Native, Intact, and Top-down Mass Spectrometry Service Summary

Overview

Native, intact, and top-down (TD) mass spectrometry (MS) are techniques that can be used to characterise important aspects of whole proteins and peptides, including higher-order structure. These techniques all analyse whole proteins in their mature form which are distinct from conventional bottom-up proteomics techniques, which rely on enzymatic digestion of proteins prior to analysis.

All three techniques form intact gas-phase protein ions. These protein ions typically carry multiple charges, creating a distribution of signals within a mass spectrum known as a charge state envelope. By measuring these specific mass-to-charge (*m/z*) ratios, the molecular mass of the original protein can be mathematically calculated (deconvoluted).

Despite the similarities between these three techniques, these techniques also differ in their scientific purposes, primarily reflected in the sample preparation and experimental workflows. A simple overview of the three techniques is shown in Figure 1.

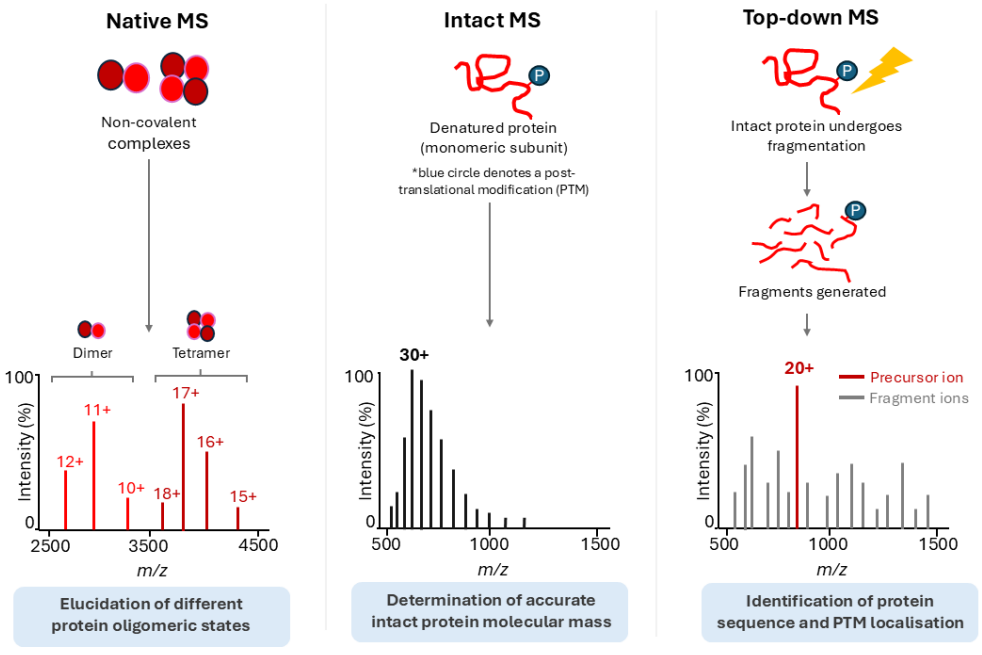

Figure 1. Native, intact and top-down[1] mass spectrometric techniques are used to study different aspects of mature and undigested proteins.

[1]*Native* top-down MS is a subset of top-down MS, in which native MS conditions are employed in a top-down MS experiment. This technique enables structural characterisation of protein complexes and proteoform subunits in a single experiment.

The following sections expound upon each technique and provide information on the recommended sample conditions, as well as examples of the data outputs typically generated for each experiment.

Note: native and TD MS are not suited for high-throughput capacities or very complex samples. It is highly recommended to employ these techniques for targeted characterisation using purified proteins or very simple protein mixtures. If working with protein mixtures with greater than approximately 10 proteins, sample pre-fractionation will be necessary.

1. Native mass spectrometry

Native MS preserves proteins in a folded native-like state. This can enable characterisation of higher-order protein structure and non-covalent protein interactions. To facilitate these measurements, gentle mass spectrometry conditions are used. These conditions include relatively low spray voltages and temperatures during protein ionisation, and the use of volatile physiological-like buffers (*e.g*. ammonium acetate) and nano-electrospray ionisation.

Native MS can be used for:

- Characterisation of non-covalent interactions and higher-order structures (protein-protein or protein-ligand complexes).
- Stoichiometric analysis of protein complexes (*e.g.* monomeric, dimeric, trimeric species).

Sample conditions:

- Recommended concentration range: 5 – 10 µM. A higher concentration is recommended to allow for further sample dilution as needed.
- Sample complexity: purified protein or simple mixture (<10 proteins).
- Recommended buffer systems: 100 mM ammonium acetate or ammonium bicarbonate (volatile buffers).
- NOT recommended buffer systems: non-volatile buffers including phosphate, Tris, glycerol or solutions that include non-volatile salts (*e.g.* NaCl). These solutions contribute to spectral noise and suppress protein ion signal of interest.

Example of native MS data output:

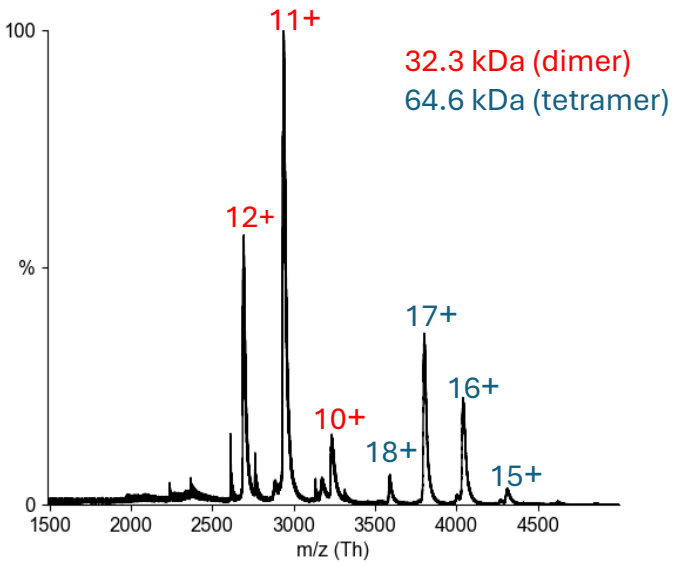

Figure 2. Native MS analysis of haemoglobin (5 µM) in 100 mM ammonium acetate. Dimeric (red) and tetrameric (blue) states of haemoglobin are captured. Charge states of each oligomeric state are as annotated and the deconvoluted molecular masses determined.

2. Intact mass spectrometry

Intact MS can often be used interchangeably with "denaturing" MS. The sample conditions are designed to disrupt non-covalent interactions, unfold the protein and maximise the sensitivity of whole protein mass measurements. Intact MS measurements can be made by directly infusing the sample into the mass spectrometer via nano-electrospray ionisation, or if higher throughput is required, via a coupled LC separation and electrospray ionisation setup.

Intact MS can be used for:

- Determination of the accurate intact protein molecular mass
- Proteoform (protein isoform) profiling at the MS1 level by coupling to LC separation for protein mixtures

Sample conditions:

- Recommended concentration range: 5 – 10 µM. A higher concentration is recommended to allow for further sample dilution as needed.
- Sample complexity: purified protein or simple mixture (<10 proteins). Note: online LC separation can be coupled to intact MS analysis.
- Recommended solvent systems: a proportion of MS-grade organic solvent (methanol/acetonitrile) and MS-grade formic acid (*e.g.* 49.5 % MilliQ water/ 49.5 % methanol/1 % formic acid). The recommended percentage of organic solvent is <50 % but this may depend on the solubility of the protein.
- NOT recommended solvent systems and additives: dimethyl sulfoxide (DMSO), *N,N*-dimethylformamide (DMF), chlorinated solvents (*e.g.* chloroform, dichloromethane (DCM)), hexane, toluene, inorganic acids (*e.g.* phosphoric, sulfuric, hydrochloric acids).

Example of intact MS data output

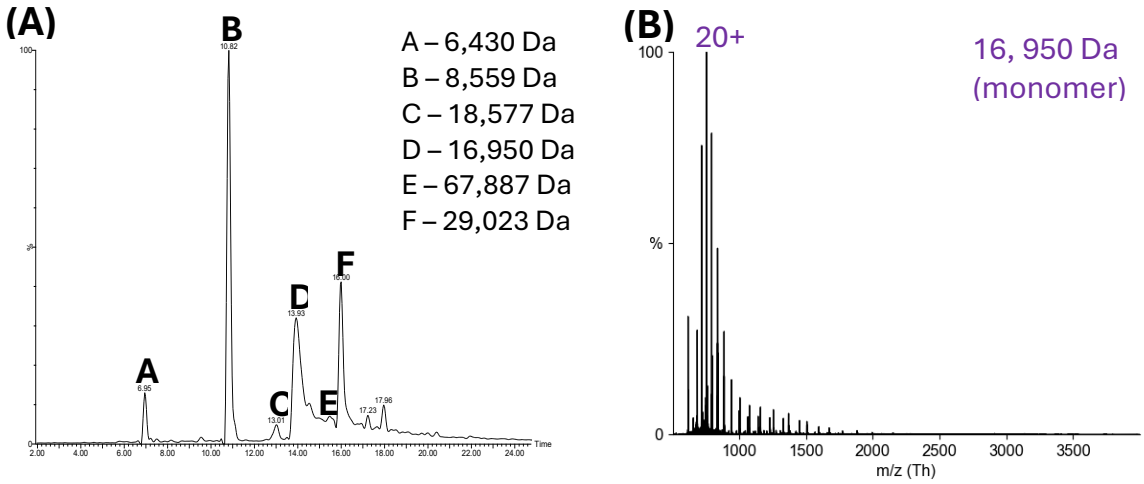

Figure 3. (A) LC-MS analysis of a simple protein mixture under denaturing conditions. A chromatogram with peaks corresponding to different proteins is shown. Protein molecular masses, as determined via spectral deconvolution of ions measured in Peaks A – F, are presented. **(B)** Intact mass spectrum of the protein at LC Peak D (haemoglobin monomeric subunit). Numerous peaks are observed, indicating that many different charge states have been generated for the single protein being measured. These high charge states are indicative of the unfolded protein structure. Molecular mass of the monomeric subunit was determined to be 16,950 Da.

3. Top-down MS

Top-down MS (TD-MS) is an advanced MS technique that allows amino acid sequences of intact proteins to be determined. It also enables the characterisation and site localisation of protein modifications, including post-translational modifications.

In TD-MS, the initial process of introducing the undigested protein into the mass spectrometer is identical to intact MS analysis. However, TD-MS is distinguished experimentally by the additional step of isolating and fragmenting protein ions in a controlled manner. This process produces additional analytical information about the protein subjected to analysis. Due to the non-uniform nature of intact undigested proteins, fragmentation methods can be very protein-dependent and often require optimisation.

This fragmentation step generates protein fragments, which are deconvoluted and matched against an existing protein sequence. This downstream data analysis workflow also differs from the intact MS analysis, as fragment assignment, deconvolution and validation are labour intensive and time-consuming processes. Due to this, high throughput workflows are not enabled and this technique more suited to purified proteins and targeted analysis (see below).

TD-MS can be used for:

- Characterisation of protein primary structures (amino acid sequence information can be obtained)
- Identification and localisation of post-translation modifications (PTMs)
- In-depth characterisation of intact proteoforms (modification states of entire proteins can be determined)

Sample conditions:

Sample preparation requirements are the same as for intact MS analysis.

- Recommended concentration range: 5 – 10 µM. A higher concentration is recommended to allow for further sample dilution as needed.
- Sample complexity: purified protein or simple mixture (<10 proteins).
- Recommended buffer systems: a proportion of MS-grade organic solvent (methanol/acetonitrile) and MS-grade formic acid (*e.g.* 49.5 % MilliQ water/ 49.5 % methanol/1 % formic acid). The recommended percentage of organic solvent is <50 % but this may depend on the solubility of the protein.

- NOT recommended solvent systems and additives: dimethyl sulfoxide (DMSO), *N,N*-dimethylformamide (DMF), chlorinated solvents (*e.g.* chloroform, dichloromethane (DCM)), hexane, toluene, inorganic acids (*e.g.* phosphoric, sulfuric, hydrochloric acids)

Example of TD-MS data output

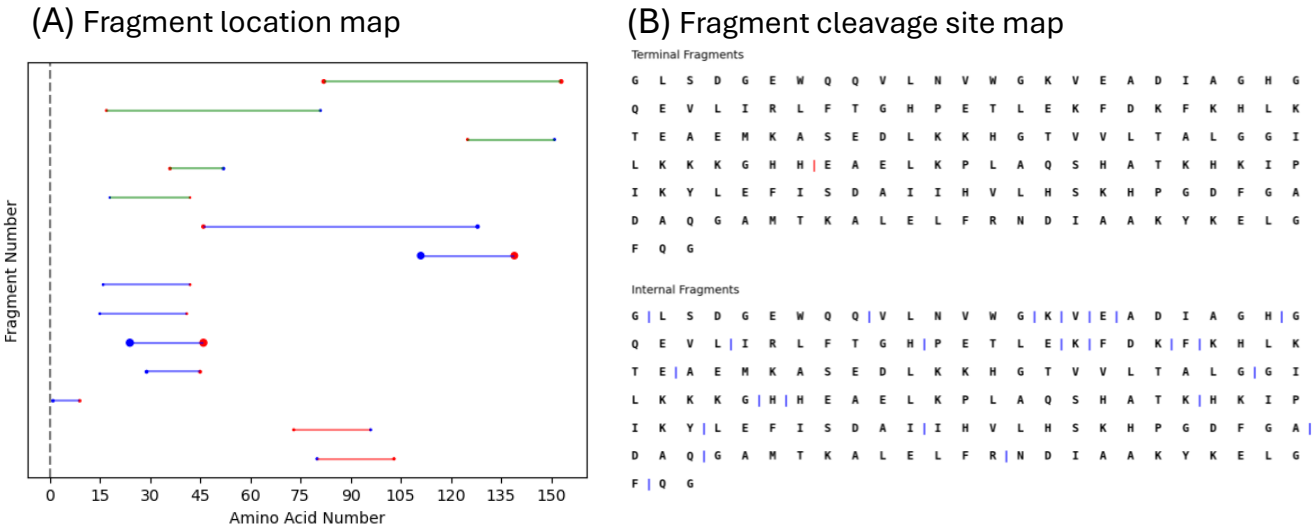

Figure 4. Example outputs from sequence matching for myoglobin [16,940 Da +21H]$^{21+}$. **(A)** The fragment location map indicates the protein sequence regions that are covered by the fragments. **(B)** The fragment cleavage site map indicates the specific inter-residue cleavage sites for terminal fragments and internal fragments.