), 829–834.
13. Fort, K.L.; van de Waterbeemd, M.; Boll, D.; Reinhardt-Szyba, M.; Belov, M.E.; Sasaki, E.; Zschoche, R.; Hilvert, D.; Makarov, A.A.; Heck, A.J.R., Expanding the structural analysis capabilities on an Orbitrap-based mass spectrometer for large macromolecular complexes. *Analyst* **2017**, *143*(1), 100–5.
14. Brochure: Q Exactive UHMR Hybrid Quadrupole-Orbitrap Mass Spectrometer – Go Beyond What You Thought Possible with Native MS.

Find out more at thermofisher.com/nativeMS-proteomics