

Published in final edited form as:

Q Rev Biophys. 2016 January ; 49: . doi:10.1017/S0033583516000160.

Mass Spectrometry: A Technique of Many Faces

Maya A. Olshina and Michal Sharon*

Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot 7610001, Israel

Abstract

Protein complexes form the critical foundation for a wide range of biological process, however understanding the intricate details of their activities is often challenging. In this review we describe how mass spectrometry plays a key role in the analysis of protein assemblies and the cellular pathways which they are involved in. Specifically, we discuss how the versatility of mass spectrometric approaches provides unprecedented information on multiple levels. We demonstrate this on the ubiquitin-proteasome proteolytic pathway, a process that is responsible for protein turnover. We follow the various steps of this degradation route and illustrate the different mass spectrometry workflows that were applied for elucidating molecular information. Overall, this review aims to stimulate the integrated use of multiple mass spectrometry approaches for analyzing complex biological systems.

1 Introduction

1.1 The proteasomal protein degradation pathway

The maintenance of cellular homeostasis requires the concerted action of multiple interconnected biochemical pathways, enabling the cell to perform its required functions and undergo appropriate developmental transitions, while responding to changing environments and stressors. Specifically, the proteins that comprise the cellular machinery, as well as the regulatory proteins that mediate their activity, must be maintained at appropriate levels to ensure proper functioning of the pathways they dictate. A major contributor to the flux of the proteome is the proteasome protein degradation system, which includes a complex network of ubiquitin conjugating enzymes and the proteasome proteolytic machinery, that selectively degrades the tagged proteins into peptides, which can then be recycled for new protein synthesis (Glickman & Ciechanover, 2002; Goldberg, 2003). Several factors influence if a protein is targeted for degradation, such as genetic mutations and errors in protein synthesis (Kostova & Wolf, 2003), or damage as a result of different external or endogenous stresses, such as oxidation or aging (Pickering & Davies, 2012). If misfolded, mutated and damaged proteins are not rapidly removed from the cell, they can cause further damage by forming toxic aggregates, which are known to contribute to neurodegenerative diseases such as Alzheimer's and Parkinson's disease (Schwartz & Ciechanover, 2009). In addition to abnormally folded proteins, the levels of many important regulatory proteins, such as cell

* Author for correspondence: Prof. Michal Sharon, Department of Biomolecular Sciences, Weizmann Institute of Science, 234 Herzl St, Rehovot 7610001, Israel. Tel: +972-8-934-3947, Fax: +972-8-934-6010, michal.sharon@weizmann.ac.il.

Conflict of Interest: None

cycle regulators and tumor suppressors, must be carefully balanced during normal development, as disruption of their equilibrium can contribute to tumor development and cancers (Schwartz & Ciechanover, 2009).

Protein degradation predominantly centers around the 26S proteasome, a 2.5 MDa protein complex responsible for recognizing and degrading proteins that have been targeted for destruction (Schwartz & Ciechanover, 2009). The 26S proteasome complex is comprised of the 19S regulatory particle (RP), responsible for the recognition and unfolding of proteasomal substrates, and the 20S core particle (CP), a barrel shaped catalytic entity that accepts the unfolded substrates and degrades them into small peptides (Fig. 1). Proteins are targeted for degradation by the 26S proteasome by the covalent linkage of ubiquitin (Ub), a small 8.5 kDa regulatory protein, onto the substrate via a repetitive enzymatic cascade (Fang & Weissman, 2004). Once tagged with a poly-Ub chain, substrates are recognized by Ub receptors (UbRs); either ‘shuttling factors’, which direct the tagged proteins to the 26S proteasome, or integral subunits of the 19S RP itself. The poly-Ub chains are then removed by various deubiquitinating enzymes (DUBs), allowing recycling of Ub monomers, while the substrate is unfolded and directed to the 20S CP for degradation (Finley, 2009) (Fig. 1).

Our understanding of the finer details of this complicated and essential pathway has greatly improved in recent years, due in large to dramatic improvements in the technologies available for studying complex biological systems. Here, we will specifically discuss the contribution of mass spectrometry (MS). We will focus on distinct steps along the Ub-proteasome degradation pathway and describe the MS-based techniques that were applied to reveal novel biological input on these processes (Fig. 1).

1.2 A general overview of mass spectrometry

The fundamental principle of MS is the measurement of the mass of a molecule, from which multiple levels and types of information can be gained. The foundation of the technique involves the ionization of molecules i.e. the transformation of molecules from solution to the gas phase. The generated ions are then focused into a beam and directed into a collision cell, wherein their dissociation can be induced. Subsequently, the masses of the intact ions and/or their fragments are mass analyzed based on their resulting mass to charge ratio (m/z). Below we provide a brief overview of these steps.

1.2.1 Sample preparation—Prior to MS analysis, different sample preparation methods can be utilized depending on the question being addressed. A substantial proportion of experiments require the identification, quantification and sequence analysis of individual proteins from a heterogeneous sample, such as whole cell protein extracts. One routine approach involves the bottom-up method, which utilizes enzymatically produced peptides of complex protein mixtures that are subjected to liquid chromatography (LC) separation and two steps of MS analysis. In the first MS acquisition, the masses of the intact peptides are determined, while the tandem MS (MS/MS) fragmentation process gives rise to cleavage products that break along peptide bonds, producing information on the identity and sequence of the protein as well as its modifications.

The most commonly used enzyme for digestion is trypsin, which is known to cleave peptides exclusively at the carboxyl side of lysine and arginine residues, thus generating unique, reproducible peptides which can be identified by MS, and assigned to their parent proteins using advanced search algorithms and existing databases of known and predicted tryptic peptides (Gillet *et al.*, 2016; Vandermarliere *et al.*, 2013). Tryptic digestion is a cornerstone of modern quantitative proteomics, and is often used in combination with other techniques to create advanced MS workflows that can address complex biological questions.

Alternatively, the top-down approach considers a protein sample without in-solution digestion, introducing into the mass spectrometer intact proteins or protein complexes. The major advantage of using this approach is its ability to maintain weak non-covalent interactions between proteins and associated biomolecules such as protein partners, drugs and cofactors (Sharon, 2013). This method provides details on the existence of multiple isoforms of single proteins as well as structural data on protein assemblies, revealing their subunit stoichiometry, protein composition, interaction partners, subunit topology and overall architecture.

1.2.2 Ionization methods—Two main methods exist for the ionization of biomolecules; electrospray ionization (ESI), in which aqueous protein solutions are subjected to strong electric currents, causing the ejection of liquid droplets into the gas phase, followed by desolvation and the accumulation of charge on the protein surface (Fenn, 1989; Gaskell, 1997; Wilm & Mann, 1996), and matrix-assisted laser desorption/ionization (MALDI), where protein samples are applied to a dry crystalline matrix and volatilized by laser pulses (Karas & Hillenkamp, 1988; Zenobi & Knochenmuss, 1998). In general, ESI is the method of choice when studying complex protein samples, as it provides a more gentle ionization process which maintains non-covalent interactions (Loo, 1997). In addition, the advent of nanoESI drastically reduced the required sample size and increased the tolerance of the ionization process to the presence of salts found in complex biological buffers (Juraschek *et al.*, 1999), thus allowing the analysis of protein complexes in their native state. The added advantage of ESI is the capacity to couple an LC column to the electrospray needle, a technique known as LC-MS. This permits the separation of peptides or proteins prior to their entry into the mass spectrometer, a method that is particularly relevant in the analysis of complex peptide samples. After the sample has been ionized, the generated ions then enter the mass spectrometer under high vacuum, and are separated based on their m/z by a mass analyzer.

1.2.3 Mass analyzers and MS/MS—Several types of mass analyzers can be used, such as quadrupole, time-of-flight (TOF), Fourier transform ion cyclotrons (FT-IC) and ion traps, such as the Orbitrap. They each have different strengths and weaknesses, and have been previously reviewed in excellent detail (Aebersold & Mann, 2003; Benesch *et al.*, 2007; Domon & Aebersold, 2006). Each of these mass analyzers can be used independently, or they can be coupled together in tandem in a hybrid instrument, facilitating MS/MS analysis.

In MS/MS, after the ions have been separated in the first mass analyzer, specific ions, also known as precursor ions, can be selected and fragmented to produce product ions, which are then separated and detected by the second mass analyzer. Various fragmentation methods

can be utilized depending on the required outcome, such as surface-induced dissociation (SID) (Cooks *et al.*, 1975; Zhou & Wysocki, 2014), electron-capture dissociation (ECD) (Zubarev & Kelleher, 1998), electron transfer dissociation (ETD) (Syka *et al.*, 2004) or UV photodissociation (UVPD) (Brodbeck, 2014). However, the most commonly used method is collision-induced dissociation (CID), in which neutral gas molecules are collided with the selected ions increasing their internal energy. When analyzing peptides, as done in bottom-up proteomic approaches, this method breaks the peptide bonds, which in turn enables sequence identification. When used for the analysis of protein complexes, the CID approach leads to the disruption of non-covalent associations and consequently to the dissociation of individual subunits and generation of “stripped complexes” (Benesch, 2009). Nevertheless, in both cases, the detection and identification of the precursor and product ions provides valuable information.

1.2.4 MS workflows for investigating the proteasome system—Here we will describe in detail several MS workflows that have been specially designed and adapted for the study of protein complexes, with a focus on their application for dissecting the details of the proteasome mediated degradation pathway. Specifically, we will describe recent advances in our understanding of the ubiquitination mechanism, the regulation of proteasome assembly and degradation process, the overall structure of the 26S proteasome, as well as the heterogeneity of proteasome populations present within the cell (Fig. 1). The methodologies described in these studies include the use of quantitative proteomics approaches such as absolute quantification using the AQUA strategy (Fig. 2a), isobaric labelling of protein extracts (Fig. 2b) and label free quantification (Fig. 2c), as well as structural MS approaches like crosslinking (Fig. 2d) and native MS (Fig. 2e). The use of these MS workflows, often in combination with other biochemical and cell biology analyses, have driven our fundamental understanding of this essential pathway, highlighting the relevance and versatility of MS as an exceptional technique for addressing multiple aspects of a complex and dynamic cellular pathway.

2 The ubiquitination process

The first stage of the protein degradation pathway is the recognition and ubiquitination of proteins that are destined for destruction by the 26S proteasome. Ubiquitination is a complex procedure, involving a three step enzymatic cascade coordinated by multiple enzymes (Fang & Weissman, 2004). These enzymes can be divided into three main families; Ub activating enzymes (E1), conjugating enzymes (E2) and ligating enzymes (E3). The mechanism of function of each of these enzymes has been described in detail (Fang & Weissman, 2004), with the final result of the first round of the process being the formation of a covalent isopeptide bond between the C-terminal glycine of Ub, and a lysine residue in the substrate. Subsequent rounds of ubiquitination lead to the linkage of additional Ub monomers, attached to one of the seven lysine residues in Ub itself (K6, K11, K27, K29, K33, K48, K63) or the N-terminal methionine, forming a poly-Ub chain on the protein substrate (Fig. 3a). These linkages can either be homotypic i.e. through the same K residue in all the Ub monomers in the chain, or heterotypic i.e. a combination of multiple types of linkages resulting in mixed or branched Ub chains. Recognition of these different Ub chain structures

by Ub-binding proteins/Ub receptors determines whether the tagged substrate will be directed to the 26S proteasome for degradation, or towards other cellular pathways in which Ub-tagged proteins have been shown to play a role, such as cell cycle progression and DNA repair (Pickart & Fushman, 2004). Some of the fine details regarding the diversity of poly-Ub chains on substrates, as well as how these different poly-Ub chain types can direct a substrate to a particular fate, were described using MS, by defining the precise K-linkages on a ubiquitinated substrate along with quantifying their amount within a single sample using a technique known as AQUA (Fig. 2a, Fig. 3) (Kirkpatrick *et al.*, 2005; Kirkpatrick *et al.*, 2006).

2.1 The use of AQUA to quantify ubiquitin linkages

AQUA is a quantitative method, involving the absolute quantification of peptides within a trypsin digested heterogeneous sample (Gerber *et al.*, 2003; Kirkpatrick *et al.*, 2005). This is achieved by pre-designing signature peptides that contain the tryptic peptide sequence and modification of interest (if relevant), and synthesizing them to contain stable heavy isotopes (Fig. 3a). The signature peptides are then added into the sample under investigation at a known concentration and analyzed by LC-MS/MS. In the first stage of the MS analysis, both the heavy labelled peptide and the light analyte peptide will co-elute from the chromatography column with the same retention time, as their biochemical properties are identical (Fig. 3b). Selected reaction monitoring (SRM) is then performed, in which both the heavy and light peptide precursor ions are specifically selected and subjected to MS/MS, with the product ions from MS/MS detected and analyzed (Fig. 3c). The small mass difference between the heavy and light peptides can be distinguished by the mass spectrometer, and by comparing their relative intensities, absolute quantification of the peptide levels in the sample can be achieved (Gerber *et al.*, 2003; Kirkpatrick *et al.*, 2005).

This method was successfully applied to address the question of ubiquitin linkages by the design of isotope-labelled tryptic peptides corresponding to the types of K-linkages possible (Fig. 2a, Fig. 3) (Kirkpatrick *et al.*, 2006). The final amino acids at the C-terminus of Ub are LRRG, and as described earlier, the C-terminal G forms a covalent bond with any of the seven K residues in Ub during the formation of poly-Ub chains. Therefore, if a polyubiquitinated substrate is digested with trypsin, one of the cleavage events will occur between the R and the first G at the C-terminus of Ub. This will yield peptides in which the modified K in the substrate, and the modified K in the attached Ubs, will have a covalently linked -GG motif (Fig. 3a). This motif will therefore add additional mass to the peptide, permitting its identification by MS, and differentiation from peptides that have not been modified with Ub. Thus, by designing AQUA peptides to mimic these -GG linkages, and adding them to the sample under investigation, the quantities of each type of Ub linkage and their location on a particular substrate can be accurately determined (Fig. 2a, Fig. 3) (Kirkpatrick *et al.*, 2005; Kirkpatrick *et al.*, 2006). More recently, a monoclonal antibody designed to recognize -GG modifications on isopeptides was developed, enabling the enrichment of these modified peptides from whole cell tryptic digests prior to MS analysis, dramatically improving the number of identified ubiquitination events across the entire proteome (Kim *et al.*, 2011).

To demonstrate the utility of the AQUA method, the precise quantification of the types of Ub-linkages present on the model substrate cyclin B1, a cell cycle regulator, was performed (Kirkpatrick *et al.*, 2006). The authors demonstrated that this substrate is tagged by multiple mono-Ub and heterogeneous short chain poly-Ubs, although only through K11, K48 or K63, and that this spread of linkages is sufficient for proteasomal recognition and degradation. This was in contrast to a previous hypothesis that only homotypic K48 poly-Ub linkages target substrates for proteasomal degradation (Pickart & Fushman, 2004). The authors took this analysis further by investigating the relationship between the E2 and E3 enzymes which coordinate the ubiquitination of cyclin B1, in order to understand how the proteasomal targeting signal of the previously described ubiquitination is generated, leading to the degradation of cyclin B1.

The ubiquitination of cyclin B1 is known to be coordinated by the anaphase-promoting complex (APC), a multi-subunit complex of E3 ligases, which interacts with E2 conjugating enzymes Ubc4 or UbcH10 (Peters, 2002). By using Ub-AQUA peptides and comparing the resulting ubiquitination of cyclin B1 *in vitro*, when either Ubc4 or UbcH10 was present, it was revealed that Ubc4 produces twice the amount of K48 linkages, while UbcH10 favors the formation of K11 linkages, although both enzymes produce the same level of K63 linkages (Kirkpatrick *et al.*, 2006). Other types of K linkages i.e. K6, K27, K29 and K33, were either present in minor quantities or not detectable at all, indicating that the APC restricts the type of K-linkages formed on cyclin B1. Time dependent analysis of ubiquitination events revealed that cyclin B1 is modified by several mono-Ub linkages first, followed by their extension to short poly-Ub chains over time. These are then recognized by various UbRs which direct the substrate to the 26S proteasome for degradation. The authors concluded that while the types of K-linkages formed on a substrate is restricted by E3 enzymes, the proportions of these linkages i.e. mono/poly-Ub, depends on the E2 enzymes present. This method has since been utilized in other studies, for example in the discovery that homotypic K11 linkages are not capable of targeting a substrate for proteasomal degradation while heterotypic K11/K48 poly-Ub are (Grice *et al.*, 2015).

The fates of ubiquitinated substrates are not limited to proteasomal degradation, indeed Ub-AQUA MS played a role in the demonstration that K11 and K63 linkages are involved in endocytosis (Boname *et al.*, 2010). Many more examples than can be described here have benefited from the development of the Ub-AQUA MS method, clearly demonstrating its widespread applicability (Bennett *et al.*, 2010; Xu *et al.*, 2009). Overall, by utilizing the specificity of AQUA peptides and the high resolving power of the new generation mass spectrometers, this approach enabled the definition of the complex relationships between the various possible Ub-linkages, their outcomes, and the roles of the enzymes involved in the ubiquitination cascade, significantly increasing the cumulative knowledge surrounding the ubiquitination process. We anticipate that this methodology will continue to be used to explore the molecular details of this and other dynamic biological processes.

2.2 *In vivo* labelling to explore the effects of deubiquitination

Deubiquitination involves the enzymatic removal of the conjugated Ub moiety from a tagged substrate, providing a proofreading mechanism that enhances the fidelity of the degradation

process. DUBs also allow for recycling of the ubiquitin moieties from proteins prior to their final commitment to degradation (Finley *et al.*, 2012). To date many different members of the DUB family have been identified (Finley *et al.*, 2012; Reyes-Turcu *et al.*, 2009) and it is known that they can act either as independent entities or as part of the proteasome complex. For example the integral 19S subunit Rpn11 and proteasome associated protein Usp6 both display DUB activity, rescuing Ub from being degraded by the proteasome along with the substrate (Sakata *et al.*, 2011; Verma *et al.*, 2002). While the mechanistic details of the deubiquitination process have been determined, greater knowledge of the global role played by DUBs in a cellular context is imperative for our understanding of the diverse effects of this enzyme family.

To address this question, a multiplexed MS approach incorporating isobaric labelling and quantitative proteomics was developed, permitting the precise, comparative measurement of specific protein levels from several different cell lines in a single experiment (Isasa *et al.*, 2015). This proteomic workflow utilized Tandem Mass Tagging (TMT) of enzymatically digested whole cell protein extracts followed by LC-MS/MS (Fig. 2b). TMT tags are made up of four components: a reactive group on one end that reacts with a protein to covalently couple the tag, a reporter ion at the other end that is isotopically labeled, and between them a cleavable linker to permit the release of the reporter ion during MS/MS as well as a mass normalization tag that ensures all the TMTs used in an experiment will have the same mass (Dayon *et al.*, 2008). Due to the availability of up to 10 different isotope-reporter ions, it is possible to label and analyze 10 samples each with a different tag, thus providing a multiplexed platform from which multiple cell lines can be simultaneously compared. After tagging, the samples are pooled together and subjected to LC-MS/MS. In the first round of MS, the same peptide being analyzed across the samples will have the same mass regardless of the TMT tag attached, and therefore they will all fall into a single ion peak. However, upon selection of this ion peak followed by MS/MS, the cleavable linker will allow release of the reporter ions, and the subsequent mass shifts of each separate tag will allow discrimination between the same peptide from all the samples under investigation, as well as quantification of their relative amounts (Fig. 2b).

In order to address the question of the role of DUBs in a global context, this method was applied to analyze changes in the proteome of 9 yeast strains in which specific DUBs had been knocked out, by comparison with the proteome of wild type cells (Isasa *et al.*, 2015). The authors were able to quantify thousands of proteins with high reproducibility, noting significant changes in certain protein levels in the absence of specific DUBs. For example, they demonstrated that the global pool of Ub-conjugated proteins is differentially effected when different DUBs are knocked out, with the most drastic reduction in Ub-conjugate level seen in the Ubp6 knockout cells. This suggests that in the absence of Ubp6, the 26S proteasome degrades the Ub chain along with the substrate, indicating an essential role for Ubp6 in maintaining the overall pool of free Ub in the cell. Moreover, the ability to simultaneously quantify and compare the levels of thousands of proteins in each cell line lead to the identification of novel interaction pathways for certain DUBs. For example, aside from its known deubiquitinating activities, Ubp3 appears to be involved in mitochondrial regulation via interactions with the cytochrome c oxidase complex, as upon Ubp3 knockout significant up-regulation of the members of the cytochrome c oxidase complex was

observed. In addition, knockout of several DUBs, including Ubp3, Ubp10 and Otu2, affected the inorganic phosphate pathway, by causing downregulation of several phosphate receptors as well as certain phosphatases and kinases, indicating a clear role for these DUBs in inorganic phosphate signaling regulation (Isasa *et al.*, 2015). Taken together, these data demonstrate the power of this multiplexed MS platform, which exploits the simultaneous identification of thousands of proteins from 10 different samples to identify novel biological roles for certain DUBs. Future expansion of this technology to include an even wider range of available isobaric tags will enable the direct comparison of many more different cell states and conditions, which will undoubtedly drive forward our understanding of complex protein interaction networks.

2.3 Identifying UbR/proteasome complexes using label free quantitative MS

The recognition of Ub modifications on protein substrates is performed by UbRs, which as described earlier are either 'shuttling factors', which independently escort Ub-tagged proteins to the proteasome for degradation, or integral subunits of the 19S RP. Five different families of UbR have been described; the shuttling factors Rad23, Dsk2 and Ddi1, and the 19S subunits Rpn10 and Rpn13, with various isoforms identified in each family. These UbRs have been shown to have different affinities for certain substrates (Verma *et al.*, 2004), and the mechanism of their recognition has been previously reviewed in depth (Finley, 2009). Until recently however, the relationship between the various UbRs, their ability to recognize different substrates and direct them to the proteasome population was unclear and challenging to study, given the transitory nature of many of these interactions and the complexity of the network as a whole.

To address this issue, a label-free MS-based approach was developed to examine the interactions between the different UbRs with the 26S proteasome (Fig. 2c) (C Yu *et al.*, 2016). Label-free proteomics involves the identification, quantification and comparison of proteins between multiple samples without the use of isobaric tags or labels. The method relies on the comparison of precursor signal intensity or spectral counting. Often, proteins of interest are isolated from cells using various methods such as affinity purification, which allows the purification of intact protein complexes and their interacting partners. The purified proteins are then digested and analyzed by LC-MS/MS followed by data analysis, in which the peptide ion peaks are integrated and used to measure the relative amounts of the proteins present in each sample. This technique allows the determination of the identity of interacting proteins as well as subunit stoichiometry within a complex (Fabre *et al.*, 2014).

In order to examine the interactions between the proteasome and the UbRs, an MS workflow was designed to incorporate affinity purification of specific complexes from whole cell extracts, followed by LC-MS/MS for the identification of the proteins present in each complex. As many of the interactions between the 26S proteasome and UbRs are dynamic, the authors included a mild crosslinking step to capture these transient complexes. They utilized low levels of formaldehyde, a small membrane permeable compound capable of non-specific crosslinking of protein complexes (crosslinking is discussed in further detail below) (Sutherland *et al.*, 2008). By using this mild crosslinker, dynamic interactions between proteins can be captured, thus allowing short lived complexes to be isolated and

studied. An additional component of this MS workflow is the use of tandem affinity purification to capture specific complexes. By incorporating affinity tags onto two different subunits of the complexes under investigation, it is possible to pull out these specific complexes from the cellular milieu using antibodies, thus enabling the preparation of cleaner samples and consequently acquiring better data with reduced background. As the authors were interested in studying UbR/proteasome complexes, they incorporated affinity tags on the UbRs of interest as well as on the 26S proteasome, thus allowing tandem affinity purification to pull out UbR/proteasome complexes as opposed to each entity separately (C Yu *et al.*, 2016). They generated 7 cell lines, each expressing tagged Rpn11 to pull out the 26S proteasome, as well a different tagged UbR. After the complexes were purified, they were enzymatically digested and the generated peptides were separated and identified using LC-MS/MS. Quantification of the complexes revealed that while each of the UbRs bound to the 26S proteasome, distinct differences in their binding stoichiometries were apparent. For example, hHR23B, an isoform of the Rad23 family, bound significantly more 26S proteasome compared with the other shuttling factors, indicating that this UbR may be the primary UbR responsible for escorting Ub-tagged substrates to the 26S proteasome for degradation. In addition, the authors identified several proteasome interacting proteins (PIPs) that co-purified with the UbR/proteasome complexes. While some of these PIPs were identified across all the samples, indicating that they interact with all the UbRs under investigation, distinct subsets of these PIPs selectively engaged with specific UbRs. For example, subunits of the CCT complex, a large protein folding chaperonin, were shown to interact only with Ddi2, indicating that Ddi2 is responsible for shuttling either misfolded CCT substrates, or the CCT subunits themselves to the 26S for degradation. By combining tandem affinity purification with MS, the authors were able to describe the interaction network between UbRs, the 26S proteasome, and other PIPs, revealing connections between different cellular pathways and specific UbRs, thus providing novel insights into the coordination of substrate recognition and targeting to the 26S proteasome (C Yu *et al.*, 2016).

3 The structure of the 26S proteasome

Understanding the function and behavior of a protein complex relies heavily on detailed structural information, in order to determine subunit composition, contact interfaces, conformational changes upon ligand binding and the effects of mutations. While structural techniques such as X-ray crystallography and NMR spectroscopy are the gold standard for determining protein structures of individual proteins or small uniform complexes, their utility for examining large protein complexes are often limited. This is partly due to the requirements for large volumes of highly purified sample, as well as the dynamic nature of many protein complexes which leads to heterogeneous populations comprised of varying subunits and interacting proteins. The development of alternative hybrid-MS based approaches for extracting structural information from large protein complexes, such as crosslinking coupled MS (XL-MS), has therefore proven invaluable for examining protein complex structure (Calabrese & Pukala, 2013; Chait *et al.*, 2016; Leitner *et al.*, 2016; Rappsilber, 2011).

While the crystal structure of the 20S CP from the Archaea *Thermoplasma acidophilum* was determined more than two decades ago (Lowe *et al.*, 1995), followed by the structure of the yeast (Groll *et al.*, 1997) and mammalian 20S CP (Unno *et al.*, 2002), the structure of the 19S RP has been much more elusive, likely due to the dynamic nature of its subunits. The 19S RP is composed of 19 different subunits in both yeast and mammals, which are organized into two distinct subcomplexes; the base, which contacts the 20S CP, and the lid, which attaches to the base (Finley, 2009; Finley *et al.*, 2016). The base consists of 10 subunits, six of which are ATPases organized into a hexameric ring (Rpt1-Rpt6) as well as four non-ATPase subunits (Rpn1, Rpn2, Rpn10, Rpn13). The base is responsible for the recognition and docking of Ub-tagged substrates via the UBRs Rpn10 and Rpn13, as well as unfolding the substrates and translocating them through the central pore of the hexameric ATPase ring into the 20S CP for degradation. The lid is comprised of 9 non-ATPase subunits (Rpn3, Rpn5–Rpn9, Rpn11, Rpn12, and Rpn15/Sem1), with Rpn11 performing the key activity of deubiquitination at the proteasome itself (Finley, 2009; Finley *et al.*, 2016). In recent years the topology and organization of all of these subunits was determined and validated using both XL-MS and cryoelectron microscopy (cryoEM), overcoming the challenges associated with this large dynamic structure (Bohn *et al.*, 2010; Kao *et al.*, 2012; Lander *et al.*, 2012; Lasker *et al.*, 2012; Luan *et al.*, 2016). Here we will discuss the details and outcomes of the XL-MS studies, which have provided insights into the structure of the entire 26S proteasome complex.

3.1 Determining the structure of the 26S proteasome using crosslinking

Chemical crosslinking involves the covalent linking of two amino acid residues that are located close to each other within the native structure of a protein or protein complex (Fig 2d) (Rappsilber *et al.* 2000). The crosslinker itself is comprised of two reactive groups that are joined by a spacer arm of defined length, between 0–30 Å (Leitner *et al.*, 2016). The reactive groups form covalent bonds with amino acids that are positioned within the spatial constraints of the spacer arm. Depending on the particular chemistry of the crosslinker and its spacer arm length, multiple layers of information can be gained regarding amino acid contacts within and between proteins, subunit interfaces within complexes as well as spatial constraints that provide detail regarding orientation and subunit topology. The most commonly used crosslinkers, such as bis(sulfosuccinimidyl) suberate (BS³) and disuccinimidyl (DSS), are amine reactive, linking together K residues via their amino group. Other types of crosslinkers can react with acidic residues or cysteine residues, although these have other associated challenges, such as the low abundance of cysteines and reduced reactivity, and therefore have a more limited application (Leitner *et al.*, 2010; Leitner *et al.*, 2014). In order to gain detailed structural information from a crosslinked protein complex, the sample is enzymatically digested, resulting in a mixture of crosslinked and uncrosslinked peptides (Fig 2d). The crosslinked peptides can then be enriched using various chromatography methods to reduce the background levels of uncrosslinked peptides, followed by LC-MS/MS for sequence identification. During MS/MS, the peptide ions are selected, fragmented and analyzed to locate the position of the covalently attached crosslinker, which will add a known mass to the peptide thus allowing the exact identification of the amino acid that has been modified on both peptides involved. This is achieved using sophisticated search algorithms and pre-generated databases of all the

potential crosslinked peptides that could exist in the protein complex under investigation. Some crosslinkers are also designed to incorporate an MS cleavable spacer arm, facilitating MS/MS fragmentation and peptide identification, as described below (Fig. 4). Once the location of the crosslinks on the peptides has been determined, this information can then be used in combination with the known spatial constraints of the crosslinker to create an interaction map between the various subunits of the complex, thus producing a topological model of the quaternary structure of the complex (Calabrese & Pukala, 2013; Leitner *et al.*, 2010; Leitner *et al.*, 2016; Rappsilber *et al.*, 2000; Rappsilber, 2011).

The first analysis of the composition of the 19S RP using XL-MS was performed a decade ago (Sharon *et al.*, 2006), revealing the unit stoichiometry of the complex and detailing the first interaction map between the subunits, however, the exact interaction sites between subunits were not identified. Since then several studies have used XL-MS to further describe the interactions between the 19S base and the 20S, as well as the positioning and assembly of the 19S lid subunits (Bohn *et al.*, 2010; Kao *et al.*, 2012; Lasker *et al.*, 2012), which have been further validated by cryoEM analysis (Lander *et al.*, 2012; Luan *et al.*, 2016). These studies used amine-reactive crosslinkers such as BS³ or DSS, with the exception of Kao *et al.*, who described the use of a novel MS cleavable crosslinker, disuccinimidyl sulfoxide (DSSO) (Kao *et al.*, 2011). DSSO is also amine reactive, however it contains a C-S bond in the spacer arm that breaks more readily than the peptide backbone when subjected to CID (Fig. 4a). The generated peptide fragment ions contain a thiol group on the α -peptide, and an alkene group on the β -peptide, or vice versa. As a consequence, the generated pair of fragments have a characteristic mass relationship (+86 Da and +54 Da respectively) and are therefore readily identifiable by MS (Fig. 4b). By applying three successive steps of MS (MS/MS/MS) analysis, these peptide fragments can then define the peptide sequence, thus mapping the exact location of the crosslink, and hence the K-K interaction site between proteins (Fig. 4c). The simplicity of the crosslinking chemistry, in combination with the integrated data from the multiple stages of MS analysis, facilitated unambiguous, high confidence peptide identification using a rapid workflow and conventional database searches (Kao *et al.*, 2011). After validation of DSSO as a viable and effective crosslinker on the 20S CP, the authors applied this method to determine with higher confidence the 19S RP structure, achieving hundreds of inter- and intra-subunit crosslinks. 24 unique interactions between subunits were identified, both within and between the lid and base. For example, a crosslink between K195 in Rpn8 and K208 in Rpn11 was unambiguously identified with high confidence (Fig. 4c), confirming the close proximity and interaction between these two subunits, which have been shown to form a heterodimer (Lander *et al.*, 2012; Lasker *et al.*, 2012; Luan *et al.*, 2016). In addition, the authors were also able to identify novel interactions between 19S subunits, such as that between two base subunits - Rpn2, a critical scaffolding protein, and Rpn13, a UBR.

To generate an overall model of the complexes topology from the identified interactions, rather than relying on previously obtained structural data from cryoEM studies, Kao *et al.* also developed a probabilistic modelling approach, in which only crosslink data was used to generate maximum likelihood models of the relationships between the subunits, ultimately producing the highest confidence maximum likelihood model. They generated accurate models for the 19S base and lid separately, both of which were in excellent agreement with

previously published structural data (Bohn *et al.*, 2010; Lander *et al.*, 2012; Lasker *et al.*, 2012). Thus, the cumulative knowledge gained from XL-MS, in combination with cryoEM, have led to the current consensus model for the 19S RP (Bohn *et al.*, 2010; Kao *et al.*, 2012; Lander *et al.*, 2012; Lasker *et al.*, 2012; Luan *et al.*, 2016). Overall, the intricate details of the subunit topology of the 19S as determined by XL-MS and cryoEM studies has greatly expanded our understanding of the mechanisms underlying substrate recognition and processing by the 26S proteasome.

4 The heterogeneity of proteasome populations

The global pool of proteasomes within cells is not just restricted to the classical 26S complex that has been the major focus of the structural studies outlined above. It is well established that there are several different variants of the 20S CP (Ben-Nissan & Sharon, 2014). The standard 20S CP is composed of four stacked hexameric rings; the two outer rings formed by subunits α 1-7, and the two inner rings formed by subunits β 1-7 (Groll *et al.*, 1997; Unno *et al.*, 2002). The proteolytic activity of the 20S is localized to three catalytic β subunits, β 1, β 2 and β 5, each possessing caspase-like, trypsin-like and chymotrypsin-like activity, respectively (Dick *et al.*, 1998). These three subunits can be replaced by β 1i, β 2i and β 5i, forming the immunoproteasome, which is upregulated in immune cells and in response to interferon stimulation (Krüger & Kloetzel, 2012). In addition, there exist two intermediate proteasomes, formed by a mixed assortment of the immune subunits and standard subunits (β 1- β 2- β 5i and β 1i- β 2- β 5i) (Guillaume *et al.*, 2010). Beyond the variation in the 20S CP, there are also several proteasome regulatory complexes that can modulate the activity of the 20S CP. Aside from the 19S RP, there are the heterohexameric PA28 $\alpha\beta$, the homohexameric PA28 γ and the PA200 protein activator, as well as the inhibitory PI31 protein (Chu-Ping *et al.*, 1992; Schmidt *et al.*, 2005; Shibatani *et al.*, 2006; Ustrell *et al.*, 2002). However, the distribution and quantities of these proteasome subtypes, as well as their connections with the various regulators, was until recently undetermined due to the difficulties associated with distinguishing and quantifying subtle changes in proteasome complex composition. This was overcome in studies which used a combination of *in vivo* crosslinking, affinity purification and label-free MS quantification to determine the cellular localizations and levels of the different 20S proteasome populations and their interacting regulatory complexes and proteins, as well as the plasticity of these populations upon changing cellular conditions.

4.1 Exploring proteasome heterogeneity using label free quantitative MS

As described earlier, the combination of *in vivo* crosslinking with formaldehyde, affinity purification and LC-MS/MS is a powerful approach to quantify protein complexes of interest. The studies described here used a single affinity purification step, pulling out proteasomes using an antibody directed to the α 2 subunit of the human proteasome, which does not vary across the different proteasome subtypes. The first study to use this approach also incorporated cellular fractionation after *in vivo* crosslinking to isolate specific populations of proteasomes that exist in different cellular compartments (Fabre *et al.*, 2013). They were able to identify varying levels of standard, immuno and intermediate 20S proteasomes in the cytoplasm, nucleus and microsomes of two different cell types, as well as

measuring differences in proteasome activity between the cellular compartments. The same group continued their analysis, comparing the diversity of proteasome and PIPs across 9 different cell types, revealing significant differences in the amount of free 20S CP, associated regulators and proteasome assembly chaperones across the different cell types, indicating that there exists careful modulation of proteasome heterogeneity depending on the metabolic needs and cellular functions of the cell lines under investigation (Fabre *et al.*, 2014b). Following on from these results, Fabre *et al.* significantly extended the scope of their approach, combining the aforementioned MS workflow with glycerol sedimentation to isolate different proteasome complex populations prior to LC-MS/MS, and protein correlation profiling (PCP), an in-depth statistical analysis of the absolute quantification of each of the proteasome complexes, regulatory complexes and other interacting proteins, to determine the specificity of interactions with the proteasome (Fig. 2c) (Fabre *et al.*, 2015).

PCP requires the calculation of the protein abundance index (PAI) of each protein under investigation e.g. individual proteasome subunits, PIPs etc. This is achieved by analyzing the three most intense peptides validated for each protein of interest during LC-MS/MS and quantifying the absolute amount of protein based on these peptides using AQUA reference peptides as described earlier (Fabre *et al.*, 2015; Silva *et al.*, 2006). The PAIs of different proteins can then be compared to determine if there is a correlation between their absolute quantities in a particular sample, for example after fractionation to separate discrete protein complexes from one another. A strong correlation between two proteins indicates that they exist in comparable quantities in the same sample, and are therefore likely part of the same complex. Validation of this approach was achieved by comparing the levels of the standard proteasome subunits, e.g. $\alpha 6$ and $\alpha 7$, which showed an extremely strong correlation, and are known to be integral and unchanging 20S subunits, as well as subunits of the 19S, e.g. Rpn1 and Rpn3 (Fabre *et al.*, 2015).

MS datasets were collected from different cell lines that had been subjected to fractionation and LC-MS/MS using the workflow described above, and PCP was performed on the identified proteasome subunits, regulatory complex subunits and interacting proteins. The authors were able to unambiguously define strong correlations between the standard and immunoproteasomes with specific regulatory complexes across multiple cell types. They demonstrated that while the 19S is equally associated with both proteasome subtypes, the immunoproteasome is preferentially associated with PA28 $\alpha\beta$ complex, while the standard proteasome is preferentially associated with PA200 and PI31 – a previously unknown specificity of association. These associations were further validated by altering the ratio of immunoproteasome to standard proteasome, using known methods such as interferon stimulation of cells, which is known to increase the levels of immunoproteasome present (Fabre *et al.*, 2015; Seifert *et al.*, 2010). The absolute quantification of protein levels in isolated proteasomes using this robust MS-PCP approach revealed that there is a defined spread of heterogeneous proteasome complex populations within cells, which show distinct associations with regulatory complexes and regulators. This indicates that there is a mechanism through which proteasome function is fine-tuned depending on the cell type and in response to changing cellular requirements.

5 The regulation of proteasome assembly and degradation

The formation of the 26S proteasome involves a complex assembly pathway in which the 20S CP, 19S base and 19S lid each form as defined entities prior to uniting to become the intact 26S complex. While the assembly pathway of the 20S CP has been well defined (Kunjappu & Hochstrasser, 2014), the assembly of the 19S base and lid and their incorporation into the 26S complex remains under debate. Two assembly pathways have been proposed; either the 19S lid and base form independently of the 20S CP (Funakoshi *et al.*, 2009; Kaneko *et al.*, 2009; Saeki *et al.*, 2009; Sakata *et al.*, 2011; Tomko *et al.*, 2015; Tomko *et al.*, 2010; Tomko & Hochstrasser, 2011), or the 20S CP acts as a scaffold for the formation of the 19S base (Hendil *et al.*, 2009; Park *et al.*, 2009; Roelofs *et al.*, 2009; Z Yu *et al.*, 2015) followed by association of the intact 19S lid (for review (Bar-Nun & Glickman, 2011)). Investigating this incredibly complex pathway, which involves multiple co-existing intermediate precursor complexes and transiently associating chaperones, is particularly difficult. Some of these challenges were overcome using a well-designed hybrid MS approach, incorporating both native MS and quantitative proteomics to determine the organization and sequential combination of assembly precursors which form the 19S base complex (Sakata *et al.*, 2011).

5.1 Using native MS to probe the assembly pathway of the proteasome

While quantitative MS workflows, such as those described above, are excellent at determining the quantity and identity of proteins present in a sample, they are not able to provide details regarding the exact size, oligomeric status and specific stoichiometries of protein complexes in the same sample. Native MS is an effective approach that is capable of overcoming these limitations, in that it does not require the enzymatic digestion of samples into peptides prior to MS analysis. Instead, it involves the transfer of intact proteins and protein complexes into the gas phase, maintaining both biomolecular interactions and quaternary structures of the assemblies (for reviews see (Chorev *et al.*, 2015; Heck, 2008; Konijnberg *et al.*, 2013; Mehmood *et al.*, 2015; Sharon, 2013)). The introduction of extended mass range mass analyzers capable of transmitting, selecting and detecting large m/z values, was vital to the progression of the technique (Benesch *et al.*, 2007; Sharon & Robinson, 2007; Snijder & Heck, 2014). This method can also be used in hybrid MS workflows, such as in combination with XL-MS to define specific structural interactions within a complex (Sinz *et al.*, 2015).

Upon application of a sample for native MS analysis, the measured spectrum will immediately yield information regarding the size of the intact complexes present in the sample, as well as their charge distribution series. Desired peaks can then be selected, isolated, and subject to MS/MS fragmentation, causing the dissociation and release of intact subunits from the complex, which can then be individually analyzed to determine their size and thus identity (Fig. 2e). This powerful technique can be used to determine not only the subunits of a complex, but also additional associating proteins that are difficult to detect by other means. Given that all co-existing states of an assembly can be detected in a single spectrum, native MS is well suited for monitoring the intermediate steps along the proteasome assemble process.

To understand the nature of the assembly precursors of the 19S RP base, 19S RP base subcomplexes were isolated using affinity purification of tagged Rpn1, an integral base subunit. As this isolation technique will pull out all assembly precursors that contain Rpn1 with different combinations of the other subunits, which are en route to forming the intact base complex, the sample was also subjected to a sucrose gradient to isolate individual subcomplexes, allowing for the interrogation of their particular compositions (Sakata *et al.*, 2011). In addition to native MS to define the size and stoichiometries of the complexes, the authors also performed quantitative proteomics to unambiguously define the quantities of proteins present in each subcomplex. They identified the association of several assembly chaperones, Nas6, Nas2, Rpn14 and Hsm3, with defined combinations of 19S base subunits, forming the assembly precursors of the intact 19S base complex. Importantly, it was revealed that the deubiquitinating enzyme, Ubp6, is a component of one of these assembly precursors, the Hsm3 module; a complex of Rpn1, Rpt1, Rpt2 and Hsm3. While the other chaperones could not be detected in the intact 19S base complex, indicating that they dissociate from the 19S base during assembly of the precursors, Ubp6 remains bound to Rpn1, until after binding of the 19S lid to form the intact 19S complex. Further exploration of the continued association of Ubp6 throughout assembly revealed that the deubiquitinating activity of Ubp6 was essential for proteasome assembly, likely due to its continual trimming of ubiquitin chains which are recognized by Rpn1. The use of native MS in combination with quantitative MS in this study was essential for defining not only the composition of the 19S base complex assembly precursors, but also the concerted role of their chaperones to complete the assembly process.

5.2 Understanding the regulation of 20S proteasomal degradation using native MS

Another aspect of the proteasomal pathway that has been investigated using native MS is the regulation of protein degradation performed by the 20S CP, independently of the 19S RP. It has been established that the 20S is capable of degrading partially unfolded proteins in an ATP and Ub independent manner (Baugh *et al.*, 2009; Ben-Nissan & Sharon, 2014; Hwang *et al.*, 2011). Substrates of this pathway are proteins that contain unstructured elements due to aging, mutation, or oxidation, as well as intrinsically disordered regions. While under basal conditions, proteolysis by the 20S proteasome may not constitute a dominant route for protein turnover, under oxidative stress conditions in which many proteins are oxidatively damaged, the 20S proteasome degradation capacity is significantly enhanced. This occurs through disassembly of the 26S proteasome into its 20S and 19S components, as well as de novo synthesis of 20S proteasomes (Aiken *et al.*, 2011; Pickering & Davies, 2012). However, given that degradation by the 20S proteasome is a non-selective process, due to the absence of the ubiquitination step, appropriate control of the 20S proteasome is essential to avoid the unnecessary degradation of proteins.

The first identified regulator of Ub-independent 20S degradation was NAD(P)H:quinone-oxidoreductase-1 (NQO1) (Asher, 2005). NQO1 was shown to inhibit the 20S proteasomal degradation of tumor suppressors' p53 and p73, although the exact mechanism of this inhibition was unclear. Several years later, the application of native MS led to the revelation that NQO1 exists in a double negative feedback loop with the 20S proteasome. Native mass spectra of purified NQO1 confirmed that it forms a homodimer and is bound to the cofactor

FAD (Fig. 5a) (Moscovitz *et al.*, 2012), in agreement with the published crystal structure (Asher *et al.*, 2006). Additional native MS techniques were then applied, such as ion mobility MS (IM-MS), in which intact proteins are directed through a field of inert gas molecules with which they collide, and the time taken to traverse the field is proportional to the size and shape of the protein. This can provide information regarding the folding state of the protein, as more compact molecules will move faster through the field, while unfolded proteins will move slower. Analysis of apo and FAD-NQO1 demonstrated that FAD-NQO1 is compact and stable, while in the apo-state NQO1 adopts a more unfolded structure. Interestingly, the 20S proteasome can degrade apo-NQO1 but not FAD-NQO1, likely due to the partial unfolding of the protein in the absence of FAD. Furthermore, native MS/MS experiments were able to establish that NQO1 can bind directly to the 20S proteasome. Further cellular experiments indicated that in the presence of increased levels of FAD and consequently increased levels of stable NQO1, 20S proteasomal degradation of substrates such as p53 were reduced, indicating that NQO1 inhibited degradation (Moscovitz *et al.*, 2012).

Another inhibitor of the 20S proteasome, DJ-1, was discovered via its structural similarities to NQO1, including homodimerization and the presence of a flavodoxin like fold, as well as their reported involvement in the cellular response to oxidative stress (Moscovitz *et al.*, 2015). *In vitro* and cellular characterization of its ability to regulate 20S mediated degradation of substrates was performed, demonstrating that it efficiently inhibits the degradation of multiple substrates. Using a similar native MS approach as was utilized for NQO1, the authors revealed that DJ-1 physically binds to the 20S proteasome, as it is released from the selected 20S ions upon the induction of collisional activation (Fig. 5b). Taken together, the use of advanced native MS techniques in these studies established that NQO1 and DJ-1 are both inhibitors of 20S proteasomal degradation, able to physically bind to the 20S proteasome complex and affect its activity.

Recent advances in mass spectrometer technology have led to the development of the extended mass range (EMR) Orbitrap system, an instrument capable of examining large protein complexes in their native state at high spectral resolution and mass accuracy (Rose *et al.*, 2012). The power of this platform was demonstrated by the in-depth analysis of missense mutants of DJ-1, which lead to early-onset Parkinson's disease (Ben-Nissan *et al.*, 2016). Two mutants were analyzed, DJ-1_{A104T} and DJ-1_{D149A}, each displaying a small mass shift from the wild type protein of +44 Da and -30 Da respectively. Despite these very minor mass differences, both mutants were readily distinguishable from DJ-1_{WT} on the Orbitrap EMR when mixed together and analyzed from the same needle (Fig. 5c), while they could not be distinguished from one another using other mass spectrometers. Further analysis of these mutants demonstrated subtle changes in their structural stability which affect their inhibitory influences on the 20S proteasome. The ability to analyze all three proteins concurrently not only reduced data acquisition time but also permitted direct comparison between the proteins within the same experiment, thus removing sample to sample variation which could obfuscate subtle differences. This compelling native MS methodology will allow for the examination of other missense mutations, numbering in the tens of thousands, that have been associated with other diseases (Stefl *et al.*, 2013).

Conclusions and future directions

In this review we have described the multiple contributions of MS to the study of a complex biological pathway, namely the proteasomal protein degradation system. MS workflows incorporating several techniques in tandem, such as *in vivo* and *in vitro* labelling, enzymatic digestion, crosslinking, chromatography and native MS, all coupled with advanced data analysis and statistical methods, have greatly advanced our cumulative knowledge regarding this dynamic and essential system. The ability to scrutinize the identity, quantity, interactions and structures of the multiple proteins and protein complexes involved at every stage of the degradation pathway in such unprecedented detail, exemplifies the use of MS as a critical tool for addressing complicated biological systems.

As described here, we now have a much clearer understanding of the ubiquitination process, assembly pathway and overall structure of the proteasome, as well as details regarding its regulation and population heterogeneity within the cell. However, much remains to be discovered about this fascinating process. For example, deciphering the influence of ubiquitin linkage diversity on substrate selection, recognition and proteasomal targeting will provide important insights into the broad reach of the ubiquitin-proteasome network. Similarly, unraveling the conformational transitions of proteasome subunits as substrates are recognized and deciphering the biological impact of proteasome population heterogeneity across the cellular compartments, especially under stress conditions, will provide essential information for future therapeutic benefits considering that the proteasome is an important target in neurodegenerative disease and cancer. We anticipate that further advances in MS methodologies and technologies will undoubtedly drive our future investigations into this and other biological pathways, as we continually strive to advance our knowledge of these dynamic and enigmatic systems.

Acknowledgements

We would like to acknowledge Gili Ben-Nissan for critically appraising the manuscript during its preparation. We are grateful for the support of a Starting Grant from the European Research Council (ERC) (Horizon 2020)/ERC Grant Agreement no. 636752, and a Minerva Foundation Grant, with funding from the Federal Ministry for Education and Research, Germany. M.S. is an incumbent of the Aharon and Ephraim Katzir Memorial Professorial Chair. MO is thankful for the support of the Deans Fellowship from the Faculty of Life Sciences, Weizmann Institute of Science, and a Postdoctoral Fellowship from the Israel Cancer Research Fund.

References

- Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature*. 2003; 422:198–207. [PubMed: 12634793]
- Aiken CT, Kaake RM, Wang X, Huang L. Oxidative stress-mediated regulation of proteasome complexes. *Molecular & Cellular Proteomics*. 2011; 10 R110.006924.
- Asher G. A mechanism of ubiquitin-independent proteasomal degradation of the tumor suppressors p53 and p73. *Genes & Development*. 2005; 19:316–321. [PubMed: 15687255]
- Asher G, Dym O, Tsvetkov P, Adler J, Shaul Y. The crystal structure of NAD(P)H Quinone Oxidoreductase 1 in complex with its potent inhibitor dicoumarol. *Biochemistry*. 2006; 45:6372–6378. [PubMed: 16700548]
- Bar-nun S, Glickman MH. Proteasomal AAA-ATPases: Structure and function. *Biochimica et Biophysica Acta*. 2011; 1823:67–82. [PubMed: 21820014]

- Baugh JM, Viktorova EG, Pilipenko EV. Proteasomes can degrade a significant proportion of cellular proteins independent of ubiquitination. *Journal of Molecular Biology*. 2009; 386:814–827. [PubMed: 19162040]
- Ben-nissan G, Chotiner A, Tarnavsky M, Sharon M. Structural characterization of missense mutations using high resolution mass spectrometry: A case study of the Parkinson's-related protein, DJ-1. *Journal of the American Society for Mass Spectrometry*. 2016; 27:1062–1070. [PubMed: 27048419]
- Ben-nissan G, Sharon M. Regulating the 20S proteasome ubiquitin independent degradation pathway. *Biomolecules*. 2014; 4:862–884. [PubMed: 25250704]
- Benesch JLP. Collisional activation of protein complexes: Picking up the pieces. *Journal of the American Society for Mass Spectrometry*. 2009; 20:341–348. [PubMed: 19110440]
- Benesch JLP, Ruotolo BT, Simmons DA, Robinson CV. Protein complexes in the gas phase: Technology for structural genomics and proteomics. *Chemical Reviews*. 2007; 107:3544–3567. [PubMed: 17649985]
- Bennett EJ, Rush J, Gygi SP, Harper JW. Dynamics of cullin-RING ubiquitin ligase network revealed by systematic quantitative proteomics. *Cell*. 2010; 143:951–965. [PubMed: 21145461]
- Bohn S, Beck F, Sakata E, Walzthoeni T, Beck M, Aebersold R, Förster F, Baumeister W, Nickell S. Structure of the 26S proteasome from *Schizosaccharomyces pombe* at subnanometer resolution. *Proceedings of the National Academy of Sciences*. 2010; 107:20992–20997.
- Boname JM, Thomas M, Stagg HR, Xu P, Peng J, Lehner PJ. Efficient internalization of MHC I requires lysine-11 and lysine-63 mixed linkage polyubiquitin chains. *Traffic*. 2010; 11:210–220. [PubMed: 19948006]
- Broadbelt JS. Photodissociation mass spectrometry: New tools for characterization of biological molecules. *Chemical Society Reviews*. 2014; 43:2757–2783. [PubMed: 24481009]
- Calabrese AN, Pukala TL. Chemical cross-linking and mass spectrometry for the structural analysis of protein assemblies. *Australian Journal of Chemistry*. 2013; 66:749–759.
- Chait BT, Cadene M, Olinares PD, Rout MP, Shi Y. Revealing higher order protein structure using mass spectrometry. *Journal of the American Society for Mass Spectrometry*. 2016; 27:952–965. [PubMed: 27080007]
- Chorev DS, Ben-nissan G, Sharon M. Exposing the subunit diversity and modularity of protein complexes by structural mass spectrometry approaches. *Proteomics*. 2015; 15:2777–2791. [PubMed: 25727951]
- Chu-Ping M, Slaughter CA, Demartino GN. Purification and characterization of a protein inhibitor of the 20S proteasome (macropain). *Biochimica et Biophysica Acta*. 1992; 1119:303–311. [PubMed: 1312359]
- Cooks RG, Terwilliger DT, Ast T, Beynon JH, Keough T. Surface modified mass spectrometry. *Journal of the American Chemical Society*. 1975; 97:1583–1585.
- Dayon L, Hainard A, Licker V, Turck N, Kuhn K, Hochstrasser DF, Burkhard PR, Sanchez J-C. Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-Plex isobaric tags. *Analytical Chemistry*. 2008; 80:2921–2931. [PubMed: 18312001]
- Dick TP, Nussbaum AK, Deeg M, Heinemeyer W, Groll M, Schirle M, Keilholz W, Stevanovi S, Wolf DH, Huber R, Rammensee H-G, Schild H. Contribution of proteasomal β -subunits to the cleavage of peptide substrates analyzed with yeast mutants. *Journal of Biological Chemistry*. 1998; 273:25637–25646. [PubMed: 9748229]
- Domon B, Aebersold R. Mass spectrometry and protein analysis. *Science*. 2006; 312:212–217. [PubMed: 16614208]
- Fabre B, Lambour T, Bouyssié D, Menneteau T, Monsarrat B, Burlet-schiltz O, Bousquet-Dubouch M-P. Comparison of label-free quantification methods for the determination of protein complexes subunits stoichiometry. *EuPA Open Proteomics*. 2014; 4:82–86.
- 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. *Molecular & Cellular Proteomics*. 2013; 12:687–699. [PubMed: 23242550]
- Fabre B, Lambour T, Garrigues L, Amalric F, Vigneron N, Menneteau T, Stella A, Monsarrat B, Van den Eynde B, Burlet-Schiltz O, Bousquet-Dubouch MP. Deciphering preferential interactions

- within supramolecular protein complexes: The proteasome case. *Molecular Systems Biology*. 2015; 11:771–771. [PubMed: 25561571]
- Fabre B, Lambour T, Garrigues L, Ducoux-Petit M, Amalric F, Monsarrat B, Burlet-Schiltz O, Bousquet-Dubouch M-P. Label-free quantitative proteomics reveals the dynamics of proteasome complexes composition and stoichiometry in a wide range of human cell lines. *Journal of Proteome Research*. 2014b; 13:3027–3037. [PubMed: 24804812]
- Fang S, Weissman AM. A field guide to ubiquitylation. *Cellular and Molecular Life Sciences*. 2004; 61:1546–1561. [PubMed: 15224180]
- Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. *Science*. 1989; 246:64–71. [PubMed: 2675315]
- Finley D. Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annual Review of Biochemistry*. 2009; 78:477–513.
- Finley D, Chen X, Walters KJ. Gates, channels, and switches: Elements of the proteasome machine. *Trends in Biochemical Sciences*. 2016; 41:77–93. [PubMed: 26643069]
- Finley D, Ulrich HD, Sommer T, Kaiser P. The ubiquitin-proteasome system of *Saccharomyces cerevisiae*. *Genetics*. 2012; 192:319–360. [PubMed: 23028185]
- Funakoshi M, Tomko RJ Jr, Kobayashi H, Hochstrasser M. Multiple assembly chaperones govern biogenesis of the proteasome regulatory particle base. *Cell*. 2009; 137:887–899. [PubMed: 19446322]
- Gaskell SJ. Electrospray: Principles and practice. *Journal of Mass Spectrometry*. 1997; 32:677–688.
- Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proceedings of the National Academy of Sciences*. 2003; 100:6940–6945.
- Gillet LC, Leitner A, Aebersold R. Mass spectrometry applied to bottom-up proteomics: Entering the high-throughput era for hypothesis testing. *Annual Review of Analytical Chemistry*. 2016; 9:449–472.
- Glickman MH, Ciechanover A. The ubiquitin-proteasome proteolytic pathway: Destruction for the sake of construction. *Physiological Reviews*. 2002; 82:373–428. [PubMed: 11917093]
- Goldberg AL. Protein degradation and protection against misfolded or damaged proteins. *Nature*. 2003; 426:895–899. [PubMed: 14685250]
- Grice GL, Lobb IT, Weekes MP, Gygi SP, Antrobus R, Nathan JA. The proteasome distinguishes between heterotypic and homotypic lysine-11-linked polyubiquitin chains. *Cell Reports*. 2015; 12:545–553. [PubMed: 26190103]
- Groll M, Ditzel L, Löwe J, Stock D, Bochtler M, Bartunik HD, Huber R. Structure of 20S proteasome from yeast at 2.4Å resolution. *Nature*. 1997; 386:463–471. [PubMed: 9087403]
- Guillaume B, Chapiro J, Stroobant V, Colau D, Van Holle B, Parvizi G, Bousquet-Dubouch M-P, Théate I, Parmentier N, Van den Eynde BJ. Two abundant proteasome subtypes that uniquely process some antigens presented by HLA class I molecules. *Proceedings of the National Academy of Sciences*. 2010; 107:18599–18604.
- Heck AJR. Native mass spectrometry: A bridge between interactomics and structural biology. *Nature Methods*. 2008; 5:927–933. [PubMed: 18974734]
- Hendil KB, Kriegenburg F, Tanaka K, Murata S, Lauridsen A-MB, Johnsen AH, Hartmann-Petersen R. The 20S proteasome as an assembly platform for the 19S regulatory complex. *Journal of Molecular Biology*. 2009; 394:320–328. [PubMed: 19781552]
- Hwang J, Winkler L, Kalejta RF. Ubiquitin-independent proteasomal degradation during oncogenic viral infections. *Biochimica et Biophysica Acta*. 2011; 1816:147–157. [PubMed: 21664948]
- Isasa M, Rose CM, Elsasser S, Navarrete-Perea J, Paulo JA, Finley DJ, Gygi SP. Multiplexed, proteome-wide protein expression profiling: Yeast deubiquitylating enzyme knockout strains. *Journal of Proteome Research*. 2015; 14:5306–5317. [PubMed: 26503604]
- Juraschek R, Dülcks T, Karas M. Nano-electrospray—More than just a minimized-flow electrospray ionization source. *Journal of the American Society for Mass Spectrometry*. 1999; 10:300–308. [PubMed: 10197351]

- Kaneko T, Hamazaki J, Iemura S-I, Sasaki K, Furuyama K, Natsume T, Tanaka K, Murata S. Assembly pathway of the mammalian proteasome base subcomplex is mediated by multiple specific chaperones. *Cell*. 2009; 137:914–925. [PubMed: 19490896]
- Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry*. 1988; 60:2301–2303. [PubMed: 2977069]
- Kao A, Chiu C-L, Vellucci D, Yang Y, Patel VR, Guan S, Randall A, Baldi P, Rychnovsky SD, Huang L. Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes. *Molecular & Cellular Proteomics*. 2011; 10 M110.002212.
- Kao A, Randall A, Yang Y, Patel VR, Kandur W, Guan S, Rychnovsky SD, Baldi P, Huang L. Mapping the structural topology of the yeast 19S proteasomal regulatory particle using chemical cross-linking and probabilistic modeling. *Molecular & Cellular Proteomics*. 2012; 11:1566–1577. [PubMed: 22550050]
- Kim W, Bennett EJ, Huttlin EL, Guo A, Li J, Possemato A, Sowa ME, Rad R, Rush J, Comb MJ, Harper JW, et al. Systematic and quantitative assessment of the ubiquitin-modified proteome. *Molecular Cell*. 2011; 44:325–340. [PubMed: 21906983]
- Kirkpatrick DS, Gerber SA, Gygi SP. The absolute quantification strategy: A general procedure for the quantification of proteins and post-translational modifications. *Methods*. 2005; 35:265–273. [PubMed: 15722223]
- Kirkpatrick DS, Hathaway NA, Hanna J, Elsasser S, Rush J, Finley D, King RW, Gygi SP. Quantitative analysis of in vitro ubiquitinated cyclin B1 reveals complex chain topology. *Nature Cell Biology*. 2006; 8:700–710. [PubMed: 16799550]
- Konijnenberg A, Butterer A, Sobott F. Native ion mobility-mass spectrometry and related methods in structural biology. *Biochimica et Biophysica Acta*. 2013; 1834:1239. [PubMed: 23246828]
- Kostova Z, Wolf DH. For whom the bell tolls: Protein quality control of the endoplasmic reticulum and the ubiquitin–proteasome connection. *The EMBO Journal*. 2003; 22:2309–2317. [PubMed: 12743025]
- Krüger E, Kloetzel P-M. Immunoproteasomes at the interface of innate and adaptive immune responses: Two faces of one enzyme. *Current Opinion in Immunology*. 2012; 24:77–83. [PubMed: 22296715]
- Kunjappu MJ, Hochstrasser M. Assembly of the 20S proteasome. *Biochimica et Biophysica Acta*. 2014; 1843:2–12. [PubMed: 23507199]
- Lander GC, Estrin E, Matyskiela ME, Bashore C, Nogales E, Martin A. Complete subunit architecture of the proteasome regulatory particle. *Nature*. 2012; 482:186–191. [PubMed: 22237024]
- 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. *Proceedings of the National Academy of Sciences*. 2012; 109:1380–1387.
- Leitner A, Faini M, Stengel F, Aebersold R. Crosslinking and mass spectrometry: An integrated technology to understand the structure and function of molecular machines. *Trends in Biochemical Sciences*. 2016; 41:20–32. [PubMed: 26654279]
- Leitner A, Joachimiak LA, Unverdorben P, Walzthoeni T, Frydman J, Förster F, Aebersold R. Chemical cross-linking/mass spectrometry targeting acidic residues in proteins and protein complexes. *Proceedings of the National Academy of Sciences*. 2014; 111:9455–9460.
- Leitner A, Walzthoeni T, Kahraman A, Herzog F, Kahraman A, Rinner O, Herzog F, Beck M, Aebersold R. Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Molecular & Cellular Proteomics*. 2010; 9:1634–1649. [PubMed: 20360032]
- Loo JA. Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrometry Reviews*. 1997; 16:1–23. [PubMed: 9414489]
- Lowe J, Stock D, Jap B, Zwickl P, Baumeister W, Huber R. Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution. *Science*. 1995; 268:533–539. [PubMed: 7725097]
- Luan B, Huang X, Wu J, Mei Z, Wang Y, Xue X, Yan C, Wang J, Finley DJ, Shi Y, Wang F. Structure of an endogenous yeast 26S proteasome reveals two major conformational states. *Proceedings of the National Academy of Sciences*. 2016; 113:2642–2647.

- Mehmood S, Allison TM, Robinson CV. Mass spectrometry of protein complexes: From origins to applications. *Annual Review of Physical Chemistry*. 2015; 66:453–474.
- Moscovitz O, Ben-Nissan G, Fainer I, Pollack D, Mizrahi L, Sharon M. The Parkinson's-associated protein DJ-1 regulates the 20S proteasome. *Nature Communications*. 2015; 6:6609 1-13.
- Moscovitz O, Tsvetkov P, Hazan N, Michalevski I, Keisar H, Ben-Nissan G, Shaul Y, Sharon M. A mutually inhibitory feedback loop between the 20S proteasome and its regulator, NQO1. *Molecular Cell*. 2012; 47:76–86. [PubMed: 22793692]
- park S, Roelofs J, Kim W, Robert J, Schmidt M, Gygi SP, Finley D. Hexameric assembly of the proteasomal ATPases is templated through their C termini. *Nature*. 2009; 459:866–870. [PubMed: 19412160]
- Peters J-M. The anaphase-promoting complex: Proteolysis in mitosis and beyond. *Molecular Cell*. 2002; 9:931–943. [PubMed: 12049731]
- Pickart CM, Fushman D. Polyubiquitin chains: Polymeric protein signals. *Current Opinion in Chemical Biology*. 2004; 8:610–616. [PubMed: 15556404]
- Pickering AM, Davies KJA. Degradation of damaged proteins: The main function of the 20S proteasome. *Progress in Molecular Biology and Translational Science*. 2012; 109:227–248. [PubMed: 22727423]
- Rappsilber J, Siniosoglou S, Hurt EC, Mann M. A generic strategy to analyze the spatial organization of multi-protein complexes by cross-linking and mass spectrometry. *Analytical Chemistry*. 2000; 72:267–275. [PubMed: 10658319]
- Rappsilber J. The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *Journal of Structural Biology*. 2011; 173:530–540. [PubMed: 21029779]
- Reyes-Turcu FE, Ventii KH, Wilkinson KD. Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes. *Annual review of Biochemistry*. 2009; 78:363–397.
- Roelofs J, Park S, Haas W, Tian G, Mcallister FE, Huo Y, Lee B-H, Zhang F, Shi Y, Gygi SP, Finley D. Chaperone-mediated pathway of proteasome regulatory particle assembly. *Nature*. 2009; 459:861–865. [PubMed: 19412159]
- Rose RJ, Damoc E, Denisov E, Makarov A, Heck AJR. High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. *Nature Methods*. 2012; 9:1084–1086. [PubMed: 23064518]
- Saeki Y, Toh-E A, Kudo T, Kawamura H, Tanaka K. Multiple proteasome-interacting proteins assist the assembly of the yeast 19S regulatory particle. *Cell*. 2009; 137:900–913. [PubMed: 19446323]
- Sakata E, Stengel F, Fukunaga K, Zhou M, Saeki Y, Förster F, Baumeister W, Tanaka K, Robinson CV. The catalytic activity of Ubp6 enhances maturation of the proteasomal regulatory particle. *Molecular Cell*. 2011; 42:637–649. [PubMed: 21658604]
- Schmidt M, Haas W, Crosas B, Santamaria PG, Gygi SP, Walz T, Finley D. The HEAT repeat protein Blm10 regulates the yeast proteasome by capping the core particle. *Nature Structural & Molecular Biology*. 2005; 12:294–303.
- Schwartz AL, Ciechanover A. Targeting proteins for destruction by the ubiquitin system: Implications for human pathobiology. *Annual Review of Pharmacology and Toxicology*. 2009; 49:73–96.
- Seifert U, Bialy LP, Ebstein F, Bech-Otschir D, Voigt A, Schröter F, Prozorovski T, Lange N, Steffen J, Rieger M, Kuckelkorn U, et al. Immunoproteasomes preserve protein homeostasis upon interferon-induced oxidative stress. *Cell*. 2010; 142:613–624. [PubMed: 20723761]
- Sharon M. Structural MS pulls its weight. *Science*. 2013; 340:1059–1060. [PubMed: 23723227]
- Sharon M, Robinson CV. The role of mass spectrometry in structure elucidation of dynamic protein complexes. *Annual Review of Biochemistry*. 2007; 76:167–193.
- Sharon M, Taverner T, Ambroggio XI, Deshaies RJ, Robinson CV. Structural organization of the 19S proteasome lid: Insights from MS of intact complexes. *PLoS Biology*. 2006; 4:1314–1323.
- Shibatani T, Carlson EJ, Larabee F, McCormack AL, Früh K, Skach WR. Global organization and function of mammalian cytosolic proteasome pools: Implications for PA28 and 19S regulatory complexes. *Molecular Biology of the Cell*. 2006; 17:4962–4971. [PubMed: 16987959]
- Silva JC, Gorenstein MV, Li G-Z, Vissers JPC, Geromanos SJ. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Molecular & Cellular Proteomics*. 2006; 5:144–156. [PubMed: 16219938]

- Sinz A, Sharon M, Chorev D, Arlt C. Chemical cross-linking and native mass spectrometry: A fruitful combination for structural biology. *Protein Science*. 2015; 24:1193–1209. [PubMed: 25970732]
- Snijder J, Heck AJR. Analytical approaches for size and mass analysis of large protein assemblies. *Annual Review of Analytical Chemistry*. 2014; 7:43–64.
- Stefl S, Nishi H, Petukh M, Panchenko AR, Alexov E. Molecular mechanisms of disease-causing missense mutations. *Journal of Molecular Biology*. 2013; 425:3919–3936. [PubMed: 23871686]
- Sutherland BW, Toews J, Kast J. Utility of formaldehyde cross-linking and mass spectrometry in the study of protein–protein interactions. *Journal of Mass Spectrometry*. 2008; 43:699–715. [PubMed: 18438963]
- Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proceedings of the National Academy of Sciences*. 2004; 101:9528–9533.
- Tomko RJ, Funakoshi M, Schneider K, Wang J, Hochstrasser M. Heterohexameric ring arrangement of the eukaryotic proteasomal ATPases: Implications for proteasome structure and assembly. *Molecular Cell*. 2010; 38:393–403. [PubMed: 20471945]
- Tomko RJ, Hochstrasser M. Incorporation of the Rpn12 subunit couples completion of proteasome regulatory particle lid assembly to lid-base joining. *Molecular Cell*. 2011; 44:907–917. [PubMed: 22195964]
- Tomko RJ, Taylor DW, Chen ZA, Wang H-W, Rappsilber J, Hochstrasser M. A single α helix drives extensive remodeling of the proteasome lid and completion of regulatory particle assembly. *Cell*. 2015; 163:432–444. [PubMed: 26451487]
- Unno M, Mizushima T, Morimoto Y, Tomisugi Y, Tanaka K, Yasuoka N, Tsukihara T. The structure of the mammalian 20S proteasome at 2.75 Å resolution. *Structure*. 2002; 10:609–618. [PubMed: 12015144]
- Ustrell V, Hoffman L, Pratt G, Rechsteiner M. PA200, a nuclear proteasome activator involved in DNA repair. *The EMBO Journal*. 2002; 21:3516–3525. [PubMed: 12093752]
- Vandermarliere E, Mueller M, Martens L. Getting intimate with trypsin, the leading protease in proteomics. *Mass Spectrometry Reviews*. 2013; 32:453–465. [PubMed: 23775586]
- Verma R, Aravind L, Oania R, McDonald WH, Yates JR, Koonin EV, Deshaies RJ. Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. *Science*. 2002; 298:611–615. [PubMed: 12183636]
- Verma R, Oania R, Graumann J, Deshaies RJ. Multiubiquitin chain receptors define a layer of substrate selectivity in the ubiquitin-proteasome system. *Cell*. 2004; 118:99–110. [PubMed: 15242647]
- Wilm M, Mann M. Analytical properties of the nanoelectrospray ion source. *Analytical Chemistry*. 1996; 68:1–8. [PubMed: 8779426]
- Xu P, Duong DM, Seyfried NT, Cheng D, Xie Y, Robert J, Rush J, Hochstrasser M, Finley D, Peng J. Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation. *Cell*. 2009; 137:133–145. [PubMed: 19345192]
- Yu C, Yang Y, Wang X, Guan S, Fang L, Liu F, Walters KJ, Kaiser P, Huang L. Characterization of dynamic UbR-proteasome subcomplexes by in vivo cross-linking (X) assisted bimolecular tandem affinity purification (XBAP) and label-free quantitation. *Molecular & Cellular Proteomics*. 2016; 15:2279–2292. [PubMed: 27114451]
- Yu Z, Livnat-Levanon N, Kleifeld O, Mansour W, Nakasone MA, Castañeda CA, Dixon EK, Fushman D, Reis N, Pick E, Glickman MH. Base-CP proteasome can serve as a platform for stepwise lid formation. *Bioscience Reports*. 2015; 35:1–14.
- Zenobi R, Knochenmuss R. Ion formation in MALDI mass spectrometry. *Mass Spectrometry Reviews*. 1998; 17:337–366.
- Zhou M, Wysocki VH. Surface induced dissociation: Dissecting noncovalent protein complexes in the gas phase. *Accounts of Chemical Research*. 2014; 47:1010–1018. [PubMed: 24524650]
- Zubarev RA, Kelleher NL. Electron capture dissociation of multiply charged protein cations. A nonergodic process. *Journal of the American Chemical Society*. 1998; 120:3265–3266.

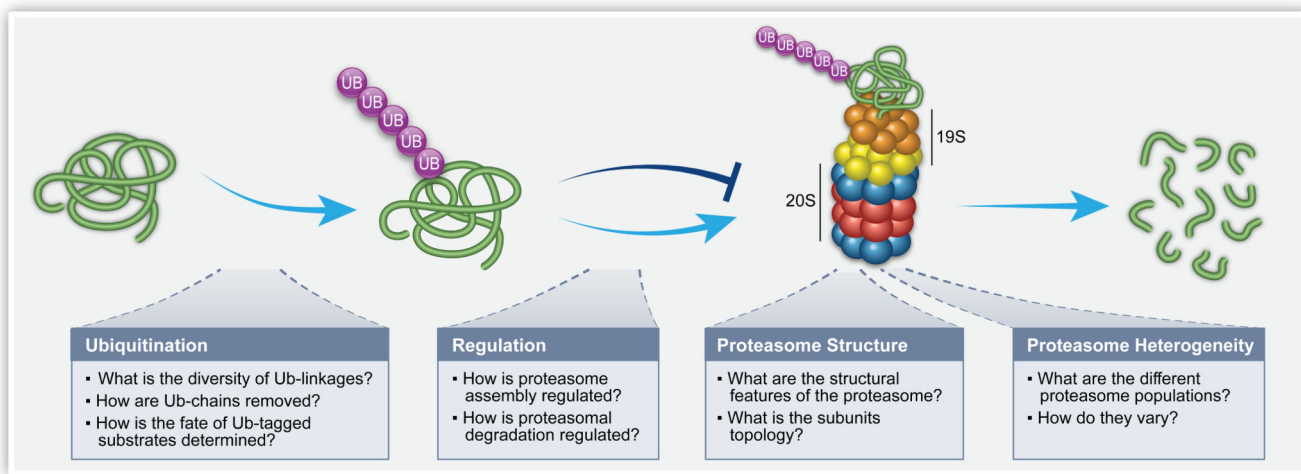
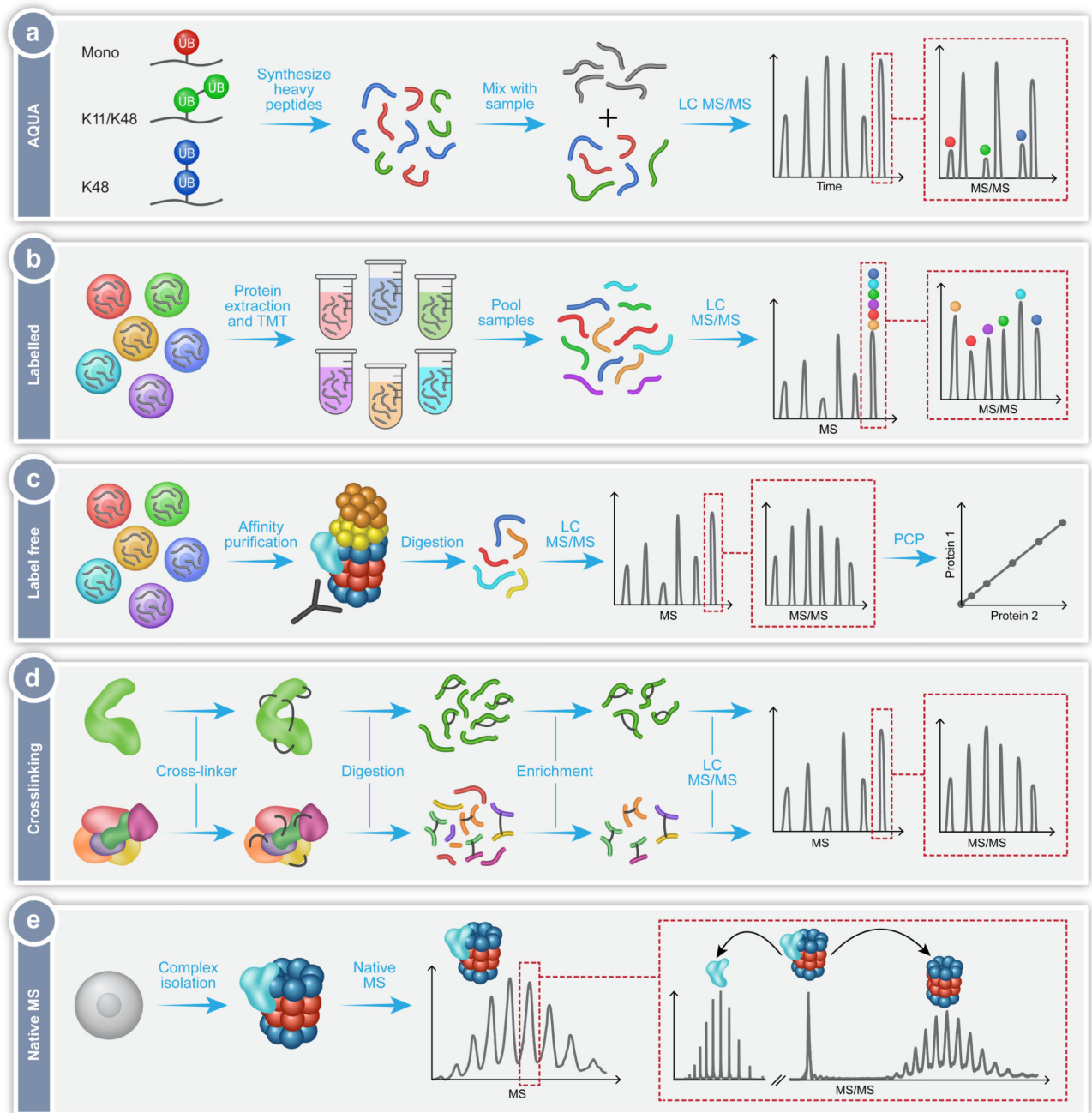


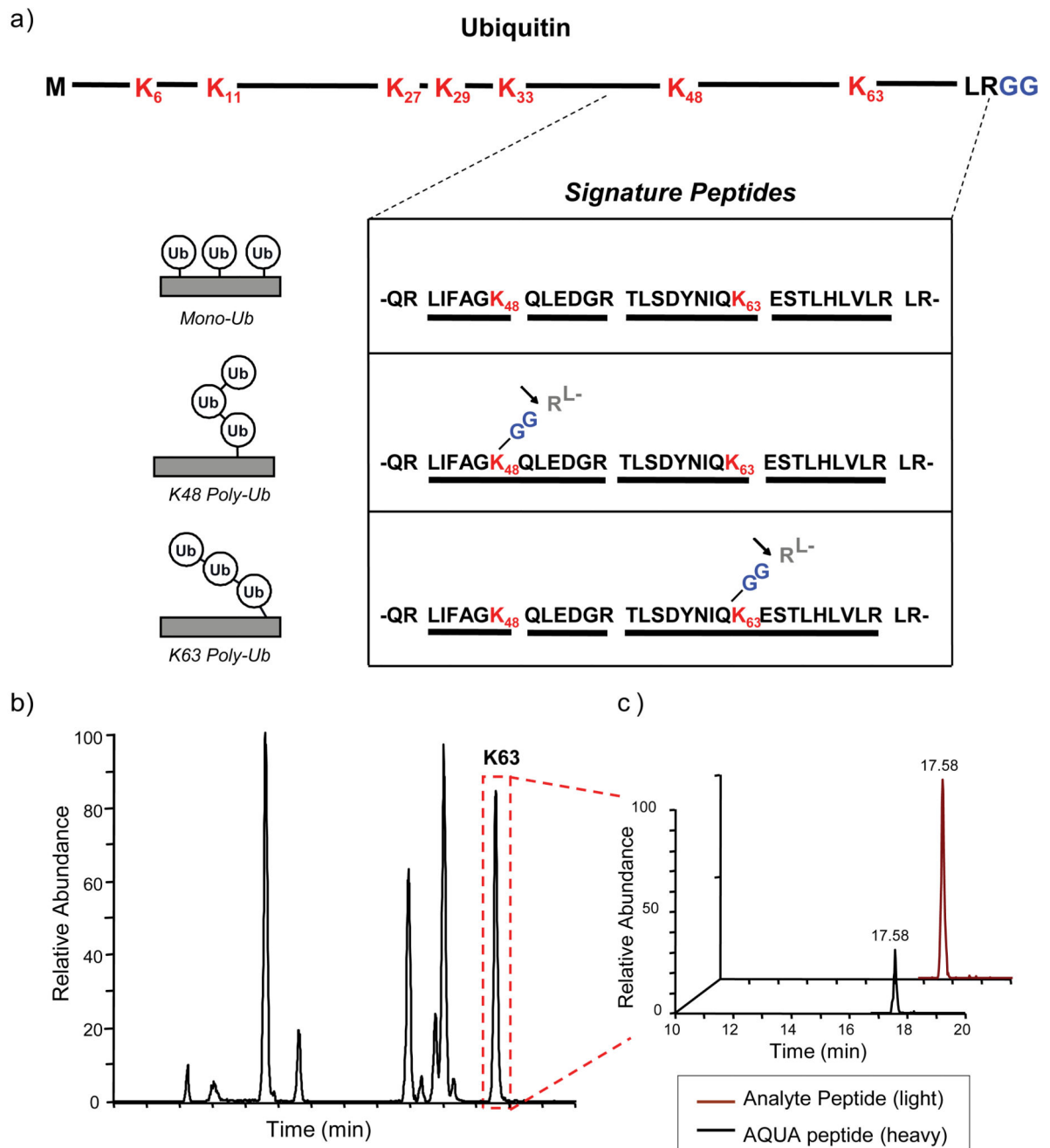
Fig. 1. Overview of the 26S proteasome protein degradation pathway. Proteins are tagged multiple times with Ub, leading to the formation of covalently attached poly-Ub chains on the substrate. These tagged substrates are then directed to the 26S proteasome, where they are recognized, deubiquitinated and unfolded by the 19S RP, followed by translocation to the 20S CP where they are degraded into peptides. For each of these stages the biological questions that were addressed using MS techniques are highlighted.

**Fig. 2.**

MS workflows utilized in the exploration of the proteasomal protein degradation pathway.

(a) Absolute quantification of ubiquitination using the AQUA technique, in which synthetic heavy peptides corresponding to the types of Ub-linkages are produced, for example mono-Ub (red), heterogeneous poly-Ub containing K11 and K48 linkages (green), or homotypic poly-Ub containing only K48 linkages (blue). Known quantities of these peptides are then combined with the peptides from the sample under investigation (gray) and analyzed by LC-MS/MS. The synthetic heavy peptide and light analyte peptide will have the same retention

time during chromatography, thus eluting in the same peak. The peaks can be individually selected and subject to MS/MS, with the mass difference between the light and heavy peptides now distinguishable. Consequently, the ratio between the native peptide and the isotope labeled AQUA standard peptide enables absolute quantification. (b) Labelling of protein extracts using Tandem Mass Tagging (TMT). Proteins are extracted from several cell lines under investigation, followed by the differential labelling of each extract with a unique isobaric tag. The samples are digested and the produced peptides from each sample pooled together. This mixture is then analyzed by LC-MS/MS with the same peptide from each sample falling under the same m/z peak. Selection of this peak followed by MS/MS leads to fragmentation and separation of the differently tagged peptides. The relative peptide levels from each sample can then be quantified and compared. (c) Label free quantification of isolated protein complexes. In this method the different samples are analyzed separately and compared. Often, the protein complex of interest is extracted from a cell lysate using affinity purification. The complex is then digested and subjected to LC-MS/MS allowing identification of the components of the complex and any modifications they contain. By measuring the intensity or ion counts of the precursor ions and comparing this value across samples, relative quantification is achieved. Protein correlation profiling (PCP) can then be performed to determine the composition of the protein complexes under investigation. (d) Crosslinking coupled MS of proteins (top) and protein complexes (bottom). The addition of crosslinkers to the proteins of interest will cause the formation of covalent bonds between the crosslinker and two amino acids (black lines). These can be within the same protein or between two different proteins. The crosslinked samples are then digested to produce peptides, some of which will contain the crosslink. Uncrosslinked peptides can be removed during an enrichment step. The crosslinked peptides are then subjected to LC-MS/MS, with specific peak selection and fragmentation allowing the identification of the location of the crosslink, and consequently the determination of the proximity of the amino acids involved. (e) Native MS of intact protein complexes. Whole protein complexes are isolated, purified and subjected to MS analysis. The intact mass of the complex reflects the homogeneity and stability of the sample as well as its subunit stoichiometry. Isolation of peaks followed by MS/MS fragmentation leads to the dissociation of distinct associating proteins or subunits, which can then be individually identified. This step also enables differentiation between core and peripheral subunits as well as revealing protein-protein connectivities.

**Fig. 3.**

Quantification of Ub-linkages using AQUA and LC-MS/MS. (a) Ubiquitin contains seven lysine (K) residues in its primary sequence (highlighted in red) and a C-terminal diglycine (GG) motif (highlighted in blue). Tryptic digestion will cleave after K and R residues, producing unique peptides for mono-Ub (top), K48 poly-Ub (middle) and K63 poly-Ub (bottom) linkages, as well as after the R at the C-terminus of Ub (indicated by arrow), leading to the identifiable –GG motif on K48 and K63 linked peptides. These signature peptides are isotopically labelled and added to a sample at known concentrations, followed

by analysis by LC-MS/MS. (b) A representative analysis of a peptide sample containing free K63-linked poly-Ub chains, showing the retention time of K63 linked Ub peptides (boxed in red) during LC-MS. Both the heavy signature peptides and light analyte peptides have the same retention time. (c) SRM (MS/MS) analysis of the K63 peak showing the mass difference between the heavy AQUA peptide and the light analyte peptide. Due to the known concentration of the AQUA peptide added to the sample, the amount of analyte peptide can be quantified. Reproduced and adapted with permission from (Kirkpatrick *et al.*, 2006).

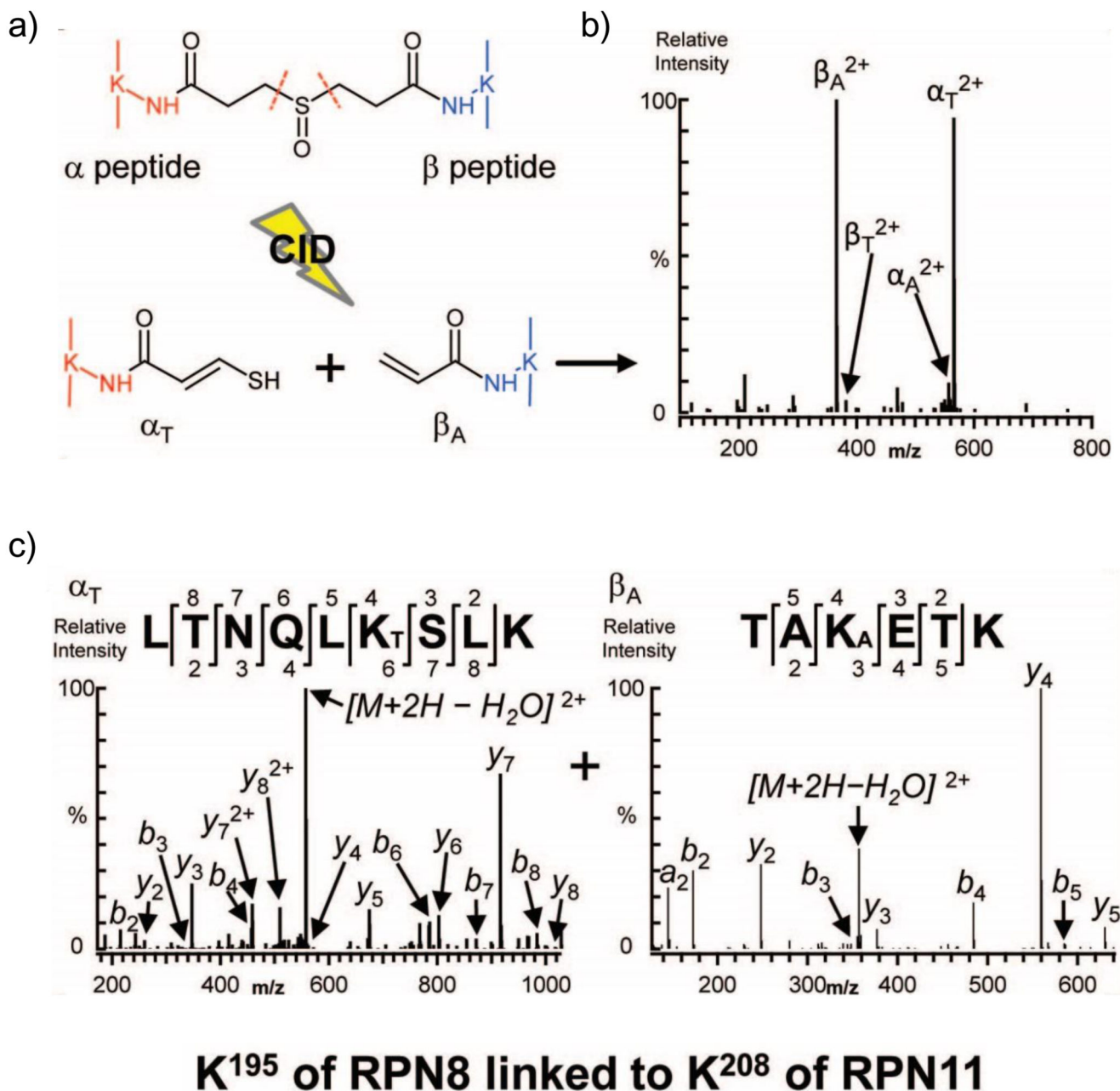
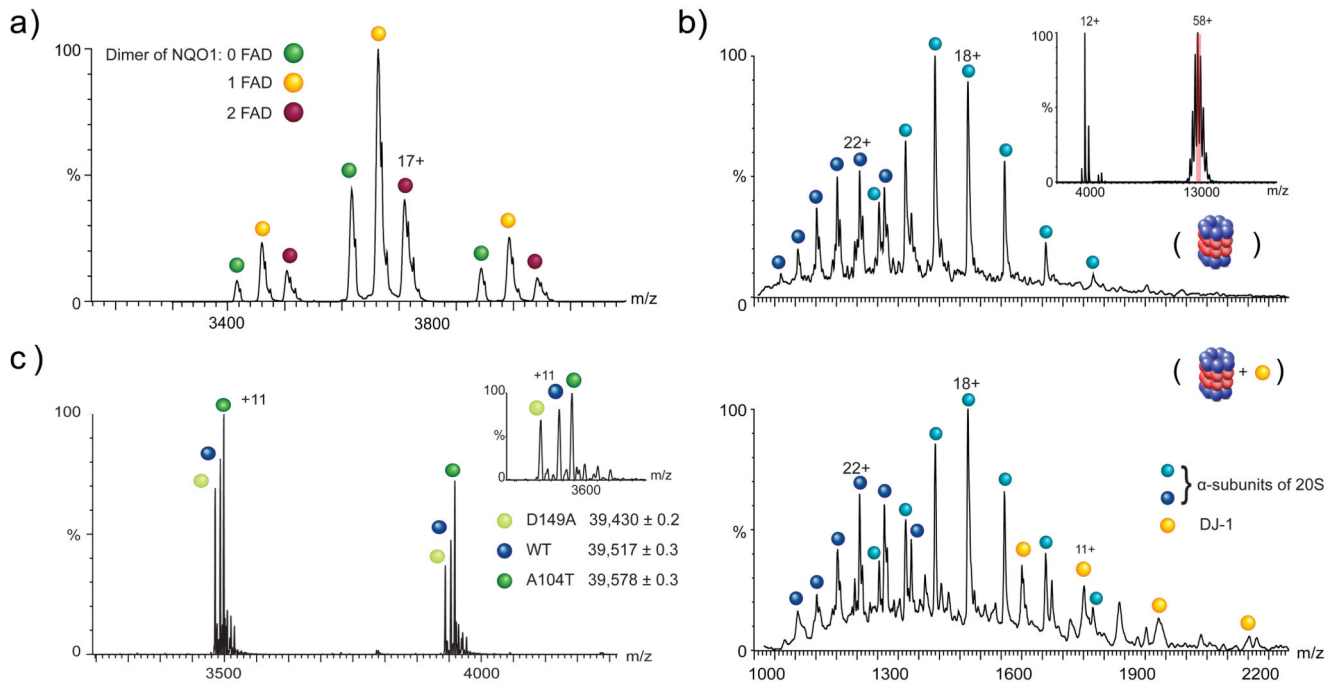


Fig. 4.

Use of the MS-cleavable crosslinker DSSO to identify a crosslinked peptide between Rpn8 and Rpn11. (a) Structure of DSSO demonstrating the covalent bond formed between each reactive group and K residues from two different proteins (orange and blue). The C-S bonds in the spacer arm which can break during CID are highlighted (dashed orange line). The generated peptide fragments after CID are shown, demonstrating the presence of a thiol (-SH) group on the α -peptide (left) and an alkene (=C) on the β -peptide (right). (b) The MS/MS fragmentation pattern of the produced peptides. The presence of the thiol on the α -peptide (denoted α_T^{2+}) and the alkene on the β -peptide (denoted β_A^{2+}) is distinguishable

due to their characteristic mass relationship (+86 Da and +54 Da respectively). (c) Sequence determination of each crosslinked peptide in (b) by MS/MS/MS, unambiguously defining the site of the crosslink as K195 of Rpn8 to K208 of Rpn11. Reproduced and adapted with permission from (Kao *et al.*, 2012).

**Fig. 5.**

Native MS spectra of 20S proteasome regulators. (a) MS spectra of NQO1 confirming the presence of a homodimer. Each of the charge states is split into three reflecting the presence of three different homodimer populations, namely apoNQO1 or the NQO1 bound to one or two FAD molecules. (b) MS/MS spectra of 20S proteasome alone (top), showing the identification of individual α -subunits of the 20S (blue balls) that have been ejected from the intact 20S proteasome during the CID process. Inset shows the mass spectrum of the intact 20S proteasome centered at 13,000 m/z , and the peak selected for further MS/MS analysis is highlighted in red. Comparison with the 20S in the presence of DJ-1 (bottom) reveals additional peaks that correspond in mass to the monomeric form of DJ-1 (yellow balls). By extrapolation, we can therefore conclude that before MS/MS analysis, DJ-1 binds to the 20S proteasome. (c) The high resolving power of the Orbitrap mass spectrometer can distinguish between the small mass differences of the Parkinson's disease missense mutants of DJ-1. Compared with the WT DJ-1, the D149A and A104T DJ-1 mutants differ by -30 Da and +44 Da respectively. Inset shows the clear separation of the peaks in a single charge state. Reproduced and adapted with permission from (Ben-Nissan *et al.*, 2016; Moscovitz *et al.*, 2012; Moscovitz *et al.*, 2015).