Methods 55 (2011) 330–336

Contents lists available at SciVerse ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Review Article

Profiling of integral membrane proteins and their post translational modifications using high-resolution mass spectrometry

Puneet Souda [a], Christopher M. Ryan [a], William A. Cramer [b], Julian Whitelegge [a,*]

[a] The Pasarow Mass Spectrometry Laboratory, NPI–Semel Institute, David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA
[b] Department of Biological Sciences, Purdue University, IN, USA

ARTICLE INFO

Article history:
Available online 29 September 2011

Keywords:
Integral membrane proteins
Top-down mass spectrometry
Membrane protein complexes
Intact protein mass spectrometry
High-resolution mass spectrometry

ABSTRACT

Integral membrane proteins pose challenges to traditional proteomics approaches due to unique physicochemical properties including hydrophobic transmembrane domains that limit solubility in aqueous solvents. A well resolved intact protein molecular mass profile defines a protein's native covalent state including post-translational modifications, and is thus a vital measurement toward full structure determination. Both soluble loop regions and transmembrane regions potentially contain post-translational modifications that must be characterized if the covalent primary structure of a membrane protein is to be defined. This goal has been achieved using electrospray-ionization mass spectrometry (ESI-MS) with low-resolution mass analyzers for intact protein profiling, and high-resolution instruments for top-down experiments, toward complete covalent primary structure information. In top-down, the intact protein profile is supplemented by gas-phase fragmentation of the intact protein, including its transmembrane regions, using collisionally activated and/or electron-capture dissociation (CAD/ECD) to yield sequence-dependent high-resolution MS information. Dedicated liquid chromatography systems with aqueous/organic solvent mixtures were developed allowing us to demonstrate that polytopic integral membrane proteins are amenable to ESI-MS analysis, including top-down measurements. Covalent post-translational modifications are localized regardless of their position in transmembrane domains. Top-down measurements provide a more detail oriented high-resolution description of post-transcriptional and post-translational diversity for enhanced understanding beyond genomic translation.

1. Introduction

Membrane proteins are high value targets for over half of all marketed drugs and represent 20–30% of all coded proteins in sequenced genomes making them important for both structure determination and mass spectrometric characterization. Both transmembrane and loop regions may contain post-translational modifications of both functional and structural significance, and must be well understood if we are to collectively define the native covalent state of membrane proteins [1]. Mass spectrometry can be used to obtain sequence identification, deliver molecular mass profiles and define post-translational modifications (PTMs), for both soluble and membrane proteins.

Bottom-up mass spectrometry techniques involve approaches where the intact protein is enzymatically cleaved to peptides before measurements via tandem mass spectrometry. Liquid chromatography with tandem mass spectrometry (LC–MSMS) is one of the most common workflows employed for separation and identification of peptides. Tandem mass spectrometry data includes both parent ion and product ion fragment masses, and are frequently good enough to assign sequence identity to short peptides (10–30 residues) based on comparison to translated gene sequences. Though progress has been made with technical improvements in digestion and chromatography, sequence coverage can still be marginal, and this is especially true for the transmembrane domains of integral membrane proteins [2,3]. Typically, a handful of easily recovered peptides known as 'proteotypic' peptides are routinely observed in these tandem mass spectrometry experiments such that bottom-up approaches are typically biased with incomplete sequence coverage and PTM information [4]. Integral membrane proteins are not ideally suited for bottom-up proteomics due to their unique physiochemical properties, yielding some peptides with poor solubility and/or ability to be ionized, especially from transmembrane domains. Another caveat with bottom-up approaches is that they are heavily dependent on underlying genomic

* Corresponding author.
E-mail address: jpw@chem.ucla.edu (J. Whitelegge).

1046-2023/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymeth.2011.09.019

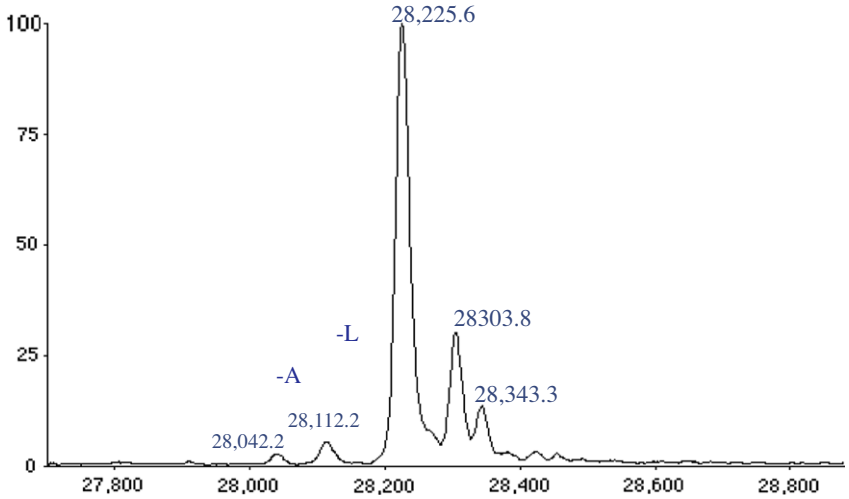

Fig. 1. Zero-charge intact protein molecular mass profile (MMP) of bovine major intrinsic protein (MIP). The data was collected using a low-resolution triple quad mass spectrometer and transformed to obtain the zero-charge molecular mass profile shown. The molecular heterogeneity of different protein species is clearly visible with phosphorylation (80) at 28303.8, cysteinylation (119) at 28,343.3 and also C-terminal processing by removal of Ala and Leu/Ile residues. Mass calculated from the MIP gene sequence was 28,223.1 which reflects a delta of 2.1% or 0.0075%, coincident within experimental error. Bovine MIP is one of very few eukaryotic proteins whose mature primary sequence is exactly as predicted from the genomic translation without post-translation processing.

information thus ignoring molecular heterogeneity not immediately predictable from gene sequences.

Top-down mass spectrometry addresses many of the problems of the bottom-up approach by targeting intact proteins rather than peptides for analysis. The goal is to define a protein's primary structure by providing highly accurate structural assignment of fragments. High-resolution Fourier-transform mass spectrometry (FT-MS) is most frequently used for top-down measurements due to the need to accurately assign product ions [5–7]. The whole intact protein can be dissociated using multiple dissociation mechanisms including CAD or ECD toward full sequence and PTM coverage. Complete interrogation of the primary structure via top-down mass spectrometry usually requires larger quantities of proteins than bottom-up experiments, and is thus well suited for protein crystallography experiments where both purity and abundance are typically attained prior to MS analysis. Much progress has been made in top-down MS as proteins of increasing size and complexity are being resolved [1]. Aqueous conditions suitable for mass spectrometry of soluble proteins are often inadequate for integral membrane proteins requiring specialized sample preparation and chromatography protocols, which we will discuss presently.

In conclusion, the bottom-up approach is suitable if an overall picture of a complex proteome is required, while top-down offers more valuable information if PTMs, protein heterogeneity and complete information about the primary structure is desired.

2. General considerations

The challenges associated with proteomics of membrane proteins arises due to their amphipathicity, a combination of polar soluble domains and apolar transmembrane domains, complicated by the presence of free thiols in the bilayer [8]. The coupling of ESI to MS has turned out be an essential breakthrough in intact protein analysis by mass spectrometry [12]. ESI is preferred over MALDI as it produces multiply charged intact protein ions that dissociate with high efficiency for information-rich spectra that can be analyzed to deduce the protein sequence and PTMs. Liquid chromatography is easily interfaced with electrospray-ionization sources yielding a versatile, robust analytical platform for protein and peptide mass spectrometry.

2.1. LC–MS+ approach and solvent systems for integral membrane proteins

Integral membrane proteins were first analyzed by MALDI-TOF in 1992 and ESI in 1993 [13–15]. In 1998 we successfully used high formic acid concentrations with liquid chromatography and demonstrated that integral membrane proteins could be analyzed with mass accuracy similar to that achievable for soluble proteins [9]. However, high concentrations of formic acid could also lead to sporadic and unpredictable problems associated with protein formylation (+28 Da adducts). In newer and more improved approaches a high concentration of formic acid (up to 90%) is still preferred owing to its unrivaled capability to solvate proteins, but to reduce adducts, formic acid is introduced just seconds (<120 s) before mass spectrometry analysis. Tri-fluoro acetic acid (TFA) also has excellent solubilizing properties but routinely suppresses electrospray ionization, adds +114 Da adducts to proteins as well as presenting safety issues.

In order to obtain intact protein profiles such as the one shown (Fig. 1), a methodology known as LC–MS+ was developed. LC–MS+ refers to liquid chromatography with mass spectrometry and concomitant fraction collection. The technique employs a flow splitter between the HPLC and a low-resolution electrospray-ionization mass spectrometer so that half of the column eluent is diverted to collect fractions that can be used later for downstream experiments involving protein identification and PTM characterization on high-resolution Fourier transform mass spectrometers (FT-MS), if such a detailed analysis is required. Mass data from the initial LC–MS+ experiment is used to guide the subsequent top-down experiments. The LC–MS+ protocol is limited by the complexity of the protein sample and the capacity of the separations used. Both size-exclusion and reversed phase chromatography are used in the LC–MS+ protocol, depending upon sample complexity (Fig. 2). We have successfully used our size-exclusion LC–MS+ protocol to analyze a wide range of integral membrane proteins containing up to 15 transmembrane helices [1,9,16,17] in circumstances where a single protein or a modest mixture were available after prior fractionation. LC–MS+ has also been widely applied using a reversed-phase protocol involving volatile aqueous/organic solvent mixtures