Positional proteomics for identification of secreted proteoforms released by site-specific processing of membrane proteins

Stefan Niedermaier, Pitter F. Huesgen*

Forschungszentrum Jülich

Central Institute for Engineering, Electronics and Analytics, ZEA-3 Analytics

Wilhelm-Johnen-Str.

52425 Jülich

* Corresponding author: Pitter F. Huesgen

Email: p.huesgen@fz-juelich.de

Phone: +49 2461 61-6873

Fax: +49 2461 61-2560

This is the author version of the article after peer review. The final version of this review article has been published by Biochimica et Biophysica Acta – Proteins and Proteomics, © 2018 Elsevier B.V., and is available at doi.org/10.1016/j.bbapap.2018.09.004.

Abstract (max 150 words)

Proteolytic processing shapes cellular interactions with the environment. As a pathway of unconventional protein secretion, ectodomain shedding releases soluble proteoforms of membrane-anchored proteins. This can trigger subsequent cleavage within the membrane stub and the release of additional soluble fragments to intra- and extracellular environments. Distinct membrane-bound proteases, or sheddases, may cleave the same membrane proteins at different sites. Determination of these precise cleavage sites is important, as differently processed proteoforms may exhibit distinct physical properties and execute antagonistic paracrine and endocrine signaling functions. Conventional quantitative proteomic approaches reliably identify shed proteoforms, but typically not their termini and are thus not able distinguish between functionally different proteoforms differing only by a few amino acids. Dedicated positional proteomics overcomes this challenge and enables proteome-wide identification of protein N- and C-termini. Here, we review positional proteomics techniques, summarize their application to ectodomain shedding and discuss current challenges and developments.

Keywords

positional proteomics; proteolysis; protein termini enrichment; degradomics; ectodomain shedding; sheddase

1. Introduction

Limited proteolytic processing fundamentally contributes to cellular interactions with their environment by modulation of the functional cell surface proteome. Proteolytic release of membrane protein extracellular domains, a process known as ectodomain shedding [1], generates new soluble proteoforms with distinct biological functions. Ectodomain shedding is not a rare form of unconventional protein secretion, as a large proportion of all membrane-bound proteins, including cell adhesion molecules, growth factors, immunomodulators, and their respective receptors, are known to be subjected to proteolytic cleavage close to the cell surface [1-3]. Ectodomain shedding regulates cell-cell and cell-matrix interactions [4], releases cytokines, chemokines and growth factors with autocrine, paracrine and endocrine signaling functions [5], activates or inactivates receptor proteins [6, 7], and regulates enzymatic activities such as proteolysis at the cell membrane and in the extracellular environment [8]. Furthermore, ectodomain shedding often constitutes the rate-limiting step of a signaling mechanism called regulated intramembrane proteolysis (RIP) [9]. In RIP,ectodomain shedding is followed by a second proteolytic cleavage within the target protein transmembrane segment, releasing additional fragments to the extracellular environment and into the cytosol (Fig. 1). All fragments may be further processed by a wide range of extracellular endo- and exopeptidases, resulting in a large variety of different proteoforms that may exhibit different biological functions and activities. Notably, all of these unconventionally secreted proteoforms are clearly defined by their unique protease-generated (neo-)N- or C-termini (Fig. 1).

The majority of ectodomain shedding events identified to date are executed by metalloproteases of the ADAM and MMP families [2, 3], but also other membrane-bound proteases such as meprin β [10] and the aspartic proteases BACE1 and BACE2 contribute to shedding [11]. Additionally, soluble extracellular proteases can cleave within extracellular domains or even peripherally attach to cell membranes for cleavage close to the membrane, resulting in sheddase activity [12]. Importantly, different sheddases cleave the same membrane substrates at distinct sites and thereby release distinct proteoforms that may lead to dramatically different physiological consequences (Fig. 1). Arguably the most prominent example is the amyloid precursor protein (APP) in the brain, where sequential processing by beta site cleaving enzyme 1 (BACE1) and the y-secretase complex releases the soluble ectodomain sAPPβ and the aggregation-prone Aβ peptides to the extracellular environment [11]. Formation of neurotoxic Aβ oligomers, caused by increased Aβ production or impaired clearance, is considered as a very early step in the pathogenesis of Alzheimer´s disease [13]. However, in normal conditions APP is mostly processed by ADAM10 at the α-site within the Aβ sequence, leading to the release of a larger soluble APPα ectodomain and the non-amyloidogenic P3 peptide [11]. Also meprin β has been shown to cleave APP at and close to the β-site, generating more toxic forms of Aβ [14]. Additional APP cleavages by various proteases at more membrane-

distant sites have been described [11]. This illustrates that precise definition of different proteoforms released by alternative proteolytic processing can be important to understand their biological function and physiological consequences.

2. Proteomic identification of proteolytically released proteoforms

Mass spectrometry-based proteomics is the method of choice for the characterization of proteins and their post-translational modifications in biological systems [15]. Standard proteomics workflows require digestion of proteins with sequence-specific proteases such as trypsin to generate predictable peptides amenable to mass spectrometry-based identification. Peptide ions are fragmented by MS/MS and the acquired spectra are matched to theoretical spectra of peptide sequences predicted from proteome databases using the digestion protease specificity as constraint. Protein identities are computationally inferred from the identified peptide sequences. The observed peptide coverage depends on the specificity of the digestion protease and, in principle, covers the entire protein sequence. Therefore, differentially processed proteoforms cannot be distinguished by this approach in whole cell or tissue lysates. Consequently, additional experimental constraints such as analysis of secreted proteomes collected from cell culture-conditioned media need to be implemented to distinguish between intact and sheddase-processed, soluble proteoforms [16]. In gel-based approaches, visual information on apparent molecular weight of a protein is used to differentiate shed proteoforms from their larger, membrane-bound precursor proteins [17, 18]. Shedding-inducers such as lipopolysaccharide (LPS) or 12-O-tetradecanoylphorbol 13-acetate (TPA) have been used to elucidate potential shedding candidates [16]. Subcellular fractionation and/or pre-enrichment of secreted proteins reduces sample complexity and provides additional evidence that the observed peptides actually derive from secreted proteoforms. Since most surface membrane proteins are glycosylated or predicted to be glycosylated, enrichment of this modification is a promising approach to define shed proteoforms. For example, enrichment of glycopeptides using the established hydrazide chemistry [19] enabled identification of 328 glycoproteins in the cell culture supernatants of N1E-115 neuroblastoma cells, including the ectodomains of 18 proteins that were shed in a metalloprotease-dependent manner [20]. However, most secretome analyses require work with serum-depleted media, which is not compatible with many sensitive cell lines including primary culture of neurons. An elegant method that overcomes this limitation is **S**ecretome **P**rotein **E**nrichment with **C**lick **S**ugars (SPECS), which introduces a synthetic azido sugar as metabolic glycan label for click chemistry-mediated biotinylation [21]. In this approach only newly synthesized glycoproteins are labeled and biotinylated, enabling distinction between cellular and serum glycoproteins. In screens for novel sheddase substrates, SPECS identified 34 candidate BACE1 substrates [21] and 91 candidate ADAM10 substrates [22]. Recent refinements of the procedure allowed increased coverage with half the amount of input material [23]. However, identification of shed proteoforms by

glycoprotein enrichment remains limited to cell culture systems due to the requirements for metabolic labeling and/or secretome collection and provides only limited information on processing sites.

3. Terminomics defines distinct proteolytic proteoforms

Sheddase-generated neo-N and neo-C-termini unambiguously distinguish distinct proteoforms generated by cleavage at different sites, even if they differ only by a few amino acids (Fig. 1). However, information on N- and C-terminal peptides is missing in standard "bottom-up" proteomics workflows that typically only consider peptides delimited by the specificity of the protease used for digestion of the proteome to peptides. In addition, protein N- and C-terminal peptides constitute only a minor fraction among all peptides in complex digest mixtures and are therefore rarely identified even in semi-specific searches, which require a cleavage site of the digestion protease only on one side of the peptide. To overcome this challenge, which is equivalent to the analysis of other post-translational protein modifications, dedicated enrichment protocols for protein termini have been developed. Two basic strategies can be distinguished: Positive selection protocols directly access the termini of interest, while negative selection protocols chemically modify proteins before digestion and use digestion-generated functionalities to remove the protein internal peptides. In all cases, the terminal peptides originally present in the analyzed proteome are modified by a chosen modification, which adds confidence in the identification and excludes non-specific hydrolysis during sample preparation.

3.1. Enrichment of protein N termini by positive selection

A major challenge in N-terminal peptide enrichment is the fact that a protein N termini of interest shares the same functional primary α-amine group with internal peptides arising from proteolytic digestion during sample preparation. Furthermore, side-chains of lysine residues contain Ɛ-amines with similar reactivity as primary α-amine groups. Therefore, positive enrichment protocols need to be carried out under conditions that distinguish the reactivity of those amine groups to avoid co-enrichment of lysine-containing peptides. A very elegant solution for this problem is selective enzyme-mediated modification of protein N-terminal α-amines using subtiligase, a rationally engineered peptide ligase derived from subtilisin BPN [24]. Incubation with subtiligase and a synthetic peptide ester substrate selectively marks protein N-termini, typically with a biotinylated peptide containing a cleavage site for the highly sequence specific TEV-protease **(Fig. 2A)**[25]. After proteolytic digestion of the labeled proteins, the biotin-peptide-tagged N-terminal peptides are captured with immobilized avidin and separated from internal peptides by stringent washing. N-terminal peptides are then released from the avidin beads by TEV protease cleavage for identification by LC-MS/MS. Several alternative peptide ester tags have been developed [26] and a step-by-step protocol has been published [27]. Recently, a library of subtiligase variants has been developed to overcome biases introduced by subtiligase sequence specificity [28].

An alternative method employs selective chemical biotinylation of protein N termini [29]. In this protocol, the slight difference in α- and ε-amine reactivity is exploited to selectively convert lysine residues to homoarginine using o-methylisourea. N-terminal α-amines remain unblocked, are then covalently modified with NHS-SS-biotin and captured and enriched by immobilized streptavidin **(Fig. 2B)**.

While conceptually appealing for the identification of protease-generated neo-N termini, positive selection techniques suffer from inefficient labeling attributed to peptide ester hydrolysis [27] or required experimentally challenging selective ε-amine modification [29]. Furthermore, quantitative comparison of N-termini from different samples requires metabolic labeling as stable isotopes are not readily introduced during the procedures.

3.2. Enrichment of protein N termini by negative selection

Strategies for enrichment of protein N-terminal peptides by negative selection overcome the need to discriminate between α- and ε-amine reactivity and allow flexible incorporation of several stable isotope labels [30]. In a first step, all primary amines are chemically modified prior to proteolytic digestion **(Fig. 3A)**. In most protocols, stable isotope labels are simultaneously introduced, e.g. by reductive dimethylation with formaldehyde isotopes or other amine-reactive reagents such as iTRAQ (isobaric tags for absolute and relative quantification) or TMT (tandem mass tag). Differentially labeled samples are subsequently combined and analyzed together, which prevents differential losses during the enrichment procedure and enables confident relative quantification across different conditions and/or genotypes. Labeled proteomes are then digested with sequence-specific endoproteases such as trypsin or GluC. The resulting internal and C-terminal peptides exhibit new α-amines generated by the digestion protease. Several alternative methods either covalently capture the α-amine-containing non-N-terminal peptides or introduce a chemical modification that alters their chemical or physical properties, allowing separation from the previously modified, inert protein N-terminal peptides. All of these negative selection techniques result in high enrichment of protein N-terminal peptides that are further identified by LC-MS/MS.

3.2.1. Enrichment of protein N termini by sequential chromatography

COmbined **FRA**ctional **DI**agonal **C**hromatography (COFRADIC) uses two subsequent chromatography steps to enrich N-terminal peptides [31-33]. Primary amines of proteins are acetylated, digested with trypsin and the resulting peptides are separated by a first RP-HPLC fractionation **(Fig. 3B)**. Each fraction contains a mix of N-terminal peptides with blocked primary amines and digestion-generated peptides with free primary amines, which are then modified with a trinitrophenyl group (TNBS). Each fraction is subjected to a second RP-HPLC separation, where TNBS-modified peptides exhibit a strong retention time shift while N-terminal peptides retain their previous retention-time **(Fig. 3B)**. A modified version of the COFRADIC

protocol introduced an additional SCX-based pre-enrichment, exploiting the fact that N-terminally acetylated N-terminal peptides carry one less charge than internal peptides generated by tryptic digest [34]. A detailed step-by-step description of the COFRADIC protocol is published [32]. Inspired by COFRADIC, **Cha**rge-base **FRA**ctional **DI**agonal **C**hromatography (ChaFRADIC) similarly uses two sequential chromatography steps, but applies SCX chromatography for separation and modulates the net charge of digestion-generated internal and C-terminal peptides before the second chromatography run [35, 36]. In a recent modification, the requirement for HPLC has been substitutes with pipette tip-based SCX-tips, enabling confident N-termini identification with less than 50 µg starting material [37].

3.2.2 Enrichment of protein N termini by covalent capture of non-N-terminal peptides

In **T**erminal **A**mine **I**sotope **L**abeling of **S**ubstrates (TAILS), primary amines can be modified with any amine-reactive reagent, which is frequently used to simultaneously introduce a stable isotope label. After proteome-digest, a soluble high-molecular weight, aldehyde-functionalized polymer (HPG-ALD) is used to covalently capture internal and C-terminal peptides exposing digestion protease-generated primary amines. Subsequent ultrafiltration separates the polymer with bound peptides, while blocked N-terminal peptides elute in the flow-through **(Fig. 3C).** TAILS has been extensively optimized and detailed step-by-step protocols [38, 39] as well as protocols for application to plant proteomes [40] and body fluids are available [41].

Several alternative strategies for depletion of internal peptides have been developed, but not widely applied yet. One of the earliest procedures for negative selection of N-terminal peptides used commercial scavenger materials [42]. In the first variant of the protocol, α-and Ɛ-amines on the intact proteins were blocked by acetylation, yielding a mixture of acetylated N-terminal peptides and non-acetylated internal and C-terminal peptides that were covalently modified with NHS-biotin. Biotinylated peptides were then removed by passing the mixture over an immobilized streptavidin column, leaving N-terminal peptide in the flow through. In a further simplified protocol **(Fig. 3D)**, internal and C-terminal peptides generated by digestion were scavenged by covalent reaction of their free primary amine with activated NHS-Sepharose beads, allowing their depletion by centrifugation [43]**.** In a variation of this theme, the popular Filter-Aided Sample Preparation (FASP) protocol [44] was used for digestion of proteins after amine modification. Designated "NRich", internal peptides are then captured by NHS-activated agarose and removed using a second filtration with spin filters [45].

3.2.3 Additional methods for depletion of non-N-terminal peptides

A completely different approach modified the free primary amine group of internal and C-terminal peptides with a phospho-tag (PTAG) [46]. The phospho-tagged peptides were then retained on TiO_2-beads, which

are usually used for phosphopeptide-enrichment, while N-terminal peptides eluting in the flow-through **(Fig. 3E).** In a charge-reversal approach, free amines of peptides generated by proteome digest are modified with 4-formyl-1,3-benzenedisulfonic acid, adding multiple negative charges [47]. A subsequent SCX fractionation retains N-terminal peptides, while negatively charged peptides pass through the column in the flow through **(Fig. 3F)**. In combination with miniature SCX pipette tip columns, N-terminal peptides could be identified from minute amounts of sample present in excised gel slices. Recently, an updated step-by-step protocol of the charge reversal method has been published [48]. In hydrophobic tagging-assisted N-termini enrichment (HYTANE), primary amines of internal peptides are modified with a very hydrophobic alkyl-chain tag to subsequently deplete them on a C18 trap column **(Fig. 3G)** [49]**.** A distinct advantage of this approach is that the internal-peptide depletion step by offline C18 trap column is also desalting the sample, avoiding additional desalting steps such as offline C18-tip cleanup and therefore minimizing sample loss. Another innovative method termed "STagAu" isolated N-termini by depletion of internal peptides with gold-nanoparticles [50]. In this procedure, digestion protease-generated primary amines are modified with a sulfydryl-moiety by Traut's reagent. The sulfydrylated peptides are then bound to gold-nanoparticle composite materials and separated from inert N-terminal peptides by magnetic separation **(Fig. 2H).**

Additional steps can be included into negative enrichment protocols, e.g. pre-enrichment of termini by SCX at acidic pH [32], enzymatic removal of N-terminal pyro-Glu that can be formed during sample preparation [32, 49], depletion of abundant peptides by random peptide depletion-libraries [41], and peptide-level chromatographic SCX or high-pH-reverse phase fractionation [38, 45].

3.3. Global enrichment of protein C-terminal peptides

Knowledge of the protein C termini can contribute crucial complementary information on C-terminal PTMs and proteolytic processing by endopeptidases and carboxypeptidases. C termini also uniquely identify ectodomain shedding events that are not accessible by analysis of protein N-termini. For example, C termini of the soluble ectodomain protein of the prevalent type I membrane-spanning proteins, which are oriented with their NH_2 termini in the lumen and their COOH termini in the cytosol, directly reveal the sheddase cleavage site in the released ectodomain (Fig. 1). N terminome analysis provides this information for type II membrane protein cleavage sites (extracellular C terminus), but reveals type I membrane protein cleavage sites only if the N-termini of the remaining membrane stubs are identified, for example after release by y-secretase (Fig. 1). Hence, several methods for the investigation of protein C-termini have been introduced. Different approaches to label or identify C termini include ^{18}O-labeling [51-53], comparison of LysC/LysN digestion profiles [53], cyanogenbromide digestion [54-56] and isotopic arginine labeling based on the oxazolone chemistry [57, 58]. Although the viability of these labeling protocols was proven in several

studies, they lack a distinct C-terminal enrichment step which limits the depth of C-terminal coverage in complex samples. For further discussion on the history, advantages and disadvantages of C-terminal labeling strategies we recommend the review by Tanco et al. [59].

An enzymatic protocol for enrichment of C termini by positive selection termed ProC-TEL uses the transpeptidase activity of carboxypeptidase Y to conjugate protein C-termini with a biotin affinity-tag for subsequent enrichment with avidin beads (**Fig. 4A**) [60]. Recently, a step-by-step protocol has been published [61]. An alternative chemical labeling protocol for positive enrichment of C termini is based on the oxazolone chemistry, where formic acid and acetic anhydride are used to specifically activate the α-carboxy group of C-terminal amino acids [57]. Modification of the activated C termini with a bifunctional peptide containing an N-terminal Arginine and a C-terminal Biotin allowed selective enrichment with streptavidin-coupled magnetic beads (**Fig. 4B**) [62].

An early method for proteome-wide enrichment of C termini used SCX chromatography to pre-select N and C termini based on their predictable charge states in tryptic digests, where C-terminal peptides typically carry only their N-terminal alpha-amine as they lack basic residues [63]. A modified COFRADIC protocol later incorporated this idea to achieve simultaneous enrichment of N-terminal and C-terminal peptides [64]. As in the N-terminal version, primary amines of intact proteins are acetylated in C-terminal COFRADIC before proteome digestion with trypsin (**Fig. 4C**). A SCX pre-fractionation step at low pH removes the vast majority of internal peptides carrying two or three positive charges, while acetylated N-terminal peptides and C-terminal peptides, which lack basic residues after tryptic digest, elute earlier as single charged peptides. This is followed by a first RP-HPLC fractionation, after which an N-hydroxysuccinimide (NHS) ester of butyrate is used to modify the remaining C- and N-terminal peptides at their free α-amino group. This allows separation of N- and C-terminal peptides based on the retention time shift of the modified C-terminal peptides. An updated, detailed protocol has recently been published [65].

A major challenge for enrichment of C-terminal peptides by negative selection is the comparatively low chemical reactivity of the C-terminal carboxyl group which requires activation before reaction with carboxyl-modifying reagents. To avoid cross-reactivity with N-terminal and lysine side-chain primary amines, a modified TAILS protocol therefore requires subsequent modification of both amino- and carboxyl functionalities [66]. In C-TAILS (**C**arboxy-**T**erminal **A**mine-**b**ased **I**sotope **L**abeling of **S**ubstrates), all primary amines are therefore dimethylated before the carboxyl groups are protected with ethanolamine (**Fig. 4D**). Proteolytic digestion is performed and newly generated internal peptides are again dimethylated at their free N-terminus. The internal and N-terminal peptides are then removed by covalent coupling to a high molecular weight polyallylamine polymer via their unblocked C-termini and subsequent filtration.

Detailed step-by-step protocols have been published [67, 68], but the incompatible chemistry of the two protection steps remained a challenge. In efforts to optimize the protocol, modification of protein primary amines by acetylation rather than dimethylation appeared beneficial and inclusion of single charged ions during mass spectrometry analysis increased the number of identified C termini [69]. Furthermore, the first amine protection step may also be omitted without obvious detrimental effect, as the concentration of protein primary amines appeared negligible compared to the excess amidation reagent used for carboxyl modification [70].

An alternative negative selection strategy, amidated protein C termini with a methylamine tag before digestion with LysC, which results in C-terminal peptides devoid of Lys residues [71]. The resulting peptides were then selectively α-amine-modified at medium acidity, followed by phospho-tagging of Lys Ɛ-amines, which allowed removal of internal and N-terminal peptides with TiO2-beads.

4. Identification of proteolytically released proteoforms by termini enrichment

Positional proteomics thus identifies protein N- and C termini on a proteome-wide level and has therefore the unique potential to precisely define proteoforms that are proteolytically released to the extracellular environment. Consequently, several studies harnessed N-termini enrichment techniques to determine novel shedding substrates.

4.1. Sheddase substrate identification

Meprin β is a predominantly membrane-bound extracellular astacin metalloprotease involved in maintenance of extracellular homeostasis and regulation of inflammatory processes (reviewed in [10]). Extensive TAILS screens of cell-culture conditioned media of four different cell lines treated with recombinant soluble meprin α or β and three cell lines expressing full-length meprin β identified over 150 high-confidence cleavage sites, including many receptor proteins, growth factors, hormones, extracellular matrix proteins as well as several proteases and protease inhibitors [8, 72, 73]. Detailed follow-up studies showed that meprin β activated ADAM10, which in turn shed meprin β from the membrane thus forming regulatory feedback loop that regulates proteolytic activity of these two proteases at the cell surface [8]. Interestingly, soluble and membrane-bound forms of meprin β are activated by different proteases and exhibit distinct substrate specificity. Membrane-bound meprin β is activated by matriptase 2, whereas soluble meprin β is activated by soluble serine proteases after shedding [10]. TAILS with soluble recombinant meprin β identified several cleavages in the extracellular domain of APP that result in nontoxic APP fragments [72]. In contrast, full-length membrane-bound cleaved APP at and close to the β-site results in the release of toxic Aβ, including the highly aggregation-prone $A\beta_{2\text{-}40}$ [74]. Similarly, the interleukin-6

receptor protein could only be shed by membrane-bound meprin β, whereas the soluble form did not cleave this protein [10].

Mice with keratinocyte-specific deletion of the sheddase ADAM17 developed epidermal barrier effects after birth, which was phenocopied by mice lacking epidermal growth factor receptor (EGFR) in keratinocytes, suggesting that ADAM17 released EGFR ligands [75]. A shotgun proteome and TAILS N-terminome profiling of back skin epidermis from these mice found widespread changes in proteolytic processing of structural and regulatory components necessary for epidermal barrier formation, as well as differential accumulation of several soluble proteases [76].

Negative correlation of transcript abundance with brain amyloid load suggested that the membrane-bound metalloprotease ADAM30 might modulate APP metabolism [77]. Modulated expression of ADAM30 in cell cultures indeed affected Aβ peptide level. A COFRADIC analysis of HEK cells expressing active or catalytically inactive ADAM30 revealed cathepsin D (CTSD) as substrate. Increased activation of CTSD by ADAM30 resulted in increased APP sorting to lysosomes and thus indirectly increased release of Aβ peptides.

An example for ectodomain shedding by a soluble protease was presented by time-resolved analysis of recombinant MMP10 substrates. *Mmp10*$^{-/-}$ MEF cells were incubated with active recombinant MMP10 over a period of 16 h and were then analyzed by 8plex-iTRAQ-TAILS, which revealed ectodomain shedding of platelet-derived growth factor receptor alpha (PDGFRα) among several high-confidence substrates [78]. An extension of this study to keratinocyte secretomes and mouse epidermal tissue identified further MMP10 ectodomain shedding of cell adhesion proteins desmocollin-2, syndecan-4 and integrin alpha 6 [79].

4.2. Shedding events revealed by tissue profiling

Additionally, several large-scale N-terminome profiling experiments of isolated cells and whole tissues identified site-specific ectodomain shedding. TAILS was used to determine proteolytic proteome alterations in human platelet during 9 days storage under blood-banking conditions in the presence or absence of marimastat as a broad-spectrum metalloprotease inhibitor [80]. 7503 unique N-terminal peptide sequences from 2938 proteins were identified, of which up to 37 % changed significantly in abundance during storage. Most of the changes pointed to intracellular cleavage events, but also provided evidence for metalloproteinase-dependent shedding of several transmembrane glycoproteins, including GPIbα, a ligand-binding component of the von Willebrand factor receptor complex, and GPVI, a collagen receptor on the platelet surface. Together, these proteolytic modifications of the platelet surface during storage may contribute to reduced platelet function post-transfusion [80].

Ectodomain shedding is also observed in other tissue terminome analyses including human red blood cells [81], dental pulp [82], murine skin [83] and murine glomeruli [84], but not discussed in detail. Notably, ectodomain shedding is not limited to multicellular organisms. A recent investigation of the *Mycoplasma hyopneumoniae* N-terminome by charge-reversal enrichment of N-terminal peptides identified 669 unique N-terminal peptide sequences, of which many affected adhesins, lipoproteins and other cell surface proteins [85].

5. Current challenges in termini enrichment

Depending on sample composition, terminal enrichment protocols face numerous challenges. To date, most termini enrichment procedures required much more starting material than conventional shotgun proteomics approaches. One reason is that all termini enrichment strategies are relatively complex, multi-step sample preparation procedures that require careful handling to avoid contaminations and unspecific sample losses, for example by unspecific adhesion to reaction vessels particularly after the depletion/enrichment step.

5.1. Challenges in sample preparation

Protein termini enrichment greatly simplifies the proteome after digestion as only one peptide for each proteoform is retained, however it still reflects the full dynamic range of the sample [31]. Compared to standard shotgun proteomics, this allowed identification of lower abundance proteins within the same mass spectrometry time [86]. However, many proteolytic proteoforms are likely to exist only as instable intermediates and/or are derived from a fraction of the precursor protein, so identification of these forms remains a challenge in terminome studies. Additional (pre-)fractionation or partial enrichment of protein termini [34] should therefore be considered in the experimental design, as well as peptide-level fractionation after enrichment [38]. The extended and more complex experimental procedures compared to standard shotgun proteomics workflows also increase the likelihood of unwanted and/or unanticipated chemical modifications such as N-terminal PyroGlu-formation, oxidation or deamidation, which reduce the efficiency of the enrichment procedures and need to be removed [32] or considered during data analysis.

The established methods show that C-terminomic studies are feasible, but they remain a challenging task as the enrichment efficiency lacks far behind their N-terminal counterpart. This is exacerbated by the fact that C-terminal peptides in tryptic digest lack basic residues and thus are less likely to ionize. The alternative digestion protease LysargiNase, which mirrors trypsin specificity and thus results in C-terminal peptides with N-terminal of Lys and Arg residues [87], appears promising to overcome this challenge [68, 88]. Also chemical charge-reversal by derivatization with basic amines such as *N*,*N*-dimethylethylenediamine [89] and inclusion of single-charged precursor ions during mass spectrometric analysis improved identification of C-terminal peptides from complex proteomes [69].

5.2. Challenges in positional proteomics data analysis

By their very nature, only one side of neo-N or neo-C terminal peptides can be predicted from the specificity of the protease applied for proteome digestion. Positional proteomics thus requires semi-specific database searches for spectra-to-sequence matching, constrained only at the C- or N-terminal end, respectively, of the predicted peptides. This vastly increases the search space compared to specific searches where both ends of the predicted peptides are delimited by the digestion enzyme specificity. Depending on the digestion enzyme, semi-specific searches of the *Arabidopsis thaliana* proteome, for example, need to consider 9-fold to 33-fold more theoretical peptides with a length of 7 to 40 amino acids compared to "specific" full digest searches [90]. As a consequence, the proportion of MS/MS spectra matched to peptide sequences at a chosen false discovery rate (FDR) is typically lower compared to shotgun proteome datasets. In many cases, particularly with low resolution instruments, a higher FDR threshold (<0.05) has been accepted for identification of N-terminal peptides [76, 91].

Furthermore, the choice of digestion enzyme will determine the subset of protein termini that can be observed. Depending on the proximity to the terminus, digestion may result in terminal peptides that are too short, too long or contain an unfortunate amino acid composition that increases likelihood of unspecific losses, e.g. by adherence to reaction vessel surfaces, or impairs ionization, all of which renders identification by MS/MS unlikely. Hence, parallel digest with digestion enzymes differing in sequence specificity, for example trypsin, GluC, subtilisin and chymotrypsin, is necessary to achieve greater terminome coverage [35, 70, 81, 92]. In addition, these replicate experiments provide an opportunity for validation of single terminal peptide identifications as the same terminus may be identified by peptides with distinct sequence in the different assays.

Positional annotation, i.e. matching the identified terminus to a define position in a distinct protein model, is necessary to further understand the biological significance of the identified terminus. However, in many cases the identified sequences match several protein isoforms or even products of different genes. In these cases, parallel analysis of samples withdrawn before enrichment or parallel shotgun proteome analysis can provide additional evidence for protein expression. Termini are then first assigned to the protein with the strongest evidence in these assays, but ambiguities may remain [93]. Dedicated informatics tools have been developed to assist in choosing a preferred protein entry and assigning positional information. CLIPPER, a dedicated add-on tool for the Trans Proteomic Pipeline software suit [94], uses information from pre-enriched samples that are acquired in parallel to identify the most likely protein model for an identified N-terminal peptide, annotates the position in relation to the selected protein entry and adds information from the UniProt database [93]. The peptide-centric database TOPFIND, accessible for individual protein queries via a web interface and bulk queries via API, provides access to integrated positional, domain and PTM

annotation information from UniProt, protease specificity and known substrate cleavage site information from the protease database MEROPS [95] and curated termini identifications collected from the literature or submitted by users [96, 97]. Additional new functionalities in the current TOPFIND release 3.0 enable annotation of user-provided termini identifications and insights into direct and indirect connections to a protease of interest within the human or murine protease web [98]. Other bioinformatics tools, including ImproViser [99], QARIP [100] and Protter [101], visualize the position of identified peptides in relation to the protein sequence. This provides intuitive peptide maps pointing to shedding events in shotgun proteomics data and helps to create testable hypotheses on processed proteoforms function by highlighting protein domains affected by the identified cleavages. Finally, a dedicated database, SheddomeDB, was set up to include manually curated evidence for ectodomain shedding from the literature [102]. This database could facilitate cross-checking for prior evidence of ectodomain shedding, but unfortunately is currently not accessible.

6. Conclusions

Proteolytic processing of membrane-bound proteins is a widespread mechanism of unconventional protein secretion that has emerged as a key process for cellular interactions with their environment. A variety of positional proteomics techniques enable identification of protein termini on a proteome-wide scale and thus provide a unique opportunity to define proteolytic proteoforms. First applications have shown their utility for the identification of sheddase substrates and identification of shed proteoforms in a variety of cells and tissues. However, positional proteomics has not been widely implemented yet. This is most likely due to the requirements for relatively large amounts of sample, frequently disappointing proteome coverage particularly when a single digestion protease is employed and the complexity of the procedures and data analysis. Current developments already promise simplified methods and improved identification of protein termini from small sample amounts. Together with increasingly convenient data analysis tools this will hopefully lead to more widespread use and a greater appreciation of the different proteolytic proteoforms present in complex samples.

Acknowledgments

We gratefully acknowledge funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (project ProPlantStress, grant agreement No 639905, to P.F.H.).

Figure legends

Fig. 1: Unconventional protein secretion by proteolytic processing of membrane proteins. During ectodomain shedding, membrane-tethered proteases, also termed sheddases, cleave transmembrane proteins at specific sites to release a soluble proteoform to the extracellular environment. Different sheddases (indicated green and red) may cleave the same substrate protein at different sites resulting in distinct proteolytic proteoforms that may be further processed by soluble exo- and endopeptidases (grey) in the extracellular environment. Membrane shedding may also trigger regulated intramembrane proteolysis (RIP), where cleavage by an intramembrane protease (blue) releases additional fragments/proteoforms to the extra- and intracellular environment.

Fig. 2: Strategies for enrichment of N termini by positive selection. A) The rationally designed peptide ligase subtiligase is used to enzymatically attach a biotinylated peptide-tag containing a TEV protease cleavage-site to free protein N-terminal alpha-amines. After tryptic digest, biotin-tagged N-terminal peptides are enriched by streptavidin pullout, while internal peptides remain in the flow-through. Cleavage with TEV protease releases the N-terminal peptides for subsequent MS analysis. B) For selective chemical biotinylation, ε-amines at lysine side chains are first protected by guanidination before coupling a disulfide-linked biotin to free N-terminal α-amines with an NHS-based reagent. After tryptic digest, the biotin containing N-terminal peptides are enriched in a streptavidin pullout and recovered by reduction of the disulfide link to biotin.

Fig. 3: Strategies for enrichment of N-terminal peptide by negative selection. A) In a first step all primary amines are blocked. After proteolytic digest, the reactivity of newly formed N-terminal amine groups of internal and C-terminal peptides is utilized to separate them from the original, blocked protein N-termini before optional peptide level fractionation and mass spectrometry-based identification. B) In COFRADIC, peptides are separated into distinct fractions by HPLC. Free primary amines of internal peptides in the collected fractions are then individually modified with TNBS. This results in a retention time shift that separates these peptides in a second HPLC from the original blocked N-termini, which keep their retention

time. C) Internal peptides are modified with biotin and subsequently removed by immobilized streptavidin while the original blocked N-termini elute. D) In TAILS, primary amines of internal peptides are captured with a high molecular weight dendritic hyperbranched polyglycerol aldehyde and are subsequently removed by filtration. E) Modification of internal peptides with the PTAG reagent glyceraldhyde-3-phosphate allows their depletion through binding onto TiO_2. F) In charge reversal, sulfonic acids coupled to internal peptides to introduce a high negative charge, which prevents binding to SCX while original blocked peptides are retained. G) In hydrophobic tagging-assisted N-termini enrichment (HYTANE), primary amines of internal peptides are derivatized with a long-chain aldehyde that allows depletion on a reverse phase trap column. H) The STagAu approach modifies the primary amines of internal peptides with a sulfhydryl-tag that binds to gold particles which can magnetically be removed from the original N-termini

Fig. 4: Strategies for enrichment of C-terminal peptides from complex samples. A) ProC-TEL protects side-chain carboxygroups by esterification before ligating an activated biotin tag to the protein C-terminus using the transpeptidase activity of carboxypeptidase Y. This allows positive selection of C-terminal peptides by avidin pullout. B) Oxazolone-based chemistry allows derivatization of C-termini with Arginine, which may be used to couple a dual functionalized Arg-biotin-peptide. Streptavidin pullout enables positive selection of the biotin-tagged C-terminal peptides. C) COFRADIC approach for negative selection of C-terminal peptides. Amine-protected proteins are digested with trypsin, followed by SCX pre-enrichment of N- and C-terminal peptides at low pH. The positive charged primary amine of internal peptides promotes binding to SCX while mostly terminal peptides elute in the flow-through. In a first HPLC run, peptides are separated into distinct fractions that are subsequently modified with NHS-butyrate, which reacts with the free primary amine of the C-terminal peptide. A second HPLC run separates N-terminal peptides with identical retention time from C-terminal peptides with altered retention time. D) In C-TAILS, primary amines and carboxyl groups are modified before enzymatic digest. A second amine protection is performed to block primary amines of digestion-generated peptides to avoid cross-reactivity in the final capture of internal peptides with polyallyamine. C-terminal peptides containing blocked carboxy groups do not react and are negatively selected by filtration. In all approaches, enriched C-terminal peptides are identified by mass spectrometry.

Literature

[1] S.F. Lichtenthaler, M.K. Lemberg, R. Fluhrer, Proteolytic ectodomain shedding of membrane proteins in mammals-hardware, concepts, and recent developments, The EMBO journal, 37 (2018).

[2] K. Hayashida, A.H. Bartlett, Y. Chen, P.W. Park, Molecular and cellular mechanisms of ectodomain shedding, Anat Rec (Hoboken), 293 (2010) 925-937.

[3] F. Zunke, S. Rose-John, The shedding protease ADAM17: Physiology and pathophysiology, Biochimica et biophysica acta, 1864 (2017) 2059-2070.

[4] J.W. van Kilsdonk, L.C. van Kempen, G.N. van Muijen, D.J. Ruiter, G.W. Swart, Soluble adhesion molecules in human cancers: sources and fates, Eur J Cell Biol, 89 (2010) 415-427.

[5] B. Singh, G. Carpenter, R.J. Coffey, EGF receptor ligands: recent advances, F1000Res, 5 (2016).

[6] M.A. Miller, R.J. Sullivan, D.A. Lauffenburger, Molecular Pathways: Receptor Ectodomain Shedding in Treatment, Resistance, and Monitoring of Cancer, Clin Cancer Res, 23 (2017) 623-629.

[7] E.K. Petutschnig, M. Stolze, U. Lipka, M. Kopischke, J. Horlacher, O. Valerius, W. Rozhon, A.A. Gust, B. Kemmerling, B. Poppenberger, G.H. Braus, T. Nurnberger, V. Lipka, A novel Arabidopsis CHITIN ELICITOR RECEPTOR KINASE 1 (CERK1) mutant with enhanced pathogen-induced cell death and altered receptor processing, The New phytologist, 204 (2014) 955-967.

[8] T. Jefferson, U. Auf dem Keller, C.L. Bellac, V.V. Metz, C. Broder, J. Hedrich, A. Ohler, W. Maier, V. Magdolen, E. Sterchi, J.S. Bond, A. Jayakumar, H. Traupe, A. Chalaris, S. Rose-John, C. Pietrzik, R. Postina, C.M. Overall, C. Becker-Pauly, The substrate degradome of meprin metalloproteases reveals an unexpected proteolytic link between meprin b and ADAM10, Cell. Mol. Life Sci., 70 (2013) 309-333.

[9] A.J. McCarthy, C. Coleman-Vaughan, J.V. McCarthy, Regulated intramembrane proteolysis: emergent role in cell signalling pathways, Biochemical Society transactions, 45 (2017) 1185-1202.

[10] P. Arnold, A. Otte, C. Becker-Pauly, Meprin metalloproteases: Molecular regulation and function in inflammation and fibrosis, Biochimica et biophysica acta, 1864 (2017) 2096-2104.

[11] S.A. Muller, S.D. Scilabra, S.F. Lichtenthaler, Proteomic Substrate Identification for Membrane Proteases in the Brain, Front Mol Neurosci, 9 (2016) 96.

[12] S.R. Van Doren, T.C. Marcink, R.K. Koppisetti, A. Jurkevich, Y.G. Fulcher, Peripheral membrane associations of matrix metalloproteinases, Biochimica et biophysica acta, 1864 (2017) 1964-1973.

[13] D.J. Selkoe, J. Hardy, The amyloid hypothesis of Alzheimer's disease at 25 years, EMBO Mol Med, 8 (2016) 595-608.

[14] C. Becker-Pauly, C.U. Pietrzik, The Metalloprotease Meprin beta Is an Alternative beta-Secretase of APP, Front Mol Neurosci, 9 (2016) 159.

[15] R. Aebersold, M. Mann, Mass-spectrometric exploration of proteome structure and function, Nature, 537 (2016) 347-355.

[16] K. Shirakabe, Y. Shibagaki, A. Yoshimura, S. Koyasu, S. Hattori, A proteomic approach for the elucidation of the specificity of ectodomain shedding, Journal of Proteomics, 98 (2014) 233-243.

[17] K. Shirakabe, S. Hattori, M. Seiki, S. Koyasu, Y. Okada, VIP36 protein is a target of ectodomain shedding and regulates phagocytosis in macrophage Raw 264.7 cells, The Journal of biological chemistry, 286 (2011) 43154-43163.

[18] J.L. Tacchi, B.B. Raymond, V.M. Jarocki, I.J. Berry, M.P. Padula, S.P. Djordjevic, Cilium adhesin P216 (MHJ_0493) is a target of ectodomain shedding and aminopeptidase activity on the surface of Mycoplasma hyopneumoniae, Journal of proteome research, 13 (2014) 2920-2930.

[19] H. Zhang, X.-J. Li, D.B. Martin, R. Aebersold, Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry, Nature biotechnology, 21 (2003) 660-666.

[20] K. Tsumagari, K. Shirakabe, M. Ogura, F. Sato, Y. Ishihama, A. Sehara-Fujisawa, Secretome analysis to elucidate metalloprotease-dependent ectodomain shedding of glycoproteins during neuronal differentiation, Genes to Cells, 22 (2017) 237-244.
[21] P.H. Kuhn, K. Koroniak, S. Hogl, A. Colombo, U. Zeitschel, M. Willem, C. Volbracht, U. Schepers, A. Imhof, A. Hoffmeister, C. Haass, S. Rossner, S. Brase, S.F. Lichtenthaler, Secretome protein enrichment identifies physiological BACE1 protease substrates in neurons, The EMBO journal, 31 (2012) 3157-3168.
[22] P.H. Kuhn, A.V. Colombo, B. Schusser, D. Dreymueller, S. Wetzel, U. Schepers, J. Herber, A. Ludwig, E. Kremmer, D. Montag, U. Muller, M. Schweizer, P. Saftig, S. Brase, S.F. Lichtenthaler, Systematic substrate identification indicates a central role for the metalloprotease ADAM10 in axon targeting and synapse function, eLife, 5 (2016).
[23] A. Serdaroglu, S.A. Müller, U. Schepers, S. Bräse, W. Weichert, S.F. Lichtenthaler, P.H. Kuhn, An optimised version of the secretome protein enrichment with click sugars (SPECS) method leads to enhanced coverage of the secretome, Proteomics, 17 (2017) 1600423.
[24] T.K. Chang, D.Y. Jackson, J.P. Burnier, J.A. Wells, Subtiligase: a tool for semisynthesis of proteins, Proceedings of the National Academy of Sciences, 91 (1994) 12544-12548.
[25] S. Mahrus, J.C. Trinidad, D.T. Barkan, A. Sali, A.L. Burlingame, J.A. Wells, Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini, Cell, 134 (2008) 866-876.
[26] H.A. Yoshihara, S. Mahrus, J.A. Wells, Tags for labeling protein N-termini with subtiligase for proteomics, Bioorg Med Chem Lett, 18 (2008) 6000-6003.
[27] A.P. Wiita, J.E. Seaman, J.A. Wells, Global Analysis of Cellular Proteolysis by Selective Enzymatic Labeling of Protein N-Termini, Methods in enzymology, 544 (2014) 327-358.
[28] A.M. Weeks, J.A. Wells, Engineering peptide ligase specificity by proteomic identification of ligation sites, Nature chemical biology, 14 (2018) 50-57.
[29] J.C. Timmer, M. Enoksson, E. Wildfang, W. Zhu, Y. Igarashi, J.B. Denault, Y. Ma, B. Dummitt, Y.H. Chang, A.E. Mast, A. Eroshkin, J.W. Smith, W.A. Tao, G.S. Salvesen, Profiling constitutive proteolytic events in vivo, Biochem J, 407 (2007) 41-48.
[30] P.F. Huesgen, C.M. Overall, N- and C-terminal degradomics: new approaches to reveal biological roles for plant proteases from substrate identification, Physiologia plantarum, 145 (2012) 5-17.
[31] K. Gevaert, M. Goethals, L. Martens, J. Van Damme, A. Staes, G.R. Thomas, J. Vandekerckhove, Exploring proteomes and analyzing protein processing by mass spectrometric identification of sorted N-terminal peptides, Nat Biotech, 21 (2003) 566-569.
[32] A. Staes, F. Impens, P. Van Damme, B. Ruttens, M. Goethals, H. Demol, E. Timmerman, J. Vandekerckhove, K. Gevaert, Selecting protein N-terminal peptides by combined fractional diagonal chromatography, Nat. Protocols, 6 (2011) 1130-1141.
[33] K. Gevaert, P. Van Damme, B. Ghesquiere, J. Vandekerckhove, Protein processing and other modifications analyzed by diagonal peptide chromatography, Biochimica et biophysica acta, 1764 (2006) 1801-1810.
[34] A. Staes, P. Van Damme, K. Helsens, H. Demol, J. Vandekerckhove, K. Gevaert, Improved recovery of proteome-informative, protein N-terminal peptides by combined fractional diagonal chromatography (COFRADIC), Proteomics, 8 (2008) 1362-1370.
[35] A.S. Venne, F.A. Solari, F. Faden, T. Paretti, N. Dissmeyer, R.P. Zahedi, An improved workflow for quantitative N-terminal charge-based fractional diagonal chromatography (ChaFRADIC) to study proteolytic events in Arabidopsis thaliana, Proteomics, 15 (2015) 2458-2469.
[36] A.S. Venne, F.N. Vogtle, C. Meisinger, A. Sickmann, R.P. Zahedi, Novel highly sensitive, specific, and straightforward strategy for comprehensive N-terminal proteomics reveals unknown substrates of the mitochondrial peptidase Icp55, Journal of proteome research, 12 (2013) 3823-3830.

[37] G. Shema, M.T.N. Nguyen, F.A. Solari, S. Loroch, A.S. Venne, L. Kollipara, A. Sickmann, S.H.L. Verhelst, R.P. Zahedi, Simple, scalable, and ultrasensitive tip-based identification of protease substrates, Mol Cell Proteomics, 17 (2018) 826-834.
[38] O. Kleifeld, A. Doucet, A. Prudova, U. auf dem Keller, M. Gioia, J.N. Kizhakkedathu, C.M. Overall, Identifying and quantifying proteolytic events and the natural N terminome by terminal amine isotopic labeling of substrates, Nature protocols, 6 (2011) 1578-1611.
[39] A. Doucet, O. Kleifeld, J.N. Kizhakkedathu, C.M. Overall, Identification of proteolytic products and natural protein N-termini by Terminal Amine Isotopic Labeling of Substrates (TAILS), Methods in molecular biology, 753 (2011) 273-287.
[40] F. Demir, S. Niedermaier, J.N. Kizhakkedathu, P.F. Huesgen, Profiling of Protein N-Termini and Their Modifications in Complex Samples, Meth Mol Biol, 1574 (2017) DOI 10.1007/1978-1001-4939-6850-1003_1004.
[41] F. Sabino, O. Hermes, U. auf dem Keller, Body Fluid Degradomics and Characterization of Basic N-Terminome, (2016).
[42] L. McDonald, D.H.L. Robertson, J.L. Hurst, R.J. Beynon, Positional proteomics: selective recovery and analysis of N-terminal proteolytic peptides, Nat Meth, 2 (2005) 955-957.
[43] L. McDonald, R.J. Beynon, Positional proteomics: preparation of amino-terminal peptides as a strategy for proteome simplification and characterization, Nature protocols, 1 (2006) 1790-1798.
[44] J.R. Wisniewski, Filter-Aided Sample Preparation: The Versatile and Efficient Method for Proteomic Analysis, Methods Enzymol, 585 (2017) 15-27.
[45] J. Yeom, S. Ju, Y. Choi, E. Paek, C. Lee, Comprehensive analysis of human protein N-termini enables assessment of various protein forms, Scientific Reports, 7 (2017) 6599.
[46] G.P.M. Mommen, B. van de Waterbeemd, H.D. Meiring, G. Kersten, A.J.R. Heck, A.P.J.M. de Jong, Unbiased Selective Isolation of Protein N-terminal Peptides from Complex Proteome Samples Using Phospho Tagging (PTAG) and TiO(2)-based Depletion, Molecular & cellular proteomics : MCP, 11 (2012) 832-842.
[47] Z.W. Lai, A. Gomez-Auli, E.J. Keller, B. Mayer, M.L. Biniossek, O. Schilling, Enrichment of protein N-termini by charge reversal of internal peptides, Proteomics, 15 (2015) 2470-2478.
[48] Z.W. Lai, O. Schilling, Identification of Protease Cleavage Sites by Charge-Based Enrichment of Protein N-Termini, Methods in molecular biology, 1579 (2017) 199-207.
[49] L. Chen, Y. Shan, Y. Weng, Z. Sui, X. Zhang, Z. Liang, L. Zhang, Y. Zhang, Hydrophobic Tagging-Assisted N-Termini Enrichment for In-Depth N-Terminome Analysis, Analytical chemistry, 88 (2016) 8390-8395.
[50] L. Li, R. Wu, G. Yan, M. Gao, C. Deng, X. Zhang, A novel method to isolate protein N-terminal peptides from proteome samples using sulfydryl tagging and gold-nanoparticle-based depletion, Analytical and bioanalytical chemistry, 408 (2016) 441-448.
[51] S. Julka, D. Dielman, S.A. Young, Detection of C-terminal peptide of proteins using isotope coding strategies, Journal of Chromatography B, 874 (2008) 101-110.
[52] T. Kosaka, T. Takazawa, T. Nakamura, Identification and C-Terminal Characterization of Proteins from Two-Dimensional Polyacrylamide Gels by a Combination of Isotopic Labeling and Nanoelectrospray Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, Analytical Chemistry, 72 (2000) 1179-1185.
[53] T. Kishimoto, J. Kondo, T. Takai-Igarashi, H. Tanaka, Accurate mass comparison coupled with two endopeptidases enables identification of protein termini, PROTEOMICS, 11 (2011) 485-489.
[54] P.P. Moerman, K. Sergeant, G. Debyser, B. Devreese, B. Samyn, A new chemical approach to differentiate carboxy terminal peptide fragments in cyanogen bromide digests of proteins, Journal of Proteomics, 73 (2010) 1454-1460.
[55] P.P. Moerman, K. Sergeant, G. Debyser, I. Timperman, B. Devreese, B. Samyn, Automation of C-terminal sequence analysis of 2D-PAGE separated proteins, EuPA Open Proteomics, 3 (2014) 250-261.

[56] Z. Wang, T.L. Hilder, K. van der Drift, J. Sloan, K. Wee, Structural characterization of recombinant alpha-1-antitrypsin expressed in a human cell line, Analytical biochemistry, 437 (2013) 20-28.
[57] M. Liu, L. Zhang, L. Zhang, J. Yao, P. Yang, H. Lu, Approach for Identification and Quantification of C-Terminal Peptides: Incorporation of Isotopic Arginine Labeling Based on Oxazolone Chemistry, Analytical Chemistry, 85 (2013) 10745-10753.
[58] T. Nakazawa, M. Yamaguchi, K. Nishida, H. Kuyama, T. Obama, E. Ando, T.a. Okamura, N. Ueyama, K. Tanaka, S. Norioka, Enhanced responses in matrix-assisted laser desorption/ionization mass spectrometry of peptides derivatized with arginine via a C-terminal oxazolone, Rapid Communications in Mass Spectrometry, 18 (2004) 799-807.
[59] S. Tanco, K. Gevaert, P. Van Damme, C-terminomics: Targeted analysis of natural and posttranslationally modified protein and peptide C-termini, Proteomics, 15 (2015) 903-914.
[60] G. Xu, S.B. Shin, S.R. Jaffrey, Chemoenzymatic labeling of protein C-termini for positive selection of C-terminal peptides, ACS chemical biology, 6 (2011) 1015-1020.
[61] W. Duan, G. Xu, ProC-TEL: Profiling of Protein C-Termini by Enzymatic Labeling, Methods Mol Biol, 1574 (2017) 135-144.
[62] M. Liu, C. Fang, X. Pan, H. Jiang, L. Zhang, L. Zhang, Y. Zhang, P. Yang, H. Lu, Positive enrichment of C-terminal peptides using oxazolone chemistry and biotinylation, Analytical chemistry, 87 (2015) 9916-9922.
[63] W. Dormeyer, S. Mohammed, B.v. Breukelen, J. Krijgsveld, A.J.R. Heck, Targeted Analysis of Protein Termini, Journal of Proteome Research, 6 (2007) 4634-4645.
[64] P. Van Damme, A. Staes, S. Bronsoms, K. Helsens, N. Colaert, E. Timmerman, F.X. Aviles, J. Vandekerckhove, K. Gevaert, Complementary positional proteomics for screening substrates of endo- and exoproteases, Nature methods, 7 (2010) 512-515.
[65] S. Tanco, F.X. Aviles, K. Gevaert, J. Lorenzo, P. Van Damme, Identification of Carboxypeptidase Substrates by C-Terminal COFRADIC, Methods in molecular biology, 1574 (2017) 115-133.
[66] O. Schilling, O. Barre, P.F. Huesgen, C.M. Overall, Proteome-wide analysis of protein carboxy termini: C terminomics, Nature methods, 7 (2010) 508-511.
[67] O. Schilling, P.F. Huesgen, O. Barre, C.M. Overall, Identification and relative quantification of native and proteolytically generated protein C-termini from complex proteomes: C-terminome analysis, Methods in molecular biology, 781 (2011) 59-69.
[68] N. Solis, C.M. Overall, Identification of Protease Cleavage Sites and Substrates in Cancer by Carboxy-TAILS (C-TAILS), in: S. Cal, A.J. Obaya (Eds.) Proteases and Cancer: Methods and Protocols, Springer New York, Place Published, 2018, pp. 15-28.
[69] Y. Zhang, Q. He, J. Ye, Y. Li, L. Huang, Q. Li, J. Huang, J. Lu, X. Zhang, Systematic Optimization of C-Terminal Amine-Based Isotope Labeling of Substrates Approach for Deep Screening of C-Terminome, Analytical chemistry, 87 (2015) 10354-10361.
[70] Y. Zhang, Q. Li, J. Huang, Z. Wu, J. Huang, L. Huang, Y. Li, J. Ye, X. Zhang, An Approach to Incorporate Multi-Enzyme Digestion into C-TAILS for C-Terminomics Studies, PROTEOMICS, 18 (2018) 1700034.
[71] L. Chen, Y. Shan, Y. Weng, H. Yuan, S. Zhang, R. Fan, Z. Sui, X. Zhang, L. Zhang, Y. Zhang, Depletion of internal peptides by site-selective blocking, phosphate labeling, and TiO2 adsorption for in-depth analysis of C-terminome, Analytical and bioanalytical chemistry, 408 (2016) 3867-3874.
[72] T. Jefferson, M. Causevic, U. auf dem Keller, O. Schilling, S. Isbert, R. Geyer, W. Maier, S. Tschickardt, T. Jumpertz, S. Weggen, J.S. Bond, C.M. Overall, C.U. Pietrzik, C. Becker-Pauly, Metalloprotease meprin beta generates nontoxic N-terminal amyloid precursor protein fragments in vivo, The Journal of biological chemistry, 286 (2011) 27741-27750.
[73] C. Becker-Pauly, O. Barre, O. Schilling, U. auf dem Keller, A. Ohler, C. Broder, A. Schuette, R. Kappelhoff, W. Stöcker, C.M. Overall, Proteomic Analyses Reveal an Acidic Prime Side Specificity for the Astacin Metalloprotease Family Reflected by Physiological Substrates, Molecular & cellular proteomics : MCP, 10 (2011) M111.009233.

[74] J. Bien, T. Jefferson, M. Causevic, T. Jumpertz, L. Munter, G. Multhaup, S. Weggen, C. Becker-Pauly, C.U. Pietrzik, The metalloprotease meprin beta generates amino terminal-truncated amyloid beta peptide species, The Journal of biological chemistry, 287 (2012) 33304-33313.
[75] C.W. Franzke, C. Cobzaru, A. Triantafyllopoulou, S. Loffek, K. Horiuchi, D.W. Threadgill, T. Kurz, N. van Rooijen, L. Bruckner-Tuderman, C.P. Blobel, Epidermal ADAM17 maintains the skin barrier by regulating EGFR ligand-dependent terminal keratinocyte differentiation, J Exp Med, 209 (2012) 1105-1119.
[76] S. Tholen, C. Wolf, B. Mayer, J.D. Knopf, S. Loffek, Y. Qian, J.N. Kizhakkedathu, M.L. Biniossek, C.W. Franzke, O. Schilling, Skin Barrier Defects Caused by Keratinocyte-Specific Deletion of ADAM17 or EGFR Are Based on Highly Similar Proteome and Degradome Alterations, Journal of proteome research, 15 (2016) 1402-1417.
[77] F. Letronne, G. Laumet, A.M. Ayral, J. Chapuis, F. Demiautte, M. Laga, M.E. Vandenberghe, N. Malmanche, F. Leroux, F. Eysert, Y. Sottejeau, L. Chami, A. Flaig, C. Bauer, P. Dourlen, M. Lesaffre, C. Delay, L. Huot, J. Dumont, E. Werkmeister, F. Lafont, T. Mendes, F. Hansmannel, B. Dermaut, B. Deprez, A.S. Herard, M. Dhenain, N. Souedet, F. Pasquier, D. Tulasne, C. Berr, J.J. Hauw, Y. Lemoine, P. Amouyel, D. Mann, R. Deprez, F. Checler, D. Hot, T. Delzescaux, K. Gevaert, J.C. Lambert, ADAM30 Downregulates APP-Linked Defects Through Cathepsin D Activation in Alzheimer's Disease, EBioMedicine, 9 (2016) 278-292.
[78] P. Schlage, F.E. Egli, P. Nanni, L.W. Wang, J.N. Kizhakkedathu, S.S. Apte, U. auf dem Keller, Time-resolved analysis of the matrix metalloproteinase 10 substrate degradome, Molecular & cellular proteomics : MCP, 13 (2014) 580-593.
[79] P. Schlage, T. Kockmann, F. Sabino, J.N. Kizhakkedathu, U. Auf dem Keller, Matrix Metalloproteinase 10 Degradomics in Keratinocytes and Epidermal Tissue Identifies Bioactive Substrates With Pleiotropic Functions, Molecular & cellular proteomics : MCP, 14 (2015) 3234-3246.
[80] A. Prudova, S. Serrano, U. Eckhard, N. Fortelny, D.V. Devine, C.M. Overall, TAILS N-terminomics of human platelets reveals pervasive metalloproteinase-dependent proteolytic processing in storage, Blood, 124 (2014) e49-60.
[81] P.F. Lange, P.F. Huesgen, K. Nguyen, C.M. Overall, Annotating N termini for the human proteome project: N termini and nalpha-acetylation status differentiate stable cleaved protein species from degradation remnants in the human erythrocyte proteome, Journal of proteome research, 13 (2014) 2028-2044.
[82] S.R. Abbey, U. Eckhard, N. Solis, G. Marino, I. Matthew, C.M. Overall, The Human Odontoblast Cell Layer and Dental Pulp Proteomes and N-Terminomes, J Dent Res, 97 (2018) 338-346.
[83] U. auf dem Keller, A. Prudova, U. Eckhard, B. Fingleton, C.M. Overall, Systems-level analysis of proteolytic events in increased vascular permeability and complement activation in skin inflammation, Science signaling, 6 (2013) rs2.
[84] M.M. Rinschen, A.-K. Hoppe, F. Grahammer, M. Kann, L.A. Voelker, E.-M. Schurek, J. Binz, M. Höhne, F. Demir, M. Malisic, T.B. Huber, C. Kurschat, J.N. Kizhakkedathu, B. Schermer, P.F. Huesgen, T. Benzing, N-Degradomic Analysis Reveals a Proteolytic Network Processing the Podocyte Cytoskeleton, J Am Soc Nephrol, in press (2017).
[85] I.J. Berry, V.M. Jarocki, J.L. Tacchi, B.B.A. Raymond, M. Widjaja, M.P. Padula, S.P. Djordjevic, N-terminomics identifies widespread endoproteolysis and novel methionine excision in a genome-reduced bacterial pathogen, Scientific Reports, 7 (2017) 11063.
[86] O. Kleifeld, A. Doucet, U. auf dem Keller, A. Prudova, O. Schilling, R.K. Kainthan, A.E. Starr, L.J. Foster, J.N. Kizhakkedathu, C.M. Overall, Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products, Nat Biotech, 28 (2010) 281-288.
[87] P.F. Huesgen, P.F. Lange, L.D. Rogers, N. Solis, U. Eckhard, O. Kleifeld, T. Goulas, F.X. Gomis-Ruth, C.M. Overall, LysargiNase mirrors trypsin for protein C-terminal and methylation-site identification, Nature methods, 12 (2015) 55-58.

[88] Y. Zhang, Q. Li, J. Huang, Z. Wu, J. Huang, L. Huang, Y. Li, J. Ye, X. Zhang, An Approach to Incorporate Multi-Enzyme Digestion into C-TAILS for C-Terminomics Studies, Proteomics, 18 (2018) 1700034.
[89] P. Somasundaram, T. Koudelka, D. Linke, A. Tholey, C-Terminal Charge-Reversal Derivatization and Parallel Use of Multiple Proteases Facilitates Identification of Protein C-Termini by C-Terminomics, Journal of proteome research, (2016).
[90] P. Willems, E. Ndah, V. Jonckheere, S. Stael, A. Sticker, L. Martens, F. van Breusegem, K. Gevaert, P. Van Damme, N-terminal Proteomics Assisted Profiling of the Unexplored Translation Initiation Landscape in Arabidopsis thaliana, Molecular & cellular proteomics : MCP, 16 (2017) 1064.
[91] C.L. Bellac, A. Dufour, M.J. Krisinger, A. Loonchanta, A.E. Starr, U. Auf dem Keller, P.F. Lange, V. Goebeler, R. Kappelhoff, G.S. Butler, L.D. Burtnick, E.M. Conway, C.R. Roberts, C.M. Overall, Macrophage matrix metalloproteinase-12 dampens inflammation and neutrophil influx in arthritis, Cell reports, 9 (2014) 618-632.
[92] F.N. Vogtle, S. Wortelkamp, R.P. Zahedi, D. Becker, C. Leidhold, K. Gevaert, J. Kellermann, W. Voos, A. Sickmann, N. Pfanner, C. Meisinger, Global analysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability, Cell, 139 (2009) 428-439.
[93] U. auf dem Keller, C.M. Overall, CLIPPER: an add-on to the Trans-Proteomic Pipeline for the automated analysis of TAILS N-terminomics data, Biological chemistry, 393 (2012) 1477-1483.
[94] E.W. Deutsch, L. Mendoza, D. Shteynberg, J. Slagel, Z. Sun, R.L. Moritz, Trans-Proteomic Pipeline, a standardized data processing pipeline for large-scale reproducible proteomics informatics, Proteomics. Clinical applications, (2015).
[95] N.D. Rawlings, A.J. Barrett, P.D. Thomas, X. Huang, A. Bateman, R.D. Finn, The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database, Nucleic acids research, 46 (2018) D624-D632.
[96] P.F. Lange, P.F. Huesgen, C.M. Overall, TopFIND 2.0--linking protein termini with proteolytic processing and modifications altering protein function, Nucleic acids research, 40 (2012) D351-361.
[97] P.F. Lange, C.M. Overall, TopFIND, a knowledgebase linking protein termini with function, Nature methods, 8 (2011) 703-705.
[98] N. Fortelny, S. Yang, P. Pavlidis, P.F. Lange, C.M. Overall, Proteome TopFIND 3.0 with TopFINDer and PathFINDer: database and analysis tools for the association of protein termini to pre- and post-translational events, Nucleic acids research, 43 (2015) D290-D297.
[99] P. Videm, D. Gunasekaran, B. Schroder, B. Mayer, M.L. Biniossek, O. Schilling, Automated peptide mapping and protein-topographical annotation of proteomics data, BMC bioinformatics, 15 (2014) 207.
[100] D.N. Ivankov, N.S. Bogatyreva, P. Honigschmid, B. Dislich, S. Hogl, P.H. Kuhn, D. Frishman, S.F. Lichtenthaler, QARIP: a web server for quantitative proteomic analysis of regulated intramembrane proteolysis, Nucleic acids research, 41 (2013) W459-464.
[101] U. Omasits, C.H. Ahrens, S. Muller, B. Wollscheid, Protter: interactive protein feature visualization and integration with experimental proteomic data, Bioinformatics, 30 (2014) 884-886.
[102] W.S. Tien, J.H. Chen, K.P. Wu, SheddomeDB: the ectodomain shedding database for membrane-bound shed markers, BMC bioinformatics, 18 (2017) 42.