Journal of Proteomics xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Journal of Proteomics

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

Review

Structural characterisation of medically relevant protein assemblies by integrating mass spectrometry with computational modelling

Argyris Politis [a,*], Carla Schmidt [b,*]

[a] Department of Chemistry, Kings College London, London, United Kingdom
[b] Interdisciplinary Research Center HALOmem, Martin Luther University Halle-Wittenberg, Halle/Saale, Germany

ARTICLE INFO

Article history:
Received 20 December 2016
Received in revised form 13 April 2017
Accepted 18 April 2017
Available online xxxx

Keywords:
Proteomics
Structural mass spectrometry
Computational modelling
Integrative modelling

ABSTRACT

Structural mass spectrometry with its various techniques is a powerful tool for the structural elucidation of medically relevant protein assemblies. It delivers information on the composition, stoichiometries, interactions and topologies of these assemblies. Most importantly it can deal with heterogeneous mixtures and assemblies which makes it universal among the conventional structural techniques. In this review we summarise recent advances and challenges in structural mass spectrometric techniques. We describe how the combination of the different mass spectrometry-based methods with computational strategies enable structural models at molecular levels of resolution. These models hold significant potential for helping us in characterizing the function of protein assemblies related to human health and disease.

Significance: In this review we summarise the techniques of structural mass spectrometry often applied when studying protein-ligand complexes. We exemplify these techniques through recent examples from literature that helped in the understanding of medically relevant protein assemblies. We further provide a detailed introduction into various computational approaches that can be integrated with these mass spectrometric techniques. Last but not least we discuss case studies that integrated mass spectrometry and computational modelling approaches and yielded models of medically important protein assembly states such as fibrils and amyloids.

Contents

* Corresponding author.
E-mail addresses: argyris.politis@kcl.ac.uk (A. Politis), carla.schmidt@biochemtech.uni-halle.de (C. Schmidt).

http://dx.doi.org/10.1016/j.jprot.2017.04.019

1. Introduction

Proteins are key players in life executing critical tasks for cellular functions. These functions are often linked with their structural arrangements as well as their dynamic interactions with ligands such as proteins/peptides, nucleotides, carbohydrates or lipids. The structural analysis of proteins and the complexes they form with their interaction partners is therefore of paramount importance. Recent improvements in mass spectrometry made it an ideal tool for the structural analysis of proteins and their ligands [1–5]. It delivers information on the composition, stoichiometries, protein interactions and interaction networks as well as ligand binding [6]. There is a variety of techniques linked with mass spectrometry including proteomics, (chemical or UV) cross-linking, labelling techniques (such as hydrogen/deuterium-exchange) or mass spectrometry of intact protein assemblies, the latter is often coupled with ion mobility (IM) spectroscopy (see Table 1) [6].

Mass spectrometric techniques target either the peptides produced by hydrolysis of the proteins using proteases (i.e. bottom-up approaches) or the intact proteins. The latter distinguishes between top-down proteomics, i.e. the fragmentation of the intact proteins during analysis, or the analysis of the intact proteins using "native" mass spectrometry [7]. The advantage of bottom-up approaches is the opportunity to unambiguously identify specific sites such as modified amino acid residues (proteomics) or binary interaction sites (cross-linking). A top-down approach, on the other hand, allows the quantitative determination of these sites in the same experiment. Native mass spectrometry determines protein complex' stoichiometries along with protein interactions and specificity of the ligands.

As the structure of a protein-ligand complex is directly linked with function, its malfunction is just as dependent. Nowadays mass spectrometry is playing an important role in structure elucidation of pharmaceutically relevant proteins as well as drug binding [8]. However, the information obtained from the various mass spectrometric techniques often yields low-resolution structures. One way to increase the degree of resolution and gain information unattainable by a single method is to bring together the information derived from various techniques [9,10]. This has prompted researchers to develop computational tools and methods that allowed encoding structural information derived from various mass spectrometry-based experiments (as well as other techniques) into diverse sets of modelling restraints, i.e. establishing integrative modelling [11,12]. The integration of these restraints into sophisticated algorithms for 3D model generation enabled determining the structure and in many cases the conformational dynamics of proteins and their complexes [11]. The wealth of computational tools and algorithms for structure-based prediction are exemplified by homology modelling [13], and docking strategies [14], for instance in drug discovery [15] or techniques such as molecular dynamics simulations [16] and de novo modelling.

In this review we summarise the advances of structural mass spectrometry in the study of protein-ligand complexes. We exemplify the different techniques through recent examples from literature that helped in the understanding of medically relevant protein assemblies. We further provide a detailed introduction into various computational approaches that can be used to bring together the information from the various mass spectrometric techniques. Last but not least we discuss case studies that integrated mass spectrometry and computational modelling approaches and yielded models of medically important protein assembly states such as fibrils and amyloids.

2. Structural mass spectrometry

2.1. Proteomics reveals changes in protein expression and post-translational modifications

The proteome is the entire set of proteins in a cell, an organism or a tissue at a defined time point and under defined conditions [17]. Proteomics, consequently, is the large-scale study of the proteome and nowadays mostly involves mass spectrometry-based protein identification [18–21]. For this, proteins are usually hydrolysed using specific endoproteinases. The specific peptide mixture is then analysed in a mass spectrometer. Peptide and fragment masses lead to the identification of the protein(s) after database searching [22]. Depending on the sample complexity separation techniques on the peptide or protein level are included in the proteomics workflow [23–25]. Recent improvements in mass spectrometry enabled the identification of post-translational modifications [26,27] and relative or absolute quantification of proteins [28,29] on a routine basis. Notably, entire proteomes can nowadays be explored in very short time spans [30,31]. As such several proteomes and their interactomes have been described to a high standard. Examples include the studies on *Saccharomyces cerevisiae* [32–34], *Drosophila melanogaster* [35,36] and even humans [37–39].

Of the available techniques, quantitative mass spectrometry plays a major role in clinical proteomics [40–42]. In particular, the quantification of post-translational modifications, which are the regulatory

Table 1
Overview on MS techniques discussed in this review. The principle, experimental workflows, commonly used instrumentation, limitations as well as the outcome are given for each technique.

MS technique	Principle	Experimental procedure	Instrumentation	Limitations	Outcome
Proteomics	Identification of proteins by fragmentation/sequencing of peptides	(i) Digestion of proteins, (ii) LC-MS/MS of peptides, (iii) database searching	High speed, high sensitivity mass spectrometers (mostly Orbitrap, Q-Tof or Q-Trap instruments)	Protein interactions only identified indirectly	Identification of proteins and post-translational modifications; quantification; interactomes
Cross-linking	Identification of binary protein interactions by fragmentation/sequencing of cross-linked di-peptides	(i) Cross-linking of proteins, (ii) digestion of proteins, (iii) LC-MS/MS of peptides, (iv) identification of cross-linked di-peptides by specialised software; (*) optional pre-fractionation of proteins or peptides	High sensitivity, high resolution mass spectrometers (mostly Orbitrap instruments)	Only binary interactions; cross-linking yield dependent on protein sequence and cross-linking chemistry	Identification of protein interactions; protein networks; distance restraints
H/D-X	Exchange of backbone protons by deuterium to study solvent accessibility	(i) Exchange of protons by deuterium, (ii) analysis of intact proteins or peptides after digestion, (iii) increase/decrease of protein/peptide masses, (iv) structural analysis of changes in H/D-X (peptides)	Mostly Q-ToF instruments; in some cases automated procedure using pipetting robots	Back exchange of deuterium; structural information of protein required; data analysis difficult	Solvent accessibility; structural changes/dynamics; ligand binding
Native ion-mobility MS	Transfer of intact protein assemblies into the gas phase of a mass spectrometer	(i) Exchanging purification buffer to aqueous, volatile buffer, (ii) manual MS analysis of protein complexes, (iii) manual data analysis (using specialised software)	Modified instruments, usually Q-ToF mass spectrometers (in some cases modified Orbitrap instruments)	Protein complexes need to be kept intact; modified instruments required; no direct interaction sites	Identification of protein stoichiometries and interactions; topology; ligand binding; collision cross-sections; shape restraints

Please cite this article as: A. Politis, C. Schmidt, Structural characterisation of medically relevant protein assemblies by integrating mass spectrometry with computational modelling, J Prot (2017), http://dx.doi.org/10.1016/j.jprot.2017.04.019