Trends in Plant Science

CellPress
REVIEWS

Opinion

N-Degron Pathways in Plastids

Imen Bouchnak[1] and Klaas J. van Wijk[1,2,*,@]

Protein amino (N) termini are major determinants of protein stability in the cytosol of eukaryotes and prokaryotes, conceptualized in the N-end rule pathway, lately referred to as N-degron pathways. Here we argue for the existence of N-degron pathways in plastids of apicomplexa, algae, and plants. The prokaryotic N-degron pathway depends on a caseinolytic protease (CLP) S recognin (adaptor) for the recognition and delivery of N-degron-bearing substrates to CLP chaperone-protease systems. Diversified CLP systems are found in chloroplasts and nonphotosynthetic plastids, including CLPS homologs that specifically interact with a subset of N-terminal residues and stromal proteins. Chloroplast N-terminome data show enrichment of classic stabilizing residues [Ala (A), Ser (S), Val (V), Thr (T)] and avoidance of charged and large hydrophobic residues. We outline experimental test strategies for plastid N-degron pathways.

Highlights

Arabidopsis thaliana plastids contain a CLP protease system, including CLPS1, a homolog of the bacterial **N-recognin** CLPS. CLPS1 interacts with several stromal proteins. This interaction requires the presence of two conserved degron-binding residues. CLPS1 shows specific *in vitro* affinity for several type 2 N termini, especially F and W.

A. thaliana CLPS1 interacts with CLPF, a protein unique to higher plant plastids. CLPF is likely to be a coadaptor of CLPS1, together involved in substrate delivery to the CLPC chaperones.

The *A. thaliana* stromal N terminome is enriched for the small, uncharged, classical stabilizing residues A, V and T [often with N-terminal (Nt) acetylation], and S, whereas other residues at the N terminus of identified proteins are absent or highly underrepresented. Stromal Nt maturation is likely to be a multistep process involving SPP and uncharacterized peptidases and Nt acetylases.

Current experimental evidence suggests the existence of a unique plastid N-degron pathway.

Plastid Proteostasis during Differentiation, Development, and Adaptation

Plastids in photosynthetic organisms undergo various developmental transitions and adaptations [1]. Each plastid type contains its own specific proteome through the coordinated actions of the **proteostasis** (see Glossary) network, involving the transcription, translation, protein folding, and degradation machineries. The remodeling and stability of these proteomes during plastid differentiation and adaptation must occur through selective protein synthesis and proteolysis. Understanding the chloroplast proteolytic hierarchies and **degrons** should be pursued with high priority to enable understanding and influencing of plastid differentiation, adaptation, and function.

The Bacterial Origin of Plastid and Mitochondrial Proteolytic Machineries

Plastids contain an array of proteolytic systems in each of the plastid compartments (envelope, thylakoid, lumen, stroma, plastoglobules) [2–6]. Several of the plastid proteases are also targeted to plant mitochondria [2,7]. **Primary plastids** arose monophyletically from the endosymbiosis of a cyanobacterium of unknown phylum [8–11], whereas mitochondria originate from an α-proteobacterium in an older endosymbiotic event [12]. Most of the original bacterial genes transferred to the nucleus during evolution, but the organelles retained a residual genome and bacterial features of the proteostasis network. Collectively, the plant plastid and mitochondrial proteolytic machineries include ~100 different proteins mostly of bacterial origin, but they have undergone diversification through evolution [3]. An example is the ATP-dependent caseinolytic protease (CLP) system [2,4]. The plastidial CLP system includes CLPC/D AAA+ chaperones, CLPPR proteases, CLPS **adapter** homologs, and, in higher plants only, CLPT1 and 2 and CLPF (Figure 1A) (Table 1). Importantly, and central to this Opinion article, bacterial CLPS homologs and the CLP chaperone-protease machinery are involved in the **N-degron pathway** [13–15].

Protein Amino Termini as Major Determinants of Protein Stability in Bacteria and Eukaryotes

Protein amino termini are prone to modifications and are major determinants of protein stability in the cytosol of prokaryotes and eukaryotes, and perhaps also in chloroplasts and nonphotosynthetic plastids. This is conceptualized in the **N-end rule**, which states that certain amino acids, when exposed at the N terminus of a protein, act as triggers (N-degrons)

[1]Section of Plant Biology, School of Integrative Plant Sciences (SIPS), Cornell University, Ithaca, NY 14850, USA
[2]http://blogs.cornell.edu/vanwijk/

*Correspondence:
kv35@cornell.edu (K.J. van Wijk).
@Twitter: @vanwijklab (K.J. van Wijk).

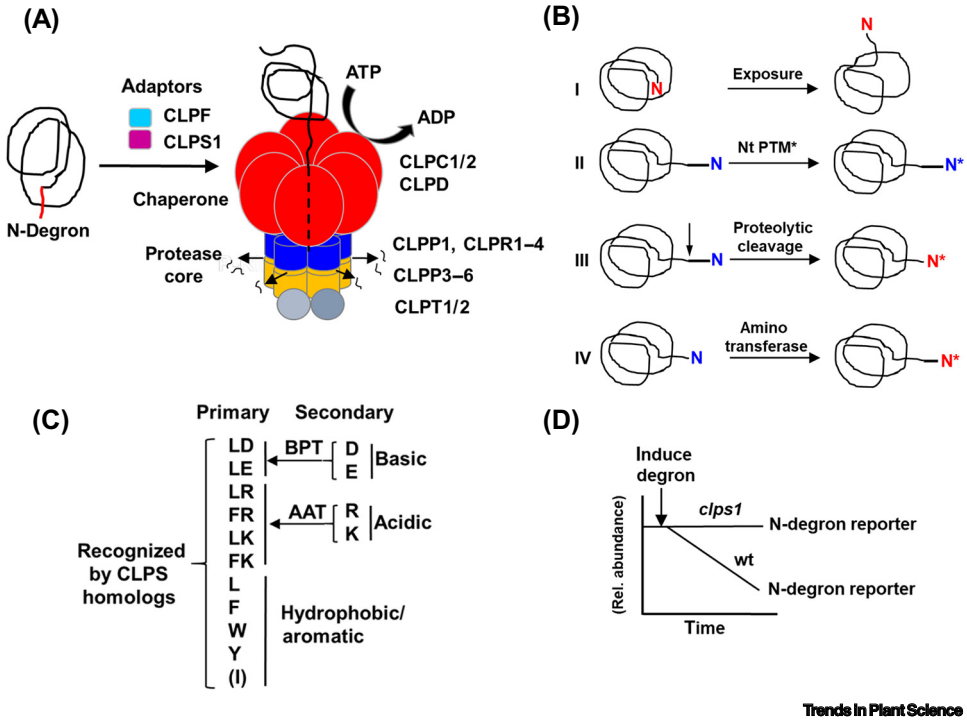

Figure 1. The Chloroplast Caseinolytic Protease (CLP) System and N-Degrons. (A) The multicomponent CLP system in chloroplasts. Substrates with N-degrons are targeted by CLPS1 to the chaperone-protease system. CLPF is postulated to be a coadaptor of CLPS1. (B) Schematic view of conditional N-degrons. N termini can be converted into N-degrons: (I) by a conformational change that results in exposure of the existing N-degron – this can involve dynamic interactions with other proteins or cofactors; (II) by a post-translational modification (PTM) of the existing N-terminal (Nt) residue, in particular deformylation of formyl-methionine (fMet), Nt-acetylation or Nt-deacetylation, or deamidation of Gln (Q) or Asn (N); (III) by Nt cleavage, with the new N terminus acting as an N-degron (* marks the new N terminus); or (IV) through the post-translational addition of a primary N-degron residue to the existing N terminus (^ marks the new N terminus). (C) Classic bacterial primary and secondary N-degrons recognized by the bacterial CLPS homologs and delivered to CLP chaperone-protease complexes for degradation. (D) Schematic view of a hypothetical experiment demonstrating CLPS1-dependent degradation of an N-degron reporter.

for degradation [16–20]. Because many residues can act as N-degrons and because of the wide diversity in N-degrons across organisms, we refer now to N-degron pathways rather than the N-end rule [21]. In the cytosol of eukaryotic cells, the N-degron pathway involves ubiquitination through E3 ligases and subsequent degradation of polyubiquitinated proteins via the proteasome [22–24]. Gram-negative bacteria contain a ubiquitin-independent version of the N-degron pathway, designated the Leu/N-end pathway [25–28]. This Leu/N-end pathway was initially discovered in the Gram-negative γ-Proteobacteria *Escherichia coli* [27] and *Vibrio vulnificus* [25]. This pathway involves the CLPS adaptor and the CLP chaperone-protease machinery (Figure 1A). Most degrons, including N-degrons, are conditional; that is, they are generated or exposed only following specific cellular events, including protein cleavage, conformational changes, or post-translational modification to the Nt residue (Figure 1B).

CLPS-Dependent N-Degron Recognition and Delivery in Bacteria

The sequence of events and molecular details for N-degron recognition and degradation by the CLP system has been resolved particularly in *E. coli* [29–31]. The first step involves N-degron recognition of hydrophobic residues [Leu (L), Phe (F), Tyr (Y), Trp (W), and, in some cases, Ile (I)] through a hydrophobic pocket in the Nt region of CLPS, followed by docking of the

Glossary

Adaptor: a protein that acts as a scaffold to selectively bind cargo and tether substrates to their cognate proteases for degradation or primarily activates the target protease. Adaptors for bacterial CLP systems include CLPS, SppB, MecA, McsB, and NblA.

Aminoacyl protein transferases: enzymes that transfer a specific amino acid from a charged tRNA to specific Nt residues of proteins. This activity occurs post-translationally without the involvement of ribosomes. Several of these enzymes are important in the N-degron pathways in the cytosol of both eukaryotes and prokaryotes, because they generate N-degrons. Examples are the L/F-tRNA-protein transferase AAT that transfers L or F to the basic residue R or K, the L-tRNA-protein transferase BPT that transfers L to the acidic residue D or E, and R-tRNA-protein transferases (ATEs) that transfer R to D or E residues.

Deamidation: loss of the Nt α-amino group of proteins resulting in the conversion of Gln (Q) and Asn (N) to Glu (E) and Asp (D), respectively. These reactions can occur nonenzymatically or enzymatically by amidohydrolases. The amidohydrolases NTAQ1 and NTAN1 in *Arabidopsis thaliana* have been demonstrated to be important for the N-degron pathways in the cytosol of plants.

Degron: short amino acid sequence that when exposed to the protein surface will act as a recognition site for protein degradation. These degrons can be at the N terminus (N-degron) or C terminus (C-degron) or internal to the protein.

N-Degron pathways: protein degradation pathways that are triggered by N-degrons. This term is an alternative for the N-end rule and better reflects the complexity of protein stability regulation through protein N termini.

N-End rule: states that the Nt region of proteins, when exposed at the N terminus of a protein, acts as a trigger (degrons) for degradation. This involves a destabilizing Nt residue and information immediately downstream of the N terminus. However, N-end rules differ across organisms and subcellular compartments and can also involve Nt post-translational modifications (e.g., phosphorylation, formylation, acetylation).

CLPS–substrate complex on the N-domain of CLPA AAA+ chaperones. Next, the CLPS–substrate complex is then 'pulled' into the CLPA pore in an ATP-dependent fashion, and the resulting distortion of the CLPS structure allows release of the substrate inside the CLPA pore. CLPS is subsequently released from CLPA and unfolding and degradation of the substrate by the CLPAP complex is completed. Notably, CLPS is necessary and sufficient for the recognition and delivery of N-degron substrates in bacterial CLP systems. Interestingly, CLPS homologs in bacteria are mostly confined to Proteobacteria (e.g., *E. coli*) and typically absent in Gram-positive bacteria (e.g., *Bacillus subtilis*) (Table 1).

Primary, Secondary, and Tertiary Degrons in Bacteria and Eukaryotes

The CLPS-driven degradation pathways in bacteria involve two types of N-degrons; namely, primary and secondary degrons (Figure 1C) (Table 1). Primary degrons are typically the aromatic residues (Y, W, F) and leucine (L) and in some cases also isoleucine (I). In addition, *E. coli* and *V. vulnificus* have also two secondary destabilizing residues Arg (R) and Lys (K), which are not directly recognized by CLPS but are modified by an L/F-tRNA protein transferase (AAT) (Figure 1C) (Table 1). AAT installs L or F residues on the Nt R or K, thus generating primary N-degrons [32,33]. *V. vulnificus* has two more secondary destabilizing residues; namely, aspartate (D) and glutamate (E), which are modified by an L-tRNA protein transferase (BPT) that conjugates L to Nt D and Nt E (Figure 1C) (Table 1). *V. vulnificus* contains both AAT and BPT. So far no such transferases have been identified in any plastid types. The N-degron pathways in the cytosol of eukaryotes also involve primary, secondary, and even tertiary N-degrons. Primary N-degrons are generally similar to those in bacteria, but also include the basic residues [R, K, and His (H)]. Furthermore, the secondary destabilizing residues D and E serve as acceptors for arginyl-tRNA-protein transferases (ATEs) that attach an R residue to these Nt acidic residues, thereby generating primary degrons (Table 1). Plants such as *Arabidopsis thaliana* have one or more ATE homologs, but they appear to localize only to the cytosol [17,18]. The apparent confinement of BPT and AAT to prokaryotes and ATE homologs to eukaryotes suggests that the primary R-degron was a late evolutionary invention [32]. Furthermore, enzymatic Nt protein **deamidation** in the cytosol of eukaryotes can convert Nt glutamine (Q) and asparagine (N) to E and D, respectively. Therefore, Q and N are tertiary N-degrons. *A. thaliana* contains both specific cytosolic Q- and N-deaminases (NTAQ1 and NTAN1), but no plastid homologs are known (Table 1) [34]. It was suggested that these tertiary N-degrons appeared after the divergence of the fungal and metazoan lineages [32]. The phylogenetic analysis in [32] indicates that prokaryotes could have many different types of **aminoacyl-transferases**. Perhaps some of these evolved to be part of plant genomes, increasing the potential for secondary destabilizing residues in chloroplasts. Finally, in plants cysteine can also act as a tertiary N-degron when enzymatically dioxygenated to sulfinic acid by cysteine dioxygenases (PCOs) [35–39] and then recognized by ATE1 and ATE2 for arginylation [38]. However, only ATE1 has been shown to be active and not ATE2 [38]. *A. thaliana* encodes five PCOs that are all likely to locate to the cytosol (Table 1). So far there is no evidence for secondary or tertiary N-degrons in plastids, but it would be premature to conclude that N-degron pathways in plastids concern primary N-degrons only. Systematic analysis of the subcellular localization of these plant Nt modifiers (ATE1, ATE2, NTAQ1, NTAN1, PCO1–5) and other homologs in the superfamily of acyl-CoA *N*-acyltransferases (IPR016181) could potentially discover new N-degron-generating pathways in plastids, including amino-transferase addition of residues.

Evolution and Diversification of the CLP Proteolytic System and Its CLPS Adaptors

Similar to nonphotosynthetic bacteria, cyanobacteria as well as plastids in algae and plants contain CLP protease systems [40] (Figure 1A). Previous phylogenetic analyses showed the

N-Recognin: protein that recognizes N-degrons or Nt primary destabilizing residues. N-Recognins are known in eukaryotic cytosolic proteasomal degradation pathways (E3 ligases) and CLP degradation systems (CLPS homologs).

Nt-α-Acetylation or N-α-acetylation (NtAc): an enzymatic post-translational modification of Nt residues involving the addition of an acetyl (-$COCH_3$) onto the Nt amine. NtAc results in the loss of the positive charge of the Nt amine and a 42.010 6-Da mass increase. This post-translational modification is catalyzed by NATs that transfer the acetyl from acetyl-CoA.

Nt-Acetyl transferases (NATs): enzymes that catalyze the transfer of an acetyl group from acetyl-coA to the Nt α-amine of proteins. There is a closely related class of enzymes that catalyze the transfer of an acetyl group from acetyl-CoA to the ε-amino of lysine residues; these are named KATs.

Primary plastids: plastids can be distinguished based on their evolutionary origin and include primary, secondary, and tertiary plastids. Primary plastids in higher plants and green algae arose monophyletically from the endosymbiosis of a nitrogen-fixing cyanobacterium of unknown phylum in a heterotrophic eukaryote.

Proteostasis: the process that regulates the protein composition and its functional state within cells. Proteostasis involves the regulated integration of protein synthesis, sorting, folding, maturation, and degradation.

Table 1. Components and Features of N-Degron Pathways and CLP Systems in Nonphotosynthetic Bacteria, Cyanobacteria, *Arabidopsis thaliana* Cytosol and Chloroplasts, and Apicoplasts of *Plasmodium falciparum*

N-Degron component or feature	Nonphotosynthetic bacteria[a,b]	Cyanobacteria[c]	*Arabidopsis thaliana*		*Plasmodium falciparum*
	Cytosol	Cytosol	Cytosol	Chloroplast	Apicoplast
CLPS homologs[d]	S1; S2 (exception)	S1, S2	No	S1	S1
CLPS N-degron affinity[e]	L, F, Y, W	L, F, Y, I	N/A	F, W, (L)	F, W, L, I
CLP AAA+ chaperones[f]	A, X	C, X	No	C1, 2; D	C
CLPP/R protease subunits[g]	P	P, R	No	P1, P3–6; R1–4	P, R
CLPT subunits	No	No	No	T1, 2	No
CLPF	No	No	No	CLPF	No
Other CLPA/C adaptors[h]	MecA, McsbB, YpbH	NblA/B	N/A	TBD	TBD
Tags for Clp substrates	ssrA	TBD	TBD	TBD	TBD
Deformylase (PDF)	Yes	Yes	No	Yes	Yes
MetAP	Yes	Yes	Yes	Yes	TBD
L/F to D/E transferase	AAT	TBD	No	TBD	TBD
L to K/R transferase	BTP	TBD	No	TBD	TBD
R to D/E transferase	TBD	TBD	ATE1, 2	TBD	TBD
NAT	Yes	Yes	Multiple	NAA80, likely others	TBD
N/Q-deaminases	TBD	TBD	NTAQ1, NTAN1	TBD	TBD
PCOs	No	No	PCO1–5	TBD	No

[a]Examples of nonphotosynthetic (Gram-negative) Proteobacteria in which the CLP system and/or N-degron pathways have been studied: *Escherichia coli* (γ), *Vibrio vulnificus* (γ), *Salmonella enterica* (γ), *Caulobacter crescentus* (α), *Agrobacterium tumefaciens* (α).
[b]The CLP system and/or N-degron pathways have also been studied in several Gram-positive species: *Bacillus subtilis* (Firmicutes), *Streptococcus* mutants (Firmicutes), *Mycobacterium tuberculosis* (Actinobacteria), *Staphylococcus aureus* (Firmicutes).
[c]Examples of Cyanobacteria in which the CLP system and/or N-degron pathways have been studied: *Synechocystis elongatus*, *Synechococcus* PC 6803.
[d]For a deep CLPS phylogeny, see [43].
[e]Differences are observed between CLPS1 and CLPS2 within the same species for *S. elongatus* and *A. tumefaciens* [15].
[f]For a deep CLPC phylogeny, see [41].
[g]For a phylogeny of plant CLPP and CLPR proteins, see [42].
[h]Except for CLPS homologs, other adaptors appear unique to specific species.

evolutionary lineages for various CLP components including CLP AAA+ chaperones [41], CLPP/R protease subunits [42], and CLPS adaptors [43,44] as well as the plant-specific CLPT auxiliary proteins [45] and the plant-specific coadaptor CLPF [46]. CLPS homologs are found throughout all branches of the (Gram-negative) Proteobacteria (α, β, δ, ε, γ), typically in an operon with the CLPA chaperone [43]. Many α-Proteobacteria also contain a second, more distant CLPS homolog that is not in the genomic vicinity of the CLPA homolog [43]. CLPS homologs are also infrequently found in other bacterial clades, but are absent in the Archaea.

CLPS Proteins in Cyanobacteria, Plastids and Apicoplasts, and N-Degrons

CLPS homologs are predicted, and in some cases studied, in Cyanobacteria [15,47], chloroplasts [44,48,49], and apicoplasts derived from secondary endosymbiosis [15,50]. These species have CLPC chaperones instead of the CLPA chaperones found in Proteobacteria; for differences between types of CLP chaperones, see [41]. The cyanobacterium *Synechococcus elongates* contains CLPS1 and CLPS2 that dynamically associate with CLPC [47]. Using peptide affinity arrays and recombinant CLPS1 and CLPS2, differential affinity for Nt residues was observed [15]. CLPS1 showed affinity for F, Y and W, whereas CLPS2 showed affinity for L, F, and Y as well as V and I. In all cases, R or K at the P2 position increased binding whereas D or E at P2 decreased binding [15]. An AAT L/F-transferase of *Synechocystis* sp. complemented

an *E. coli aat* mutant for the degradation of N-end rule substrates [51]. Through phylogenetic analysis, we showed that green algae, moss, and angiosperms contain one or more copies of CLPS1 homologs but no CLPS2 homologs, and some species also contain CLPS1-like protein(s) [44]. Furthermore, this study showed that *A. thaliana* CLPS1 interacts with plastid CLPC chaperones. Recently, we conducted *in vitro A. thaliana* CLPS1 affinity assays using eight N-degron protein reporters containing either F, Y, L, W, I, or R in the Nt position [48]. This demonstrated that CLPS1 has a restricted N-degron specificity, recognizing proteins bearing an Nt F or W, only weakly recognizing L, but not recognizing Y or I. This affinity was dependent on two conserved residues in the CLPS1 binding pocket (see [44]) and is sensitive to FR dipeptide competition. Our findings differed from [49], which suggested *A. thaliana* CLPS1 affinity for F, Y, L, and W, with the lowest affinity for W (see comment in [48]).

Apicomplexa is a phylum of obligate parasitic protozoa including the causative agents of malaria (*Plasmodium* spp.) and other diseases. These organisms typically harbor a complex plastid, the apicoplast [52]. *Plasmodium falciparum* contains single CLPC, CLPP, CLPR, and CLPS homologs [53,54]. The structure of the CLPS homolog has been determined [50] and *in vitro* CLPS affinity assays suggested that N-degrons included classic type 2 residues (F, W, Y, L) and I [50,55]. Surprisingly, the physiological substrates, or an *in vivo* demonstration of N-degron pathways in Cyanobacteria, chloroplasts, or apicoplasts, have not yet been shown. It is good to remember that even for the cytosolic N-degron pathways in *A. thaliana* and other plants [56,57], and well-studied CLPS-dependent model systems such as *E. coli*, relatively few N-degron substrates have been identified [58]. However, using ClpS affinity chromatography to isolate interacting proteins from *E. coli* cell lysates identified a large pool of proteins bearing an N-degron [59]. Most of the proteins were N-terminally truncated by peptidases and many must have been further modified by Aat [59].

Plastid CLPS1 Interactors and Substrates

Affinity experiments using recombinant CLPS1 and stromal proteomes of *clpc1 clps1* double-null mutants in *A. thaliana* identified a number of interacting stromal proteins in dependence of conserved residues in the putative N-degron-binding pocket of CLPS1 [44]. However, these candidate substrates did not reveal any specific N-degrons, in part perhaps due to technical reasons [44]. This screen also identified CLPF, but its interactions with CLPS1 did not depend on the N-degron-binding domain in CLPS1 [46]. Based on subsequent analysis, we proposed that CLPF and CLPS1 form together a binary adaptor system. CLPF is unique to higher plants and contains bacterial uvrB/C and YccV protein domains [46]. Our *in vivo* trapping study with *A. thaliana* CLPC1 identified additional candidate CLP substrates [60] and such trapping approaches could help to identify substrates with N-degrons [61].

The N Terminome of Chloroplasts

The N termini of nucleus- and plastid-encoded proteins undergo maturation steps involving several known enzymes, including stromal processing peptidase (SPP) to cleave off the chloroplast transit peptide of imported nuclear-encoded proteins [62], methionine deformylases (PDFs) for plastid-encoded proteins, and one or more **Nt-α-acetyl transferases (NATs)** for nuclear- and plastid-encoded proteins [63,64] (Figure 2). Surprisingly, some 40% of imported chloroplast proteins are represented by multiple Nt proteoforms [65], suggesting that the maturation of chloroplast protein N termini is complex and could affect protein stability in plastids. The consensus SPP cleavage site is poorly defined and the rules, mechanisms, and enzymes for possible subsequent processing, stabilization, and other post-translational modifications are not well characterized [2]. Consequently, it is difficult to accurately predict the exact N termini of mature nucleus-encoded chloroplast proteins, making understanding N-degron pathways more complex.

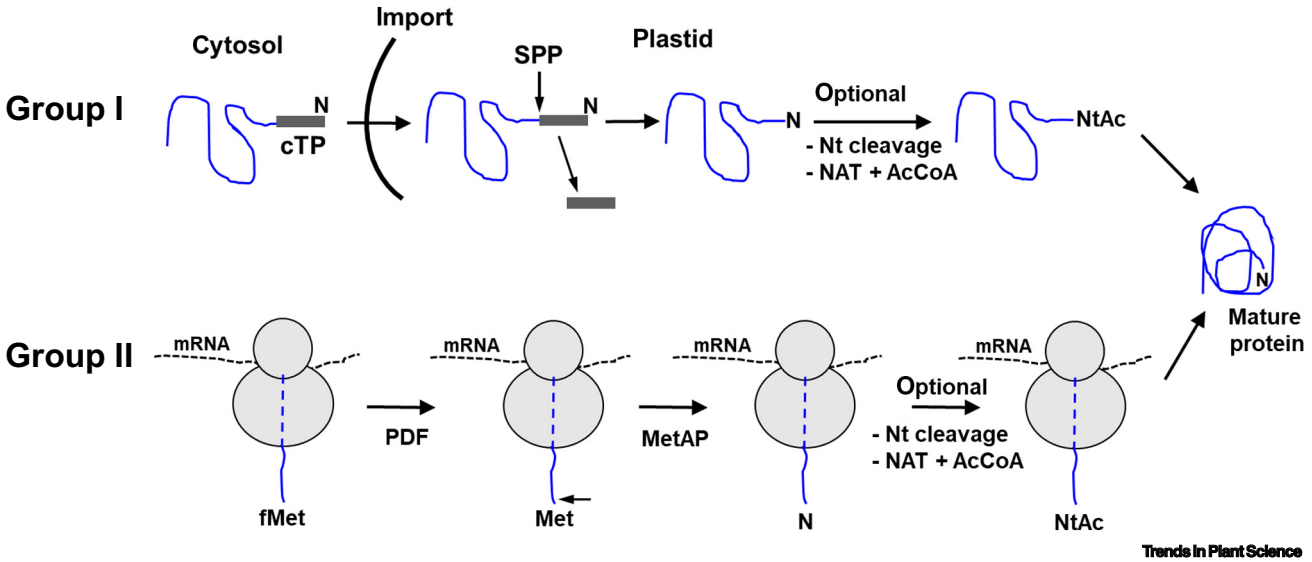

Figure 2. Maturation Pathway of Nuclear-Encoded Chloroplast Proteins (Group I) and Chloroplast-Encoded Proteins (Group II). The chloroplast (and plastid in general) proteome comprises nucleus- and chloroplast-encoded proteins, each undergoing their own maturation pathways [2]. In the case of group I, most N termini are initially generated by the activity of stromal processing peptidase (SPP), followed by possible additional N-terminal (Nt) cleavage events by other chloroplast proteases and possibly by Nt acetylation through Nt-acetyl transferase (NAT) activity. In the case of group II, proteins are synthesized in the chloroplast starting with an Nt formyl-methionine (fMet). In the case of *Escherichia coli*, fMet can act as a cotranslational degradation signal [92]. After cotranslational removal of the formyl group by methionine deformylase (PDF), the Nt Met residue is removed by methionine amino peptidase (MetAP) unless it is hindered by large, bulky residues [Trp (W), Phe (F), Tyr (Y), Leu (L), Ile (I)] in the penultimate (P2′) position [63]. Consequently, MetAP activity generally does not generate primary type II N-degrons (Figure 3B). Similar to group I, proteins in group II can undergo Nt acetylation, likely to be in both a cotranslational and a post-translational fashion, by one or more NATs [77,93].

Systematic Nt protein analysis of *A. thaliana* chloroplast proteins by terminal amine isotopic labeling of substrate (TAILS) and mass spectrometry determined the frequency of amino acids in the Nt position [65] (Figure 3A,B). This showed that A, V, T (often in N-α-acetylated form), and S were the most observed Nt residues, whereas large hydrophobic residues (e.g., W, Y, F, L) were underrepresented or even absent in stromal proteins (Figure 3A,B). This is consistent with a classical prokaryotic N-degron pathway. However, so far there are no physiological studies that support the significance of these residues as N-degrons, even if Nt residues did impact steady-state accumulation of plastid-encoded reporter proteins [66]. Experiments are in progress in our laboratory to systematically test the influence of N termini on protein stability *in vivo*, using approaches as outlined below. Several other plant studies also used the TAILS approach or other N-terminomics techniques (e.g., COFRADIC) to investigate protein degradation pathways involving the proteasome or metacaspases, and they provide a rich resource to explore experimental N-terminome data [67–69].

Nt-Acetylation and Impact on N-Degrons

Many proteins undergo **Nt-α-acetylation (NtAc)** by NATs, and NtAc is gaining recognition as a major cellular regulator [70,71]. NtAc can affect localization, function, and, in some cases, protein stability, including in the N-degron pathways [20,70]. N-terminomics studies have shown that a significant portion of chloroplast proteins undergo NtAc; in particular, a high percentage for Nt residues of the nuclear-encoded proteins V, T, A, and S (highest for V and lowest for S) [65] (Figure 3A). A few Nt residues (W and Y) are found only in NtAc form. So far there is no general evidence for the influence of NtAc on protein stability in plastids, perhaps with the exception of the ε-ATP subunit of the thylakoid ATP-synthase in watermelon [72] (see also [73]). A positive correlation between NtAc proteins and their half-life time was suggested for stromal proteins in

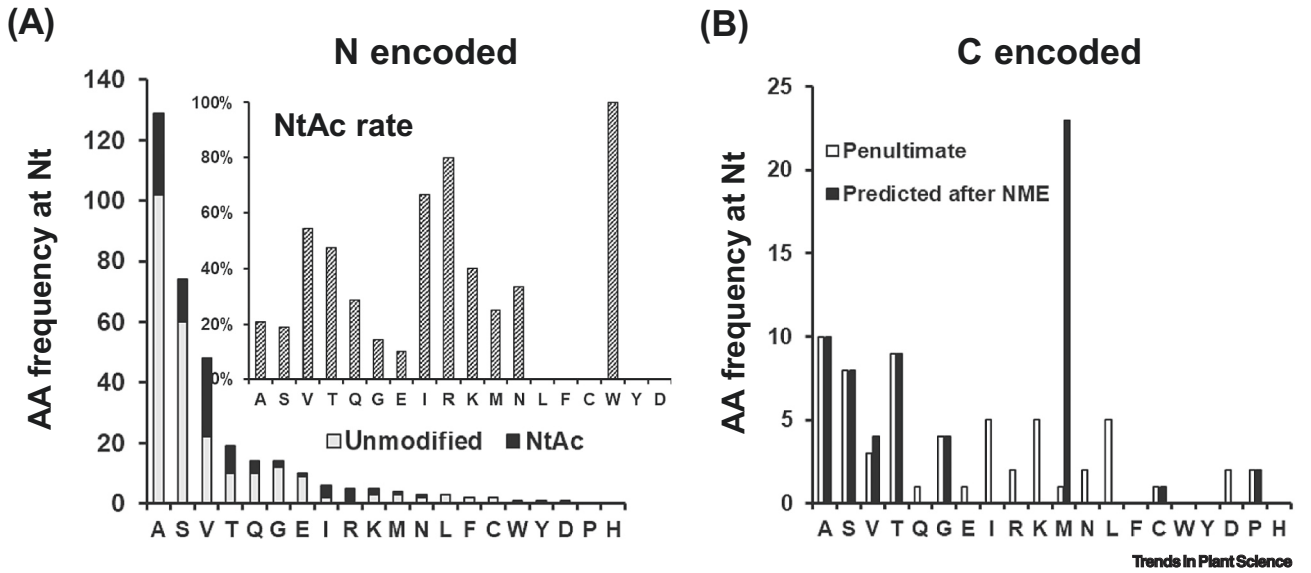

Figure 3. N-Terminal (Nt) Amino Acid Distribution of Nucleus-Encoded Stromal Proteins and Chloroplast-Encoded Stroma-Exposed Proteins. (A) Nt frequency and Nt acetylation (NtAc) rates experimentally observed in the stromal proteome using a terminal amine isotopic labeling of substrate (TAILS) approach. Redrawn from data in [59]. Residues P and H were not observed or observed only once (W, Y, D) among the 341 N termini of nucleus-encoded stroma-exposed proteins. (B) Penultimate residues [following the initiating M (iM)] of chloroplast-encoded proteins with their N termini exposed to the stromal phase, as well as the predicted Nt residues after the predicted excision of the iM (NME) by methionine amino peptidase (MetAP). We note that the predicted NME follows closely the experimentally observed N. Redrawn from data in [65]. Several residues [Phe (F), Trp (W), Tyr (Y), His (H)] are never in the penultimate position of chloroplast-encoded proteins (with N termini facing the stroma).

Chlamydomonas reinhardtii [74]. The wide range of observed NtAc residues suggests that more than one NAT operates in the chloroplast [75,76]. So far there is one confirmed chloroplast-localized NAT (Naa70; AT2G39000) with confirmed NAT activity [77]. A second chloroplast NAT (NTSI; AT1G32070) acetylates the ε-amino group of lysines rather than N termini and impacts thylakoid state transitions [78]. Other NATs are predicted to localize to chloroplasts [77], but experimental localization and substrate specificity remain to be determined.

Experimental Strategies to Test for Plastid N-Degron Pathways in Plants and Algae

To determine whether an N-degron pathway exists in chloroplasts it should be demonstrated that *in vivo* protein degradation rates (or half-life) depend on well-defined N-degrons, one or more specific adaptors/recognins, and proteolytic systems. These principles should be shown for one or more endogenous chloroplast proteins. Molecular mechanisms, including enzymatic activities, should also be determined, either *in vivo* or through *in vitro* reconstitution.

The most versatile and comprehensive approach would be to develop an *in vivo* quantitative reporter system that allows direct testing of the effect of specific Nt residues on reporter protein stability in dependence of the CLP protease system, including CLPS1 and CLPF (Figure 1D). Such a system must have three key features.

(i) Allows the generation of any Nt residue of a suitable protein reporter in the chloroplast. Following strategies employed for cytosolic proteins in yeast, plants, and animals, reporters are allowed to accumulate with a dormant N-degron. Upon the expression of a site-specific protease, the dormant N-degron is deprotected thus triggering the degradation of N-degron-containing targets. This approach was pioneered by Varshavsky using an Nt Ub-X-reporter

fusion and cleavage by deubiquitylating enzymes (DUBs) resulting in the release of a X-reporter protein bearing a specific residue, X, at its N terminus [16,79,80]. Importantly, dormant N-degrons can be activated through a range of different proteases [81], optogenetics [82,83], temperature [84,85], and auxin [86]. For a recent overview, see [87].

(ii) Allows quantification of the relative protein half-life in chloroplasts in direct dependence of the Nt residue or Nt domain. Noninvasive fluorescence reporter systems have been developed, including a tandem fluorescence protein timer (tFT) in yeast and other non-plant systems [88]. This tFT system was recently also adapted to plants [89].

(iii) Facilitates direct testing of chloroplast-localized molecular players in the recognition and possible generation/modification of the N-degrons, as well as subsequent degradation. These players would be CLPS1, CLPF, and CLP chaperones and proteases, as well as possible protein aminotransferases, NATs, and amino peptidases.

The challenge for the demonstration of a chloroplast N-degron pathway is that the N-degrons must be generated in the chloroplast stroma. If the reporter and the N-degron-generating system are expressed from the nuclear genome, they must all be properly targeted into the chloroplast. An alternative is to express some or all components from the plastid genome. The main bottleneck here is that plastid transformation is very inefficient and until recently practical only in tobacco and *C. reinhardtii*, but a recent breakthrough in *A. thaliana* plastid transformation might make this more practical in the future [90,91]. Plastid transformation in tobacco was used to generate plastid-encoded N-degron GFP reporters to test the impact on protein stability of all 20 amino acids in the position after the initiator methionine, the penultimate (P2) position [66]. This relied on the removal of the Nt methionine by an endogenous peptidase [methionine amino peptidase (MetAP)] to expose these 20 amino acids. However, because MetAP is sensitive to the residue in the P2 position, large hydrophobic/aromatic residues prevented M removal and their impact on stability could not be tested [66]. Nevertheless, this systematic approach revealed strong differences in protein stability, and it indicated an important role for the penultimate Nt residue, as well as additional sequence determinants in the Nt region, in determining the protein half-life in the chloroplast.

Concluding Remarks and Future Perspectives

There is substantial experimental and theoretical support for a CLP-driven N-degron pathway in the stroma of plant chloroplasts, as well as in Cyanobacteria and apicoplasts. The N-degron pathway is likely to extend to nonphotosynthetic plant plastids as well as plastids in algae. However, specific plastid N-degrons and the Nt-modifying enzymes involved in Nt maturation and the generation of N-degrons remain to be determined. Experimental systems are needed to test and demonstrate the influence of N-degron, CLPS1, and possibly other adaptors, as well as Nt modifiers. A full mechanistic understanding of plastid N-degron pathways will be predictive of plastid protein stability and will improve rational protein design for plastid biotechnology (see Outstanding Questions).

Acknowledgments

This research was supported by a grant from the National Science Foundation (NSF), Division of Molecular and Cellular Biosciences (MCB) #1614629 and EAGER #1834636 to K.J.v.W.

References

1. Jarvis, P. and Lopez-Juez, E. (2013) Biogenesis and homeostasis of chloroplasts and other plastids. *Nat. Rev. Mol. Cell Biol.* 14, 787–802
2. van Wijk, K.J. (2015) Protein maturation and proteolysis in plant plastids, mitochondria, and peroxisomes. *Annu. Rev. Plant Biol.* 66, 75–111
3. Majsec, K. *et al.* (2017) The plastid and mitochondrial peptidase network in *Arabidopsis thaliana*: a foundation for testing genetic interactions and functions in organellar proteostasis. *Plant Cell* 29, 2687–2710
4. Nishimura, K. *et al.* (2017) Essentials of proteolytic machineries in chloroplasts. *Mol. Plant* 10, 4–19

Outstanding Questions

Is there an N-degron pathway in chloroplasts? If yes, what are the specific N-degrons and chloroplast machineries involved? As outlined in this Opinion article, there is considerable, but mostly indirect, support for an N-degron pathway in chloroplasts. However, direct *in vivo* and/or *in vitro* demonstrations will be needed to answer this central question. Assuming that such pathways exist, more specific questions listed below need to be answered.

Is CLPS1 involved and required in a chloroplast N-degron pathway? N-degron pathways in the cytosol of bacteria appear to strictly require CLPS. Are chloroplasts the same in this respect?

Is chloroplast CLPF involved in a possible chloroplast N-degron pathway and what is the functional relationship in such a pathway with CLPS1? CLPF is a protein unique to plants and interacts with CLPS1 independent of conserved N-degron residues. Published data so far suggest that CLPS1 and CLPF mutually stimulate interactions with the CLPC chaperones.

Are there chloroplast Nt aminotransferases that can generate N-degrons? So far there are no identified homologs of bacterial AAT, BTP, or ATE in chloroplasts or nongreen plastids. However, we are not aware of exhaustive efforts to establish such activities in chloroplasts.

What is the role of Nt acetylation (NtAc) in chloroplast/plastid protein stability and N-degron pathways? Many stromal proteins are in NtAc form, in particular the frequent Nt residues V, T, A, and S, whereas Nt W and Y are found exclusively in NtAc form.

Are there different NtAc pathways for chloroplast-encoded and chloroplast imported nuclear-encoded proteins? Perhaps there is a post-translational pathway for imported nuclear-encoded proteins and a cotranslational pathway for chloroplast-encoded proteins.

Is there a role for peptidases in the maturation of imported chloroplast proteins, beyond the general cTP-cleaving SPP? Did these peptidases

5. Kato, Y. and Sakamoto, W. (2018) FtsH protease in the thylakoid membrane: physiological functions and the regulation of protease activity. *Front. Plant Sci.* 9, 855
6. Nishimura, K. *et al.* (2016) Chloroplast proteases: updates on proteolysis within and across suborganellar compartments. *Plant Physiol.* 171, 2280–2293
7. Kmiec, B. *et al.* (2018) A common peptidolytic mechanism for targeting peptide degradation in mitochondria and chloroplasts. *Mol. Plant* 11, 342–345
8. Ponce-Toledo, R.I. *et al.* (2017) An early-branching freshwater cyanobacterium at the origin of plastids. *Curr. Biol.* 27, 386–391
9. de Vries, J. and Gould, S.B. (2018) The monoplastidic bottleneck in algae and plant evolution. *J. Cell Sci.* 131, jcs203414
10. McFadden, G.I. (2014) Origin and evolution of plastids and photosynthesis in eukaryotes. *Cold Spring Harb. Perspect. Biol.* 6, a016105
11. Sanchez-Baracaldo, P. *et al.* (2017) Early photosynthetic eukaryotes inhabited low-salinity habitats. *Proc. Natl. Acad. Sci. U. S. A.* 114, E7737–E7745
12. Roger, A.J. *et al.* (2017) The origin and diversification of mitochondria. *Curr. Biol.* 27, R1177–R1192
13. Zeth, K. *et al.* (2002) Structural analysis of the adaptor protein ClpS in complex with the N-terminal domain of ClpA. *Nat. Struct. Biol.* 9, 906–911
14. Erbse, A. *et al.* (2006) ClpS is an essential component of the N-end rule pathway in *Escherichia coli*. *Nature* 439, 753–756
15. Stein, B.J. *et al.* (2016) Structural basis of an N-degron adaptor with more stringent specificity. *Structure* 24, 232–242
16. Bachmair, A. *et al.* (1986) *In vivo* half-life of a protein is a function of its amino-terminal residue. *Science* 234, 179–186
17. Gibbs, D.J. *et al.* (2016) From start to finish: amino-terminal protein modifications as degradation signals in plants. *New Phytol.* 211, 1188–1194
18. Dissmeyer, N. *et al.* (2018) Life and death of proteins after protease cleavage: protein degradation by the N-end rule pathway. *New Phytol.* 218, 929–935
19. Kim, J.M. *et al.* (2018) Formyl-methionine as an N-degron of a eukaryotic N-end rule pathway. *Science* 362, eaat0174
20. Nguyen, K.T. *et al.* (2018) Control of protein degradation by N-terminal acetylation and the N-end rule pathway. *Exp. Mol. Med.* 50, 91
21. Varshavsky, A. (2019) N-degron and C-degron pathways of protein degradation. *Proc. Natl. Acad. Sci. U. S. A.* 116, 358–366
22. Varshavsky, A. (2017) The ubiquitin system, autophagy, and regulated protein degradation. *Annu. Rev. Biochem.* 86, 123–128
23. Kwon, Y.T. and Ciechanover, A. (2017) The ubiquitin code in the ubiquitin-proteasome system and autophagy. *Trends Biochem. Sci.* 42, 873–886
24. Stary, S. *et al.* (2003) PRT1 of *Arabidopsis* is a ubiquitin protein ligase of the plant N-end rule pathway with specificity for aromatic amino-terminal residues. *Plant Physiol.* 133, 1360–1366
25. Varshavsky, A. (2011) The N-end rule pathway and regulation by proteolysis. *Protein Sci.* 20, 1298–1345
26. Varshavsky, A. (1996) The N-end rule: functions, mysteries, uses. *Proc. Natl. Acad. Sci. U. S. A.* 93, 12142–12149
27. Tobias, J.W. *et al.* (1991) The N-end rule in bacteria. *Science* 254, 1374–1377
28. Shrader, T.E. *et al.* (1993) The N-end rule in *Escherichia coli*: cloning and analysis of the leucyl, phenylalanyl-tRNA-protein transferase gene *aat*. *J. Bacteriol.* 175, 4364–4374
29. Kress, W. *et al.* (2009) Clp chaperone-proteases: structure and function. *Res. Microbiol.* 160, 618–628
30. Dougan, D.A. *et al.* (2012) The N-end rule pathway: from recognition by N-recognins, to destruction by AAA+ proteases. *Biochim. Biophys. Acta* 1823, 83–91
31. Kuhlmann, N.J. and Chien, P. (2017) Selective adaptor dependent protein degradation in bacteria. *Curr. Opin. Microbiol.* 36, 118–127
32. Graciet, E. *et al.* (2006) Aminoacyl-transferases and the N-end rule pathway of prokaryotic/eukaryotic specificity in a human pathogen. *Proc. Natl. Acad. Sci. U. S. A.* 103, 3078–3083
33. Dong, M.Q. *et al.* (2007) Quantitative mass spectrometry identifies insulin signaling targets in *C. elegans*. *Science* 317, 660–663
34. Vicente, J. *et al.* (2019) Distinct branches of the N-end rule pathway modulate the plant immune response. *New Phytol.* 221, 988–1000
35. Weits, D.A. *et al.* (2014) Plant cysteine oxidases control the oxygen-dependent branch of the N-end-rule pathway. *Nat. Commun.* 5, 3425
36. Licausi, F. *et al.* (2011) Oxygen sensing in plants is mediated by an N-end rule pathway for protein destabilization. *Nature* 479, 419–422
37. Gibbs, D.J. *et al.* (2011) Homeostatic response to hypoxia is regulated by the N-end rule pathway in plants. *Nature* 479, 415–418
38. White, M.D. *et al.* (2017) Plant cysteine oxidases are dioxygenases that directly enable arginyl transferase-catalysed arginylation of N-end rule targets. *Nat. Commun.* 8, 14690
39. White, M.D. *et al.* (2018) The plant cysteine oxidases from *Arabidopsis thaliana* are kinetically tailored to act as oxygen sensors. *J. Biol. Chem.* 293, 11786–11795
40. Nishimura, K. and van Wijk, K.J. (2015) Organization, function and substrates of the essential Clp protease system in plastids. *Biochim. Biophys. Acta* 1847, 915–930
41. Miller, J.M. *et al.* (2018) Phylogenetic analysis predicts structural divergence for proteobacterial ClpC proteins. *J. Struct. Biol.* 201, 52–62
42. Olinares, P.D. *et al.* (2011) Subunit stoichiometry, evolution, and functional implications of an asymmetric plant plastid ClpP/R protease complex in *Arabidopsis*. *Plant Cell* 23, 2348–2361
43. Lupas, A.N. and Koretke, K.K. (2003) Bioinformatic analysis of ClpS, a protein module involved in prokaryotic and eukaryotic protein degradation. *J. Struct. Biol.* 141, 77–83
44. Nishimura, K. *et al.* (2013) ClpS1 is a conserved substrate selector for the chloroplast Clp protease system in *Arabidopsis*. *Plant Cell* 25, 2276–2301
45. Kim, J. *et al.* (2015) Structures, functions, and interactions of ClpT1 and ClpT2 in the Clp protease system of *Arabidopsis* chloroplasts. *Plant Cell* 27, 1477–1496
46. Nishimura, K. *et al.* (2015) Discovery of a unique Clp component, ClpF, in chloroplasts: a proposed binary ClpF–ClpS1 adaptor complex functions in substrate recognition and delivery. *Plant Cell* 27, 2677–2691
47. Stanne, T.M. *et al.* (2007) Distinctive types of ATP-dependent Clp proteases in Cyanobacteria. *J. Biol. Chem.* 282, 14394–14402
48. Montandon, C. *et al.* (2019) N-Degron specificity of chloroplast ClpS1 in plants. *FEBS Lett.* 593, 992–970
49. Colombo, C.V. *et al.* (2018) A gatekeeper residue of ClpS1 from *Arabidopsis thaliana* chloroplasts determines its affinity towards substrates of the bacterial N-end rule. *Plant Cell Physiol.* 59, 624–636
50. AhYoung, A.P. *et al.* (2016) Structure of a putative ClpS N-end rule adaptor protein from the malaria pathogen *Plasmodium falciparum*. *Protein Sci.* 25, 689–701
51. Abramochkin, G. and Shrader, T.E. (1996) Aminoacyl-tRNA recognition by the leucyl/phenylalanyl-tRNA-protein transferase. *J. Biol. Chem.* 271, 22901–22907
52. van Dooren, G.G. and Striepen, B. (2013) The algal past and parasite present of the apicoplast. *Annu. Rev. Microbiol.* 67, 271–289
53. El Bakkouri, M. *et al.* (2010) The Clp chaperones and proteases of the human malaria parasite *Plasmodium falciparum*. *J. Mol. Biol.* 404, 456–477
54. El Bakkouri, M. *et al.* (2013) Structural insights into the inactive subunit of the apicoplast-localized caseinolytic protease complex of *Plasmodium falciparum*. *J. Biol. Chem.* 288, 1022–1031
55. Tan, J.L. *et al.* (2016) The N-end rule adaptor protein ClpS from *Plasmodium falciparum* exhibits broad substrate specificity. *FEBS Lett.* 590, 3397–3406
56. Dissmeyer, N. (2019) Conditional protein function via N-degron pathway-mediated proteostasis in stress physiology. *Annu. Rev. Plant Biol.* 70, 83–117
57. Perrar, A. *et al.* (2019) New beginnings and new ends – methods for large-scale characterization of protein termini and their use in plant biology. *J. Exp. Bot.* 70, 2021–2038
58. Mahmoud, S.A. and Chien, P. (2018) Regulated proteolysis in bacteria. *Annu. Rev. Biochem.* 87, 677–696

evolve to avoid N-degrons of newly imported proteins? About 40% of stromal chloroplast proteins accumulate with more than one N terminus.

Are N-degron pathways found in plant chloroplasts also present in plastids of nonphotosynthetic tissues? The CLP machinery appears to be expressed in most tissue types, also suggesting functional relevance in nongreen tissues.

59. Humbard, M.A. *et al.* (2013) The N-degradome of *Escherichia coli*: limited proteolysis *in vivo* generates a large pool of proteins bearing N-degrons. *J. Biol. Chem.* 288, 28913–28924
60. Montandon, C. *et al.* (2019) *In vivo* trapping of proteins interacting with the chloroplast CLPC1 chaperone; potential substrates and adaptors. *J. Proteome Res.* 18, 2585–2600
61. Rei Liao, J.Y. and van Wijk, K.J. (2019) Discovery of AAA+ protease substrates through trapping approaches. *Trends Biochem. Sci.* 44, P528–PP545
62. Trosch, R. and Jarvis, P. (2011) The stromal processing peptidase of chloroplasts is essential in *Arabidopsis*, with knockout mutations causing embryo arrest after the 16-cell stage. *PLoS One* 6, e23039
63. Breiman, A. *et al.* (2016) The intriguing realm of protein biogenesis: facing the green co-translational protein maturation networks. *Biochim. Biophys. Acta* 1864, 531–550
64. Millar, A.H. *et al.* (2019) The scope, functions, and dynamics of posttranslational protein modifications. *Annu. Rev. Plant Biol.* 70, 119–151
65. Rowland, E. *et al.* (2015) The *Arabidopsis* chloroplast stromal N-terminome: complexities of amino-terminal protein maturation and stability. *Plant Physiol.* 169, 1881–1896
66. Apel, W. *et al.* (2010) Identification of protein stability determinants in chloroplasts. *Plant J.* 63, 636–650
67. Zhang, H. *et al.* (2015) Quantitative proteomics analysis of the Arg/N-end rule pathway of targeted degradation in *Arabidopsis* roots. *Proteomics* 15, 2447–2457
68. Zhang, H. *et al.* (2018) N-Terminomics reveals control of *Arabidopsis* seed storage proteins and proteases by the Arg/N-end rule pathway. *New Phytol.* 218, 1106–1126
69. Tsiatsiani, L. *et al.* (2013) The *Arabidopsis* METACASPASE9 degradome. *Plant Cell* 25, 2831–2847
70. Aksnes, H. *et al.* (2019) Co-translational, post-translational, and non-catalytic roles of N-terminal acetyltransferases. *Mol. Cell* 73, 1097–1114
71. Aksnes, H. *et al.* (2016) First things first: vital protein marks by N-terminal acetyltransferases. *Trends Biochem. Sci.* 41, 746–760
72. Hoshiyasu, S. *et al.* (2013) Potential involvement of N-terminal acetylation in the quantitative regulation of the epsilon subunit of chloroplast ATP synthase under drought stress. *Biosci. Biotechnol. Biochem.* 77, 998–1007
73. Schmidt, C. *et al.* (2017) Acetylation and phosphorylation control both local and global stability of the chloroplast F1 ATP synthase. *Sci. Rep.* 7, 44068
74. Bienvenut, W.V. *et al.* (2011) Dynamics of post-translational modifications and protein stability in the stroma of *Chlamydomonas reinhardtii* chloroplasts. *Proteomics* 11, 1734–1750
75. Starheim, K.K. *et al.* (2012) Protein N-terminal acetyltransferases: when the start matters. *Trends Biochem. Sci.* 37, 152–161
76. Giglione, C. *et al.* (2014) N-terminal protein modifications: bringing back into play the ribosome. *Biochimie* 114, 134–146
77. Dinh, T.V. *et al.* (2015) Molecular identification and functional characterization of the first Nα-acetyltransferase in plastids by global acetylome profiling. *Proteomics* 15, 2426–2435
78. Koskela, M.M. *et al.* (2018) Chloroplast acetyltransferase NSI is required for state transitions in *Arabidopsis thaliana*. *Plant Cell* 30, 1695–1709
79. Baker, R.T. *et al.* (1992) Ubiquitin-specific proteases of *Saccharomyces cerevisiae*. Cloning of UBP2 and UBP3, and functional analysis of the UBP gene family. *J. Biol. Chem.* 267, 23364–23375
80. Levy, F. *et al.* (1996) Using ubiquitin to follow the metabolic fate of a protein. *Proc. Natl. Acad. Sci. U. S. A.* 93, 4907–4912
81. Taxis, C. and Knop, M. (2012) TIPI: TEV protease-mediated induction of protein instability. *Methods Mol. Biol.* 832, 611–626
82. Renicke, C. *et al.* (2013) A LOV2 domain-based optogenetic tool to control protein degradation and cellular function. *Chem. Biol.* 20, 619–626
83. Hasenjager, S. *et al.* (2019) Optogenetic downregulation of protein levels with an ultrasensitive switch. *ACS Synth. Biol.* 8, 1026–1036
84. Faden, F. *et al.* (2019) Modulating protein stability to switch toxic protein function on and off in living cells. *Plant Physiol.* 179, 929–942
85. Faden, F. *et al.* (2016) Phenotypes on demand via switchable target protein degradation in multicellular organisms. *Nat. Commun.* 7, 12202
86. Natsume, T. and Kanemaki, M.T. (2017) Conditional degrons for controlling protein expression at the protein level. *Annu. Rev. Genet.* 51, 83–102
87. Trauth, J. *et al.* (2019) Synthetic control of protein degradation during cell proliferation and developmental processes. *ACS Omega* 4, 2766–2778
88. Khmelinskii, A. *et al.* (2012) Tandem fluorescent protein timers for *in vivo* analysis of protein dynamics. *Nat. Biotechnol.* 30, 708–714
89. Zhang, H. *et al.* (2019) Tandem fluorescent protein timers for non-invasive relative protein lifetime measurement in plants. *Plant Physiol.* 180, 718–731
90. Yu, Q. *et al.* (2017) Efficient plastid transformation in *Arabidopsis*. *Plant Physiol.* 175, 186–193
91. Ruf, S. *et al.* (2019) High-efficiency generation of fertile transplastomic *Arabidopsis* plants. *Nat. Plants* 5, 282–289
92. Piatkov, K.I. *et al.* (2015) Formyl-methionine as a degradation signal at the N-termini of bacterial proteins. *Microb. Cell* 2, 376–393
93. Linster, E. and Wirtz, M. (2018) N-Terminal acetylation: an essential protein modification emerges as an important regulator of stress responses. *J. Exp. Bot.* 69, 4555–4568