



CHAPTER 17 Proteomics and proteoforms: Bottom-up or top-down, how to use high-resolution mass spectrometry to reach the Grail

J. Vinh

► To cite this version:

J. Vinh. CHAPTER 17 Proteomics and proteoforms: Bottom-up or top-down, how to use high-resolution mass spectrometry to reach the Grail. B. Kanawati & P. Schmitt-Kopplin. Fundamentals and Applications of Fourier Transform Mass Spectrometry, Elsevier, pp.529-567, 2019, 9780128140130. 10.1016/B978-0-12-814013-0.00017-X . hal-02358384

HAL Id: hal-02358384

<https://hal.science/hal-02358384>

Submitted on 24 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CHAPTER 17

Proteomics and proteoforms: Bottom-up or top-down, how to use high-resolution mass spectrometry to reach the Grail

Joelle Vinh

Laboratory of Biological Mass Spectrometry and Proteomics, CNRS USR 3149/FRE2032, ESPCI Paris, PSL University, Paris, France
TGE FT-ICR CNRS, France

1	Introduction: the importance of proteins in a multi-omics context	1
2	Seeking for high-performance mass spectrometry.....	2
2.1	Analytical strategies: from top-down to bottom-up and vice versa	2
2.2	Sequencing and activation methods.....	4
3	Bioinformatics	7
4	Bottom-up proteomics: analysis of peptides	8
4.1	Separative techniques.....	8
4.2	LC-MS hyphenation.....	8
4.3	Scan modes, targeted analysis, and data-independent analysis.....	9
4.4	Quantification	10
5	Top-down proteomics.....	11
5.1	Purification and separation of proteins.....	12
5.2	Specificity of MS analysis	13
6	Epilogue	15
7	Acknowledgments	15
8	References	15

1 Introduction: the importance of proteins in a multi-omics context

Proteins are key players in most biological processes. With a well-defined subcellular localization, their function is mainly correlated with their molecular structure. This specific function results from a complex combination of information from the genome that is a static entity and information from the transcriptome and the proteome, which are more dynamic, and from the physicochemical environment. The term Proteome appeared in 1995 [1] to designate the total protein complement of a genome [2]. The term was immediately widely taken up in the literature [3,4]. Even if the study of the proteome had begun more than 20 years ago with bidimensional gel electrophoresis (2D GE), it took off thanks to mass spectrometry (MS), and namely with the development of Matrix-Assisted Laser Desorption Ionization (MALDI) and Electrospray Ionization (ESI) ion sources in the mid-80s [5–7]. The use of MS has certainly led to the exponential progression of polypeptide and small protein studies [8]. The full power of ESI-MS approaches for proteomics has been achieved with the miniaturization of the ionization source in nanoESI [9–11]. Rapidly multi-omics approaches have been developed in an attempt to correlate the observed phenotype and the structure of macro-biomolecules in presence, particularly in the context of biomedical applications. The concept of Proteomics was introduced two years after the term Proteome to introduce more quantitative and dynamic proteomic approaches in order to characterize biological processes (e.g., disease processes and drug effects) and deciphering the mechanisms of gene expression control [12]. MS combined with genomic sequence information had become an indispensable tool for tracking the abundance of a protein on a given biological model correlated with a specific functional pathway [13,14]. Furthermore, since MS is not an inherently quantitative mode of detection, several approaches have been gradually designed for quantitative or semi-quantitative analysis. However, the mere identification of the gene expression products represents partial information since the expressed proteins are matured and undergo a series of variable modifications to generate the final functional proteoforms. To understand the functional diversity of proteoforms, it is essential to characterize the molecular structure. Some modifications may reverse the function of a protein, as is the case for glycosylation and phosphorylation. In this perspective, the need to study proteins' molecular diversity is essential to identify the different molecular forms of a protein encoded by a gene and not just the protein [15]. Describing protein' functions and understanding their dysfunction is a major challenge for environmental and biomedical life sciences.

Two decades later, proteomics analysis reached a stage where the quantitative analysis of functional proteoforms [16], as well as the identification of each modification site, is a major objective [17,18]. Each variant of a given primary sequence is the result of many correlated phenomena. These include point mutations on the primary sequence, alternative splicing, covalent post-translational modifications (PTMs) of the primary sequence, or even the phenomenon of proteolysis during maturation [19]. Therefore, the number of proteoforms is a combinatorial calculation of these PTMs, which explains the considerable number of functional structures that result: from the 20,000 coded human protein sequences [20], several million proteoforms are predicted *in silico* [18].

2 Seeking for high-performance mass spectrometry

2.1 Analytical strategies: from top-down to bottom-up and vice versa

From the beginning, proteome analysis has been described as a high-throughput analysis based on the model of genome studies aimed at characterizing the protein composition of a biological sample in the most comprehensive way. The most common strategy is to analyze the protein content of a mixture from the associated proteolytic peptides, and this strategy has already been the subject of many reviews (see, for example, the excellent review of the group of Yates [21]). Historically, from protein separations in 2D GE and after staining for protein detection, thousands of resolved protein-containing spots were then processed in order to analyze the proteolytic peptides generated [22]. The MS provided the peptide mass fingerprint (PMF) of the protein (the list of monoisotopic masses of proteolytic peptides) that was used to search databases for the entire protein sequence that would best match the experimental masses [23]. Very quickly, it appeared that the PMF does not allow to differentiate homologous sequences and leads to false-positive results as soon as several proteins co-migrate on the gel, which is very often the case [24]. Tandem mass spectrometry (MS/MS) has been introduced to solve this problem [25]. Collision Induced Dissociation (CID) MS/MS spectra of tryptic peptides (that are mainly doubly and triply charged in ESI and have a basic residue in the C-terminal position) allow for easy sequencing of peptides [26]. Singly charged peptides can also be sequenced, even if this requires higher collisional energy processes. Gradually, several other MS/MS modes were introduced. The analysis of complex biological samples required the coupling of other separation techniques to reduce the effect of signal suppression, to increase the dynamic range, and to record MS/MS acquisitions on a larger number of species. This coupling also allowed to free itself from 2D GE as a prerequisite. The sensitivity of the whole process has been improved by miniaturization (from the proteolytic sample to the gas-phase ions, with the nanoLC-ESI coupling [27–30]) and by technological optimization (from the ion production to signal-to-noise ratios), with new ionic optics [31,32] and the use of electronic ion-trapping systems instead of scanning ion detection [33], such as linear ion traps, orbitraps or ICR cells. In parallel, the link between mass spectrometry data (peptide mass and sequence elements) and the genome and proteome databases of a given organism has been automated by developing specific bioinformatics tools that are continually evolving to follow the technological evolution [34].

Instead of studying proteolytic peptides to identify the proteins in a bottom-up approach, why not directly study proteins in a top-down-approach [35,36]? The advantage of the bottom-up approach is the greater ease of analysis of peptides and polypeptides of molecular masses below 4 kDa (separation efficiency, sensitivity, and resolution). Of course, the analysis of proteolytic species offers a partial and fragmentary view of the proteins in the sample. The idea is not new [37], but the implementation of a combination of bottom-up, top-down, and middle-down (using limited proteolysis) strategies always remains an analytical challenge [38]. Despite this, the LC-MS analysis of the intact protein is currently considered as a complementary and essential analytical approach [17,39]. Indeed, our goal is the characterization of functional proteoforms. Generally, in bottom-up, the specific characterization of a proteoform of a protein is a challenge (or even utopia) because most of their peptides are shared between several proteoforms. Moreover, the most detailed structural characterization of proteolytic peptides does not provide an overall structural characterization of the associated protein because full sequence coverage is rarely achieved. To solve this problem, it is necessary to use a top-down strategy in which all intact proteoforms are analyzed in LC-MS/MS without proteolysis, with up to now lower proteome coverage but higher protein sequence coverage [40]. In this case, each whole proteoform with its PTMs is detected, which nevertheless does not raise all interpretation ambiguities for high molecular mass or strongly modified overlapping proteins. Moreover, there is always a lack of a technological rupture to separate our high-resolution protein species over a wide dynamic range for high precision and high sensitivity detection. In addition, sequence analysis requires an efficient fragmentation mode on long sequences to identify the modification sites. However, regardless of the strategy used (bottom-up versus top-down), high accuracy and precision and high resolution like those offered by the FT MS configurations are needed to generate robust results [41]. A straightforward example illustrates this point in Fig. 1 and demonstrates that high resolution only can provide unambiguous results. Lower resolution provides partial coverage of the protein samples.

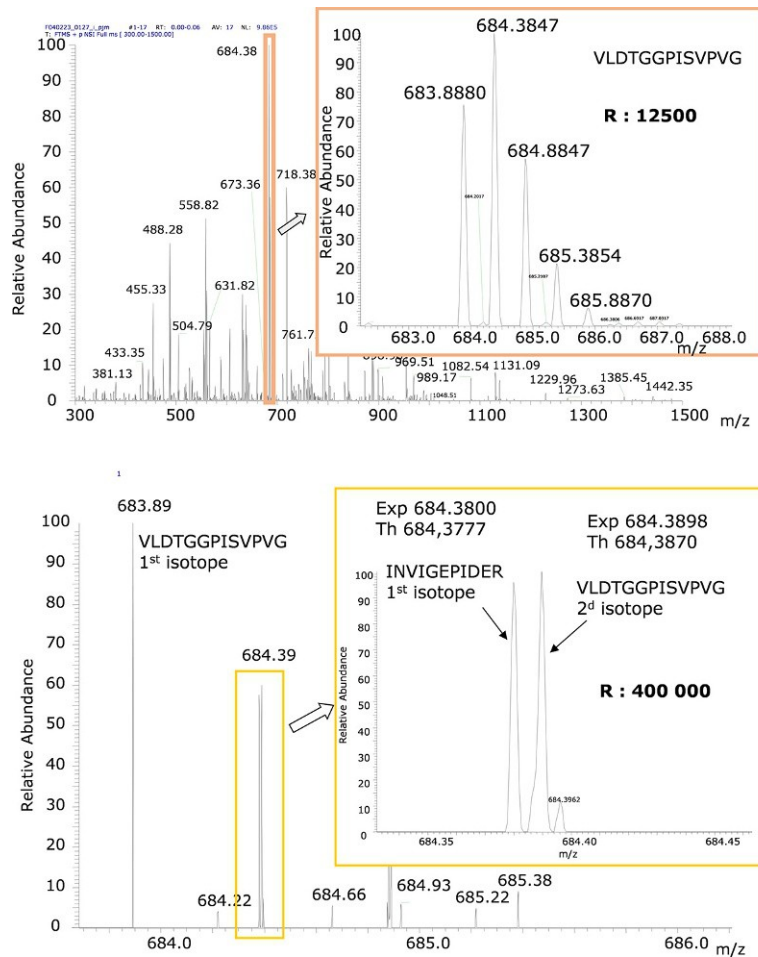


Fig. 1 High-resolution impact on bottom-up proteomics on a simple example, the tryptic digest of mitochondrial ATP synthase subunit beta from *Saccharomyces cerevisiae*. Box A: full scan MS of the tryptic digest in FT-ICR MS (LTQ-FT Thermo Fisher, resolution set at 12,500 at m/z 400), with a zoom on the most intense species. Insert of Box A shows that the isotopic profile is not correct for a peptide at m/z 1365.7 Da, because the first isotopic peak should be the most intense. Box B: a zoomed view of the same species recorded in FT-ICR MS at resolution 400,000 at m/z 400. The isotopic profile is correct, and the first isotope corresponds to peptide ¹²⁹INVIGEPIDER¹³⁹. Insert of Box B shows that the intensity of the second isotopic peak decreased because it is split into two species, one is the second isotope of the previous peptide, and the second one is the first isotope of ¹⁰⁹VLDTGGPISVPVG¹²¹. Even though the peptide masses differ from each other 1 Da only, only high-resolution FT-ICR MS can separate them by resolving the full isotopic pattern.

Since the very beginning, the search for very high technical specifications has been the golden standard for protein MS analysis. This probably explains why FT-ICR MS was from the beginning coupled to ESI sources for protein analysis [42]. Currently, FTMS approaches' limitation most certainly lies in their low acquisition rate, which is especially critical in MS/MS mode. This is a problem to work on the chromatographic/electrophoretic scale, which is now unavoidable [43]. The following will illustrate the recent evolutions that have been carried out to better characterize the proteoforms from complex biological mixtures, with some recent works that illustrate the developments realized in this direction without pretending to be exhaustive.

2.2 Sequencing and activation methods

Sequencing of peptides and proteins is based on MS/MS fragmentation of these macro-biomolecules. Depending on the activation methods used for fragmentation, the most labile bonds vary, and some of them are more suitable to evidence the PTMs, but all are based on the principle that the peptide chains fragment most of the time by preferentially disrupting the amide bonds between the amino acid residues. The ion activation methods for peptide and protein sequencing in general have already been reviewed [44]. Most are available with an FT-MS detection, such as CID, HCD, EDC, ETD, IRMPD, UVPD (those acronyms will be introduced in the following). The most frequent fragments were quickly nomenclatured to facilitate the reading of the sequences [45,46]. We call them a and x, b and y, or c and z fragments ions. Among the activation methods, the collision-induced dissociation (CID) or collisionally activated dissociation (CAD) is a low energy dissociation process and is the routine method on all ESI source instruments (Q-TOF, triple quadrupole, ion traps, FTICR, etc.). Although CID has long remained the reference method for peptide sequencing, this fragmentation mode suffers from a specific limitation in the case of ion traps, namely the impossibility of effectively trapping radio frequency signals for the entire range of useful masses of the fragment ions during the activation of the precursor ion. This induced cutting the low-mass range corresponding to about 1/3 of the precursor's mass, i.e., the mass range of the immonium ions for the peptides and the reporter ions for the isotopic labeling. This limitation has become a significant drawback with the ramp-up of very sensitive ion trap hybrid spectrometers coupled to high resolution and accuracy FT analyzer. The higher energy collisional activation dissociation (HCD) has been developed to overcome this problem in ion traps, for instance in hybrids linear quadrupole trap-orbitrap (LQT-OT), which are extremely common for bottom-up analysis. The activation is carried out analogously to what is done in the triple quadrupoles, but with detection in high-resolution FT-MS. This was done first using the C-trap, which usually allows storing the ions coming out from the ion trap and before their introduction to the orbitrap, then quickly using an additional multipole with a first application exploiting the detection of immonium ions from phospho-tyrosine [47]. Although the obtained MS/MS spectra are very similar to those obtained in the CID mode, the fragmentation profiles are slightly different in HCD [48], as are the fragmentation profiles of an ion trap vs. a hybrid Quadrupole-Time Of Flight. Indeed, HCD mode provides better identification of triply charged peptide ions, a sharp reduction of N-terminal fragments of type b, with singly charged fragments over the entire mass range, which can be very useful for targeted analyses. This HCD mode is now the exclusive fragmentation mode of the latest hybrid family Quadrupole-Orbitrap (Q-OT) [49], apart from the previously published in-source fragmentation on a quadrupole-free configuration [50]. In the latter study, the authors showed that without selecting a precursor ion, it was possible to identify 45 proteins among an equimolar mixture of 48 proteins thanks to the high resolution of 100,000 at m/z 200 and the high mass measurement accuracy of the Orbitrap mass analyzer. The efficiency of this configuration was recognized quickly, and the development efforts were such that six years after the first commercial version, we are at the fourth version of commercial Q-OT mass spectrometers, which corresponds to a new version every 18 months on average [51].

The high mass accuracy of FT-MS is a key parameter to discriminate between quasiisobaric PTMs such as trimethylation (+42.011 amu) and acetylation (+42.047 amu) as shown in Fig. 2, where dimethylation and acetylation were simultaneously localized on lysine K124 of the human Histone H3-like centromeric protein A (CENP-A). This was recently confirmed by other studies (data not shown). In CID and HCD, the regularity of fragmentation of the peptide backbone decreases in the presence of PTMs. The fragmentation of the PTMs themselves can be the major fragmentation process. It can be detrimental to the generation of C-terminal and N-terminal fragments of the peptide backbone required for sequencing.

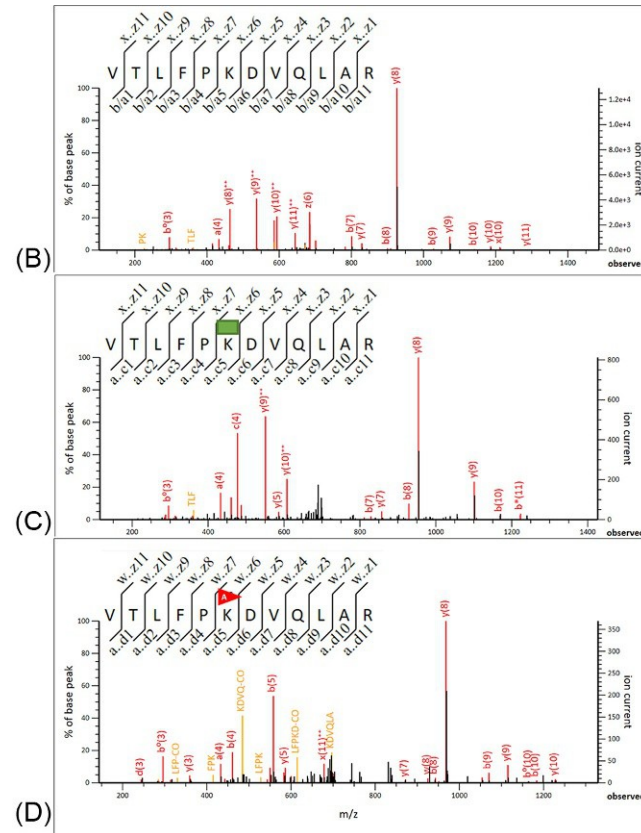
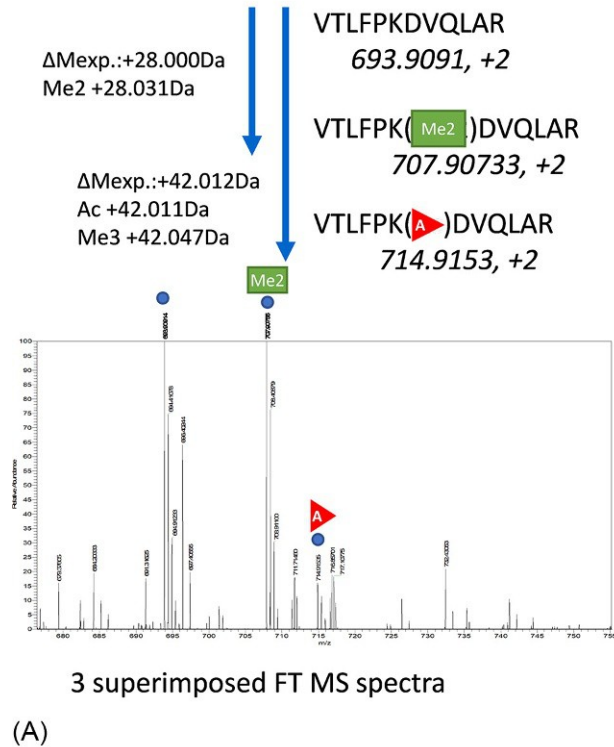


Fig. 2 Evidence for post-translational modifications on lysine K124 of human Histone H3-like centromeric protein A (CENP-A). Peptide $^{119}\text{VTLFPKDVQLAR}^{130}$ is simultaneously modified with a dimethylation and an acetylation on lysine. Accurate mass measurement discriminates acetylation and trimethylation, and MS/MS sequencing localizes the modification on residue K124. Box A: superimposed MS spectra for the three peptides eluted at different retention times. Boxes B, C, and D: MS/MS spectra of the unmodified, dimethylated, and acetylated peptides, respectively (personal communication C. Giangrande, J. Vinh).

Other activation methods have, therefore, been developed. Activation can also be achieved via an electron beam, as in electron capture dissociation (ECD) and electron transfer dissociation (ETD). ECD has been developed in FT-ICR technology in which multi-charged cations pass through a low-energy electron beam in the magnetic field of the ICR cell. The initiation of fragmentation is more random, which allows us to achieve different sequence coverages [52]. ECD has been described as particularly useful for analyzing glycosylated or phosphorylated peptides, for example, or for reducing the disulfide bridges of intact proteins. The main recent developments for peptide analysis concern two-dimensional FTMS approaches, introduced on glycosylated peptides [53]. The first analyses in bottom-up data-independent acquisition (DIA) mode were performed on cytochrome C [54] and subsequently on calmodulin and ubiquitin [55,56]. This very promising approach exploits the possibility of correlating the precursor ion and its fragments via the ion of interest's cyclotron frequency modulation by avoiding any selection step. However, it remains applied to the analysis of reference compounds, well below the complexity of a conventional proteomic mixture. Recently ECD was evaluated on the non-commercial configuration of an orbitrap mass spectrometer [57] to analyze peptides and intact proteins with a sequence coverage > 85%, for example, on ubiquitin and myoglobin without specific optimization according to the polypeptide studied. By intercalating the ECD cell between the triple quadrupole and the C-trap, the hardware modifications are minor and do not require any modification of the control of the mass spectrometer. However, the total acquisition time to map the fragmentation profile of a peptide (for example, the substance P) is 30 s and reaches 1 min for the fragmentation of a protein. In a similar way to the FT-ICRECD, the phosphorylations are maintained on the sequence of a phosphopeptide, and disulfide bridge mapping could be performed on Human Serum Albumin (HSA) after tryptic digestion with the preferential disruption of the bridges to the detriment of the classic peptide fragmentation in fragments c and z. This instrumental configuration can perform HCD, ECD experiments, and their EChcD combination. However, it still requires the accumulation of 5 microscans per scan for the LC-MS/MS coupling at 17,500 mass resolution and is limited to MS/MS data-dependent acquisition (DDA) mode, where MS/MS is performed on the five most intense precursor ions only (Top5 in ECD to compare with Top20 executed routinely in DDA HCD mode).

ETD was developed for ion traps that did not efficiently trap electrons, which are then replaced by anion radicals. [58]. Its optimization has made it possible to achieve a coupling to LC-MS/MS [59]. Comparison of the fragmentation in CID vs. in ETD mode has confirmed the possibility of maintaining intact PTMs on the peptide chain, as it is reported for the phosphorylation [60,61], methylation, acetylation, glycosylation [62], nitrosylation, sulfation but also ubiquitination [63]. The evolution of the ETD was reviewed in particular [64] as soon as it was introduced with Orbitrap technology [65]. More recently, the excellent review of Riley and Coon on the role of ETD in proteomics [66] repositioned the fundamental concepts and the latest developments of this mode of ion activation, which is used in bottom-up and top-down proteomics as a powerful complement to CID/HCD modes. ETD has expanded its range of applications with proteomics (bottom-up and deep proteome, middle down and top-down) but also peptidomics, PTMomics, or structural analysis of proteins. It has been implemented in all commercial orbitrap and FT-ICR configurations but also, of course, in ion traps and TOFs. In the orbitraps, the original configuration with an ETD cell behind the HCD cell has evolved to a larger size cell (high capacity HCD mode) close to the source, and this principle has been taken up on the quadrupoles upstream of the ICR cell ETD (such as ECD). It is well suited for large multicharged polypeptides up to a mass of 30 kDa beyond which fragmentation efficiency falls and for which ETD must be combined with other modes of activation. ETD is of particular interest for non-tryptic peptides that fragment poorly or irregularly in CID/HCD and for which sequencing is therefore problematic. Two main application domains are the analysis of endogenous peptides (peptidomics) and the study of non-tryptic proteolytic peptides (as in middle down). Finally, the possibility of mapping the links between peptide chains (SS or crosslink bridges) is also a field of study in full development [67].

Two major limitations of ECD and ETD remain relevant for their application in proteomics. First of all, the fragmentation efficiency remains low, and the multiplicity of fragmentation sites is important, making the method less sensitive, especially for ECD. The irradiation time remains significant, which reduces the peak capacity of the LC-MS/MS coupling compared to the CID/HCD modes. Within this framework, approaches without a priori but still targeting certain sub-families of peptides thanks to decision trees have been proposed to reduce this analysis flow problem, such as the Autopilot decision algorithm [68]. Autopilot uses the HCD mode by default and switches to the more time-consuming ETD mode only if necessary, or the analysis of the glycosylation only uses the HCD if an oxonium ion is detected from the fragmentation of a proteolytic peptide [69,70].

The infrared multiphoton dissociation (IRMPD) and the ultraviolet photodissociation (UVPD) modes are two activation methods, which use photons. Described for FT-ICR devices [71], IRMPD was very quickly applied to peptides and proteins [72,73]. Despite some early results on protein and peptide sequencing, IRMPD is not used alone for proteomics applications

due to the lack of sequence coverage and the low sensitivity/efficiency. However, recent studies have employed IRMPD as a supplemental activation method in bottom-up [54] or top-down [74] proteomics model studies. Instead of multiple low-energy IR photons, activation can also be performed using single high energy UV photons. Several UV pulsed laser sources can be used, but 193 nm is the most reported for proteomics because of its ease of use and because its wavelength corresponds to one of the amide absorption bands. Even if the first applications at the proteome scale are recent, UVPD has quickly gained popularity thanks to its fragmentation efficiency [75], its low duty cycle time, and its versatility toward PTMomics or proteins complex studies [76]. A study of the underlying fragmentation mechanisms was recently published [77] in order to understand how a very fast (5 ns time pulse) and high energy activation allowed to obtain the sequence coverage of the monomeric Pin1 (29 kDa) protein, and also to dissociate a protein tetramer ion into intact monomer or dimer ions. UVPD is currently the most studied method for top-down proteomics and seems to give promising and complementary results to HCD mode. [78]. Commercial configurations provide a source of UV photons at the HCD cell in orbitrap, but UVPD has also been implemented successfully in an FT-ICR cell [79]. A recent study has compared and combined the effects of photoactivation by IR and UV photons to optimize the study of phospho-, glyco- and sulfo-peptides [80,81].

Activation methods for fragmentation of peptides and proteins vary and offer complementary specificities. The HCD mode is the fastest: its fragmentation efficiency and the very high peak capacity it offers for LC-MS/MS studies make it unavoidable. The ETD and EthcD modes are now used routinely, especially for the study of PTMs on peptides in association with HCD, such as the analysis of peptides O-glycosylated in humans serum by EThcD and HCD [82]. The ECD mode remains more confidential but is very interesting for whole protein analyses, especially when used in conjunction with photodissociation methods. IRMPD was used alone in the past, particularly in FT-ICR, to analyze the PTMs on the polypeptides, but it seems less promising than UVPD. The latter mode is currently the subject of many developments and is of interest for sequencing of whole proteins and peptides on which PTMs are preserved during fragmentation, combined or not with IRMPD. Some structural biology approaches have shown that the UVPD mode may also provide valuable information for the study of protein complexes that dissociate while maintaining the intact proteins that compose it [83]. Its speed makes it an activation mode, which is very readily compatible with LC-MS/MS analyses, and this fact opens the door to perform applications for the study of the “deep proteome” in bottom-up and in top-down investigations [84].

3 Bioinformatics

The more complex the sample and the analysis workflow are, the more complex become the obtained data. A manual interpretation of automated LC-MS/MS acquisitions is not realistic, and a manual validation for only ambiguous results (the sequences inference with very low probabilistic scores) is the rule. Since the first reports on automated protein inferences from MS/MS acquisition [85], many algorithms have been developed and are being further developed. A full review of the available bioinformatics tools is a complete topic that cannot fit in this short review. Bioinformatics tools for bottom-up proteomics are regularly evaluated and compared (see for instance a benchmark of free database search algorithms [86], of protein inference validation tools [87,88], or an example of a new tool to assay the existing data processing tools [89]). The choice of possible tools that can be used is so overcrowded that no one could claim to propose an up-to-date exhaustive view of the available technical solutions [90], even if some recent reviews provide an overview [91]. However, the choice of the acquisition technique and the interpretation algorithm are key parameters for identifying peptides and for protein inference [92].

Top-down analysis generates data that are not directly interpretable with the algorithms developed for the bottom-up approach. The deconvolution of MS spectra, which has been in use for more than 20 years, is still under development, as top-down analysis concerns even more complex mixtures. Most of the algorithms perform well on a restricted mass range, and some can be very resource-demanding such as MaxEnt [93]. Recently a new algorithm has been proposed to address deconvolution when fewer peaks can be identified [94]. Another data mining effort involves protein sequencing and database search. Very few freely available resources exist: one should mention Mash Pro Suite [95], capable of processing high-resolution MS and tandem MS (MS/MS) and performing FASTA database searching, characterization of PTMs, and site mutations, with relative quantitation of multiple proteoforms. The annotation of the identified proteoforms has been standardized with the Proforma [96] to help the scientific exchange of top-down results and unify the interface with the different algorithms. Attempting to draw up a list of the available solutions would be doomed to failure because they evolve very quickly. The ongoing consensus effort by the Consortium for Top-Down Proteomics (CTDP, <http://www.topdownproteomics.org>) is therefore of particular interest.

4 Bottom-up proteomics: analysis of peptides

After a decade when the idea of a global characterization of the proteome was considered an unrealistic goal, advances in mass spectrometry and separation science paved the way for its direction [97]. Bottom-up proteomics has entered the era of “deep proteome” analysis, as reported by Olsen’s group, who recently identified 166,620 unique peptide sequences spanning 11,292 protein groups (associated with 10,284 coding genes) from 46 µg of the tryptic peptide mixture of a reference of HeLa cells analyzed in 2DLC-MS/MS [98]. To achieve this analytical depth, comparable to RNA sequencing, the authors used a hybrid quadrupole-orbitrap (Q-OT) instrument capable of acquiring MS/MS data at a high scanning rate of 20 Hz, so that more than 1,000 peptides per minute can be selected for sequencing.

4.1 Separative techniques

Whatever the performance of the downstream MS, peptide analysis also involves extremely efficient preparation and sample separation upstream. Zongo et al. have worked on the proteome of human glioblastoma: they estimated that an average of 230 proteoforms could be identified in a single 2D GE spot by nanoLC MS/MS on a hybrid LQT-OT (LTQ-Velos type) and estimated that up to 400,000 proteoforms are present on a 2D gel, that is well above the peak capacity of any LC separation [24]. Several groups have worked on the development of resolving, sensitive, accurate, and robust LC-MS couplings [99]. Because many bottom-up proteomics acquisition methods rely on the robustness of the separations, the optimization of the nanoLC-MS coupling is essential [100,101]. Different strategies have recently been proposed, ranging from the derivation of peptides between successive separation dimensions [102], to the combination of different LC principles using multidimensional nanoLC, such as ion exchange (IE) [103,104], Hydrophilic interaction chromatography (HILIC) [105], or 2D high pH reversed-phase/low pH reversed-phase chromatography (hi RP-lo RP) [106,107], or 3D RPLC-MS [108], or using electrophoresis as the capillary zone electrophoresis (CZE)-MS coupling [109] for greater complementarity [110]. A study on more specific stationary phases was proposed for applications without a priori [111,112] or for targeted applications such as phosphoproteomics [113], glycoproteomics [114], detection of angiotensin [115], or detection of amyloid peptide [116]. For example, 2D nanoLC coupling has increased the sensitivity by two orders of magnitude in MS mode to reach a sub-attomolar detection limit [117]. The absolute abundance of unmodified endogenous peptides derived from the human transcription factor MEF2C was estimated at the mean level of 180 amol/µg total cell lysate, corresponding to an average of 11,000 molecules/cell, and the phosphorylation site occupancy was estimated at 26% with 42 molecules/cell, from 1 µg of extracts isolated from 10,000 human cells. This study highlights two limiting factors: the first being the need to improve the ionization efficiency in nanoLC-MS coupling mode, the second is the problem of isolation of quasi-isobaric species that cause the signal-to-noise ratio to fall, which must be treated by an increase of the peak capacity of upstream separative methods.

4.2 LC-MS hyphenation

MS optimization must be performed in agreement with LC conditions. This comprehensive approach has been reviewed recently by Zhu et al. [118]. They demonstrated an increase in signal intensities > 3-fold (corresponding to a 32% increase in peptide identifications) using 30 µm-i.d. columns. compared to the standard 75-µm-i.d. Higher LC resolution also requires a higher sampling rate with a high frequency of MS/MS experiments and a higher efficiency of ion transmission so that less abundant species can be identified. They reported an increase of peptide identifications by a factor of 2.9 and an increase of identified protein groups by a factor of 1.7 using the latest generation of tribrid Q-LQT-OT MS instruments compared to the first one (Lumos vs. LTQ Orbitrap Fusion) due to electrodynamic ion funnel [119] and the reduction of MS duty cycle time by a factor of 3 times in Top10DDA mode (1 s vs. 3.6 s). High sensitivity can be achieved by tuning the ion fill time (optimizing the amount of a given species collected before fragmentation) that requires ion counting and optimizing the associated selection mass windows (from 1.3 to 4 mass units). Longer fill time and larger windows increase sensitivity and decrease sampling frequency and specificity [120]. The FTMS acquisition is long and more specifically for orbitraps (the transient duration is 32 ms at a resolution of 15,000 (FWHM) for m/z 200, but it can be increased to more than 1 s for achieving a resolution of 450,000 for m/z 200, and one must add the ion accumulation time on the top of it). Most of the technological changes have focused on parallelizing the different stages (filling, counting, fragmentation, and ion detection). FT-Orbitrap instruments' family has evolved into two subgroups; one has kept coupling to an LQT and has evolved with the tribrid Q-LQT-OT (Quadrupole-linear quadrupole ion trap-Orbitrap) configuration, and the other has kept coupling to an upstream quadrupole analyzer (Q-OT). In both cases, an ion counting device before detection by OT FT-MS is implemented. Similarly, FT-ICR

instruments have been equipped with a multipolar counting accumulation device to be effectively coupled with LC. Finally, a good sensitivity requires a good ionization efficiency, with a stable spray despite variations in the organic solvent content due to the LC gradient, which can be optimized by varying the voltage applied to the spray source with the instantaneous content of organic solvent [117].

4.3 Scan modes, targeted analysis, and data-independent analysis

In the quest for better sample coverage with a gain in sensitivity, dynamic range, and peak capacity, acquisition modes have evolved tremendously in recent years. The MS/MS acquisition was initially performed in DDA where the MS/MS analysis is triggered by the detection of a precursor above a defined signal intensity threshold in one MS, with the succession of cycles “1 MS survey scan, N DDA MS/MS scans” in the so-called TopN DDA acquisition mode. With this, it was recognized that many species remained undetected [121] because they either coeluted with abundant species or were too scarce to be selected [122]. Other strategies complemented this mode of operation. In order to increase the coverage of our sample, multiplex analyzes were introduced [123]. Thus, it is possible to select several precursors, either continuously over a window of more or less restricted mass or discretely by isolating various precursors of different masses, to subsequently analyze the composite MS/MS spectrum obtained from the fragmentation of these multiple precursors. It is necessary to deconvolute this composite spectrum: this has been achieved by using the retention times, in particular in MS^E or All ion fragmentation mode (AIF) [50,124] or by working with spectral databases, in DIA or SWATH [125]. Developed initially with TOF analyzers, this approach has also been proposed with FT ICR analyzers [126] and orbitraps [127]. The systematization of this strategy has given rise to data-independent acquisition (DIA) analysis. The MS/MS mode analysis is carried out sequentially and blindly on small mass windows designed to cover the whole mass range of interest. The interpretation of this type of data requires a dedicated LC MS/MS database (retention time, precursor mass, fragmentation profile) with an exact alignment of retention times. LC alignment is performed by relying on a set of standard peptides spiked in the sample of interest [128]. For example, the DIA multiplex combination has been used for histone analysis [129]. However, the same type of database can also be exploited to perform targeted analyses on some of the compounds in the scheduled Parallel Reaction Monitoring (PRM) strategy [130], according to the principle of the multiple reaction monitoring (MRM) strategy. Targeted studies are, of course, more limiting in terms of the number of compounds detected but can improve the sensitivity of the analysis [131,132]. The advantage of PRM for targeted analyses is that it does not require prior selection transitions of interest. It offers high resolution and high measurement accuracy for precursor and fragment ions, which reduces interferences.

Adding a standard set of peptides has very little influence on the result of an LC-MS/MS analysis and can align the chromatograms very efficiently within a campaign in your lab. It is not the case for inter-laboratory comparisons or even inter-experimental set-ups. In order to pool LC-MS/MS databases, it has therefore recently been proposed to introduce other types of information, such as the level of useful acetonitrile HI (% ACN) for the elution of a peptide instead of the retention time and distinguish the experimental and theoretical retention times (calculated with SSRCalc DB) [133]. Finally, the chromatographic conditions (stationary phase and mobile phase) and the type of sample injected seem to affect the peptides' retention properties. Therefore, they must be controlled: a correlation coefficient higher than 0.995 has been recommended to be able to use a bank of data from one DIA run to another.

Recently a study on the effects of ion isolation and accumulation on the sensitivity and quantitative accuracy of targeted proteomics was performed on a tribrid instrument coupled to a multidimensional nanoLC separation [117]. The authors reported a sensitivity of 1 yoctomole/ms or 100 molecules per scan for a pure standard peptide (yM, 10.E-24 mol) with a range of linearity for quantification that reaches 7 orders of magnitude. At this sensitivity level, the authors see unexpectedly no influence of the ion fill time with comparable signal intensities for ion fill times ranging between 25 and 2500 ms. The applied voltage to stabilize the electrospray is decreased with the increasing concentration of acetonitrile in the mobile phase during the LC gradient to achieve this amazing sensitivity. The used acquisition mode is a high-resolution targeted mode that derives from the PRM mode, and which is based on the indication of a list of peptides of interest with their exact masses and a window of expected retention time (AIM mode or Accumulated Ion Monitoring). In AIM mode, the tribrid configuration is fully exploited by selecting the precursor on a window of 8 amu and ion accumulation in the multipole, a high-resolution MS quantitation step in the orbitrap, and an identification of the MS/MS fragmentation sequence in the linear ion trap. The detection sensitivity drops when the analytes are presented in a more complex matrix.

While many strategies have been proposed to improve the rate of analysis in MS/MS mode, little work has been done to optimize MS detection. When acquiring the survey scan, a maximum number of ions is used to not deteriorate the spectral quality by space charge effects. This favors detecting the most abundant species, but many species are then detected with a low signal-to-noise ratio, and their identification is then compromised. A small mass range is analyzed in DIA mode each time. It is possible

to insert an MS spectrum between each MS/MS spectrum for each mass window. In this case, the analysis is analogous to what was done in Selected Ion Monitoring (SIM) mode in LTQ FT-ICR, which significantly improves the signal-to-noise ratio of the minor species and the accuracy of mass measurement at the expense of proteome coverage performed in DDA [134]. This idea was taken up in the BoxCar approach, which applies to any instrument capable of making a precursor selection upstream of the detection [135]. Briefly, the method uses the selection quadrupole upstream of the C-trap to transmit mass ranges of fixed or variable width sequentially and adapted to the number of charges contained in each range with a fixed filling rate per window (100,000 ions). These windows are grouped in two or three series, disjointed within each series, and allow recovery of 1 Da between each window in the joined series. It will generate a multiplex SIM that covers the entire mass range. A Top5 DDA MS/MS acquisition is then performed on this multiplexed SIM MS spectrum. Globally the signal MS is equalized, i.e., the windows containing the major species will be limited to the same number of ions as the windows containing less abundant species. The set of two or three acquisition series is then assembled to proceed with the data processing that has been modified from the database search engine MaxQuant [136]. The BoxCar strategy application on 1 μ g of proteolytic peptides from HeLa cell protein extract identified 60,228 single peptide sequences and 7775 protein groups, with one order of magnitude improvement in the dynamic range of the analysis. It can be adapted to other strategies such as DIA approaches. One of the limitations of the method is the time-consuming SIM approach that must be adjusted to the time required to perform the MS/MS analyses and the resolution of the LC separation.

4.4 Quantification

The identification of peptides and the resulting inferences of proteins is generally insufficient to answer the biology problems and must be supplemented by semi-quantitative information (sample comparison) or even quantitative information (protein assay). The diverse nature of the obtained information is a major obstacle since, on the one hand, proteins are not directly identified, and on the other hand, mass spectrometry is not a quantitative method of detection per se. It is possible to divide quantitative approaches in bottom-up proteomics into two categories from a methodological perspective (label-free approaches, stable isotope labeling methods) or from an application point of view (untargeted or targeted approaches). This field is continuously evolving and has been the subject of regular reviews, such as the review by Ankney et al. [137]. Here we reconsider the general principles and some recent results derived from the above conventional methods.

The quantification of a species in mass spectrometry is based on comparing the signal recorded for this species in two (or more) samples. This intensity is related to the physicochemical properties of the species, and it is therefore not possible to directly compare the signals of different species. The ionization efficiency can be modulated by the environment (matrix effect), but in general, this influence is neglected as a first approximation. Variation in intensity can also be reflected in the number of times (spectral counting) where this species exceeds a certain intensity threshold, such as the one set to switch to MS/MS mode. It is, therefore, possible to estimate it by considering the spectral counting. To compare two samples, it is possible to compare the signals of the separate acquisitions for each sample (label-free approach, non-targeted in DDA/DIA or targeted in PRM for example) or to mark this species with a label of the same molecular structure and different mass (stable isotopic labeling approach) and compare the intensities of the signals obtained in a single analysis of stoichiometrically pooled samples. This labeling can be carried out either chemically by derivation, enzymatically, or metabolically by incorporating labeled amino acids during protein synthesis. Finally, the isotopic label can be differentiated either directly at the MS level (different isotopic mass labels, for example in SILAC [138]/AACT [139] or enzymatic labeling ^{18}O [140]), either at the MS/MS level (isobaric isotopic labels but generating distinct reporters' mass in fragment ions, for example with isobaric Tags for Relative And Absolute Quantitation (iTRAQ) or with Tandem Mass Tag (TMT) labeling [141]). The use of very high-resolution mass analyzers also makes it possible to work with neutron markers (NeuCode) using amino acids whose monoisotopic mass with 6 mDa mass difference [142]. Combined with stable isotopic labeling such as SILAC, this increases the number of samples that can be quantified simultaneously [143] and enables absolute quantitative measurements using internal standards with stable isotopic labels. Each approach has its limits and its advantages [144]. The cost of the required reagents must be compared with the cost of the analyses since label-free approaches are instrument time-consuming. Metabolic labeling is not always possible since it is a priori relevant for auxotrophic models for labeled amino acids, even if models of living organisms have been used: *Drosophila* [145], *C. elegans* [146], *C. albicans* [147]. The use of ^{18}O labeling by incorporation of an H_2O molecule during proteolysis is also of interest. Although not offering a large mass increment, this strategy was applied to identify N-glycosylation sites after treatment of glycoproteins with a PNGase F [148]. The chemical derivatization of proteins has the advantage of applying to any protein. However, when the targeted reactive residues coincide with the proteolytic digestion sites, this can inhibit proteolysis. This is the case with the TMT and iTRAQ labeling that target lysine for tryptic digestion. In addition to cost, another disadvantage of this ubiquitous labeling is the risk of co-selection of multiple precursors in a complex mixture. The reporter ions resulting from the MS/MS fragmentation are then identical

regardless of the co-selected precursors, and their superposition causes a compression of the intensity ratios of these reporters [149]. For the same reason, the DIA mode is incompatible with this approach and not compatible with the SILAC strategy because, in addition to the major increase of the sample's complexity in multiplex mode, the labeled peptides can be split into several DIA windows. Reducing the mass window of the markers is then very interesting, and the NeuCode marking has been applied successfully.

More recently, a mass defect-based data-independent acquisition strategy called MdFDIA has been described, using metabolic labeling using $^{13}\text{C}_6^{15}\text{N}_2$ -lysine (+8.0142 Da, light) and D_8 -lysine (+8.0512 Da, heavy) and chemical derivatization on the C-terminal part of the peptides with light ($2^{13}\text{CD}_2\text{H}$, +34.06312 Da) and heavy (2CD_3 , +34.06896 Da) dimethyl groups, respectively. So, a peptide (and its C-terminal fragments) generated from four differentially labeled samples will be detected as a quadruplet with mass increments of 5.84, 31.16, and 5.84 mDa. The high mass resolution obtained from DIA MS/MS experiments allowed to resolve the peaks to perform a differential quantification and characterize protein changes in different breast cancer cell proteomes [150].

Absolute quantification of a species ideally requires the use of a labeled standard compound. This is not realistic for proteome-scale applications, so various algorithms have been proposed to correct the fact that two peptides of the same abundance do not give the same intensity in MS with approaches like the Exponentially Modified Protein Abundance Index (emPAI) [151] derived from spectral counting, and Intensity Based Absolute Quantification (iBAQ) [152] or iAPEX [153] based on the intensity of the peaks or the Top3 and its variant [154,155] based on the intensity of the peaks and the number of peptides identified. Recently a new LFAQ algorithm has been proposed to correct biases for minor proteins, notably by introducing the concept of a machine learning system [156] to integrate the concept of detectability in the form of the quantitative factor (Q-factor). This algorithm was evaluated on two sets of data obtained in DDA FTMS from total protein extracts of *Saccharomyces cerevisiae* strain BY4743 and *Mus musculus* RAW264.7 cells. The obtained correlation coefficients for the quantification of 48 standard UPS2 proteins that have been spiked in both complex samples at six different concentrations ranging from 0.05 fmol to 5000 fmol have shown that this Q-factor makes it possible to improve the quality of the quantification (correlation factor $R^2 \frac{1}{4}$ 0.931, 0.902, 0.888, 0.790 in yeast and $R^2 \frac{1}{4}$ 0.908, 0.897, 0.849, 0.820 in mouse respectively with LFAQ, iBAQ, Top3).

The acquisition modes are also essential for quantitative analysis. To circumvent the ratio compression effect of the reporter ions in TMT DDA MS2, the use of a tribrid in MultiNotch mode MS3 TMT-MS3 eliminates the chemical noise obtained in MS2 by considering the reporters of the fragments in MS3. The results of this type of analysis should be recovered and exported in editable format [157], which was a strong limitation for orbitrap data processing. The TMTc+ mode was introduced recently (c+ stands for the integrated deconvolution of complement reporter ion intensities). In TMTc+ mode, the complement ions formed as a result of the fragmentation of the precursor peptide are considered in a simplified strategy that does not require the use of the tribrid mass analyzer. This approach, however, requires precise deconvolution of those species whose isotopes are superimposed [158]. Moreover, TMT quantification requires fragmentation energy that can be deleterious to localize the PTMs on the peptides. Phosphorylation sites are usually lost due to neutral loss fragmentation. In the tribrid configuration, if this fragmentation energy is lowered to identify the modified peptide in MS/MS, then the fragments are submitted to MS3 analysis to perform the quantitation on the reporters in the MS3-IDQ mode as presented for example, for the phosphoproteome study [159] which can then be combined with the analytical strategies developed for the enrichment of phosphorylated peptides [160].

5 Top-down proteomics

Tandem mass spectrometry for protein analysis is not recent since the first MS/MS spectra on albumin were published in 1990 [161] with a series of multiply charged b fragments obtained in CID by fragmentation in the source on a triple Quadrupole [162]. Moreover, the identification and quantification of proteoforms directly from whole molecules is in many ways attractive: on the one hand, we can distinguish different states of modifications of a protein in order to correlate them with their function, which is impossible in bottom-up, and on the other hand, we can avoid the pitfalls of the statistical processing of data necessary for the protein inference step from the sequences of proteolytic peptides which is unavoidable in the bottom-up approach. However, the very large protein polypeptides suffer from technical difficulties that have always delayed top-down approaches compared to bottom-up approaches. Nevertheless, the term top-down proteomics was introduced first [163] before the term bottom-up proteomics. Closer to the final functional bricks, proteins and proteoforms also have a greater diversity of physicochemical properties, and their unbiased handling is, therefore, more delicate. They differ in their solubility, and their specific purification, separation, and storage are more

critical. Moreover, the analysis of these macrobiomolecules in MS and in MS/MS is also more demanding in terms of specifications. In MALDI or ESI and under non-denaturing conditions (where low charge states are obtained), the required mass range for ion detection is relatively high. In ESI under denaturing conditions, very high charge states require high mass resolution and high mass accuracy in order to discriminate the species involved correctly. Finally, data processing is more complicated because the fragmentation rules are less established for these MS/MS spectra (with the need for deconvolution of superimposed multiple charge states) and because the activation methods best suited to obtain good sequence coverage (namely ECD/ETD and UVPD modes) are still in full development as we previously reported. The development of top-down approaches is regularly reviewed [164] and gave birth to an international consortium, in a willingness to redouble its efforts to get closer to the bottom up's analytical capacity as quickly as possible.

5.1 Purification and separation of proteins

Despite the lower number of components of the protein samples, when the proteins are not proteolyzed, a separation and solubilization step is essential [17]. The solubilization of proteins often requires the addition of surfactants that can interact with stationary chromatography phases and limit the protein's ionization efficiency in MS. To extract and solubilize proteins, the most common surfactant is sodium dodecyl sulfate (SDS), although it is poorly tolerated in MS and LC [165] because it could not be effectively replaced. Various methods have been developed to remove SDS by performing acetone precipitation, membrane filtration (FASP), or electrophoresis transmembrane (TME) [166], which are accompanied by problems of aggregation of the proteins depleted in SDS. This problem of manipulation and processing of samples is probably one of the major limitations of top-down approaches.

The separation of proteins is mainly carried out either in liquid chromatography (RPLC, HILIC, Size-exclusion Chromatography SEC or IE mainly) or in electrophoresis (GELFrEE, IEF or capillary electrophoresis) [167]. RPLC is most frequently coupled to MS because of its high compatibility with the ESI ionization mode. It combines a non-polar stationary phase (silica in general) whose hydrophobicity is modulated by grafting of alkyl chains (C1 to C18 generally) with a polar acidic mobile phase containing volatile components. The length of the alkyl chain is adapted to the size of the compounds of interest. Other parameters can influence the quality of the separation, such as the particle size of the silica particles or granulometry (with a smaller size giving a lower theoretical plateau height and therefore a better resolving power), and porosity (ideally 200–450 Å, too small pores for better separation will increase the risk of clogging and cross-contamination) [168]. The length of the column directly increases the number of theoretical plates but at the cost of greater pressure drop, which can be lowered by a higher separation temperature (45–60 °C). The effect of temperature has been used to apply a chromatographic elution gradient [169]. Alternatives to silica have been proposed, with the use of polymeric support [170]. Most approaches work in nanoLC mode with small internal diameter (75 µm usually) columns to be coupled efficiently with good sensitivity. RPLC is rarely used alone because its resolving power is insufficient for proteins, and two-dimensional approaches are generally proposed. As for bottom-up proteomics, the “hi RP/lo RP” coupling [171], which exploits the orthogonality of RPLC dimensions in basic and acidic conditions, has been applied to top-down approaches. [172]. HILIC separations are less developed as final separation techniques, which is undoubtedly due to the difficulties of working in capillary mode and its more limited phase capacity and less understanding of the underlying mechanisms, which reduces the robustness of these mechanisms. However, recent developments are trying to exploit their specificities [173]. The applications for the analysis of glycoproteoforms seem very promising [174]. Separations SEC and IE are often dedicated to a first-dimension separation in on-line or off-line mode [175,176], and this is due in particular to a lack of miniaturization and the poor compatibility of the solvents with the MS.

Electrophoretic approaches are historically the first to have demonstrated the diversity of proteoforms. Thus, post-translational modifications such as glycosylation or phosphorylation are traditionally highlighted by 2D gel electrophoresis (GE) [177,178]. On the other hand, the extraction of proteins is not easy from Polyacrylamide Gel Electrophoresis (SDS-PAGE) gels [179]. Other approaches have been developed, such as GELFrEE (i.e., gel-eluted liquid fraction entrapment electrophoresis) [180] originally developed for small proteins (<25 kDa) and recently applied to identify [78] a bit less than 1000 proteoforms from a HeLa whole cell lysate in GEL-FrEE nanoLC FT MS/MS. Other electrophoretic approaches also offer great prospects, such as capillary electrophoresis. A study coupled a first preparative dimension in RPLC to a CZE-FTMS/MS analysis to identify 580 proteoforms of the yeast proteome on a hybrid Q-OT [181]. Optimizing the sample loading capacity up to 1 µL with a dynamic pH junction-based CZE-MS has increased the number of identifications to 580 proteoforms in *E. coli* [182]. Finally, the combination of chromatography and CZE in a three-dimensional SEC, RPLC, and CZE-MS approach has reached 5700 proteoforms associated with 850 proteins from the *E. coli* proteome using the SEC-

RPLC-CZE-MS/MS platform with a 1% spectrum-level FDR [183]. CZE-MS/MS coupling could achieve an analysis depth comparable to the nanoLC-MS/MS coupling used in a multidimensional approach.

5.2 Specificity of MS analysis

Since the MALDI-TOF MS of human IgM immunoglobulin, mass spectrometry techniques have evolved considerably [184]. The identification (and quantification) of proteins in the top-down approach requires high resolution mass spectrometers, making FT MS particularly suitable, with orbitrap or FT-ICR technology. The relative mass variations induced by the PTMs decrease with the length of the analyzed polypeptide. Detecting a deamidation, or distinguishing trimethylation from acetylation, for example, occurring in a compound of several kDa requires a high resolution and a high mass measurement accuracy. Fig. 3 is an example of the complementarities of peptide-centric and protein-centric strategies. The MS profile of three proteoforms of lactalbumin confirms that they are differentially modified, and the bottom-up analysis provides evidence of the modification sites, in this case, deamidation on specific sites.

As previously explained, the ESI source has been widely favored for top-down proteomics because it can be coupled on-line with separative methods and because the high charge states allow lowering the working mass range (even if higher resolution is required). The proteins analyzed in ESI under denaturing conditions can be ionized with several tens of proton attachments, and they will overlap in the spectrum with other co-eluted species. Many software programs have been developed to deconvolute these spectra, but this requires obtaining well-resolved signals with an acceptable signal-to-noise ratio, knowing that a single protein species can be distributed over multiple charge states, which lowers the intrinsic signal of each multi-charged isotopic profile. FTMS approaches are also less limited by random chemical or electronic noise often observed with TOF MS.

LC ESI FT-ICR coupling is only operational when an intermediate multipole is added to count and accumulate ions from the continuous ESI source to optimize the transmission of compounds [185]. The automatic gain control (AGC) unit is useful because it limits space charge effects when coupled to FT-Orbitrap or FT-ICR analyzers. Thus, the configuration developed by Thermo Finnigan, which was very quickly exploited for orbitrap analyzers later, allowed direct coupling thanks to its interface with an ion trap and a transfer octopole to the ICR cell [186].

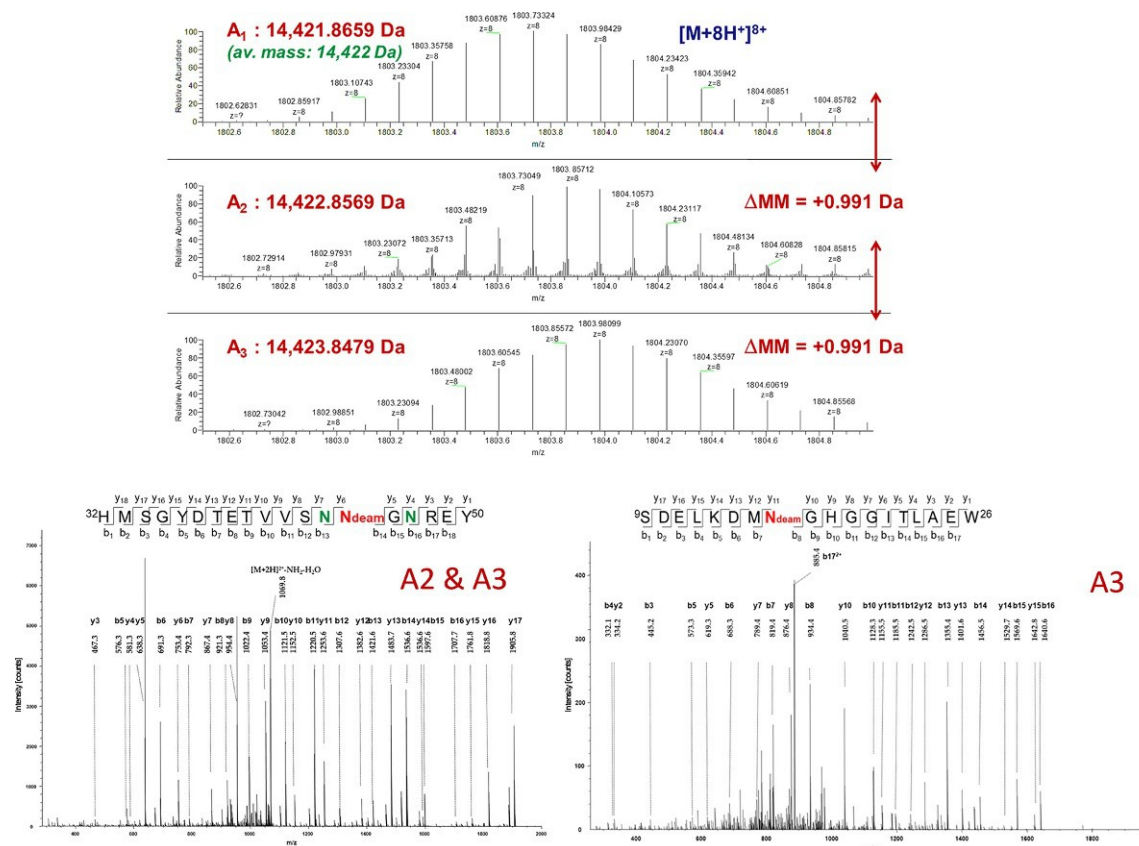


Fig. 3 Deamidation of beta lactalbumin from camel milk. During incubation beta lactalbumin (A1 form) gradually degrades to A2 and A3 forms. Intact mass measurement (top spectra of the 3 isolated forms) indicates a mass shift of +0.991 Da, in agreement with deamidation. Two sequons NG prone to deamidation are found in the protein sequence. Additional bottom-up analysis of the three proteoforms demonstrates that a first deamidation occurs on asparagine N45 (found in both A2 and A3 proteoforms, see lower MS/MS spectra on the left) and a second deamidation occurs later on asparagine N16 (found in proteoform A3 only, see lower MS/MS spectra on the right). Adapted with permission from Elsevier: S. Si Ahmed Zennia, A. Mati, F. Saulnier, Y. Verdier, G. Chiappetta, G. Mulliert, L. Miclo, J. Vinh, J.-M. Girardet, Identification by FT-ICR-MS of *Camelus dromedarius* α -lactalbumin variants as the result of nonenzymatic deamidation of Asn-16 and Asn-45, *Food Chem.* 187 (2015) 305–313.

A quadrupole working over a wide mass range is required to analyze proteins in non-denaturing conditions with low charge states [187] and is well suited for new imaging applications with a MALDI ion source. Top-down proteomics used other configurations involving orbitrap analyzers of more limited resolution [188]. Although the first generation of MS instruments suffered in the past from low transmission of large polypeptides and also from significant ion suppression when counting the multicharged ions, the latest developments can now give out interesting proteome covers [189], with useful resolutions ($>240,000$ at m/z 200) but which remain below the performance of an FT-ICR.

Fragmentation patterns for protein sequence coverage are one of the most active topics of study today. They provide additional information. Activated ion-electron transfer dissociation (AI-ETD), which combines IR irradiation with ETD, increases the sequence coverage by generating ECD/ETD-like fragments of type c and z [190], and the sequencing of proteins of higher molecular mass (up to 70 kDa) have been reported [191]. UV photodissociation [77] also produces complete sequence coverage. It is suitable for studying the composition of intact protein complexes [76,192], with the study of disulfide bridges and the localization of protonated protein species in the gas phase as was done for small proteins such as ubiquitin and beta-lactoglobulin [193]. The UVPD/HCD combination in LCMS/MS on the HeLa cell proteome demonstrated that UVPD was complementary to HCD and allowed the identification of fewer proteins but with better sequence coverage. [78]. The combined analysis of intact top-down protein and associated polypeptides from middle-down proteolysis on Ides-digested IgG1 (150 kDa), which releases fragments Fc/2 and Fab'/2 (25 and 50 kDa respectively), achieves 40% sequence coverage through the combination of UVPD and EThcD [194]. However, this approach was carried out on a monoclonal antibody without limitation due to the superposition of multiple endogenous proteoforms of IgG.

6 Epilogue

Protein analysis by mass spectrometry is not restricted to proteomic analyses (i.e., identifying and quantifying proteoforms in one given sample). Other studies are as exciting such as the structural and conformational analysis of proteins by limited proteolysis [195,196], by H/D exchange [197,198], or as protein imaging [199,200] or related derivatives imaging [201]. This often implies new developments for the sample treatment [202] and ionization [203]. The use of a mass analyzer combining ultra-high-resolution, high precision of mass measurement with a high speed of acquisition is required to address such applications, and these are newly opened fields where Fourier transform mass spectrometry could bring much invaluable information.

7 Acknowledgments

We want to thank the SMBP ESPCI group and, more especially, Chiara Giangrande for the histone code study, Giovanni Chiappetta and Yann Verdier for the lactalbumin characterization, Emmanuelle Demey for help on the preparation of the illustrations.

8 References

- [1] V.C. Wasinger, S.J. Cordwell, A. Cerpa-Poljak, J.X. Yan, A.A. Gooley, M.R. Wilkins, M.W. Duncan, R. Harris, K.L. Williams, I. Humphery-Smith, Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*, *Electrophoresis* 16 (1995) 1090–1094.
- [2] M.R. Wilkins, J.C. Sanchez, A.A. Gooley, R.D. Appel, I. Humphery-Smith, D.F. Hochstrasser, K.L. Williams, Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it, *Biotechnol. Genet. Eng. Rev.* 13 (1996) 19–50.
- [3] P. Kahn, From genome to proteome: looking at a cell's proteins, *Science* 270 (1995) 369–370.
- [4] D. Swinbanks, Government backs proteome proposal, *Nature* 378 (1995) 653.
- [5] M. Karas, D. Bachmann, U. Bahr, F. Hillenkamp, Matrix-assisted ultraviolet laser desorption of non-volatile compounds, *Int. J. Mass Spectrom. Ion Process.* 78 (1987) 53–68.
- [6] C.M. Whitehouse, R.N. Dreyer, M. Yamashita, J.B. Fenn, Electrospray interface for liquid chromatographs and mass spectrometers, *Anal. Chem.* 57 (1985) 675–679.
- [7] K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida, T. Matsuo, Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry, *Rapid Commun. Mass Spectrom.* 2 (1988) 151–153.
- [8] B. Chait, S. Kent, Weighing naked proteins: practical, high-accuracy mass measurement of peptides and proteins,

- Science 257 (1992) 1885–1894.
- [9] G.A. Valaskovic, N.L. Kelleher, D.P. Little, D.J. Aaserud, F.W. McLafferty, Attomole-sensitivity electrospray source for large-molecule mass spectrometry, *Anal. Chem.* 67 (1995) 3802–3805.
 - [10] M.R. Emmett, R.M. Caprioli, Micro-electrospray mass spectrometry: ultra-high sensitivity analysis of peptides and proteins, *J. Am. Soc. Mass Spectrom.* 5 (1994) 605–613.
 - [11] M. Wilm, M. Mann, Analytical properties of the nanoelectrospray ion source, *Anal. Chem.* 68 (1996) 1–8.
 - [12] N.L. Anderson, N.G. Anderson, Proteome and proteomics: new technologies, new concepts, and new words, *Electrophoresis* 19 (1998) 1853–1861.
 - [13] M. Mann, P. Højrup, P. Roepstorff, Use of mass spectrometric molecular weight information to identify proteins in sequence databases, *Biol. Mass Spectrom.* 22 (1993) 338–345.
 - [14] P. Roepstorff, Mass spectrometry in protein studies from genome to function, *Curr. Opin. Biotechnol.* 8 (1997) 6–13.
 - [15] N.H. Packer, M.J. Harrison, Glycobiology and proteomics: is mass spectrometry the holy grail? *Electrophoresis* 19 (1998) 1872–1882.
 - [16] The Consortium for Top Down Proteomics, L.M. Smith, N.L. Kelleher, Proteoform: a single term describing protein complexity, *Nat. Methods* 10 (2013) 186–187.
 - [17] F.E. Regnier, J. Kim, Proteins and proteoforms: new separation challenges, *Anal. Chem.* 90 (2018) 361–373.
 - [18] R. Aebersold, J.N. Agar, I.J. Amster, M.S. Baker, C.R. Bertozzi, E.S. Boja, C.E. Costello, B.F. Cravatt, C. Fenselau, B.A. Garcia, Y. Ge, J. Gunawardena, R.C. Hendrickson, P.J. Hergenrother, C.G. Huber, A.R. Ivanov, O.N. Jensen, M.C. Jewett, N.L. Kelleher, L.L. Kiessling, N.J. Krogan, M.R. Larsen, J.A. Loo, R.R. Ogorzalek Loo, E. Lundberg, M.J. MacCoss, P. Mallick, V.K. Mootha, M. Mrksich, T.W. Muir, S.M. Patrie, J.J. Pesavento, S.J. Pitteri, H. Rodriguez, A. Saghatelian, W. Sandoval, H. Schlöter, S. Sechi, S.A. Slavoff, L.M. Smith, M.P. Snyder, P.M. Thomas, M. Uhlen, J.E. Van Eyk, M. Vidal, D.R. Walt, F.M. White, E.R. Williams, T. Wohlschläger, V.H. Wysocki, N.A. Yates, N.L. Young, B. Zhang, How many human proteoforms are there? *Nat. Chem. Biol.* 14 (2018) 206–214.
 - [19] E.A. Ponomarenko, E.V. Poverennaya, E.V. Ilgisonis, M.A. Pyatnitskiy, A.T. Kopylov, V.G. Zgoda, A.V. Lisitsa, A.I. Archakov, The size of the human proteome: the width and depth, *Int. J. Anal. Chem.* 2016 (2016) 1–6.
 - [20] The UniProt Consortium, UniProt: the universal protein knowledgebase, *Nucleic Acids Res.* 45 (2017) D158–D169.
 - [21] Y. Zhang, B.R. Fonslow, B. Shan, M.-C. Baek, J.R. Yates, Protein analysis by shotgun/bottom-up proteomics, *Chem. Rev.* 113 (2013) 2343–2394.
 - [22] A. Shevchenko, O.N. Jensen, A.V. Podtelejnikov, F. Sagliocco, M. Wilm, O. Vorm, P. Mortensen, A. Shevchenko, H. Boucherie, M. Mann, Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels, *Proc. Natl. Acad. Sci. U. S. A.* 93 (1996) 14440–14445.
 - [23] J.R. Yates, S. Speicher, P.R. Griffin, T. Hunkapiller, Peptide mass maps: a highly informative approach to protein identification, *Anal. Biochem.* 214 (1993) 397–408.
 - [24] X. Zhan, H. Yang, F. Peng, J. Li, Y. Mu, Y. Long, T. Cheng, Y. Huang, Z. Li, M. Lu, N. Li, M. Li, J. Liu, P.R. Jungblut, How many proteins can be identified in a 2DE gel spot within an analysis of a complex human cancer tissue proteome? *Electrophoresis* 39 (2018) 965–980.
 - [25] J.R. Yates, The revolution and evolution of shotgun proteomics for large-scale proteome analysis, *J. Am. Chem. Soc.* 135 (2013) 1629–1640.
 - [26] K. Biemann, Sequencing of peptides by tandem mass spectrometry and high-energy collision-induced dissociation, *Methods Enzymol.* (1990) 455–479.
 - [27] A.J. Oosterkamp, E. Gelpí, J. Abian, Quantitative peptide bioanalysis using column-switching nano liquid chromatography/mass spectrometry, *J. Mass Spectrom.* 33 (1998) 976–983.
 - [28] M.T. Davis, D.C. Stahl, S.A. Hefta, T.D. Lee, A microscale electrospray interface for on-line, capillary liquid chromatography/tandem mass spectrometry of complex peptide mixtures, *Anal. Chem.* 67 (1995) 4549–4556.
 - [29] J. Abian, A.J. Oosterkamp, E. Gelpí, Comparison of conventional, narrow-bore and capillary liquid chromatography/mass spectrometry for electrospray ionization mass spectrometry: practical considerations, *J. Mass Spectrom.* 34 (1999) 244–254.
 - [30] R.H. Robins, J.E. Guido, Design, construction and application of a simple packed capillary high performance liquid chromatography/electrospray mass spectrometry system, *Rapid Commun. Mass Spectrom.* 11 (1997) 1661–1666.
 - [31] Y. Chen, F.E. Leach, N.K. Kaiser, X. Dang, Y.M. Ibrahim, R.V. Norheim, G.A. Anderson, R.D. Smith, A.G. Marshall, Improved ion optics for introduction of ions into a 9.4-T Fourier transform ion cyclotron resonance mass spectrometer, *J. Mass Spectrom.* 50 (2015) 280–284.
 - [32] M.E. Belov, S.R. Ellis, M. Dillillo, M.R.L. Paine, W.F. Danielson, G.A. Anderson, E.L. de Graaf, G.B.

- Eijkel, R.M.A. Heeren, L.A. McDonnell, Design and performance of a novel interface for combined matrix-assisted laser desorption ionization at elevated pressure and electrospray ionization with Orbitrap mass spectrometry, *Anal. Chem.* 89 (2017) 7493–7501.
- [33] A.H. Payne, G.L. Glish, Tandem mass spectrometry in quadrupole ion trap and ion cyclotron resonance mass spectrometers, *Methods Enzymol.* 402 (2005) 109–148.
- [34] M.R. Hoopmann, R.L. Moritz, Current algorithmic solutions for peptide-based proteomics data generation and identification, *Curr. Opin. Biotechnol.* 24 (2013) 31–38.
- [35] N.L. Kelleher, H.Y. Lin, G.A. Valaskovic, D.J. Aaserud, E.K. Fridriksson, F.W. McLafferty, Top down versus bottom up protein characterization by tandem high-resolution mass spectrometry, *J. Am. Chem. Soc.* 121 (1999) 806–812.
- [36] B.T. Chait, Mass spectrometry: bottom-up or top-down? *Science* 314 (2006) 65–66. [37] B. Bogdanov, R.D. Smith, Proteomics by FTICR mass spectrometry: top down and bottom up, *Mass Spectrom. Rev.* 24 (2005) 168–200.
- [38] B. Meyer, D.G. Papasotiriou, M. Karas, 100% protein sequence coverage: a modern form of surrealism in proteomics, *Amino Acids* 41 (2011) 291–310.
- [39] L.M. Smith, N.L. Kelleher, Proteoforms as the next proteomics currency, *Science* 359 (2018) 1106–1107.
- [40] T.K. Toby, L. Fornelli, N.L. Kelleher, Progress in top-down proteomics and the analysis of proteoforms, *Annu. Rev. Anal. Chem.* 9 (2016) 499–519.
- [41] M. Mann, N.L. Kelleher, Precision proteomics: the case for high resolution and high mass accuracy, *Proc. Natl. Acad. Sci.* 105 (2008) 18132–18138.
- [42] K.D. Henry, E.R. Williams, B.H. Wang, F.W. McLafferty, J. Shabanowitz, D.F. Hunt, Fourier-transform mass spectrometry of large molecules by electrospray ionization, *Proc. Natl. Acad. Sci. U. S. A.* 86 (1989) 9075–9078.
- [43] E. Shishkova, A.S. Hebert, J.J. Coon, Now, more than ever, proteomics needs better chromatography, *Cell Syst.* 3 (2016) 321–324.
- [44] J.S. Brodbelt, Ion activation methods for peptides and proteins, *Anal. Chem.* 88 (2016) 30–51.
- [45] K. Biemann, Contributions of mass spectrometry to peptide and protein structure, *Biol. Mass Spectrom.* 16 (1988) 99–111.
- [46] P. Roepstorff, J. Fohlman, Letter to the editors: proposal for a common nomenclature for sequence ions in mass spectra of peptides, *Biol. Mass Spectrom.* 11 (1984) 601.
- [47] J.V. Olsen, B. Macek, O. Lange, A. Makarov, S. Horning, M. Mann, Higher-energy C-trap dissociation for peptide modification analysis, *Nat. Methods* 4 (2007) 709–712.
- [48] C. Shao, Y. Zhang, W. Sun, Statistical characterization of HCD fragmentation patterns of tryptic peptides on an LTQ Orbitrap Velos mass spectrometer, *J. Proteomics* 109 (2014) 26–37.
- [49] A. Michalski, E. Damoc, J.-P. Hauschild, O. Lange, A. Wiegand, A. Makarov, N. Nagaraj, J. Cox, M. Mann, S. Horning, Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer, *Mol. Cell. Proteomics* 10 (2011) M111.011015.
- [50] T. Geiger, J. Cox, M. Mann, Proteomics on an Orbitrap benchtop mass spectrometer using all-ion fragmentation, *Mol. Cell. Proteomics* 9 (2010) 2252–2261.
- [51] C.D. Kelstrup, D.B. Bekker-Jensen, T.N. Arrey, A. Høgrebe, A. Harder, J.V. Olsen, Performance evaluation of the Q Exactive HF-X for Shotgun Proteomics, *J. Proteome Res.* 17 (2018) 727–738.
- [52] H.J. Cooper, K. Håkansson, A.G. Marshall, The role of electron capture dissociation in biomolecular analysis: electron capture dissociation, *Mass Spectrom. Rev.* 24 (2005) 201–222.
- [53] M.A. van Agthoven, L. Chiron, M.-A. Coutouly, M.-A. Delsuc, C. Rolando, Two-dimensional ECD FT-ICR mass spectrometry of peptides and glycopeptides, *Anal. Chem.* 84 (2012) 5589–5595.
- [54] M.A. van Agthoven, C.A. Wootton, L. Chiron, M.-A. Coutouly, A. Soulby, J. Wei, M.P. Barrow, M.-A. Delsuc, C. Rolando, P.B. O'Connor, Two-dimensional mass spectrometry for proteomics, a comparative study with Cytochrome c, *Anal. Chem.* 88 (2016) 4409–4417.
- [55] F. Floris, L. Chiron, A.M. Lynch, M.P. Barrow, M.-A. Delsuc, P.B. O'Connor, Top-down deep sequencing of ubiquitin using two-dimensional mass spectrometry, *Anal. Chem.* 90 (2018) 7302–7309.
- [56] F. Floris, M.A. van Agthoven, L. Chiron, C.A. Wootton, P.Y.Y. Lam, M.P. Barrow, M.-A. Delsuc, P.B. O'Connor, Bottom-up two-dimensional electron-capture dissociation mass spectrometry of calmodulin, *J. Am. Soc. Mass Spectrom.* 29 (2018) 207–210.
- [57] K.L. Fort, C.N. Cramer, V.G. Voinov, Y.V. Vasil'ev, N.I. Lopez, J.S. Beckman, A.J.R. Heck, Exploring ECD on a Benchtop Qexactive Orbitrap mass spectrometer, *J. Proteome Res.* 17 (2018) 926–933.
- [58] J.E.P. Syka, J.J. Coon, M.J. Schroeder, J. Shabanowitz, D.F. Hunt, Peptide and protein sequence analysis by electron

- transfer dissociation mass spectrometry, *Proc. Natl. Acad. Sci.* 101 (2004) 9528–9533.
- [59] N.D. Udeshi, P.D. Compton, J. Shabanowitz, D.F. Hunt, K.L. Rose, Methods for analyzing peptides and proteins on a chromatographic timescale by electron-transfer dissociation mass spectrometry, *Nat. Protoc.* 3 (2008) 1709–1717.
- [60] S. Tamara, R.A. Scheltema, A.J.R. Heck, A.C. Leney, Phosphate transfer in activated protein complexes reveals interaction sites, *Angew. Chem. Int. Ed.* 56 (2017) 13641–13644. [61] P. L6ssl, A.M. Brunner, F. Liu, A.C. Leney, M. Yamashita, R.A. Scheltema, A.J.R. Heck, Deciphering the Interplay among multisite phosphorylation, interaction dynamics, and conformational transitions in a tripartite protein system, *ACS Cent. Sci.* 2 (2016) 445–455.
- [62] V. Franc, Y. Yang, A.J.R. Heck, Proteoform profile mapping of the human serum complement component C9 revealing unexpected new features of N -, O -, and C-glycosylation, *Anal. Chem.* 89 (2017) 3483–3491.
- [63] F. Sobott, S.J. Watt, J. Smith, M.J. Edelman, H.B. Kramer, B.M. Kessler, Comparison of CID versus ETD based MS/MS fragmentation for the analysis of protein ubiquitination, *J. Am. Soc. Mass Spectrom.* 20 (2009) 1652–1659.
- [64] M.-S. Kim, A. Pandey, Electron transfer dissociation mass spectrometry in proteomics, *Proteomics* 12 (2012) 530–542.
- [65] G.C. McAlister, D. Phanstiel, D.M. Good, W.T. Berggren, J.J. Coon, Implementation of electron-transfer dissociation on a hybrid linear ion trap-Orbitrap mass spectrometer, *Anal. Chem.* 79 (2007) 3525–3534.
- [66] N.M. Riley, J.J. Coon, The role of electron transfer dissociation in modern proteomics, *Anal. Chem.* 90 (2018) 40–64.
- [67] F. Liu, P. L6ssl, B.M. Rabbitts, R.S. Balaban, A.J.R. Heck, The interactome of intact mitochondria by cross-linking mass spectrometry provides evidence for coexisting respiratory supercomplexes, *Mol. Cell. Proteomics* 17 (2018) 216–232.
- [68] K.R. Durbin, R.T. Fellers, I. Ntai, N.L. Kelleher, P.D. Compton, Autopilot: an online data acquisition control system for the enhanced high-throughput characterization of intact proteins, *Anal. Chem.* 86 (2014) 1485–1492.
- [69] C. Singh, C.G. Zampronio, A.J. Creese, H.J. Cooper, Higher energy collision dissociation (HCD) product ion-triggered electron transfer dissociation (ETD) mass spectrometry for the analysis of N-linked glycoproteins, *J. Proteome Res.* 11 (2012) 4517–4525.
- [70] J. Saba, S. Dutta, E. Hemenway, R. Viner, Increasing the productivity of glycopeptides analysis by using higher-energy collision dissociation-accurate mass-product-dependent electron transfer dissociation, *Int. J. Proteomics* 2012 (2012) 1–7.
- [71] J. Laskin, J.H. Futrell, Activation of large ions in FT-ICR mass spectrometry, *Mass Spectrom. Rev.* 24 (2005) 135–167.
- [72] W. Li, C.L. Hendrickson, M.R. Emmett, A.G. Marshall, Identification of intact proteins in mixtures by alternated capillary liquid chromatography electrospray ionization and LC ESI infrared multiphoton dissociation Fourier transform ion cyclotron resonance mass spectrometry, *Anal. Chem.* 71 (1999) 4397–4402.
- [73] G. van der Rest, F. He, M.R. Emmett, A.G. Marshall, S.J. Gaskell, Gas-phase cleavage of PTC-derivatized electrosprayed tryptic peptides in an FT-ICR trapped-ion cell: mass-based protein identification without liquid chromatographic separation, *J. Am. Soc. Mass Spectrom.* 12 (2001) 288–295.
- [74] S. Bourgoïn-Voillard, N. Leymarie, C.E. Costello, Top-down tandem mass spectrometry on RNase A and B using a Qh/FT-ICR hybrid mass spectrometer, *Proteomics* 14 (2014) 1174–1184.
- [75] J.B. Shaw, W. Li, D.D. Holden, Y. Zhang, J. Griep-Raming, R.T. Fellers, B.P. Early, P.M. Thomas, N.L. Kelleher, J.S. Brodbelt, Complete protein characterization using top-down mass spectrometry and ultraviolet photodissociation, *J. Am. Chem. Soc.* 135 (2013) 12646–12651.
- [76] L.J. Morrison, J.S. Brodbelt, 193 nm ultraviolet photodissociation mass spectrometry of tetrameric protein complexes provides insight into quaternary and secondary protein topology, *J. Am. Chem. Soc.* 138 (2016) 10849–10859.
- [77] R.R. Julian, The mechanism behind top-down UVPD experiments: making sense of apparent contradictions, *J. Am. Soc. Mass Spectrom.* 28 (2017) 1823–1826.
- [78] T.P. Cleland, C.J. DeHart, R.T. Fellers, A.J. VanNispen, J.B. Greer, R.D. LeDuc, W.R. Parker, P.M. Thomas, N.L. Kelleher, J.S. Brodbelt, High-throughput analysis of intact human proteins using UVPD and HCD on an Orbitrap mass spectrometer, *J. Proteome Res.* 16 (2017) 2072–2079.
- [79] J.B. Shaw, E.W. Robinson, L. Paša-Tolić, Vacuum ultraviolet photodissociation and Fourier transform-ion cyclotron resonance (FT-ICR) mass spectrometry: revisited, *Anal. Chem.* 88 (2016) 3019–3023.
- [80] M.A. Halim, M. Girod, L. MacAleese, J. Lemoine, R. Antoine, P. Dugourd, Combined infrared multiphoton dissociation with ultraviolet photodissociation for ubiquitin characterization, *J. Am. Soc. Mass Spectrom.* 27 (2016) 1435–1442.
- [81] M.A. Halim, L. MacAleese, J. Lemoine, R. Antoine, P. Dugourd, M. Girod, Ultra-violet, infrared, and high-low energy photodissociation of post-translationally modified peptides, *J. Am. Soc. Mass Spectrom.* 29 (2018) 270–283.
- [82] Y. Zhang, X. Xie, X. Zhao, F. Tian, J. Lv, W. Ying, X. Qian, Systems analysis of singly and multiply O-glycosylated peptides in the human serum glycoproteome via EThcD and HCD mass spectrometry, *J. Proteomics* 170 (2018) 14–27.

- [83] J.A. Madsen, H. Xu, M.R. Robinson, A.P. Horton, J.B. Shaw, D.K. Giles, T. S. Kaoud, K.N. Dalby, M.S. Trent, J.S. Brodbelt, High-throughput database search and large-scale negative polarity liquid chromatography–tandem mass spectrometry with ultraviolet photodissociation for complex proteomic samples, *Mol. Cell. Proteomics* 12 (2013) 2604–2614.
- [84] S.M. Greer, M. Bern, C. Becker, J.S. Brodbelt, Extending proteome coverage by combining MS/MS methods and a modified bioinformatics platform adapted for database searching of positive and negative polarity 193 nm ultraviolet photodissociation mass spectra, *J. Proteome Res.* 17 (2018) 1340–1347.
- [85] J.S. Cottrell, Protein identification using MS/MS data, *J. Proteomics* 74 (2011) 1842–1851.
- [86] K. Verheggen, L. Martens, F.S. Berven, H. Barsnes, M. Vaudel, Database search engines: paradigms, challenges and solutions, *Adv. Exp. Med. Biol.* 919 (2016) 147–156.
- [87] E. Audain, J. Uszkoreit, T. Sachsenberg, J. Pfeuffer, X. Liang, H. Hermjakob, A. Sanchez, M. Eisenacher, K. Reinert, D.L. Tabb, O. Kohlbacher, Y. Perez-Riverol, In-depth analysis of protein inference algorithms using multiple search engines and well-defined metrics, *J. Proteomics* 150 (2017) 170–182.
- [88] C. Bamberger, S. Martínez-Bartolome, M. Montgomery, S. Pankow, J.D. Hulleman, J.W. Kelly, J.R. Yates, Deducing the presence of proteins and proteoforms in quantitative proteomics, *Nat. Commun.* 9 (2018) 2320.
- [89] X.-D. Feng, L.-W. Li, J.-H. Zhang, Y.-P. Zhu, C. Chang, K.-X. Shu, J. Ma, Using the entrapment sequence method as a standard to evaluate key steps of proteomics data analysis process, *BMC Genomics* 18 (2017) 143.
- [90] B.B. Misra, Updates on resources, software tools, and databases for plant proteomics in 2016–2017, *Electrophoresis* 39 (13) (2018) 1543–1557.
- [91] R.L. Gundry, Bioinformatics for mass spectrometry-based proteomics, in: D.R.M. Graham, D.E. Ott (Eds.), *HIV-1 Proteomics*, Springer New York, New York, NY, 2016, pp. 99–112.
- [92] K.D.B. Anapindi, E.V. Romanova, B.R. Southey, J.V. Sweedler, Peptide identifications and false discovery rates using different mass spectrometry platforms, *Talanta* 182 (2018) 456–463.
- [93] D. Bergeron, A.-M.S. Tremblay, Algorithms for optimized maximum entropy and diagnostic tools for analytic continuation. *Phys. Rev. E* 94 (2016). <https://doi.org/10.1103/PhysRevE.94.023303>.
- [94] M. Bern, T. Caval, Y.J. Kil, W. Tang, C. Becker, E. Carlson, D. Kletter, K.I. Sen, N. Galy, D. Hagemans, V. Franc, A.J.R. Heck, Parsimonious charge deconvolution for native mass spectrometry, *J. Proteome Res.* 17 (2018) 1216–1226.
- [95] W. Cai, H. Guner, Z.R. Gregorich, A.J. Chen, S. Ayaz-Guner, Y. Peng, S.G. Valeja, X. Liu, Y. Ge, MASH suite pro: a comprehensive software tool for top-down proteomics, *Mol. Cell. Proteomics* 15 (2016) 703–714.
- [96] R.D. LeDuc, V. Schwämmle, M.R. Shortreed, A.J. Cesnik, S.K. Solntsev, J.B. Shaw, M.J. Martin, J.A. Vizcaino, E. Alpi, P. Danis, N.L. Kelleher, L.M. Smith, Y. Ge, J.N. Agar, J. Chamot-Rooke, J.A. Loo, L. Pasa-Tolic, Y.O. Tsybin, ProForma: a standard proteoform notation, *J. Proteome Res.* 17 (2018) 1321–1325.
- [97] J. Eriksson, D. Fenýő, Improving the success rate of proteome analysis by modeling protein-abundance distributions and experimental designs, *Nat. Biotechnol.* 25 (2007) 651–655.
- [98] D.B. Bekker-Jensen, C.D. Kelstrup, T.S. Batth, S.C. Larsen, C. Haldrup, J. B. Bramsen, K.D. Sørensen, S. Hoyer, T.F. Ørntoft, C.L. Andersen, M.L. Nielsen, J.V. Olsen, An optimized shotgun strategy for the rapid generation of comprehensive human proteomes, *Cell Syst.* 4 (2017) 587–599.e4.
- [99] H. Mirzaei, M. Carrasco (Eds.), *Modern Proteomics – Sample Preparation, Analysis and Practical Applications*, Springer International Publishing, Cham, 2016.
- [100] J. Beri, M.M. Rosenblatt, E. Strauss, M. Urh, M.S. Bereman, Reagent for evaluating liquid chromatography–tandem mass spectrometry (LC-MS/MS) performance in bottom-up proteomic experiments, *Anal. Chem.* 87 (2015) 11635–11640.
- [101] M.S. Bereman, J. Beri, V. Sharma, C. Nathe, J. Eckels, B. MacLean, M.J. MacCoss, An automated pipeline to monitor system performance in liquid chromatography–tandem mass spectrometry proteomic experiments, *J. Proteome Res.* 15 (2016) 4763–4763–4769.
- [102] K. Gevaert, F. Impens, P. Van Damme, B. Ghesquière, X. Hanouille, J. Vandekerckhove, Applications of diagonal chromatography for proteome-wide characterization of protein modifications and activity-based analyses, *FEBS J.* 274 (2007) 6277–6289.
- [103] D. Gussakovsky, H. Neustaeter, V. Spicer, O.V. Krokhn, Sequence-specific model for peptide retention time prediction in strong cation exchange chromatography, *Anal. Chem.* 89 (2017) 11795–11802.
- [104] M.-Z. Zhu, N. Li, Y.-T. Wang, N. Liu, M.-Q. Guo, B. Sun, H. Zhou, L. Liu, J.-L. Wu, Acid/salt/pH gradient improved resolution and sensitivity in proteomics study using 2D SCX-RP LC–MS, *J. Proteome Res.* 16 (2017) 3470–3475.
- [105] O.V. Krokhn, P. Ezzati, V. Spicer, Peptide retention time prediction in hydrophilic interaction liquid chromatography: data collection methods and features of additive and sequence-specific models, *Anal. Chem.* 89 (2017) 5526–5533.

- [106] V.R. Richard, D. Domanski, A.J. Percy, C.H. Borchers, An online 2D-reversed-phase–reversed-phase chromatographic method for sensitive and robust plasma protein quantitation, *J. Proteomics* 168 (2017) 28–36.
- [107] H. Lee, D.-G. Mun, J.E. So, J. Bae, H. Kim, C. Masselon, S.-W. Lee, Efficient exploitation of separation space in two-dimensional liquid chromatography system for comprehensive and efficient proteomic analyses, *Anal. Chem.* 88 (2016) 11734–11734–11741.
- [108] V. Spicer, P. Ezzati, H. Neustaeter, R.C. Beavis, J.A. Wilkins, O.V. Krokhin, 3D HPLC-MS with reversed-phase separation functionality in all three dimensions for large-scale bottom-up proteomics and peptide retention data collection, *Anal. Chem.* 88 (2016) 2847–2855.
- [109] O.V. Krokhin, G. Anderson, V. Spicer, L. Sun, N.J. Dovichi, Predicting electrophoretic mobility of tryptic peptides for high-throughput CZE-MS analysis, *Anal. Chem.* 89 (2017) 2000–2008.
- [110] M. Di Venere, S. Viglio, D. Sasser, M. Fumagalli, A. Bardoni, R. Salvini, M. Cagnone, P. Iadarola, Do the complementarities of electrokinetic and chromatographic procedures represent the “Swiss knife” in proteomic investigation? An overview of the literature in the past decade: proteomics and 2DE, *Electrophoresis* 38 (2017) 1538–1550.
- [111] T. Vehus, H. Roberg-Larsen, J. Waaler, S. Aslaksen, S. Krauss, S.R. Wilson, E. Lundanes, Versatile, sensitive liquid chromatography mass spectrometry implementation of 10 μ m OT columns suitable for small molecules, peptides and proteins, *Sci. Rep.* (2016) 37507.
- [112] P. Yu, S. Petzoldt, M. Wilhelm, D.P. Zolg, R. Zheng, X. Sun, X. Liu, G. Schneider, A. Huhmer, B. Kuster, Trimodal mixed mode chromatography that enables efficient offline two-dimensional peptide fractionation for proteome analysis, *Anal. Chem.* 89 (2017) 8884–8891.
- [113] V. Gladilovich, U. Greifenhagen, N. Sukhodolov, A. Selyutin, D. Singer, D. Thieme, P. Majovsky, A. Shirkin, W. Hoehenwarter, E. Bonitenko, E. Podolskaya, A. Frolov, Immobilized metal affinity chromatography on collapsed Langmuir-Blodgett iron(III) stearate films and iron(III) oxide nanoparticles for bottom-up phosphoproteomics, *J. Chromatogr. A* 1443 (2016) 181–190.
- [114] B. Jiang, Q. Wu, L. Zhang, Y. Zhang, Preparation and application of silver nanoparticle-functionalized magnetic graphene oxide nanocomposites, *Nanoscale* 9 (2017) 1607–1615.
- [115] M. Olkowicz, A. Radulska, J. Suraj, A. Kij, M. Walczak, S. Chlopicki, R.T. Smolenski, Development of a sensitive, accurate and robust liquid chromatography/mass spectrometric method for profiling of angiotensin peptides in plasma and its application for atherosclerotic mice, *J. Chromatogr. A* 1393 (2015) 37–46.
- [116] D. Payto, C. Heideloff, S. Wang, Sensitive, simple, and robust nano-liquid chromatography-mass spectrometry method for amyloid protein subtyping, *Methods Mol. Biol.* 1378 (2016) 55–60.
- [117] P. Cifani, A. Kentsis, High sensitivity quantitative proteomics using automated multi-dimensional nano-flow chromatography and accumulated ion monitoring on Quadrupole-Orbitrap-linear ion trap mass spectrometer, *Mol. Cell. Proteomics* 16 (2017) 2006–2016.
- [118] Y. Zhu, R. Zhao, P.D. Pichowski, R.J. Moore, S. Lim, V.J. Orphan, L. Paša-Tolić, W.-J. Qian, R.D. Smith, R.T. Kelly, Subnanogram proteomics: impact of LC column selection, MS instrumentation and data analysis strategy on proteome coverage for trace samples, *Int. J. Mass Spectrom.* 427 (2018) 4–10.
- [119] R.T. Kelly, A.V. Tolmachev, J.S. Page, K. Tang, R.D. Smith, The ion funnel: theory, implementations, and applications, *Mass Spectrom. Rev.* 29 (2) (2010) 294–312.
- [120] B. Sun, J.R. Kovatch, A. Badiou, N. Merbouh, Optimization and modeling of quadrupole Orbitrap parameters for sensitive analysis toward single-cell proteomics, *J. Proteome Res.* 16 (2017) 3711–3721.
- [121] A. Michalski, J. Cox, M. Mann, More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS, *J. Proteome Res.* 10 (2011) 1785–1793.
- [122] G. Alves, A.Y. Ogurtsov, S. Kwok, W.W. Wu, G. Wang, R.-F. Shen, Y.-K. Yu, Detection of co-eluted peptides using database search methods, *Biol. Direct.* 3 (2008) 27.
- [123] J.D. Chapman, D.R. Goodlett, C.D. Masselon, Multiplexed and data-independent tandem mass spectrometry for global proteome profiling, *Mass Spectrom. Rev.* 33 (2014) 452–470.
- [124] G.H.M.F. Souza, P.C. Guest, D. Martins-de-Souza, LC-MSE, multiplex MS/MS, ion mobility, and label-free quantitation in clinical proteomics, *Methods Mol. Biol.* 1546 (2017) 57–73.
- [125] L.C. Gillet, P. Navarro, S. Tate, H. Röst, N. Selevsek, L. Reiter, R. Bonner, R. Aebersold, Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis, *Mol. Cell. Proteomics* 11 (2012) O111.016717.
- [126] C.R. Weisbrod, J.K. Eng, M.R. Hoopmann, T. Baker, J.E. Bruce, Accurate peptide fragment mass analysis: multiplexed peptide identification and quantification, *J. Proteome Res.* 11 (2012) 1621–1632.
- [127] J.D. Egerton, A. Kuehn, G.E. Merrihew, N.W. Bateman, B.X. MacLean, Y.S. Ting, J.D. Canterbury, D.M. Marsh, M. Kellmann, V. Zabrouskov, C.C. Wu, M.J. MacCoss, Multiplexed MS/MS for improved data-independent acquisition, *Nat. Methods* 10 (2013) 744–746.

- [128] O.V. Krokhn, V. Spicer, Peptide retention standards and hydrophobicity indexes in reversed-phase high-performance liquid chromatography of peptides, *Anal. Chem.* 81 (2009) 9522–9530.
- [129] S. Sidoli, R. Fujiwara, B.A. Garcia, Multiplexed data independent acquisition (MSX-DIA) applied by high resolution mass spectrometry improves quantification quality for the analysis of histone peptides, *Proteomics* 16 (2016) 2095–2105.
- [130] A.C. Peterson, J.D. Russell, D.J. Bailey, M.S. Westphall, J.J. Coon, Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics, *Mol. Cell. Proteomics* 11 (2012) 1475–1488.
- [131] S. Gallien, A. Bourmaud, S.Y. Kim, B. Domon, Technical considerations for large-scale parallel reaction monitoring analysis, *J. Proteomics* 100 (2014) 147–159.
- [132] N. Rauniyar, Parallel reaction monitoring: a targeted experiment performed using high resolution and high mass accuracy mass spectrometry, *Int. J. Mol. Sci.* 16 (2015) 28566–28581.
- [133] O.V. Krokhn, V. Spicer, Generation of accurate peptide retention data for targeted and data independent quantitative LC-MS analysis: chromatographic lessons in proteomics, *Proteomics* 16 (2016) 2931–2936.
- [134] W. Haas, B.K. Faherty, S.A. Gerber, J.E. Elias, S.A. Beausoleil, C.E. Bakalarski, X. Li, J. Villén, S.P. Gygi, Optimization and use of peptide mass measurement accuracy in shotgun proteomics, *Mol. Cell. Proteomics* 5 (2006) 1326–1337.
- [135] F. Meier, P.E. Geyer, S. Virreira Winter, J. Cox, M. Mann, BoxCar acquisition method enables single-shot proteomics at a depth of 10,000 proteins in 100 minutes, *Nat. Methods* 15 (2018) 440–448.
- [136] J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification, *Nat. Biotechnol.* 26 (2008) 1367–1372.
- [137] J.A. Ankney, A. Muneer, X. Chen, Relative and absolute quantitation in mass spectrometry-based proteomics, *Annu. Rev. Anal. Chem.* 11 (2018) 49–77.
- [138] S.-E. Ong, B. Blagoev, I. Kratchmarova, D.B. Kristensen, H. Steen, A. Pandey, M. Mann, Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics, *Mol. Cell. Proteomics* 1 (2002) 376–386.
- [139] H. Zhu, S. Pan, S. Gu, E.M. Bradbury, X. Chen, Amino acid residue specific stable isotope labeling for quantitative proteomics, *Rapid Commun. Mass Spectrom.* 16 (2002) 2115–2123.
- [140] X. Ye, B. Luke, T. Andreasson, J. Blonder, ¹⁸O stable isotope labeling in MS-based proteomics, *Brief. Funct. Genomic. Proteomic.* 8 (2009) 136–144.
- [141] A. Thompson, J. Schaefer, K. Kuhn, S. Kienle, J. Schwarz, G. Schmidt, T. Neumann, R. Johnstone, A.K.A. Mohammed, C. Hamon, Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS, *Anal. Chem.* 75 (2003) 1895–1904.
- [142] A.E. Merrill, A.S. Hebert, M.E. MacGilvray, C.M. Rose, D.J. Bailey, J.C. Bradley, W.W. Wood, M. El Masri, M.S. Westphall, A.P. Gasch, J.J. Coon, NeuCode labels for relative protein quantification, *Mol. Cell. Proteomics* 13 (2014) 2503–2512.
- [143] G.K. Potts, E.A. Voigt, D.J. Bailey, C.M. Rose, M.S. Westphall, A.S. Hebert, J. Yin, J.J. Coon, Neucode labels for multiplexed, absolute protein quantification, *Anal. Chem.* 88 (2016) 3295–3303.
- [144] A.F.M. Altelaar, C.K. Frese, C. Preisinger, M.L. Hennrich, A.W. Schram, H.T. M. Timmers, A.J.R. Heck, S. Mohammed, Benchmarking stable isotope labeling based quantitative proteomics, *J. Proteomics* 88 (2013) 14–26.
- [145] M.D. Sury, J.-X. Chen, M. Selbach, The SILAC fly allows for accurate protein quantification in vivo, *Mol. Cell. Proteomics MCP.* 9 (2010) 2173–2183.
- [146] M. Miyagi, Monitoring protein synthesis in *Caenorhabditis elegans* using SILAC, *Methods Enzymol.* 585 (2017) 77–89.
- [147] I.N. Kaneva, J. Longworth, P.E. Sudbery, M.J. Dickman, Quantitative proteomic analysis in *Candida albicans* using SILAC-based mass spectrometry, *Proteomics* 18 (2018) e1700278.
- [148] J.M. Smekens, W. Chen, R. Wu, Mass spectrometric analysis of the cell surface N-glycoproteome by combining metabolic labeling and click chemistry, *J. Am. Soc. Mass Spectrom.* 26 (2015) 604–614.
- [149] S.Y. Ow, M. Salim, J. Noirel, C. Evans, P.C. Wright, Minimising iTRAQ ratio compression through understanding LC-MS elution dependence and high-resolution HILIC fractionation, *Proteomics* 11 (2011) 2341–2346.
- [150] Y. Di, Y. Zhang, L. Zhang, T. Tao, H. Lu, MdFDIA: a mass defect based four-plex data-independent acquisition strategy for proteome quantification, *Anal. Chem.* 89 (2017) 10248–10255.
- [151] B. Zybailov, A.L. Mosley, M.E. Sardi, M.K. Coleman, L. Florens, M.P. Washburn, Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*, *J. Proteome Res.* 5 (2006) 2339–2347.
- [152] B. Schwanhäusser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, M. Selbach, Global quantification of mammalian gene expression control, *Nature.* 473 (2011) 337–342.
- [153] P. Lu, C. Vogel, R. Wang, X. Yao, E.M. Marcotte, Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation, *Nat. Biotechnol.* 25 (2007) 117–124.

- [154] J.C. Silva, M.V. Gorenstein, G.-Z. Li, J.P.C. Vissers, S.J. Geromanos, Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition, *Mol. Cell. Proteomics* 5 (2006) 144–156.
- [155] M. Blein-Nicolas, M. Zivy, Thousand and one ways to quantify and compare protein abundances in label-free bottom-up proteomics, *Biochim. Biophys. Acta* 1864 (2016) 883–895.
- [156] C. Chang, Z. Gao, W. Ying, Y. Zhao, Y. Fu, S. Wu, M. Li, G. Wang, X. Qian, Y. Zhu, F. He, LFAQ: towards unbiased label-free absolute protein quantification by predicting peptide quantitative factors, *Anal. Chem.* (2018) Epub 2018 Dec 10.
- [157] K.A. Kovalchik, S. Moggridge, D.D.Y. Chen, G.B. Morin, C.S. Hughes, Parsing and quantification of raw Orbitrap mass spectrometer data using RawQuant, *J. Proteome Res.* 17 (6) (2018) 2237–2247.
- [158] M. Sonnett, E. Yeung, M. Wehr, Accurate, sensitive, and precise multiplexed proteomics using the complement reporter ion cluster, *Anal. Chem.* 90 (2018) 5032–5039.
- [159] M.J. Berberich, J.A. Paulo, R.A. Everley, MS3-IDQ: utilizing MS3 spectra beyond quantification yields increased coverage of the phosphoproteome in isobaric tag experiments, *J. Proteome Res.* 17 (2018) 1741–1747.
- [160] B. Bai, H. Tan, V.R. Pagala, A.A. High, V.P. Ichhaporia, L. Hendershot, J. Peng, Deep profiling of proteome and phosphoproteome by isobaric labeling, extensive liquid chromatography, and mass spectrometry, *Methods Enzymol.* 585 (2017) 377–395.
- [161] R.D. Smith, J.A. Loo, C.G. Edmonds, C.J. Barinaga, H.R. Udseth, New developments in biochemical mass spectrometry: electrospray ionization, *Anal. Chem.* 62 (1990) 882–899.
- [162] J.A. Loo, C.G. Edmonds, R.D. Smith, Tandem mass spectrometry of very large molecules: serum albumin sequence information from multiply charged ions formed by electrospray ionization, *Anal. Chem.* 63 (1991) 2488–2499.
- [163] N.L. Kelleher, S.V. Taylor, D. Grannis, C. Kinsland, H.-J. Chiu, T.P. Begley, F. W. McLafferty, Efficient sequence analysis of the six gene products (7–74 kDa) from the *Escherichia coli* thiamin biosynthetic operon by tandem high-resolution mass spectrometry, *Protein Sci.* 7 (1998) 1796–1801.
- [164] A.D. Catherman, O.S. Skinner, N.L. Kelleher, Top down proteomics: facts and perspectives, *Biochem. Biophys. Res. Commun.* 445 (2014) 683–693.
- [165] C. Kachuk, A.A. Doucette, The benefits (and misfortunes) of SDS in top-down proteomics, *J. Proteomics.* 175 (2018) 75–86.
- [166] N. Unterlander, A.A. Doucette, Membrane-based SDS depletion ahead of peptide and protein analysis by mass spectrometry, *Proteomics* 18 (2018) e1700025.
- [167] A.L. Capriotti, C. Cavaliere, P. Foglia, R. Samperi, A. Laganà, Intact protein separation by chromatographic and/or electrophoretic techniques for top-down proteomics, *J. Chromatogr. A* 1218 (2011) 8760–8776.
- [168] Y. Shen, N. Tolić, P.D. Pichowski, A.K. Shukla, S. Kim, R. Zhao, Y. Qu, E. Robinson, R.D. Smith, L. Paša-Tolić, High-resolution ultrahigh-pressure long column reversed-phase liquid chromatography for top-down proteomics, *J. Chromatogr. A* 1498 (2017) 99–110.
- [169] H. Kanazawa, T. Okano, Temperature-responsive chromatography for the separation of biomolecules, *J. Chromatogr. A* 1218 (2011) 8738–8747.
- [170] M.J. Roth, D.A. Plymire, A.N. Chang, J. Kim, E.M. Maresh, S.E. Larson, S.M. Patrie, Sensitive and reproducible intact mass analysis of complex protein mixtures with superficially porous capillary reversed-phase liquid chromatography mass spectrometry, *Anal. Chem.* 83 (2011) 9586–9592.
- [171] F. Yang, Y. Shen, D.G. Camp, R.D. Smith, High-pH reversed-phase chromatography with fraction concatenation for 2D proteomic analysis, *Expert Rev. Proteomics* 9 (2012) 129–134.
- [172] Z. Wang, H. Ma, K. Smith, S. Wu, Two-dimensional separation using high-pH and low-pH reversed phase liquid chromatography for top-down proteomics, *Int. J. Mass Spectrom.* 427 (2018) 43–51.
- [173] A.F.G. Gargano, L.S. Roca, R.T. Fellers, M. Bocxe, E. Domínguez-Vega, G.W. Somsen, Capillary HILIC-MS: a new tool for sensitive top-down proteomics, *Anal. Chem.* 90 (2018) 6601–6609.
- [174] E. Domínguez-Vega, S. Tengattini, C. Peintner, J. van Angeren, C. Temporini, R. Haselberg, G. Massolini, G.W. Somsen, High-resolution glycoform profiling of intact therapeutic proteins by hydrophilic interaction chromatography-mass spectrometry, *Talanta* 184 (2018) 375–381.
- [175] D.C. Simpson, S. Ahn, L. Pasa-Tolić, B. Bogdanov, H.M. Mottaz, A.N. Vilkov, G.A. Anderson, M.S. Lipton, R.D. Smith, Using size exclusion chromatography-RPLC and RPLC-CIEF as two-dimensional separation strategies for protein profiling, *Electrophoresis* 27 (2006) 2722–2733.
- [176] S.G. Valeja, L. Xiu, Z.R. Gregorich, H. Guner, S. Jin, Y. Ge, Three dimensional liquid chromatography coupling ion exchange chromatography/hydrophobic interaction chromatography/reverse phase chromatography for effective protein separation in top-down proteomics, *Anal. Chem.* 87 (2015) 5363–5371.
- [177] H. Kosako, K. Motani, Global identification of ERK substrates by phosphoproteomics based on IMAC and 2D-DIGE, *Methods Mol. Biol.* 1487 (2017) 137–149.
- [178] K. Motani, H. Kosako, Phosphoproteomic identification and functional characterization of protein kinase substrates by 2D-DIGE and Phos-tag PAGE, *Biochim. Biophys. Acta* 186 (1) (2019) 57–61.

- [179] B.T. Kurien, R.H. Scofield, Extraction of proteins from gels: a brief review, *Methods Mol. Biol.* 869 (2012) 403–405.
- [180] J.E. Lee, J.F. Kellie, J.C. Tran, J.D. Tipton, A.D. Catherman, H.M. Thomas, D.R. Ahlf, K.R. Durbin, A. Vellaichamy, I. Ntai, A.G. Marshall, N.L. Kelleher, A robust two-dimensional separation for top-down tandem mass spectrometry of the low-mass proteome, *J. Am. Soc. Mass Spectrom.* 20 (2009) 2183–2191.
- [181] Y. Zhao, L. Sun, G. Zhu, N.J. Dovichi, Coupling capillary zone electrophoresis to a Q exactive HF mass spectrometer for top-down proteomics: 580 proteoform identifications from yeast, *J. Proteome Res.* 15 (2016) 3679–3685.
- [182] R.A. Lubeckyj, E.N. McCool, X. Shen, Q. Kou, X. Liu, L. Sun, Single-shot top-down proteomics with capillary zone electrophoresis-electrospray ionization-tandem mass spectrometry for identification of nearly 600 *Escherichia coli* proteoforms, *Anal. Chem.* 89 (2017) 12059–12067.
- [183] E.N. McCool, R.A. Lubeckyj, X. Shen, D. Chen, Q. Kou, X. Liu, L. Sun, Deep top-down proteomics using capillary zone electrophoresis-tandem mass spectrometry: identification of 5700 proteoforms from the *Escherichia coli* proteome, *Anal. Chem.* 90 (2018) 5529–5533.
- [184] R.W. Nelson, D. Dogruel, P. Williams, Mass determination of human immunoglobulin IgM using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *Rapid Commun. Mass Spectrom.* 8 (1994) 627–631.
- [185] M.E. Belov, E.N. Nikolaev, G.A. Anderson, H.R. Udseth, T.P. Conrads, T.D. Veenstra, C.D. Masselon, M.V. Gorshkov, R.D. Smith, Design and performance of an ESI interface for selective external ion accumulation coupled to a Fourier transform ion cyclotron mass spectrometer, *Anal. Chem.* 73 (2001) 253–261.
- [186] M.E. Belov, R. Zhang, E.F. Strittmatter, D.C. Prior, K. Tang, R.D. Smith, Automated gain control and internal calibration with external ion accumulation capillary liquid chromatography-electrospray ionization Fourier transform ion cyclotron resonance, *Anal. Chem.* 75 (2003) 4195–4205.
- [187] S.M. Patrie, J.P. Charlebois, D. Whipple, N.L. Kelleher, C.L. Hendrickson, J.P. Quinn, A.G. Marshall, B. Mukhopadhyay, Construction of a hybrid quadrupole/Fourier transform ion cyclotron resonance mass spectrometer for versatile MS/MS above 10 kDa, *J. Am. Soc. Mass Spectrom.* 15 (2004) 1099–1108.
- [188] A. Makarov, E. Denisov, Dynamics of ions of intact proteins in the Orbitrap mass analyzer, *J. Am. Soc. Mass Spectrom.* 20 (2009) 1486–1495.
- [189] D. Nolting, R. Malek, A. Makarov, Ion traps in modern mass spectrometry, *Mass Spec. Rev.* 38 (2019) 50–168.
- [190] N.M. Riley, M.S. Westphall, J.J. Coon, Activated ion-electron transfer dissociation enables comprehensive top-down protein fragmentation, *J. Proteome Res.* 16 (2017) 2653–2659.
- [191] N.M. Riley, M.S. Westphall, J.J. Coon, Sequencing larger intact proteins (30–70 kDa) with activated ion electron transfer dissociation, *J. Am. Soc. Mass Spectrom.* 29 (2018) 140–149.
- [192] S. Warnke, G. von Helden, K. Pagel, Analyzing the higher order structure of proteins with conformer-selective ultraviolet photodissociation, *Proteomics* 15 (2015) 2804–2812.
- [193] L.J. Morrison, J.S. Brodbelt, Charge site assignment in native proteins by ultraviolet photodissociation (UVPD) mass spectrometry, *Analyst* 141 (2016) 166–176.
- [194] L. Fornelli, K. Srzentić, R. Huguet, C. Mullen, S. Sharma, V. Zabrouskov, R.T. Fellers, K.R. Durbin, P.D. Compton, N.L. Kelleher, Accurate sequence analysis of a monoclonal antibody by top-down and middle-down Orbitrap mass spectrometry applying multiple ion activation techniques, *Anal. Chem.* 90 (14) (2018) 8421–8429.
- [195] Y. Feng, G. De Franceschi, A. Kahraman, M. Soste, A. Melnik, P.J. Boersema, P.P. de Laureto, Y. Nikolaev, A.P. Oliveira, P. Picotti, Global analysis of protein structural changes in complex proteomes, *Nat. Biotechnol.* 32 (2014) 1036–1044.
- [196] S. Schopper, A. Kahraman, P. Leuenberger, Y. Feng, I. Piazza, O. Müller, P.J. Boersema, P. Picotti, Measuring protein structural changes on a proteome-wide scale using limited proteolysis-coupled mass spectrometry, *Nat. Protoc.* 12 (2017) 2391–2410.
- [197] G. Wang, I.A. Kaltashov, Approach to characterization of the higher order structure of disulfide-containing proteins using hydrogen/deuterium exchange and top-down mass spectrometry, *Anal. Chem.* 86 (2014) 7293–7298.
- [198] K.D. Rand, M. Zehl, T.J.D. Jørgensen, Measuring the hydrogen/deuterium exchange of proteins at high spatial resolution by mass spectrometry: overcoming gas-phase hydrogen/deuterium scrambling, *Acc. Chem. Res.* 47 (2014) 3018–3027.
- [199] E.H. Seeley, K. Schwamborn, R.M. Caprioli, Imaging of intact tissue sections: moving beyond the microscope, *J. Biol. Chem.* 286 (2011) 25459–25466.
- [200] B.M. Prentice, D.J. Ryan, R. Van de Plas, R.M. Caprioli, J.M. Spraggins, Enhanced ion transmission efficiency up to m/z 24 000 for MALDI protein imaging mass spectrometry, *Anal. Chem.* 90 (2018) 5090–5099.
- [201] M.K. Passarelli, A. Pirkl, R. Moellers, D. Grinfeld, F. Kollmer, R. Havelund, C.F. Newman, P.S. Marshall, H. Arlinghaus, M.R. Alexander, A. West, S. Horning, E. Niehuis, A. Makarov, C.T. Dollery, I.S. Gilmore, The 3D OrbiSIMS-label-free metabolic imaging with subcellular lateral resolution and high mass-

- resolving power, *Nat. Methods* 14 (2017) 1175–1183.
- [202] R. Longuespée, R. Casadonte, M. Kriegsmann, C. Pottier, G. Picard de Muller, P. Delvenne, J. Kriegsmann, E. De Pauw, MALDI mass spectrometry imaging: a cutting-edge tool for fundamental and clinical histopathology, *Proteomics Clin. Appl.* 10 (2016) 701–719.
- [203] M. Wisztorski, J. Quanico, J. Franck, B. Fatou, M. Salzert, I. Fournier, Droplet-based liquid extraction for spatially-resolved microproteomics analysis of tissue sections, *Methods Mol. Biol.* 1618 (2017) 49–63.
- [204] S. Si Ahmed Zennia, A. Mati, F. Saulnier, Y. Verdier, G. Chiappetta, G. Mulliert, L. Miclo, J. Vinh, J.-M. Girardet, Identification by FT-ICR-MS of *Camelus dromedarius* α -lactalbumin variants as the result of nonenzymatic deamidation of Asn-16 and Asn-45, *Food Chem.* 187 (2015) 305–313.