NEW BEGINNINGS AND NEW ENDS: METHODS FOR LARGE-SCALE CHARACTERIZATION OF PROTEIN TERMINI AND THEIR USE IN PLANT BIOLOGY

ANDREAS PERRAR1,1, NICO DISSMEYER2,3,4,* AND PITTER F. HUESGEN1,4,4,*

1 Forschungszentrum Jülich, Central Institute for Engineering, Electronics and Analytics, ZEA-3 Analytics, Wilhelm-Johnen-Str., D-52425 Jülich, Germany
2 Independent Junior Research Group on Protein Recognition and Degradation, Leibniz Institute of Plant Biochemistry (IPB), Weinberg 3, D-06120 Halle (Saale), Germany
3 ScienceCampus Halle – Plant-based Bioeconomy, Betty-Heimann-Str. 3, D-06120 Halle (Saale), Germany
4 Medical Faculty and University Hospital, University of Cologne, Cologne, Germany

* Correspondence: p.huesgen@fz-juelich.de or nico.dissmeyer@ipb-halle.de

Received 19 November 2018; Editorial decision 25 February 2019; Accepted 27 February 2019

Editor: Kris Gevaert, Gent University/VIB, Belgium

Abstract

Dynamic regulation of protein function and abundance plays an important role in virtually every aspect of plant life. Diversifying mechanisms at the RNA and protein level result in many protein molecules with distinct sequence and modification, termed proteoforms, arising from a single gene. Distinct protein termini define proteoforms arising from translation of alternative transcripts, use of alternative translation initiation sites, and different co- and post-translational modifications of the protein termini. Also site-specific proteolytic processing by endo- and exoproteases generates truncated proteoforms, defined by distinct protease-generated neo-N- and neo-C-termini, that may exhibit altered activity, function, and localization compared with their precursor proteins. In eukaryotes, the N-degron pathway targets cytosolic proteins, exposing destabilizing N-terminal amino acids and/or destabilizing N-terminal modifications for proteasomal degradation. This enables rapid and selective removal not only of unfolded proteins, but also of substrate proteoforms generated by proteolytic processing or changes in N-terminal modifications. Here we summarize current protocols enabling proteome-wide analysis of protein termini, which have provided important new insights into N-terminal modifications and protein stability determinants, protein maturation pathways, and protease–substrate relationships in plants.

Keywords: Acetylation, arginylation, degradomics, N-terminal modifications, positional proteomics, proteolysis, proteoform, proteostasis, termini enrichment.

Introduction

Variation at the protein level underpins the complexity, differentiation, and resilience of biological systems. Diversifying mechanisms result in multiple protein molecules with distinct sequences and/or chemical modifications, termed proteoforms, arising from a single gene (Smith and Kelleher, 2013). Many proteoforms can be distinguished by their unique termini (Fig. 1),...
including proteoforms generated by alternative RNA splicing (Cheng and Tu, 2018; Laloum et al., 2018; Szakonyi and Duque, 2018) and alternative translation initiation (Willems et al., 2017). On the protein level, proteolytic processing by a large range of endo- and exoproteases generates distinct proteoforms that are defined by their new protease-generated neo-N- and neo-C-termini (Huesgen and Overall, 2012; Klein et al., 2018). Also allelic variants can exhibit point mutations or sequence variations leading to differential splicing, new translation starts (Xu et al., 2015), and altered proteolytic processing (Fig. 1). Finally, site-specific co- and post-translational protein modifications create a dazzling space of potential proteoforms that may differ from each other in subcellular localizations, interactions, activity, and function (Smith and Kelleher, 2013).

Dynamic control of proteoform stability and abundance assures protein quality control and regulates many cellular processes. An important determinant of proteoform stability in vivo is the identity and modification of its N-terminal amino acid, which is summarized by the N-end rule (Bachmair et al., 1986; Gonda et al., 1989; Tobias et al., 1991; Patuschak et al., 1998; Graciet et al., 2009; Holman et al., 2009; Dissmeyer et al., 2018; Dissmeyer, 2019; Varshavsky, 2019). In the cytosol of eukaryotic cells, proteins exposing destabilizing N-termini can be recognized by E3 ubiquitin ligases called N-recogins, and targeted for rapid degradation by the ubiquitin—proteasome system. PROTEOLYSIS (PRT) 1 and 6 are the currently known N-recogins that target substrates initiated with primary destabilizing residues for degradation (Faden et al., 2016).
et al. 2019; Naumann et al. 2016; Dong et al. 2017; Reichman and Dismeyer, 2017; Mot et al., 2018; Vicente et al., 2019). In this context, N-terminal arginylation has been recognized as a post-translational modification that targets proteins for degradation via the N-end rule ubiquitin−proteasome pathway (Dissmeyer, 2017, 2019; White et al., 2017; Dismeyer et al., 2018), but also the autophagy−lysosomal pathway (Kim et al., 2013; Chalmolstad et al., 2015a, 2018). In animals, N-terminal arginylation is more and more considered to generate a bimodal degron that operates in both autophagic and proteasomal proteolysis (Yoo et al., 2018). N-terminal arginylation requires the activity of specific arginyl-tRNA protein arginyltransferases (or short arginyltransferases/arginine transfer enzymes, ATEs). Arabidopsis harbors two genes encoding putative ATEs, of which at least ATE1 can transfer an Arg to N-termini, exposing the secondary destabilizing acidic, negatively charged amino acids Asp, Glu, and dioxygenated or trioxgenated Cys (White et al., 2017). Besides the chemical state of the N-terminal amino acid residue, the efficiency of the arginy transfer also strongly depends on the identity and properties of the subsequent residues (Wadas et al., 2016). The immediate result of this modification is the transformation of N-termini with secondary destabilizing residues into N-termini exposing the primary destabilizing Arg, enabling their rapid degradation (White et al., 2017; Dismeyer et al., 2018). Also N-terminal acetylation is a dynamically regulated, ubiquitous protein modification in eukaryotes that affects protein−protein interactions, subcellular localization, protein folding, stability, and degradation via the Ac/N-end rule pathway (Pesaresi, 2003; Zybailov et al., 2008; Varshavsky, 2011; Bienvenut et al., 2012; Gibbs, 2015; Rathore et al., 2016; Linster and Wirtz, 2018). In plants, N-terminal acetylation by N-terminal acetyltransferases (NATs) occurs co- or post-translationally in the cytosol, the Golgi apparatus, and the chloroplasts of plants, and may be dynamically regulated by phytohormones (Linster et al., 2015; Linster and Wirtz, 2018). Since acetylated N-termini can be recognized as N-degrons, rapid and correct folding of the acetylated N-terminus into the protein’s inner structure is important for its stability. In contrast, slow or incorrect folding will expose the acetylated N-terminus, leading to its recognition and degradation (Varshavsky, 2011). N-terminal acetylation contributes to diverse aspects of a plant’s life, such as flowering time, immune response, or drought stress response (Kapos et al., 2015; Linster et al., 2015; Xu et al., 2015). However, both acetylation and arginylation also occur midchain at internal Lys or Asp and Glu residues, respectively, distant from the N-terminus (Eriste et al., 2005; Wang et al., 2014; Hoernstein et al., 2016). MS-based proteomics has become the method of choice for unbiased proteome analysis, enabling proteome−wide protein identification, quantification, and profiling of their modifications and interactions (Aebesold and Mann, 2016). In standard bottom−up workflows, proteins are first proteolytically digested into predictable peptides more amenable to chromatographic separation and analysis by tandem MS (MS/MS). Peptide fragmentation spectra are recorded and matched to theoretical spectra, computationally predicted from a sequence database, constrained by the sequence specificity of the protease used for digestion. However, N- and C-terminal peptides constitute only a very minor fraction of the extraordinarily complex peptide mixture generated by proteome digest and are therefore rarely identified. Furthermore, neo-termini generated by endogenous proteolytic processing match the sequence specificity of the digestion protease only on one side and are not considered during standard search spectra to sequence matching (Niedermaier and Huesgen, 2019). Hence, selective enrichment and dedicated data analysis are necessary for comprehensive profiling of N- and C-termini and their modifications (Klein et al., 2018).

Three principal strategies achieve proteome−wide enrichment of protein termini. First, the most basic strategy uses proteases with different sequence specificity to generate N- or C-terminal peptides containing a predictable number of primary amines, namely α-amines at peptide N-termini and ε-amin e in Lys residue side chains, that differ from the remaining digest-generated peptides. Such a differential amine content results in differential charge states at low pH that can be used for the separation of target N- or C-terminal peptides from the majority of other peptides by strong cation exchange (SCX) chromatography (Dormeyer et al., 2007). Secondly, positive selection workflows directly capture and enrich target terminal peptides while discarding all other digest−generated peptides (Fig. 2A). This is typically achieved by modification of blocked, free α-amines or carboxyl groups with a purification tag prior to digestion (Table 1) or by affinity enrichment with antibodies against specific N-terminal modifications. Thirdly, negative selection strategies take the opposite approach and achieve enrichment by selective depletion of undesired internal peptides (Fig. 2B; Table 2). In these workflows, primary amines must be chemically modified on the protein level, enabling stable isotope labeling by amine reactive isotope labeling reagents such as formaldehyde (Boersema et al., 2009), isobaric tags for relative and absolute quantification (Ross et al., 2004), or tandem mass tags (TMTs, Thompson et al., 2003) for reliable comparative analysis. Proteome digest with specific proteases then exposes new reactive α-amines at non-target peptides that are used for selective depletion while modified terminal peptides remain inert. Negative selection thus simultaneously enriches both naturally modified and in vitro modified (in vivo free) terminal peptides, facilitating discovery and comprehensive profiling of endogenous protein terminal modifications. Here, we review the different methodological solutions for proteome−wide enrichment of N- and C-termini enrichment and discuss recent and emerging applications in plant biology.

**Charge-based enrichment of protein C-termini and modified N-termini**

Endogenously modified protein N-termini can be enriched by exploiting the differential primary amine content (α-amines plus basic residues Arg, Lys, and His) in peptides generated by digestion with trypsin, which cleaves after Arg and Lys residues. Modified N-terminal peptides (e.g. acetylated and propionylated) lack a free α-amine, while C-terminal peptides lack a basic residues after tryptic digest. C-terminal and modified
N-terminal peptides thus mostly carry a single charge at pH 2.7, allowing separation from the majority of doubly charged tryptic peptides by SCX chromatography (Dormeyer et al., 2007). Charge-based enrichment is also used in stable-isotope protein N-terminal acetylation quantification (SILProNAQ), a dedicated protocol for the enrichment, identification, and quantification of acetylated N-termini (Bienvenut et al., 2012). Free α-amines are modified by stable isotope-marked (D3)-acetyl using the reagent N-acetoxy-[2H3]succinimide to distinguish between in vivo acetylated and in vitro deuto-acetylated termini, respectively. O-Acetylation of Ser, Thr, and Tyr residues occurring as a side reaction is reversed by the addition of hydroxylamine before tryptic digest and enrichment of acetylated N-termini by SCX chromatography. A customized SILProNAQ data analysis pipeline, including the EnCOUNTer tool, automatically parses database search results, quantifies matching in vivo and in vitro acetylated N-termini, and calculates the degree of endogenous acetylation (Bienvenut et al., 2017).

However, both charge-based protocols achieve only incomplete enrichment and are limited to selected digestion proteases. This excludes identification of the subset of termini with cleavage sites that result in too long, too short, or generally unfavorable peptides (Niedermaier and Huesgen, 2019). Internal peptides with a high content of acidic residues may co-elute with acetylated peptides, and N-terminal peptides with a high number of basic residues, for example by missed cleavages or containing His, may elute later within the bulk of internal peptides (Gorman and Shiell, 1993).

Enzymatic tagging of N- or C-terminal peptides for positive selection

A pioneering method for positive selection of N-termini utilizes subtiligase, a peptide ligase rationally engineered from the protease subtilisin BPN’ (Chang et al., 1994), to ligate a synthetic peptide ester bearing a biotin tag and a cleavage site for the Tobacco etch virus (TEV) protease to the α-amine of naturally unmodified and protease-generated neo-N-termini (Mahrus et al., 2008). After proteome digestion, tagged N-terminal peptides are captured by immobilized avidin, subjected to stringent washes, and released by TEV protease cleavage (Mahrus et al., 2008). Diverse synthetic peptide esters have been created for various purposes, including more water-soluble esters to push the equilibrium towards the labeled reaction (Yoshihara et al., 2008). However, typically only 10–15% of α-amines are labeled, presumably due to poor accessibility of N-termini under native conditions (Wiita et al., 2014). Subtiligase exhibits a preference for small amino acids at the P1’ position and large hydrophobic or aromatic amino acids at the P2’ position, and thus introduces a bias in N-terminal labeling (positions follow the Schechter and Berger nomenclature with protease substrate positions towards the N-terminus of the hydrolyzed peptide bond denoted P1, P2, P3, etc. and positions towards the C-terminus P1’, P2’, P3’, etc. Schechter and Berger, 1967). To address these issues, stabiligase, a subtiligase variant tolerating denaturing conditions (Chang et al., 1994), and a subtiligase mutant library with altered P1’ and P2’ selectivity have been created (Weeks and Wells, 2018). Rational selection of a combination of subtiligase variants now allows effective ligation of almost any protein N-terminal sequence (Weeks and Wells, 2018).
Enzyme-mediated biotinylation has also been used to enrich C-terminal peptides by positive selection. Profiling protein C-termini by enzymatic labeling (ProC-TEL, Xu et al., 2011) utilizes the known transeptidase activity of carboxypeptidase Y (CPY) at pH >11, where the hydrolytic activity of CPY is inhibited. First, all carboxyl groups are esterified by methanolic hydrochloric acid, followed by selective deprotection of side chain esters. CPY is then employed as a transeptidase to selectively add a biotin tag to protein C-termini. After digestion, tagged C-terminal peptides are captured by immobilized avidin, purified by stringent washes, and eluted with 50% acetonitrile, 0.1% trifluoroacetic acid (TFA) (Duan et al., 2016).

### Selective chemical tagging of N- or C-terminal peptides for positive selection

Selective biotinylation of free protein N-termini can also be achieved by chemical modification (Timmer et al., 2007). First, cysteine residues are blocked by iodoacetamide (IAA) and lysine residues converted to homoarginine by reaction with O-methylisourea under high pH conditions before α-amine labeling with sulfo-NHS-SS-biotin. Proteins are digested and tagged N-terminal peptides are captured with immobilized streptavidin and released after washes by reduction of the disulfide linker with DTT.

An alternative protocol for enrichment of N-terminal peptides by chemical tagging is N-terminomics by chemical labeling of the α-amine of proteins (N-CLAP, G. Xu et al., 2009). Here, the Edman sequencing reagent phenyl isothiocyanate is used to modify all amines. Subsequent acidification with TFA causes an intramolecular cyclization and cleavage of the peptide bond between the first and second amino acid, but not at modified ε-aminos. The resulting unblocked α-aminos are tagged with NHS-SS-biotin prior to proteome digestion. N-terminal peptides are captured by immobilized avidin and released after washes by reduction of the disulfide linker. The N-terminal peptides identified by MS/MS analysis are one amino acid shorter than naturally occurring N-termini, which needs to be considered during computational data analysis and interpretation.

A method for selective chemical biotinylation of protein C-termini has also been developed (Liu et al., 2013, 2015). This method employs the established oxazolone chemistry (Liu et al., 2013) to conjugate Arg–Arg dipeptide-linked biotin to protein C-termini. After digest, biotinylated C-terminal peptides were captured with streptavidin beads and released by boiling in 50% acetonitrile and 1% TFA before identification by MS.

**Table 1. Protocols for enrichment of N- or C-termini by positive selection**

| Name               | Target peptides and enrichment principle | Specific requirements | Starting material<sup>a</sup> | Advantages and disadvantages                                                                                   | References<sup>b</sup> |
|--------------------|------------------------------------------|-----------------------|-------------------------------|---------------------------------------------------------------------------------------------------------------|------------------------|
| Subtiligase        | Unblocked and protease-generated neo-N-termini; tagging by enzymatic biotinylation | Subtiligase, TEV protease, biotinylated peptide esters | 30-300 mg                    | + Direct enrichment of neo-N-termini<br>– sequence preference of subtiligase<br>– high amount starting material<br>– not compatible with chemical stable isotope labeling<br>– subtiligase not commercially available | Timmer et al. (2007); Duan et al. (2016); Duan and Xu (2017) |
| ProC-TEL           | Unblocked and protease-generated neo-C-termini; tagging by enzymatic biotinylation | Carboxypeptidase Y, biotinylated peptide ester | 200–300 µg per condition | + Direct enrichment of neo-C-termini<br>– compatible with chemical stable isotope labeling<br>– tricky reaction condition poses challenge for efficient labeling | Xu et al. (2011); Duan et al. (2016) |
| Chemical biotinylation of N-termini | Unblocked and protease-generated neo-N-termini; tagging by chemical biotinylation | NHS-SS-biotin, neutravidin resin | 1–10 mg | + Selective tagging of neo N-termini<br>– high amount starting material<br>– not compatible with chemical stable isotope labeling<br>– challenging differential α- and ε- amine labeling | Timmer and Salvesen (2011) |
| Chemical biotinylation of C-termini | Unblocked and protease-generated neo-C-termini; biotinylation using oxazolone chemistry | Synthetic Arg-NH-NH-biotin peptide | 10 µg per condition | + Selective tagging of neo-N-termini<br>– low amount of starting material<br>– not compatible with chemical stable isotope labeling | Liu et al. (2015) |
| N-CLAP             | Unblocked and protease-generated neo-N-termini; tagging by chemical biotinylation | NHS-SS-biotin, neutravidin resin | 2 mg | + Selective tagging of neo N-termini<br>– selective deprotection of α-amine using well-established chemistry<br>– peptides one residue shorter than true terminus<br>– not compatible with chemical stable isotope labeling | G. Xu et al. (2009); Xu and Jaffrey (2010) |

<sup>a</sup>Starting material refers to the amount of extracted and purified protein/proteome.

<sup>b</sup>References in italics are step-by-step protocols.

**Characterization of plant protein termini**
| Name       | Target peptides and enrichment principle                                                                 | Specific requirements                                                                 | Starting material | Advantages and disadvantages                                                                 | References                     |
|------------|-------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------|-------------------|-----------------------------------------------------------------------------------------------|---------------------------------|
| COFRADIC   | Modified and neo-N- and C-termini; enrichment by diagonal reverse phase chromatography                       | Capillary HPLC system with fraction collector                                         | 1–3 mg            | + proven in multiple publications<br>+ versatile enrichment of N- and C-termini<br>– labor intensive<br>– extensive MS time | Gevaert et al. (2003)<br>Staes et al. (2008, 2011); van Damme et al. (2010); Tsiatsiani et al. (2013); Tam et al. (2015); Willems et al. (2017) |
| ChaFRADIC  | Modified and neo-N-termini; enrichment by diagonal SCX chromatography                                        | Capillary HPLC system with fraction collector                                         | 50–200 µg         | + proven in multiple publications<br>+ less complex than COFRADIC<br>– labor intensive<br>+ low amount of starting material | Venne et al., 2013;<br>Carrive et al. (2015);<br>Venne et al. (2015) |
| ChaFRAtip  | Modified and neo-N-termini; enrichment by tip-based diagonal SCX chromatography                              | SCX beads                                                                             | 4.3 µg per condition | + lowest amount of starting material<br>– labor intensive<br>– prone to handling errors | Shema et al. (2018) |
| TAILS      | Modified and neo-N-termini; other peptides removed by covalent binding to polymer and ultrafiltration       | HPG-ALD polymer, size exclusion filter                                                 | 0.1–1 mg per condition | + robustness proven by independent application in multiple laboratories<br>– requires moderately expensive patented aldehyde polymer | Kleifeld et al. (2010);<br>Kleifeld et al. (2011); Köhler et al. (2015a, b); Rowland et al., (2015); H. Zhang et al. (2015, 2018); Demir et al. (2018) |
| C-TAILS    | Modified and neo-C-termini, other peptides removed with polymer and ultrafiltration Without amine blocking prior to C-terminal amidation | PAA polymer, labeling reagents, size exclusion filter                                   | 1.5–2 mg per condition | – labor- and loss-intensive | Schilling et al., (2010);<br>Schilling et al. (2011); Solis and Overall (2018); Y. Zhang et al. (2015, 2018) |
| Biotinylation of internal peptides | Modified and neo-N-termini, other peptides removed by covalent binding to NHS-activated Sepharose | NHS-biotin, streptavidin column, amine-scavenging beads | 50 µg              | + simple procedure using standard chemicals<br>– non-specific losses of N-termini on Sepharose beads | McDonald et al. (2005) |
| NHS bead capture | Modified and neo-N-termini, other peptides removed with NHS-activated Sepharose | NHS-activated Sepharose                                                              | 50 µg              | + simple procedure using standard chemicals<br>– non-specific losses of N-termini on Sepharose beads | McDonald and Beynon (2006) |
| NRich      | Modified and neo-N-termini, other peptides removed with NHS-activated Sepharose and ultrafiltration        | NHS-activated Sepharose beads                                                         | 5 mg per condition | + simple procedure<br>– high amount of starting material<br>– non-specific interactions of N-termini with bead material | Yeom et al. (2017)             |
| Charge reversal | Modified and neo-N-termini, other peptides removed by disulfonate modification + SCX chromatography | SCX-packed tips or HPLC system with SCX columns and fraction collector | 1.5–3 mg per condition | + flexible format allows enrichment from gel slices | Lai et al. (2015);<br>Lai and Schilling (2017) |
Table 2. Continued

| Name   | Target peptides and enrichment principle | Specific requirements | Starting material | Advantages and disadvantages | Referencesa |
|--------|------------------------------------------|-----------------------|------------------|-------------------------------|-------------|
| HYTANE | Modified and neo-N-termini, other peptides removed by hydrophobic tagging C18 trap depletion | C18 trap columns | 100 µg per condition | + well-established chemicals – removal of excess aldehyde can be challenging | Chen et al. (2016a) |
| PTAG   | Modified and neo-N-termini, other peptides removed by addition of phosphotag and TiO2 depletion | TiO2 affinity columns or beads | 100 µg per condition | + Chemistry used for depletion established in many laboratories – depletion sensitive to sample amount and reaction conditions | Mommen et al. (2012) |
| C-PTAG | Modified and neo-C-termini, other peptides removed by addition of phosphotag and TiO2 depletion | TiO2 affinity columns or beads | 200 µg per condition | + TiO2 chemistry used for depletion well established – limited to LysC for digestion – challenging differential α-amine labeling | Chen et al. (2016b) |
| StagAu | Modified and neo-N-termini, other peptides removed by sulphydryl coupling to gold particles | Custom-made gold-covered graphene particles | 10 µg per sample | + low starting material – gold nanoparticles not readily available | Li et al. (2016) |

aStarting material refers to amount of extracted and purified protein/proteome
bReferences printed in italics are step-by-step protocols, and those in bold list applications in plants.

Anti-N-arginyl- antibodies can be used in two ways (Fig. 3). First, they are applied in pipelines to positively enrich N-terminally arginylated proteins by capture and at the same time to deplete non-arginylated protein from the sample (Fig. 3A–F). The protocols make use of two peptides initiated with Arg–Asp or Arg–Glu, respectively, to generate ‘pan-arginylation’ antibodies (Table 3). Their random peptide backbone is unrelated to any known sequence but consists of a highly immunogenic stretch of five bulky, charged residues (composed of the amino acids Arg, Asp, Glu, and Lys), followed by seven small uncharged residues with predicted low immunogenicity (of the group of Asn, Gln, His, Ser, and Val) (Wong et al., 2007; Kashina, 2015). Therefore, these antibodies should generally recognize proteins initiating with either of these two arginylated residues rather than specific targets. To reduce background from byproduct antibodies reactive against shorter versions of the peptide, for example lacking the N-terminal Arg, the specific antibodies are negatively selected against ‘scrambled’ synthetic peptides with similar amino acid content but different sequence (Fig. 3A–D) (Wong et al., 2007; T. Xu et al., 2009; Kashina, 2015). Challenges when working with N-terminal arginylation are potential cleavage of Arg due to the proteases used in sample preparation and identification in bottom–up proteomics, as standard database search parameters render peptides bearing an additional Arg as modification invisible. For identification, variable modification of N-terminal Asp, Glu, and Cys residues by arginylation (Ebhardt, 2015; White et al., 2017) or dimethyl-arginylation (Hoernstein et al., 2016) needs to be considered.

The same antibodies enable identification and quantification of modified N-termini by classical dot or western blot and immunocytochemistry (Fig. 3G). For this, a much broader portfolio of different and highly specific antibodies was raised against peptides according to the sequences of N-termini of putative ATE target proteins with an additional Arg residue at the N-terminus (Table 3). Thus, peptides were mimicking N-terminally arginylated target proteoforms. Also these antibodies are generated by classical immunization, negatively selected using the non-arginylated peptides to eliminate non-specific IgGs, followed by positive affinity purification using the arginylated peptides (Piatkov et al., 2012; Cha-Molstad et al., 2015a, b). These antibodies have in the past exclusively been used to qualitatively determine the presence of targets against which they were raised. Similarly, antibodies specific for N-terminal acetylated proteoforms were raised against specific targets with the help of acetylated peptides (Table 4; Shemorry et al., 2013). Specific antibodies have also been raised to detect selected protease-generated neo-N-termini and their modified (arginylated) versions. These must be raised individually for each target proteoform and were shown to be highly specific for the cleaved (and modified) versus uncleaved sites. This is usually calibrated by using either recombinant protein with or without the N-terminal modification or the uncleaved versus the cleaved form or by synthetic peptides in dot blots (Klecker...
Ratios of the amounts of proteolysis product to precursor can be experimentally determined by comparing signals derived from a neo-N-terminus-specific antibody versus a generic one that detects the target irrespective of cleavage and/or modification. However, these highly specific antibodies are not yet applicable for general enrichment (Huesgen et al., 2014).

**Negative selection of N- or C-terminal peptides by diagonal chromatography**

A pioneering negative selection strategy, combined fractional diagonal chromatography (COFRADIC), enriches N-terminal peptides with two consecutive reversed phase (RP) chromatography runs (Gevaert et al., 2003). First, primary amines are blocked by acetylation prior to proteome digestion, which may include differential stable isotope labels to distinguish between endogenously and in vitro acetylated N-termini. A first chromatography run separates digested peptides into >12 fractions. Each fraction is treated with 2,4,6-trinitrobenzenesulfonic acid (TNBS) to modify unblocked α-amines, resulting in increased hydrophobicity of digest-generated peptides. Each fraction is then subjected to a second chromatography run, where TNBS-modified peptides shift towards later retention times while N-terminal peptides are collected at unchanged retention times. Further refinements of the COFRADIC protocol included the pre-enrichment of acetylated N-terminal peptides by SCX chromatography and enzymatic recovery of digest-generated peptide α-amines that were blocked by spontaneous pyroglutamate formation from peptide N-terminal Glu and Gln residues (Staes et al., 2008). Based on the depletion of internal peptides by SCX, a version of COFRADIC was developed that allows subsequent separation of N- and C-terminal peptides, increasing the identification for both types of terminal peptides (van Damme et al., 2010). In short, after removing the majority of internal peptides by SCX chromatography, the N- and C-terminal peptides are separated and fractionated via a first RP chromatography run, after which each fraction contains a subset of N- and C-terminal peptides. Unblocked α-amines of the C-terminal peptides are then butyrylated to increase their hydrophobicity. Each fraction is then subjected to a second RP chromatography separation, where acetylated N-terminal peptides will elute at the same retention time as

---

**Fig. 3.** Enrichment and detection via N-terminally specific antibodies. Antibodies produced against N-terminal modifications can be used for enrichment and/or detection of modified proteins in various protocols and assays. (A) Antibodies are raised against peptides mimicking the N-termini of the desired target proteins or, as in the case of the ‘pan-arginylation’ antibodies, a random, highly immunogenic sequence. (B) Antisera are derived from different, subsequent bleedings to acquire a range of differentially performing antibodies and select ideal reagents. The resulting antisera need to be purified against unspecific (C) and specific peptides (D) lacking or showing the N-terminal modification (arginylation/additional Arg residue at the N-terminus of the natural protein sequence, acetyl- or formyl-groups). (E) Positive selection of N-terminally modified proteins or peptides by capture with the specific antibodies and wash-out of non-target proteins/peptides, followed by elution for identification by MS (F). The same antibodies can be utilized in immunological assays such as western blots (G) or immunocytochemistry (not shown). Red, unspecific antibody populations; green, specific antibody populations; R, Arg; AcM, acetylated Met; fM, formylated Met; C, Cys included at the C-terminus of antigenic peptide to increase antigenicity.
in the first run while the butyrylated C-terminal peptides will display a shift towards later elution times.

In a variation of the diagonal chromatography, charge-based fractionation of SCX chromatography (ChaFRADIC) employs two consecutive SCX chromatography steps for enrichment of N-terminal peptides (Venne et al., 2013, 2015). N-termini are blocked by stable isotope labeling reagents prior to digestion and SCX fractionation into five fractions of peptides carrying defined charges (+1, +2, +3, +4, and >+4). Next, α-aminides in each fraction are modified through for subsequent MS-based identification.

Table 3. Antibodies raised against N-terminal arginylated sequences used for enrichment or immunological detection.

| Target | Use | Application | Sequence | Reference |
|--------|-----|-------------|----------|-----------|
| Generic | Enrichment | Mass spectrometry, IP-MS | (R/D/EHKHANQHMSVC | Wong et al. (2007); T. Xu et al. (2009); Hoernstein et al. (2016) |
| Generic | Enrichment | Mass spectrometry, IP-MS | (R/D/EHKHANQHMSVC | Saha and Kashina (2011) |
| Generic | Enrichment | Mass spectrometry, IP-MS | (R/D/EHKHANQHMSVC | Hoernstein et al. (2016) |
| Arg-CRT (calretilcin; different epitope, see below) | Detection | Western blot, immunocytochemistry | (R/DPAVYFK | Decca et al. (2007) |
| Arg-BRCA1 (different epitope, see below) | Detection | Western blot | (R/DVEIGQHTSCF | Piatkov et al. (2012) |
| Arg-β actin | Detection | Western blot | (R/DDIAAL | Saha et al. (2012) |
| Arg-BIP (ER chaperone Bip, also known as GRP78 and HSPA5, heat shock 70 kDa protein 5) | Detection | Western blot | (R/D/EEKDKEVGC | Cha-Molstad et al. (2015b); Jiang et al. (2016); Yoo et al. (2018) |
| Arg-CRT (calretilcin) | Detection | Western blot | (R/EPAVYFKEQ | Cha-Molstad et al. (2015b); Yoo et al. (2018) |
| Arg-PDI (protein disulfide isomerase) | Detection | Western blot | (R/DAGP30 | Cha-Molstad et al. (2015); Yoo et al. (2018) |
| Arg-CDC6 | Detection | Western blot | (R/DEPTKASPPK | Yoo et al. (2018) |
| Arg-BRCA1 | Detection | Western blot | (R/DGEIKEDTSAF | Yoo et al. (2018) |
| Generic (MD-D2-eK-ha) | Detection | Western blot | (R/DMMKIPKDLNLNC | Hwang et al. (2010) |
| Cog1 | Detection | Dot blot (verification), western blot | Ac-MDKLPLFRDS | Shernony et al. (2013) |
| MATn2 | Detection | Dot blot (verification), western blot | f-MDIAGLTYG | Piatkov et al. (2015); Kim et al. (2018) |

Table 4. Antibodies raised against N-terminally acetylated or formylated sequences used for immunological detection

| Target | Use | Application | Sequence | Reference |
|--------|-----|-------------|----------|-----------|
| Generic | Detection | Dot blot (verification), western blot | Ac-MNKIPKDLNLNC | Hwang et al. (2010) |
| Cog1 | Detection | Dot blot (verification), western blot | Ac-MDKLPLFRDS | Shernony et al. (2013) |
| Generic (MD-D2-eK-ha) | Detection | Dot blot (verification), western blot | f-MDIAGLTYG | Piatkov et al. (2015); Kim et al. (2018) |

Negative selection of terminal peptides using polymer-based depletion of non-target peptides

Terminal amine isotope labeling of substrates (TAILS) enriches N-terminal peptides by covalent binding of undesired internal and C-terminal peptides to a high molecular weight, hyper-branched polyglycerol (HPG) polymer functionalized with aldehyde groups (Kleifeld et al., 2010). Primary amines are blocked by reductive dimethylation with formaldehyde isoto-poses or other amine-reactive stable isotope labeling reagents such as iTRAQ (Prudova et al., 2010) or TMTs, enabling multiplexing of up to 10 samples (Klein et al., 2018). Labeled pro- teomes are combined and excess reagents removed by protein precipitation before proteolytic digestion. Unblocked, digest-generated peptides with free α-aminides are then captured by the aldehyde functionalized HPG polymer using cyanoborohydrine. The polymer with bound peptides is removed by ultrafiltration, leaving only N-terminal peptides in the flow through for subsequent MS-based identification.

C-terminal amine-based isotope labeling of substrates (C-TAILS) follows a similar strategy to select for C-terminal peptides (Schilling et al., 2010). However, the C-TAILS re- quires more sample preparation steps than its N-terminal counterpart. To avoid cross-reactivity, protein primary amines are first protected by reductive dimethylation before modification of the carboxyl groups at acidic residues and protein C-termini by EDC [1-ethyl-3-(3-dimethylaminopropyl)
Carboxydiimide]-mediated condensation of ethanolamine. Proteome digestion is followed by a second labeling step of the primary amines, including the protease-generated α-amines of the C-terminal peptides. Finally, C-terminal carboxyl groups of digest-generated peptides are covalently linked to the primary amines of high molecular weight linear polyallylamine polymer (PAA) with EDC. Further optimization of the C-TAILS protocol improved amidation conditions, evaluated several alternative carboxyl and amine blocking reagents, and protected amines by acetylation instead of dimethylation (Y. Zhang et al., 2015). Reportedly, amine protection prior to amidation can be omitted, saving a labor-intensive precipitation step and allowing the use of Lys-specific digestion enzymes (Y. Zhang et al., 2018).

**Alternative methods for negative selection of N-terminal peptides**

A variety of different chemical reactions have been applied to exploit the digest-generated α-amine functionality for depletion of digest-generated peptides (Table 2). As a prerequisite, all free α-amines are modified on the protein level by acetylation (McDonald and Beynon, 2006; Yeom et al., 2017), propionylation (Yeom et al., 2017), or dimethylation (Mommen et al., 2012; Lai et al., 2015; Chen et al., 2016a; Li et al., 2016). A pioneering protocol modified α-amines with NHS-biotin for subsequent depletion with immobilized streptavidin (McDonald et al., 2005). In an optimized version, the same group directly captured the digest-generated peptides with NHS-activated Sepharose (McDonald and Beynon, 2006). A further modified protocol termed NRich combined internal peptide depletion via NHS-activated Sepharose with the popular filter-aided sample preparation (FASP) (Wiśniewski, 2017; Yeom et al., 2017). The charge reversal approach adds two covalently linked disulfonate groups to digest-generated α-amines, resulting in a strong negative charge that enables their depletion by SCX chromatography (Lai et al., 2015). This procedure can be performed in tip-based format for limited samples such as excised gel slices. Hydrophobic tagging-assisted N-termini enrichment (HYTANE) modifies digest-liberated α-amines with hexadecanoyl (Chen et al., 2016a). This increases peptide hydrophobicity and enables depletion using C18 reverse phase trap columns, where hexadecanoyl-alkylated peptides remain bound even during washes with 80% acetonitrile. In STAgAu, α-amines of digest-generated peptides are modified with sulphydryl groups using Traut’s Reagent (Li et al., 2016). This enables efficient depletion by the tight interaction of sulphydryl groups with gold-coated nanoparticles (Li et al., 2016). The phosho-tag (PTAG) approach modifies digest-generated α-amines with glyceraldehyde–3–phosphate, followed by depletion using TiO2 affinity that is well established for the enrichment of phosphorylated peptides (Mommen et al., 2012).

The PTAG strategy was further adapted for the enrichment of C-terminal peptides (Chen et al., 2016b). Carboxyl groups are amidated by EDC-mediated reaction with methylamine, followed by LysC digestion. The resulting C-terminal peptides thus do not possess a Lys residue, whereas internal and N-terminal peptides carry a C-terminal Lys. In the next steps, α-amines are selectively protected by dimethylation at acidic conditions before selective attachment of the PTAG to Lys ε-amines of the LysC-digested peptides and depletion of the phosphorylated peptides by TiO2 affinity. However, a major limitation of the method is the strict requirement for LysC as digestion enzyme, which prevents identification of C-termini where no Lys residue is suitably placed to result in MS-identifiable peptides.

**A snapshot of the plant terminome landscape**

Termini-centric studies are still rare in plant sciences despite the variety of different techniques and their wide application and impact in medical sciences (Demir et al., 2018; Klein et al., 2018). However, emerging applications are providing new and broader insights into protein synthesis, maturation, modification, and degradation in various aspects of plant life (Fig. 4).

An extensive 6-plex iTRAQ ChaFRADIC analysis of Arabidopsis seedlings employing multiple proteases for digestion identified 2791 N-terminal peptides of 2249 unique N-termini from 1270 proteins (Venne et al., 2015). The data experimentally confirmed sequence determinants for N-terminal initiator Met excision (NME) and many transit, signal, and propeptide cleavage events annotated in UniProt.
A similar study characterized the N-terminome of Arabidopsis chloroplasts (Rowland et al., 2015). TAILS N-terminome analysis of whole leaf and isolated chloroplast proteomes identified a total of 894 N-termini from 577 proteins, including 544 N-termini from 250 chloroplast proteins. Many of the nuclear-encoded proteins were observed to be ragged with two to three different N-termini, suggesting that these proteins undergo further processing directly after chloroplast import. Interestingly, for all 16 nucleus-encoded proteins with known dual targeting to chloroplasts, mitochondria, and/or the cytosol, only one N-terminus representing the chloroplast-localized form was observed. However, this may result from a bias towards the more abundant plastid proteoforms of these proteins in the investigated material. For 126 plastid proteins, only the mature N-terminus was observed, while for 100 proteins multiple N-termini were identified, including (i) N-termini located before the transit peptide cleavage site including precursor proteoforms; (ii) N-termini in close proximity to the predicted transit peptide cleavage site representing the mature plastid proteoform; and (iii) N terminal reflecting degradation products. An analysis of the N-terminal residue occurrence found an over-representation of stabilizing residues (A, V, T, and S) for nuclear-encoded proteins. A similar distribution was found for the plastid-encoded proteins, with additional frequent occurrence of intact initiator Met. Destabilizing residues were significantly under-represented as N-terminal residues, providing indirect evidence for the existence of a plastid version of the N-end rule.

**N-terminomics for identification of alternative translation initiation sites**

Identification of bottom-up-based proteomics data depends greatly on well-annotated protein models, which in turn rely on correct identification of the protein translation initiation sites. N-terminome analysis is therefore ideally suited to obtain experimental evidence for predicted protein models (Hartmann and Armengaud, 2014). A novel proteogenomics pipeline combined computational genome analysis with ribosome profiling and N-termini profiling by COFRADIC to identify novel translation initiation sites (TISs) in Arabidopsis (Willems et al., 2017). In the first stage, N-terminal peptides were enriched from Arabidopsis cell culture after parallel digest with complementary proteases to increase the coverage depth. Peptides were identified from a six-frame translation Arabidopsis database with three different search engines. In the second stage, unmatched MS/MS spectra were searched against a customized database consisting of N-terminal peptide sequences derived from translation ribosome profiling data, computationally predicted gene models, and a six-frame translation of the genomic data. Using strict thresholds criteria, experimental evidence for translation at 117 novel TIS locations was identified, of which 50 mapped to intergenic regions, 44 partly overlapped with existing protein models, 2 were located in pseudogenes, and 21 were found in transposable elements. Furthermore, 23 N-termini matched alternative translation initiation sites identified by ribosome profiling, providing experimental evidence for the physiological relevance of such alternative ribosome-binding sites. Importantly, the results obtained from Arabidopsis further enabled prediction of novel protein-coding genes in other species, with both well-annotated and poorly annotated genomes.

**Profiling of N-terminal protein modifications**

A landmark study compared protein N-terminal modifications such as NME, acetylation, and N-myristoylation in human and Arabidopsis cell cultures. Comparative SILProNAQ analysis identified 1072 N-terminal peptides from 1007 Arabidopsis proteins, and 717 N-terminal peptides from 715 human proteins (Bienvenut et al., 2012). In both data sets, ~70% of the proteins were subject to NME and 90% of proteins were N-terminally acetylated in the steady state. Only five and three myristoylated N-termini were observed in the Arabidopsis and human sample, respectively. The vast majority of acetylated peptides represented annotated expected N-termini. In Arabidopsis, 148 N-terminal peptides of plastid-located proteins were post-translationally acetylated after signal cleavage, indicating the existence of a plastidial N-terminal acetyltransferase (NAT). Three years later, the same group identified NAA70 as the first plastid-located NAT (Dinh et al., 2015). Recombinant NAA70 was able to acetylate both protein termini starting with the initiator Met and NME-processed termini in vitro, with the sequence specificity matching the known acetylated termini of plastid proteins.

SILProNAQ was also instrumental for the identification of a critical role for the cytosolic NatA complex in drought stress tolerance in Arabidopsis (Linster et al., 2015). As in human or yeast, the NatA complex contains two subunits, the catalytic subunit NAA10 and the auxiliary subunit NAA15. Comparison of the global acetylation profiles of wild-type Arabidopsis and artificial miRNA (ami) lines of NAA10 and NAA15, respectively, revealed that both NAA10 and NAA15 are necessary for NatA activity and responsible for the acetylation of almost half of the soluble proteins in the leaves. Interestingly, both ami-NAA10 and amiNAA15 lines displayed drought stress tolerance, which could be traced to closure of stomatal apertures and altered root morphology. The steady-state transcriptomes of both mutant lines highly correlated with the transcriptome of drought-stressed wild-type plants. Abscisic acid, an important phytohormone in drought stress signaling, induced rapid degradation of NatA, resulting in an increase of free, unacylated N-termini in drought-stressed wild-type plants.
A very recent study mapped N-terminal myristoylation in Arabidopsis and human cells in vitro on a proteome-wide scale and using SILProNAQ in vivo (Castrec et al., 2018). First, the substrate preference of human N-myristoyltransferase (NMT) beyond the known requirement for N-terminal Gly was deduced from the crystal structure with co-crystalized substrate peptides and used to predict potential NMT substrates in the human and, based on the high NMT sequence similarity, also in Arabidopsis proteomes. A subset of several hundred termini including both predicted substrates and non-substrates were validated in vitro with recombinant protein. Hundreds of protein starting with Gly termini were thus shown to be myristoylated in both organisms in vitro, while SILProNAQ identified myristoylation of 46 human proteins and 75 Arabidopsis proteins in vitro, predominantly in membrane-enriched fractions. Interestingly, myristoylation and acetylation appear to compete for common substrates, providing a novel co-translational regulatory mechanism possibly affecting subcellular localization.

N-terminome profiling reveals conserved multi-step proteolytic maturation of nuclear-encoded proteins imported in plastids and mitochondria

Plant cells harbor two organelles of endosymbiotic origin, mitochondria and plastids, that rely on the import of most of their protein complement by N-terminal targeting sequences to mitochondria and plastids, that rely on the import of most of their protein complement by N-terminal targeting sequences. The Arabidopsis study further investigated the consequences of impaired plastid protein import in the tic56-1 and ppi2 mutants lacking critical subunits of the translocase supercomplex in the inner and outer envelope membrane, respectively (Köhler et al., 2015b). Both mutants showed albino phenotypes and required an external carbon source for growth. TAILS N-terminome analysis identified N-terminal peptides of 348, 481, and 280 proteins from the wild type, tic56-1, and ppi2, respectively. Despite the severe phenotype, only two and four N-termini from nuclear-encoded plastid proteins matched unprocessed proteoforms with intact transit peptides in tic56-1 and ppi2, whereas the majority of plastid protein termini indicated unaltered proteolytic processing to the mature, imported proteoforms. A great overlap of imported proteins was observed in both mutant lines, although two different sites of import were affected, and the proteins showed no significant enrichment of specific functional categories. This further suggested the existence of redundant or alternative translocase complexes, as import of a subset of plastid proteins appeared unaffected.
Global analysis of N-terminal protein stability determinants

In the Arg/N-end rule pathway, basic N-terminal residues are targeted for proteasomal degradation by the Arabidopsis E3 ligase PROTEOLYSIS 6 (PRT6). Acidic residues and oxidized Cys may be N-terminally arginylated by the arginyltransferases ATE1 and 2, effectively converting N-termini with these residues to PRT6 substrates (Dissmeyer et al., 2018; Dissmeyer, 2019). Triplex dimethyl TAILS N-terminome profiling of prt6 and ate1/2 identified 1465 unique N-termini, of which 864 were acetylated and 601 dimethylated, corresponding to free N-termini in vivo (H. Zhang et al., 2015). Acetylated peptides predominantly exposed stabilizing residues Ala and Ser and mapped to the expected translation start site with and without subsequent NME and chloroplast transit peptide cleavage sites. The majority of the protease-generated neo-N-termini also exposed stabilizing residues (Ala, Ser, Thr, and Val), but a moderate number of primary (Phe, Leu, and Ile) and secondary (Asp and Glu) destabilizing residues were also observed. However, none of these N-terminal peptides accumulated significantly in the N-end rule mutants. Quantitative root proteome analysis found only 17 out of almost 3300 quantified proteins significantly up-regulated in both lines prt6 and ate1/2. Most of these reflected transcriptional changes, including genes controlled by the known Arg/N-end rule substrates, the group VII ethylene response factors (ERFVII). This suggested that the Arg/N-end rule pathway targets only a small number of substrates under standard growth conditions.

In a follow-up study, the same group assessed the impact of PRT6 in etiolated Arabidopsis seedlings (H. Zhang et al., 2018). Six-plex TMT TAILS analysis of the wild type and prt6 loss-of-function mutants with two complementary digestion enzymes resulted in a rich N-terminome data set of 5004 unique N-terminal peptides from 2396 protein groups. The N-termini of several proteases including RD21A were down-regulated, while the N-termini of 45 protein groups including cruciferins were significantly increased in the mutant line, most of which were directly or indirectly regulated by ERFVII transcription factors. Activity-based protein profiling confirmed reduced RD21 activity and increased cathepsin B activity, indicating that PRT6 affects protease activity and mobilization of seed storage proteins in a complex manner.

Identification of N-terminal arginylation by immunodetection

Prominent examples of post-translationally arginylated, proteolytic proteoforms (protein cleavage products) in the animal field are the endoplasmic reticulum (ER) chaperone BiP (also known as GRP78 and HSPA5, heat shock 70 kDa protein 5) (Cha-Molstad et al., 2015b; Jiang et al., 2016; Yoo et al., 2018), CRT (calreticulin) (Decca et al., 2007; Cha-Molstad et al., 2015b; Yoo et al., 2018), BRCA1 (Piatkov et al., 2012; Yoo et al., 2018), β actin (Saha et al., 2012), PDI (protein disulfide isomerase) (Cha-Molstad et al., 2015a; Yoo et al., 2018), and CDC6 (Yoo et al., 2018). In these studies, important roles of N-terminal arginylation in autophagy and recognition of N-terminally arginylated targets by the N-recogin p62 were shown (Cha-Molstad et al., 2015a, b, 2016, 2017, 2018).

There are other examples for such antibodies raised against a plant-derived arginylated epitope, however, not in the context of N-terminal arginylation (Havé et al., 2018). The strength of the antibody signals can also be related to in vitro arginylation assays where recombinant ATEs transfer Arg to short acceptor peptides initiated with Asp, Glut, or Cys sulfonic acid (Klecker and Dissmeyer, 2016; White et al., 2017). Here, the ideal antibody dilutions needs to be precisely determined using control arginylated versus non-arginylated peptides.

Antibodies detecting co-translationally modified Met

N-terminal acetylation of cellular proteins creates specific degradation signals termed Ac/N-degrons that can be targeted by the Ac/N-end rule pathway (Hwang et al., 2010; Gibbs, 2015). For example, the N-terminally acetylated Met residue of MAT02 was found to act as a degradation signal. It is similar to N-terminally acetylated Leu, G1y, AlA, Val, Pro, Ser, Thr, and Cys, recognized by the Dna10 ubiquitin ligase in yeast (Hwang et al., 2010). Then, Cog1, a subunit of the Golgi-associated COG complex in yeast, is an example for such a target and an example where an antibody specific for N-terminal acetylated proteins was used to determine levels of N-terminally modified protein in comparison with the non-acetylated proteoform (Shemorry et al., 2013). The use of anti-acetylation antibodies for enrichment of N-terminally acetylated proteins

Profiling of N-terminal arginylation by IP-MS

The previously mentioned ‘pan-arginylation’ antibodies were used to enrich arginylated proteins by immunoprecipitation from mouse tissue extracts, followed by MS-based identification (IP-MS). Here, 43 diverse proteins with important physiological roles were arginylated in vivo at highly specific sites. These appeared to be located both at the N-terminus of Asp- and Glu-initiated proteins and at the side chains of midchain Glu residues (Wong et al., 2007; T. Xu et al., 2009). In the moss Physcomitrella patens, several candidate substrates of ATE1 (Schuessele et al., 2016) were found by an IP-MS approach using the described ‘pan-arginylation’ antibodies (Wong et al., 2007; T. Xu et al., 2009). The identified arginylated peptides belonged to an uncharacterized protein and a putative AAA-type ATPase. Arginylation in the ABC transporter family protein PpABCB20 was found to be likely by using a less stringent search algorithm. Experimental evidence for arginylation occurring at an N-terminal Asp was reported for the acylamino-acid-releasing enzyme PpAARE (Hoernstein et al., 2016). However, it remains unclear how such an N-terminal Asp residue could be generated because it is in position 2, preceded by the initiator Met that should be retained under these circumstances.
is not common, probably mainly due to the large fraction of the proteome that is actually undergoing this co-translational modification. The lysine acetylome can be treated differently as it can indeed be characterized and quantified by enrichment with antibodies against acetylated lysine residues followed by MS (Choudhary et al., 2009; Guan et al., 2010; Finkemeier et al., 2011; Hartl et al., 2015; Lasowskat et al., 2017; Ree et al., 2018). In contrast to the clear situation in yeast, a recent study showed that the N-terminal acetylation can have antagonistic effects on the stability of two related proteoforms (Xu et al., 2015). The Nod like receptor (NLR) protein SUPPRESSOR OF NPR1, CONSTITUTIVE 1 (SNC1), has two distinct proteoforms due to alternative translation which are selectively acetylated either by NATA or by NATB, respectively. The acetylation by NATB leads to the degradation of SNC1 while the acetylation by NATB enhances SNC's stability (Xu et al., 2015).

Very recently, antibodies specific for N-terminal formylation were also produced (Kim et al., 2018). The yeast formyltransferase Fmt1 is imported into mitochondria and, nonetheless, produces N-terminal formylated proteins in the cytosol. These can be targets of the novel, so-called fMet/N-end rule pathway that aids degradation of N-terminal formylated proteins.

### N-terminomics for unbiased identification of plant protease substrates

Plants contain hundreds of proteases with mostly unknown substrates and function (van der Hoorn, 2008; Demir et al., 2018; Paulus and van der Hoorn, 2019). In a pioneering study, COFRADIC was used for complementary in vitro and in vivo approaches to identify physiological substrates of Arabidopsis metacaspase 9 (MC9) (Tsiatsiani et al., 2013). For the in vitro approach, a proteome extracted from 2-day-old seedlings of a mutant line lacking MC9 was incubated with active recombinant MC9 (rMC9). COFRADIC identified 3050 N-terminal peptides from 1138 proteins, of which 332 N-terminal peptides were significantly more abundant after rMC9 treatment. Analysis of these 332 cleavage sites confirmed the known MC9 preference for cleavage after basic residues at the P1 site. Additionally, a strong preference for acidic residues at the P1’ site was revealed. In the in vitro approach, proteome samples of 2-day-old mc9 loss-of-function seedlings were compared with the wild type or with plants overexpressing MC9 under control of the 35S promoter. COFRADIC analysis revealed 3781 N-terminal peptides from 1705 proteins in the comparison of mc9 and the wild type, and 2879 N-terminal peptides from 1407 proteins in the comparison of mc9 with the MC9 overexpressor. In total, 99 cleavage events matching the MC9 sequence specificity in 74 candidate substrate proteins were observed in at least two of the three experiments. Randomly picked candidate cleavage events were confirmed by rMC9 cleavage of synthetic peptides representing the cleavage site-spanning region and with recombinant proteins produced by in vitro transcription and translation. From those experiments, the phosphoenolpyruvate carboxykinase 1 (PEPCK1) emerged as a highly likely physiological substrate of MC9. Further functional validation revealed that MC9 activated PEPCK1 in vivo, leading to impaired gluconeogenesis and shorter hypocotyls in dark-grown mc9 mutants that resembled the phenotype of pepck1 mutants.

Similarly, COFRADIC was used to define the cleavage specificity and identify candidate substrate of the quality control proteases HhoA, HhoB, and HtrA in the model cyanobacterium Synechocystis sp. PCC 6803 (Tam et al., 2015). COFRADIC analysis of Synechocystis proteome extracts incubated with recombinant HhoA, HhoB, and HtrA revealed similar broad specificity profiles. Combined 2D-difference gel electrophoresis (DIGE) proteome and COFRADIC N-terminome analysis showed that inactivation of each single protease had a similar impact, with differential expression and processing of enzymes involved in major metabolic pathways. Taken together, the in vivo and in vitro data sets provide evidence for RbcS as a physiological HhoA substrate, PsbO as a HhoB and HtrA substrate, and Pbp8 as a HtrA substrate.

### Conclusions

Protein termini provide important information about the functional state of the proteome. A multitude of N-termini enrichment techniques have been developed, but are still rarely applied in plant sciences. Positive selection procedures enable targeted selection of one group of either unmodified or specifically modified N- or C-termini, but often require more sophisticated reagents. Negative selection strategies are often based on established chemical modifications routinely used in many laboratories, and enrich all N- or C-terminal peptides simultaneously, and are therefore ideally suited for unbiased profiling of protein termini and their endogenous modifications. Only a few proteomes have been studied by more than one technique, which generally showed low overlap between the results of different protocols (Chen et al., 2016a). However, inherent biases could not yet be identified as the observed low overlap may also be attributed to stochastic differences due to low proteome coverage or differences in sample amounts or digest enzymes and efficiency.

Most MS-based termini enrichment techniques still require expensive reagents and considerable technical experience that may limit their appeal to non-expert laboratories. Non-specific losses incurred during clean-up and depletion steps, such as by adsorption to plastic wear, filters, and/or bead material, pose challenges for sensitivity and reproducibility that practically exclude analysis of microscale samples such as microdissected tissues. Ongoing efforts in technology development strive to improve sensitivity, robustness, and ease of use of both experimental and data annotation workflows. In addition, termini-centric MS approaches face the challenge that proteoform identification and quantification are based on single peptides. Replicate experiments with stringent statistical analysis should be performed, and termini identifications that are based on single, spurious spectrum matches must be excluded. Nevertheless, single peptide quantification is inherently more variable than protein quantification by standard shotgun proteomics approaches, where data from several peptides are
averaged for each protein. However, targeted MS and antibody-based assays can be developed to enable more accurate quantification and sensitive monitoring of known proteoforms (Huesgen et al., 2014; Savickas and auf dem Keller, 2017). Such targeted assays with more accurate quantification will also enable more sensitive detection of small but meaningful changes between different genotypes or samples, especially in larger sample series. Finally, identification of proteolytically cleaved proteoforms also enables individual quantification of the processed and unprocessed forms (Fahlman et al., 2014). This in turn allows calculation of the degree of processing, similar to site occupancy for other post-translational modifications (Arsova et al., 2018). This may reveal threshold amounts for cellular switches triggering specific irreversible physiological responses or commitment to cell fate decisions.

In summary, studies of plant protein termini have already addressed a broad range of biological problems. These included plant protease substrate identification, profiling of protein terminal modifications, characterization of N-terminal protein stability determinants in various subcellular compartments, protein maturation in organelles of endosymbiotic origin in diverse organisms, and identification of alternative translation initiation sites (Fig. 4). Taken together, these studies have already revealed striking similarities in the N-terminal identity, modification, and maturation, and highlighted the extraordinary abundance of proteolytic proteoforms across evolutionarily distant organisms. With continuous improvement and widespread applications, we expect that more studies of the plant terminomes will emerge in the near future and provide unprecedented insights particularly into proteolytic processes in plants.

Acknowledgements

AP thanks the organizers of the Plant Protease and Programmed Cell Death symposium 2018 in Ghent for supporting his attendance at the meeting with a travel grant. ND and PFH acknowledge support by the European Cooperation in Science and Technology (COST) Action BM1307 – “European network to integrate research on intracellular proteolysis pathways in health and disease (PROTEOSTASIS)’. This work was supported by a starting grant of the European Research Council (ERC) starting grant 639905, to PFH. This work was supported by a grant for setting up the junior research group of the ScienceCampus Halle – Plant-based Bioeconomy and by DFG M111.015131. Molecular & Cellular Proteomics 11, M111.015131.

References

Aebpersold R, Mann M. 2016. Mass-spectrometric exploration of proteome structure and function. Nature 537, 347–355.
Arsova B, Watt M, Usadel B. 2018. Monitoring of plant protein post-translational modifications using targeted proteomics. Frontiers in Plant Science 9, 1168.
Bachmair A, Finley D, Varshavsky A. 1986. In vivo half-life of a protein is a function of its amino-terminal residue. Science 234, 179–186.
Bienvenut WV, Scarpelli JP, Dumestier J, Meinnel T, Giglione C. 2017. EnCOUNTER: a parsing tool to uncover the mature N-terminus of organello-targeted proteins in complex samples. BMC Bioinformatics 18, 182.

Bienvenut WV, Sumpton D, Martinez A, Lilla S, Espagne C, Meinnel T, Giglione C. 2012. Comparative large scale characterization of plant versus mammal proteins reveals similar and idiosyncratic N-acetylation features. Molecular & Cellular Proteomics 11, M111.015131.
Boersema PJ, Rajmakers R, Lemeer S, Mohammed S, Heck AJ. 2009. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. Nature Protocols 4, 484–494.
Carrie C, Venne AS, Zahedi RP, Soll J. 2015. Identification of cleavage sites and substrate proteins for two mitochondrial intermediate peptides in Arabidopsis thaliana. Journal of Experimental Botany 66, 2691–2708.
Castrec B, Dian C, Ciccone S, Ebert CL, Bienvenut WV, Le Caer JP, Steyaert JM, Giglione C, Meinnel T. 2018. Structural and genomic decoding of human and plant myristoylomes reveals a definitive recognition pattern. Nature Chemical Biology 14, 671–679.
Cha-Molstad H, Kwon YT, Kim BY. 2015a. Amino-terminal arginylation as a degradation signal for selective autophagy. EMBO Reports 48, 487–488.
Cha-Molstad H, Lee SH, Kim JG, et al. 2018. Regulation of autophagic proteolysis by the N-recognin SQSTM1/p62 of the N-end rule pathway. Autophagy 14, 359–361.
Cha-Molstad H, Sung KS, Hwang J, et al. 2015b. Amino-terminal arginylation targets endoplasmic reticulum chaperone BIP for autophagy through p62 binding. Nature Cell Biology 17, 917–929.
Cha-Molstad H, Yu JE, Feng Z, et al. 2017. p62/SQSTM1/Sequestosome-1 is an N-recognin of the N-end rule pathway which modulates autophagosome biogenesis. Nature Communications 8, 102.
Cha-Molstad H, Yu JE, Lee SH, et al. 2016. Modulation of SQSTM1/p62 activity by N-terminal arginylation of the endoplasmic reticulum chaperone HSP5/GRP78/BIP. Autophagy 12, 426–428.
Chang TK, Jackson DY, Burnier JP, Wells JA. 1994. Subtiligase: a tool for semisynthesis of proteins. Proceedings of the National Academy of Sciences, USA 91, 12544–12548.
Chen L, Shan Y, Weng Y, Sui Z, Zhang X, Liang Z, Zhang Y. 2016a. Hydrophobic tagging-assisted N-termini enrichment for in-depth N-terminome analysis. Analytical Chemistry 88, 8390–8395.
Chen L, Shan Y, Weng Y, Yuan H, Zhang S, Fan R, Sui Z, Zhang X, Zhang L, Zhang Y. 2016b. 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, 3867–3874.
Cheng YL, Tu SL. 2018. Alternative splicing and cross-talk with light signal- ing. Plant & Cell Physiology 59, 1104–1110.
Choudhary C, Kumar C, Gnaf N, Nielsen ML, Rehman M, Walther TC, Olsen JV, Mann M. 2009. Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science 325, 834–840.
Decca MB, Carpio MA, Bosc C, Galiano MR, Job D, Andreux A, Hallak ME. 2007. Post-translational arginylation of calreticulin: a new isospecies of calreticulin component of stress granules. Journal of Biological Chemistry 282, 8237–8245.
Demir F, Niedermaier S, Kizhakkedathu JN, Huesgen PF. 2017. Profiling of protein N-termini and their modifications in complex samples. Methods in Molecular Biology 1574, 35–50.
Demir F, Niedermaier S, Villamor JG, Huesgen PF. 2018. Quantitative proteomics in plant protease substrate identification. New Phytologist 218, 936–943.
Dinh TV, Bienvenut WV, Linster E, Feldman-Salit A, Jung VA, Meinnel T, Hell R, Giglione C, Wirtz M. 2015. Molecular identification and functional characterization of the first Nε-acetyltransferase in plastids by global acetylamino profiling. Proteomics 15, 2426–2435.
Disbmayery N. 2017. Conditional modulation of biological processes by low-temperature degrons. Methods in Molecular Biology 1669, 407–416.
Disbmayery N. 2019. Conditional protein function via N-degron pathway-mediated proteostasis in stress physiology. Annual Review of Plant Biology 70 (In press).
Disbmayery N, Rivas S, Graciet E. 2018. Life and death of proteins after protease cleavage: protein degradation by the N-end rule pathway. New Phytologist 218, 929–935.
Dong H, Dumennil J, Lu FH, et al. 2017. Ubiquitylation activates a peptide that promotes cleavage and destabilization of its activating E3 ligases and diverse growth regulatory proteins to limit cell proliferation in Arabidopsis. Genes & Development 31, 197–208.

Characterization of plant protein termini | 2035
Dormeyer W, Mohammed S, Breukelen Bv, Krijgsvejd J, Heck AJ. 2007. Targeted analysis of protein termini. Journal of Proteome Research 6, 4654–4654.

Duan W, Xu G. 2017. ProC-TEL: profiling of protein C-termini by enzymatic labeling. Methods in Molecular Biology 1574, 135–144.

Duan W, Zhang Y, Xu G. 2016. Optimization and application of protein C-terminal labeling by carboxypeptidase Y. Chinese Journal of Biotechnology 32, 135–148.

Ehhardt HA. 2015. Applying arginylation for bottom-up proteomics. Methods in Molecular Biology 1337, 129–138.

Eriste E, Norberg A, Nepomuceno D, et al. 2005. A novel form of neurotensin post-translationally modified by arginylation. Journal of Biological Chemistry 280, 35089–35097.

Faden F, Mielke S, Dissmeyer N. 2019. Modulating protein stability to switch toxic protein function on and off in living plants. Plant Physiology 179, 929–942.

Faden F, Ramezani T, Mielke S, et al. 2016. Phenotypes on demand via switchable target protein degradation in multicellular organisms. Nature Communications 7, 12202.

Fahnle RP, Chen W, Overall CM. 2014. Absolute proteomic quantification of the activity state of proteases and proteolytic cleavages using proteolytic signal peptides and isotopic tags. Journal of Proteomics 100, 73–91.

Finkemeier I, Laxa M, Miguel L, Howden AJ, Sweetlove LJ. 2011. Proteins of diverse function and subcellular location are lysine acetylated in Arabidopsis. Plant Physiology 155, 1779–1790.

Gevaert K, Goethals M, Martens L, Van Damme J, Staes A, Thomas GR, Vandenckhove J. 2003. Exploring proteomes and analyzing protein processing by mass spectrometric identification of sorted N-terminal peptides. Nature Biotechnology 21, 566–569.

Ghifari AS, Gill-Hille M, Murcha MW. 2018. Plant mitochondrial protein import: the ins and outs. Biochemical Journal 475, 2191–2208.

Gibbs DJ. 2015. Emerging functions for N-terminal protein acetylation in plants. Trends in Plant Science 20, 599–601.

Gonda DK, Bachmair A, Wünning I, Tobias JW, Lane WS, Varshavsky A. 1989. Universality and structure of the N-end rule. Journal of Biological Chemistry 264, 16700–16712.

Gorman JJ, Shiell BJ. 1993. Isolation of carboxyl-termini and blocked amino-termini of viral proteins by high-performance cation-exchange chromatography. Journal of Chromatography 646, 193–205.

Graciet E, Walter F, O’Malloieidigh DS, Pollmann S, Meyerowitz EM, Varshavsky A, Wellmer F. 2009. The N-end rule pathway controls multiple functions during Arabidopsis shoot and leaf development. Proceedings of the National Academy of Sciences, USA 106, 13618–13623.

Gruber A, Rocap G, Kroth PG, Armbrust EV, Mock T. 2015. Plastic proteome prediction for diatoms and other algae with secondary plastids of the red lineage. The Plant Journal 81, 13618–13623.

Hartl M, König AC, Finkemeier I. 2015. Identification of lysine-acetylated mitochondrial proteins and their acetylation sites. Methods in Molecular Biology 1308, 107–121.

Hartmann EM, Armengaud J. 2014. N-terminomics and proteogenomics, getting off to a good start. Proteomics 14, 2637–2646.

Havé M, Ballian T, Cottin-Bolte B, et al. 2018. Increases in activity of proteasome and papain-like cysteine protease in Arabidopsis autophagy mutants: back-up compensatory effect or cell-death promoting effect? Journal of Experimental Botany 69, 1369–1385.

Hoernstein SN, Mueller SJ, Fiedler K, et al. 2016. Identification of targets and interaction partners of arginyl-tRNA protein transferase in the moss Physcomitrella patens. Molecular & Cellular Proteomics 15, 1808–1822.

Holman TJ, Jones PD, Russell L, et al. 2009. The N-end rule pathway promotes seed germination and establishment through removal of ABA sensitivity in Arabidopsis. Proceedings of the National Academy of Sciences, USA 106, 4549–4554.

Huesgen PF, Alami M, Lange PF, Foster LJ, Schröder WP, Overall CM, Green BR. 2013. Proteomic amino-termini profiling reveals targeting information for protein import into complex plastids. PLoS One 8, e74483.

Huesgen PF, Lange PF, Overall CM. 2014. Ensembles of protein termini and specific proteolytic signatures as candidate biomarkers of disease. Proteomics. Clinical Applications 8, 338–350.

Huesgen PF, Overall CM. 2012. N- and C-terminal degradomics: new approaches to reveal biological roles for plant proteases from substrate identification. Physiologia Plantarum 145, 5–17.

Hwang CS, Shemorry A, Varshavsky A. 2010. N-terminal acetylation of cellular proteins creates specific degradation signals. Science 327, 973–977.

Jiang Y, Lee J, Lee JH, et al. 2016. The arginylation branch of the N-end rule pathway positively regulates cellular autophagic flux and clearance of proteotoxic proteins. Autophagy 12, 2197–2212.

Kapos P, Xu F, Meinnel T, Giglione C, Li X. 2015. N-terminal modifications contribute to flowering time and immune response regulations. Plant Signaling & Behavior 10, e1073874.

Kashina AS. 2015. Development of new tools for the studies of protein arginylation. Methods in Molecular Biology 1337, 139–145.

Kim J-M, Seok O-H, Su S, Heo J-E, Yeom J, Kim D-S, Yoo J-Y, Varshavsky A, Lee C, Hwang C-S. 2018. Formyl-methionine as an N-degron of a eukaryotic N-end rule pathway. Science 362, 1014.

Kim ST, Tasaki T, Zakrzewska A, et al. 2013. The N-end rule proteolytic system in autophagy. Autophagy 9, 1100–1103.

Klecker M, Dissmeyer N. 2016. Peptide arrays for binding studies of E3 ubiquitin ligases. Methods in Molecular Biology 1450, 85–94.

Kleifeld O, Doucet A, auf dem Keller U, Prudova A, Schilling O, Kainthan RK, Starr AE, Foster LJ, Kizhakkedathu JN, Overall CM. 2010. Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products. Nature Biotechnology 28, 281–288.

Kleifeld O, Doucet A, Prudova A, auf dem Keller U, Gioia M, Kizhakkedathu JN, Overall CM. 2011. Identifying and quantifying proteolytic events and the natural N-terminome by terminal amine isotopic labeling of substrates. Nature Protocols 6, 1578–1611.

Klein T, Eckhard U, Dufour A, Solis N, Overall CM. 2018. Proteolytic cleavage—mechanisms, function, and ‘omic’ approaches for a near-ubiquitous posttranslational modification. Chemical Reviews 118, 1157–1168.

Klein T, Fung SY, Renner F, et al. 2015. The paracaspase MALT1 cleaves HOIL1 reducing linear ubiquitination by LUBAC to dampen lymphocyte NF-κB signalling. Nature Communications 6, 8777.

Köhler D, Dobritzsch D, Hoehenwarter W, Helm S, Steiner JM, Baginsky S. 2015. Identification of protein N-termini in Cyanophora paradoxa cyanelles: transit peptide composition and sequence determinants for precursor maturation. Frontiers in Plant Science 6, 559.

Köhler D, Montandon C, Hause G, Majovsky P, Kessler F, Baginsky S, Agne B. 2015. Characterization of chloroplast protein import without Tic56, a component of the 1-megadalton translocon at the inner envelope membrane of chloroplasts. Plant Physiology 167, 972–990.

Lai ZW, Gomez-Auli A, Keller EJ, Mayer B, Biniossek ML, Schilling O. 2015. Enrichment of protein N-termini by charge reversal of internal peptides. Proteomics 15, 2470–2478.

Lai ZW, Schilling O. 2017. Identification of protease cleavage sites by charge-based enrichment of protein N-termini. Methods in Molecular Biology 1579, 199–207.

Laloum T, Martin G, Duque P. 2018. Alternative splicing control of abiotic stress responses. Trends in Plant Science 23, 140–150.

Lassowskat I, Hartl M, Hosp F, Boersema PJ, Mann M, Finkemeier I. 2017. Dimethyl-labeling-based quantification of the lysine acetylome and proteome of plants. Methods in Molecular Biology 1653, 65–81.

Li L, Wu R, Yan G, Gao M, Deng C, Zhang X. 2016. 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, 441–448.

Linster E, Stephan I, Bienvenut WV, et al. 2015. Downregulation of N-terminal acetylation triggers ABA-mediated drought responses in Arabidopsis. Nature Communications 6, 7640.

Linster E, Wirtz M. 2018. N-terminal acetylation: an essential protein modification emerges as an important regulator of stress responses. Journal of Experimental Botany 69, 4555–4568.
Liu M, Fang C, Pan X, Jiang H, Zhang L, Zhang L, Zhang Y, Yang P, Lu H. 2015. Positive enrichment of C-terminal peptides using oxazolone chemistry and biotinylaton. Analytical Chemistry 87, 9916–9922.

Liu M, Zhang L, Zhang L, Yao J, Yang P, Lu H. 2013. Approach for identification and quantification of C-terminal peptides: incorporation of isotopic arginine labeling based on oxazolone chemistry. Analytical Chemistry 85, 10745–10753.

Mahrus S, Trinidad JC, Barkan DT, Sali A, Burlingame AL, Wells JA. 2008. Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini. Cell 134, 866–876.

McDonald L, Beynon RJ. 2006. Positional proteomics: preparation of amino-terminal peptides as a strategy for proteome simplification and characterization. Nature Protocols 1, 1790–1798.

McDonald L, Robertson DH, Hurst JL, Beynon RJ. 2005. Positional proteomics: selective recovery and analysis of N-terminal proteolytic peptides. Nature Methods 2, 955–957.

Mommen GP, van de Waterbeemd B, Kersten G, Heck AJ, de Jong AP. 2016. APTAG and TriAD: real-time detection of N-end rule-mediated ubiquitination via fluorescently labeled substrate probes. New Phytologist 217, 613–624.

Niedermaier S, Huesgen PF. 2019. Positional proteomics for identification of secreted proteoforms released by site-specific membrane protein processing. Biochimica et Biophysica Acta (in press).

Paulus JK, van der Hoorn RAL. 2019. Do proteolytic cascades exist in plants? Journal of Experimental Botany 70. doi:10.1093/jxb/erz016

Paresi P, Gardner NA, MASiero S, Dietzmann A, Eichacker L, Wickner R, Salamini F, Leister D. 2005. Cytoplasmic N-terminal protein acetylation is required for efficient photosynthesis in Arabidopsis. The Plant Cell 17, 1817–1832.

Potuschak T, Staray S, Schögelhofer P, Becker F, Nejinskaia V, Huesgen PF. 1991. The N-end rule pathway counteracts cell death by destroying proapoptotic protein fragments. Proceedings of the National Academy of Sciences, USA 109, E1839–E1847.

Potuschak T, Vuyt T, Hwang CS, Varshavsky A. 2015. Formyl-methionine as a degradation signal at the N-terminus of bacterial proteins. Microbial Cell 2, 376–393.

Rathore OS, Faustino A, Prudêncio P, Van Damme P, Cox CJ, Martinho RG. 2016. Absence of N-terminal acetyltransferase diversification during evolution of eukaryotic organisms. Scientific Reports 6, 21304.

Ree R, Varland S, Arnesen T. 2018. Spotlight on protein N-terminal acetylation. Experimental & Molecular Medicine 50, 90.

Reichman P, Dissmeyer N. 2017. In vivo reporters for protein half-life. Methods in Molecular Biology 1669, 387–406.

Reichman P, Dissmeyer N. 2017. In vivo reporters for protein half-life. Methods in Molecular Biology 1669, 387–406.

Ross PL, Huang YN, Marchese JN, et al. 2004. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Molecular & Cellular Proteomics 3, 1154–1168.

Rowland E, Kim J, Bhuysan NH, van Wijk KJ. 2015. The Arabidopsis chloroplastic stromal N-terminome: complexities of amino-terminal protein maturation and stability. Plant Physiology 169, 1881–1896.

Saha S, Kashina A. 2011. Posttranslational arginylation as a global biological regulator. Developmental Biology 358, 1–8.

Saha S, Wang J, Buckley B, Wang Q, Lilly B, Chernov M, Kashina A. 2012. Small molecule inhibitors of arginyltransferase regulate arginylation-dependent protein degradation, cell motility, and angiogenesis. Biochemical Pharmacology 83, 866–873.

Savickas S, auf dem Keller U. 2017. Targeted degradomics in protein terminomics and protease substrate discovery. Biological Chemistry 398, 959–969.

Schechter I, Berger A. 1967. On the size of the active site in proteases. I. Papain. Biochemical and Biophysical Research Communications 27, 157–162.

Schilling O, Barré O, Huesgen PF, Overall CM. 2010. Proteome-wide analysis of protein carboxy termini: C terminomics. Nature Methods 7, 508–511.

Schilling O, Huesgen PF, Barré O, Overall CM. 2011. Identification and relative quantification of native and proteolytically generated protein C-termini from complex proteomes: C-terminome analysis. Methods in Molecular Biology 781, 59–69.

Schussele C, Hoernstein SN, Mueller SJ, Rodriguez-Franco M, Lorenz T, Lang D, Igloi GL, Reski R. 2016. Spatio-temporal patterning of arginyl-tRNA protein transferase (ATE) contributes to gametophytic development in a moss. New Phytologist 209, 1014–1027.

Sheiner L, Striepen B. 2013. Protein sorting in complex plastids. Biochimica et Biophysica Acta 1833, 352–359.

Shema G, Nguyen MTN, Solari FA, Loroch S, Venne AS, Kollipara L, Sickmann A, Verhelck SJ, Zahedi RP. 2018. Simple, scalable, and ultrasensitive tip-based identification of protease substrates. Molecular & Cellular Proteomics 17, 826–834.

Shemorry A, Hwang CS, Varshavsky A. 2013. Control of protein quality and stoichiometries by N-terminal acetylation and the N-end rule pathway. Molecular Cell 50, 540–551.

Smith LM, Kelleher NL. Consortium for Top Down Proteomics. 2013. Proteoform: a single term describing protein complexity. Nature Methods 10, 186–187.

Solis N, Overall CM. 2018. Identification of protease cleavage sites and substrates in cancer by carboxy-TAILS (C-TAILS). Methods in Molecular Biology 1731, 15–28.

Staes A, Impens F, Van Damme P, Ruttens B, Goethals M, Demol H, Timmerman E, Vandekerckhove J, Gevaert K. 2011. Selecting protein N-terminal peptides by combined fractional diagonal chromatography. Nature Protocols 6, 1130–1141.

Staes A, Van Damme P, Helsens K, Demol H, Vandekerckhove J, Gevaert K. 2008. Improved recovery of proteome-informative, protein N-terminal peptides by combined fractional diagonal chromatography (COFRADIC). Proteomics 8, 1362–1370.

Szakonyi D, Duque P. 2018. Alternative splicing as a regulator of early plant development. Frontiers in Plant Science 9, 1174.

Tam LX, Aigner H, Timmerman E, Gevaert K, Funk C. 2015. Proteomic approaches to identify substrates of the three Deg/HtrA proteases of the cyanobacterium Synechocystis sp. PCC 6803. Biochemical Journal 468, 373–384.

Thompson A, Schäfer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, Neumann T, Johnstone R, Mohammed AK, Hamon C. 2003. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Analytical Chemistry 75, 1895–1904.

Timmer JC, Enoksson M, Wildfang E, et al. 2007. Profiling constitutive proteolytic events in vivo. The Biochemical Journal 407, 41–48.

Timmer JC, Salvesen GS. 2011. N-terminomics: a high-content screen for protease substrates and their cleavage sites. Methods in Molecular Biology 753, 243–255.

Tobias JW, Shrader TE, Rocap G, Varshavsky A. 1991. The N-end rule in bacteria. Science 254, 1374–1377.

Tsatsianis L, Timmerman E, De Bock PJ, et al. 2013. The Arabidopsis metacaspase9 degradation. The Plant Cell 25, 2831–2847.

Van Damme P, Staes A, Bronsoms S, Helsens K, Colaert N, Timmerman E, Aviles FX, Vandekerckhove J, Gevaert K. 2010. Complementary positional proteomics for screening substrates of endo- and exoproteases. Nature Methods 7, 512–515.

van der Hoorn RA. 2008. Plant proteases: from phenotypes to molecular mechanisms. Annual Review of Plant Biology 59, 191–223.

Varshavsky A. 2011. The N-end rule pathway and regulation by proteolysis. Protein Science 20, 1298–1345.
Varshavsky A. 2019. N-degron and C-degron pathways of protein degradation. Proceedings of the National Academy of Sciences, USA 116, 358–366.

Venne AS, Solari FA, Faden F, Parettil T, Dismeyer N, Zahedi RP. 2015. An improved workflow for quantitative N-terminal charge-based fractional diagonal chromatography (ChaFRADIC) to study proteolytic events in Arabidopsis thaliana. Proteomics 15, 2458–2469.

Venne AS, Vogtle FN, Meisinger C, Sickmann A, Zahedi RP. 2013. Novel highly sensitive, specific, and straightforward strategy for comprehensive N-terminal proteomics reveals unknown substrates of the mitochondrial peptidase Lcp55. Journal of Proteome Research 12, 3823–3830.

Vicente J, Mendiondo GM, Pauwels J, et al. 2019. Distinct branches of the N-end rule pathway modulate the plant immune response. New Phytologist 221, 988–1000.

Vogtle FN, Wortelkamp S, Zahedi RP, et al. 2009. Global analysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability. Cell 139, 426–439.

Wadas B, Piatkov KI, Brower CS, Varshavsky A. 2016. Analyzing N-terminal arginylation through the use of peptide arrays and degradation assays. Journal of Biological Chemistry 291, 20976–20992.

Wang J, Han X, Wong CC, et al. 2014. Arginyltransferase ATE1 catalyzes midchain arginylation of proteins at side chain carboxylates in vivo. Chemistry & Biology 21, 331–337.

Weeks AM, Wells JA. 2018. Engineering peptide ligase specificity by proteomic identification of ligation sites. Nature Chemical Biology 14, 50–57.

White MD, Klecker M, Hopkinson RJ, et al. 2017. Plant cysteine oxidases that directly enable arginyl transferase-catalysed arginylation of N-end rule targets. Nature Communications 8, 14690.

Wiita AP, Seaman JE, Wells JA. 2014. Global analysis of cellular proteolysis by selective enzymatic labeling of protein N-termini. Methods in Enzymology 544, 327–358.

Willems P, Ndah E, Jonckheere V, Stael S, Sticker A, Martens L, Van Breusegem F, Gevaert K, Van Damme P. 2017. N-terminal proteomics assisted profiling of the unexplored translation initiation landscape in Arabidopsis thaliana. Molecular & Cellular Proteomics 16, 1064–1080.

Wiśniowski JR. 2017. Filter-aided sample preparation: the versatile and efficient method for proteomic analysis. Methods in Enzymology 585, 15–27.

Wong CC, Xu T, Rai R, Bailey AO, Yates JR 3rd, Wolf YI, Zebroski H, Kashina A. 2007. Global analysis of posttranslational protein arginylation. PLoS Biology 5, e258.

Xu F, Huang Y, Li L, et al. 2015. Two N-terminal acetyltransferases antagonistically regulate the stability of a nod-like receptor in Arabidopsis. The Plant Cell 27, 1547–1562.

Xu G, Jaffrey SR. 2010. N-CLAP: global profiling of N-termini by chemoselective labeling of the alpha-amine of proteins. Cold Spring Harbor Protocols 2010, pdb.p05528.

Xu G, Shin SB, Jaffrey SR. 2009. Global profiling of protease cleavage sites by chemoselective labeling of protein N-termini. Proceedings of the National Academy of Sciences, USA 106, 19310–19315.

Xu G, Shin SB, Jaffrey SR. 2011. Chemoenzymatic labeling of protein C-termini for positive selection of C-terminal peptides. ACS Chemical Biology 6, 1015–1020.

Xu T, Wong CC, Kashina A, Yates JR 3rd. 2009. Identification of N-terminally arginylated proteins and peptides by mass spectrometry. Nature Protocols 4, 325–332.

Yeom J, Ju S, Choi Y, Paek E, Lee C. 2017. Comprehensive analysis of human protein N-termini enables assessment of various protein forms. Scientific Reports 7, 6599.

Yoo YD, Mun SR, Ji CH, et al. 2018. N-terminal arginylation generates a bimodal degron that modulates autophagic proteolysis. Proceedings of the National Academy of Sciences, USA 115, E2716–E2724.

Yoshihara HA, Mahrus S, Wells JA. 2008. Tags for labeling protein N-termini with subtiligase for proteomics. Bioorganic & Medicinal Chemistry Letters 18, 6000–6003.

Zhang H, Deery MJ, Gannon L, Powers SJ, Lilley KS, Theodoulou FL. 2015. Quantitative proteomics analysis of the Arg/N-end rule pathway of targeted degradation in Arabidopsis roots. Proteomics 15, 2447–2457.

Zhang H, Gannon L, Hassall KL, Deery MJ, Gibbs DJ, Holdsworth MJ, van der Hoorn RAL, Lilley KS, Theodoulou FL. 2018. N-terminomics reveals control of Arabidopsis seed storage proteins and proteases by the Arg/N-end rule pathway. New Phytologist 218, 1106–1126.

Zhang Y, He Q, Ye J, Li Y, Huang L, Li Q, Huang J, Lu J, Zhang X. 2015. Systematic optimization of C-terminal amine-based isotope labeling of substrates approach for deep screening of C-terminome. Analytical Chemistry 87, 10354–10361.

Zhang Y, Li Q, Huang J, Wu Z, Huang L, Li Y, Ye J, Zhang X. 2018. An approach to incorporate multi-enzyme digestion into C-TAILS for C-terminomics studies. Proteomics 18, 10.1002/pmic.201700334.

Zybailev B, Rutschow H, Friso G, Rudella A, Emanuelsson O, Sun Q, van Wijk KJ. 2008. Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. PLoS One 3, e1994.