

# Structure, function, and substrates of Clp AAA+ protease systems in cyanobacteria, plastids, and apicoplasts: A comparative analysis

Received for publication, November 24, 2020, and in revised form, January 22, 2021. Published, Papers in Press, January 23, 2021.

<https://doi.org/10.1016/j.jbc.2021.100338>

Imen Bouchnak and Klaas J. van Wijk\*<sup>✉</sup>

From the Section of Plant Biology, School of Integrative Plant Sciences (SIPS), Cornell University, Ithaca, New York, USA

Edited by Joseph Jez

ATPases Associated with diverse cellular Activities (AAA+) are a superfamily of proteins that typically assemble into hexameric rings. These proteins contain AAA+ domains with two canonical motifs (Walker A and B) that bind and hydrolyze ATP, allowing them to perform a wide variety of different functions. For example, AAA+ proteins play a prominent role in cellular proteostasis by controlling biogenesis, folding, trafficking, and degradation of proteins present within the cell. Several central proteolytic systems (e.g., Clp, Deg, FtsH, Lon, 26S proteasome) use AAA+ domains or AAA+ proteins to unfold protein substrates (using energy from ATP hydrolysis) to make them accessible for degradation. This allows AAA+ protease systems to degrade aggregates and large proteins, as well as smaller proteins, and feed them as linearized molecules into a protease chamber. This review provides an up-to-date and a comparative overview of the essential Clp AAA+ protease systems in Cyanobacteria (e.g., *Synechocystis* spp), plastids of photosynthetic eukaryotes (e.g., *Arabidopsis*, *Chlamydomonas*), and apicoplasts in the nonphotosynthetic apicomplexan pathogen *Plasmodium falciparum*. Recent progress and breakthroughs in identifying Clp protease structures, substrates, substrate adaptors (e.g., NblA/B, ClpS, ClpF), and degrons are highlighted. We comment on the physiological importance of Clp activity, including plastid biogenesis, proteostasis, the chloroplast Protein Unfolding Response, and metabolism, across these diverse lineages. Outstanding questions as well as research opportunities and priorities to better understand the essential role of Clp systems in cellular proteostasis are discussed.

## Proteostasis

Protein homeostasis or 'proteostasis' is the process that regulates proteins within the cell to maintain the health of both the cellular proteome and the organism itself. Proteostasis requires integration of protein synthesis, sorting, folding, maturation, and degradation. Proteases play a very central role in proteostasis because they contribute in many different ways, including (i) removal of presequences that are required for

intracellular sorting but must be removed after sorting is completed, (ii) removal of N-terminal methionines, (iii) additional N- or C-terminal cleavages for maturation and stabilization and possibly activation of proteins, (iv) removal of miss-folded, damaged, or aggregated proteins, (v) removal of unwanted proteins in response to environmental or developmental transitions, and (vi) metabolism of proteins as alternate respiratory substrates in times of stress. Proteolysis is therefore central to proteostasis and fitness of the cell. A key challenge in understanding cellular proteostasis is determining how substrates are recognized by proteases and how the activity of proteases is regulated. This recognition typically involves modification of proteins resulting in the formation of a protein degradation signal, commonly assigned as a degron (1). Degrons can be the unfolded domain of a protein or can involve one or more post-translational modifications (PTMs). These degrons can be located at the N-terminus or C-terminus of a protein but also anywhere in between. Much research has focused on N-terminal degrons and the so-called N-end rule degradation pathway, initially discovered in yeast (2) but subsequently shown to be relevant for other eukaryotes and prokaryotes (3, 4). Mechanisms through which proteases recognize their substrates are hard to predict and therefore require experimentation.

## Protease systems with AAA+ proteins

There are many types of proteases, and a universal classification system for peptidases has been developed in which each protein is assigned to a family based on similarities in amino acid sequence (5). Many prominent and well-studied proteases are members of the protein superfamily ATPases Associated with diverse cellular Activities (AAA+). AAA+ proteins perform a wide variety of different functions essential to cell physiology, including control of protein homeostasis, DNA replication, recombination, chromatin remodeling, ribosomal RNA processing, molecular targeting, organelle biogenesis, and membrane fusion (6, 7). They typically form hexameric complexes and act as motors to remodel other proteins. Several central and prominent proteolytic systems use AAA+ proteins to unfold protein substrates to make them accessible for degradation. This allows AAA+ protease systems to degrade aggregates and large proteins, as well as smaller

\* For correspondence: Klaas J. van Wijk, [kv35@cornell.edu](mailto:kv35@cornell.edu).

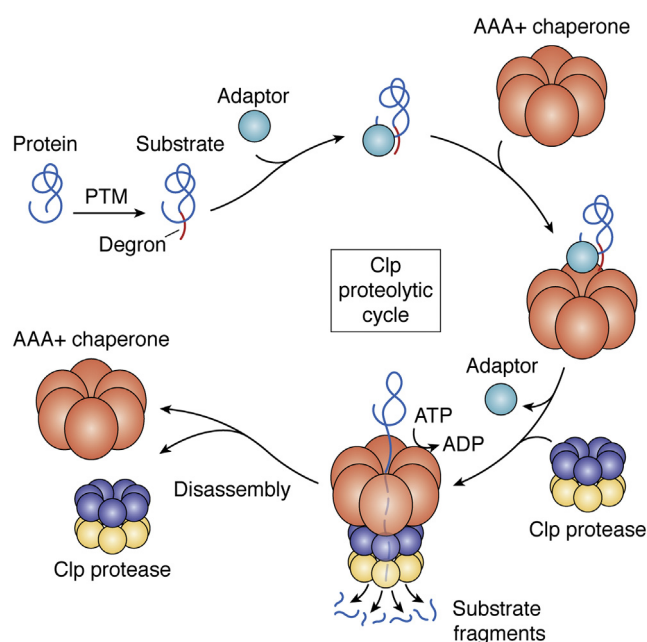
proteins, by feeding them as linearized molecules into a protease chamber protected from the cellular environment. Examples of AAA+ protease systems are the 26S proteasome (8) and the caseinolytic protease (Clp), FtsH, Deg, and Lon protease systems (9, 10). AAA+ proteins share an AAA+ domain with two canonical motifs (Walker A and B) required for ATP binding and hydrolysis (6, 7). The P-loop of the Walker A motif directly interacts with the phosphates of ATP, whereas a specific lysine residue within this P-loop (in the consensus sequence GXXXXGK[T/S]) is crucial, and its mutation prohibits nucleotide binding and inactivates the AAA+ protein (6, 7). The Walker B motif also interacts with the adenine nucleotide, and acidic residues (D and E) within a hydrophobic sequence (hhhhDE) are crucial for ATPase activity. The aspartate residue (D) recruits Mg<sup>2+</sup> required for ATP hydrolysis, whereas the glutamate residue (E) is essential to activate water for the hydrolysis reaction. Mutations in the Walker B motif can be used as effective tools to trap substrates that will bind but cannot be released, as reviewed (11).

### Plastids are essential organelles in algae, plants, and most apicomplexan

Plastids are essential organelles in photosynthetic eukaryotes (algae and plants) and most apicomplexan. Plastids in algae and plants are best known for their role in photosynthesis, but they also carry out the biosynthesis of many primary and secondary metabolites such as fatty acids, amino acids, vitamins, nucleotides, tetrapyrroles, and hormones (12–14). Plastids in photosynthetic organisms undergo various developmental transitions and adaptations, from nonphotosynthetic plastids in roots to photosynthetic chloroplasts in green tissues (13). Each plastid type contains its specific proteome through the coordinated actions of the proteostasis network, involving 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 proteolytic hierarchies and degrons is therefore essential to understand plastid differentiation, adaptation, and function (15–17). Apicoplasts are nonphotosynthetic plastids, derived from secondary endosymbiosis (Fig. 1), found in most Apicomplexa, including *Toxoplasma gondii*, *Plasmodium falciparum*, and other *Plasmodium* spp., but not in others such as *Cryptosporidium* (18). Apicoplasts contribute to amino acid, fatty acid, and isoprenoid metabolism quite comparable with the functions of plastids in plants and algae, but they are never photosynthetic.

### AAA+ protease systems in cyanobacteria, plastids, and apicoplasts

The Clp, FtsH, Deg, and Lon AAA+ protease systems in plastids and apicoplasts are derived from cyanobacterial progenitors (Fig. 1). However, after endosymbiosis, these protease systems diversified and expanded with additional components, as reviewed (15). Examples are the dramatic increase in the number of FtsH, ClpP, and ClpR proteins, as well as novel



**Figure 1. Evolution of the Clp protease system from cyanobacteria to chloroplasts in higher plants to apicoplasts in apicomplexan pathogens.** Primary plastids evolved from cyanobacteria that were engulfed by a eukaryotic cell (primary endosymbiosis). Apicoplasts evolved from a plastid containing eukaryote that was engulfed by another nonphotosynthetic eukaryotic cell (secondary endosymbiosis). The so far studied cyanobacterial Clp systems consist of (1) ClpS1 and ClpS2 and NblA as the adaptor proteins and (2) ClpC and ClpX (yellow and blue) as the chaperone subunits for ClpP3R and ClpP1P2 heterotetradecameric protease core complexes, respectively. The plant plastid Clp system consists of (1) ClpS1 and ClpF as the adaptor proteins, (2) ClpC1, C2, and D (yellow) as the chaperone subunits, (3) heterotetradecameric protease core complexes consisting of the so-called R-ring with ClpP1, R1-4, and the P-ring with ClpP3-6 subunits, (4) ClpT1 and ClpT2 (pink) as the accessory proteins associated with the P-ring. The apicoplast Clp system consists of (1) ClpS as the adaptor protein, (2) ClpC (yellow) as the chaperone, and tetradecameric ClpP3 (gray) and ClpR (black) likely assembled as a single heterotetradecameric core complex. A list of Clp subunits and their functional role for the different biological systems is provided in Table 1. Clp, caseinolytic protease.

auxiliary proteins such as ClpT in plant and algal plastids (Table 1). Over the last decade, there has been substantial progress in identifying structures, functional roles, and substrates of these AAA+ protease systems in cyanobacteria, plastid, and apicoplasts, as summarized in several general reviews on plastid proteases (15, 16, 19) as well as specific reviews for the Clp system (20, 21), FtsH (22), Deg (23), and Lon (24). Each of these AAA+ proteolytic systems is important to biogenesis and function of cyanobacteria and plastids. Loss of function of these proteases can result in lethality in the case of the Clp system as will be discussed further in this review, whereas loss of other AAA+ proteases generally results in milder phenotypes that, however, can affect biogenesis or fitness under specific stress conditions such as heat or excess light.

The FtsH systems are bound to membranes and are important in the removal of membrane proteins within the membranes where these FtsH proteases reside. The Clp, Lon, and Deg protease systems are soluble protein complexes degrading, however, not only soluble proteins but likely also membrane proteins in particular after they have been partially released from the membrane by FtsH or other membrane-

**Table 1**  
Clp proteins in cyanobacteria, plants, algae, and apicomplexes

Functional components of the Clp system	Clp subunit type	Cyanobacteria <sup>a</sup>	Plant plastids <sup>a</sup>	Algal plastids <sup>a</sup>	Apicomplexes <sup>a</sup>
AAA+ chaperones	ClpC <sup>b</sup>	×	×	×	×
	ClpD <sup>c</sup>		×		
	ClpX <sup>d</sup>	×			
Catalytically active protease	ClpP <sup>e</sup>	×	×	×	×
	ClpR <sup>f</sup>	×	×	×	×
Catalytically inactive protease	ClpT <sup>g</sup>		×	×	
	ClpS <sup>h</sup>	×	×	×	×
Auxiliary	ClpF <sup>i</sup>		×		
	NbLA <sup>j</sup>	×			

AAA+, ATPases Associated with diverse cellular Activities; Clp, caseinolytic protease.

<sup>a</sup> The analysis of the Clp system is mostly based on research in a limited set of species. For cyanobacteria, these are *Synechocystis elongatus* PCC 7942, *Synechocystis* sp. PCC803, and *Anabaena* PCC 7120. For algae, this is *Chlamydomonas reinhardtii*. For plants, this is mostly *Arabidopsis thaliana* and to a lesser extent also tobacco, tomato, maize, and rice. For apicomplexan, this is *Plasmodium falciparum*.

<sup>b</sup> Cyanobacteria and apicomplexes have one ClpC homolog, whereas *Arabidopsis* has two ClpC homologs (ClpC1 and ClpC2).

<sup>c</sup> ClpD is a chaperone specific to plants and lacks a UVR domain. Its expression is induced by drought and during senescence.

<sup>d</sup> ClpX is not found in plastids, but cyanobacteria have one ClpX homolog. It should be noted that mitochondria have one or more ClpX homologs. *Arabidopsis* has three ClpX homologs.

<sup>e</sup> ClpP is the serine-type proteolytic subunit with a conserved catalytic triad. Cyanobacteria have three ClpP subunits (P1, P2, and P3). *Arabidopsis* has six ClpP subunits (P1–P6), of which ClpP1 is plastid-encoded and the others are nuclear-encoded. Only ClpP2 is located in mitochondria. *Chlamydomonas* has three plastid-localized Clp subunits (P1, P4, and P5), one of which is plastid-encoded (ClpP1). Apicomplexes have just one plastid-localized Clp subunit.

<sup>f</sup> ClpR is a homolog of ClpP but lacks the catalytic triad and is therefore not catalytically active. Cyanobacteria have one ClpR subunit, *Chlamydomonas* and *Arabidopsis* have each four ClpR subunits (R1–R4), and apicomplexes have one ClpR subunit.

<sup>g</sup> ClpT is unique to algae and higher plants; however, they show very low sequence similarity. ClpT subunits associate with the ClpP ring and likely play a role in the activation of the Clp core or perhaps substrate delivery independent of Clp chaperones. *Arabidopsis* and *Chlamydomonas* each have two ClpT subunits.

<sup>h</sup> ClpS is a highly conserved adaptor or recognin acting specifically in recruiting substrates with an N-terminal degron. Cyanobacteria have ClpS1 and ClpS2 homologs, whereas as plastids in algae and plants only have ClpS1 homologs (typically just one). Apicomplexes have one ClpS homolog.

<sup>i</sup> ClpF is an adaptor function (or perhaps also anti-adaptor) and is unique to higher plants. It associates with both ClpS and ClpC homologs and it has both a UVR and YCCV domain.

<sup>j</sup> Cyanobacterial NbLA (nonbleaching protein A) is a specialized ClpC adaptor to deliver phycobilisome proteins for degradation. Some species have two NbLA homologs.

bound proteases. Substrate recognition and delivery mechanisms, as well as degrons, remain poorly understood for these systems; arguably, understanding the rules for substrate recognition should be the top priority for studying protein degradation by these AAA+ protease systems. These AAA+ protease system functions as a protease network together with additional proteases (25), similarly as postulated for other cellular proteases (26) and proteostasis components including chaperones (27). Not much is known to what extent the different proteases in a cell or subcellular compartment act in parallel or in series to degrade proteins to small peptides or individual amino acids. mRNA-based coexpression analysis and genetic interaction studies using loss-of-function protease mutant lines can provide insight into these networks for plastids and mitochondria as we discussed previously (25).

In the present review, we provide an update on the Clp AAA+ protease systems found in cyanobacteria, plastids of algae and plants, as well as apicomplexes in the nonphotosynthetic apicomplexan pathogen *P. falciparum*, the causal agent of malaria. We focus on the Clp system rather than the FtsH, Lon, or Deg proteases because of the high number of recent publications and findings for the Clp system. Comparison of Clp systems across cyanobacteria and plastid organelles allows the study of the diversification and adaptation of these AAA+ protease systems (21, 28). We compare and discuss progress and challenges in identifying cyanobacterial, plastid, and apicomplexan Clp protease structures, substrates, substrate adaptors (e.g., NbLA, ClpS), and their degrons.

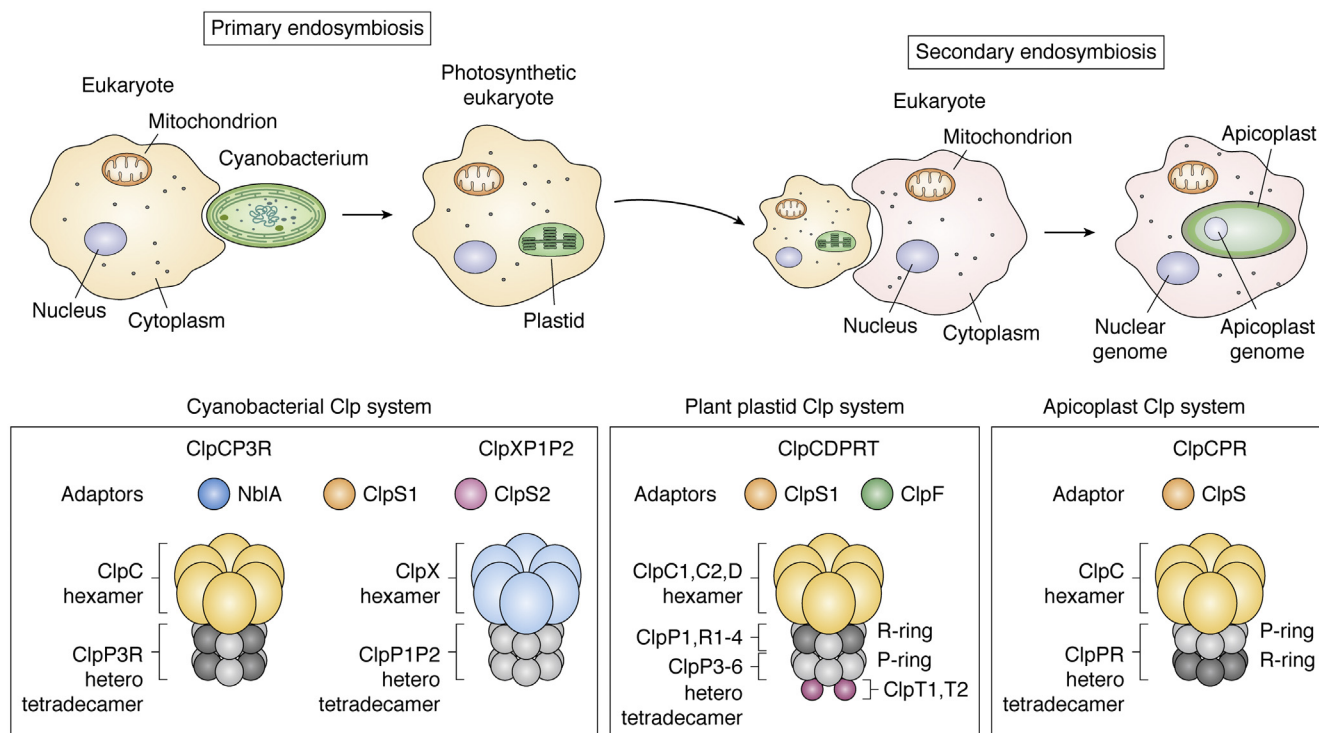
### The cycle of substrate selection and degradation by the Clp system

Clp-dependent proteolysis is an ATP-dependent multistep-regulated process that involves one or more Clp chaperones

assembled into hexamers and a Clp protease core complex consisting of two stacked heptameric rings forming a tetradecameric complex (Figs. 1 and 2; Table 1). Proteins become substrates after undergoing PTMs that must result in the generation of a degron. The degron-containing substrate is now recognized directly by the Clp chaperone(s), but this can also involve active recruitment by so-called adaptor proteins or recognins, or even other chaperones, for example, members of the Hsp70 family in case of misfolded or aggregated protein substrates (Fig. 2). Upon interaction of the substrate with the Clp chaperone, the ATP-dependent substrate unfolding process starts and the Clp protease core complex is recruited to the substrate–chaperone assembly. ATP binding and hydrolysis is required for substrate unfolding and mostly likely also for Clp chaperone oligomerization. In contrast, the actual proteolytic cleavage by the catalytic Clp protease core does not require ATP, but specific and dynamic interactions (docking) between the Clp chaperone hexamer and the Clp protease core are required for catalytic cleavage of peptidyl bonds in the substrate. Small substrate fragments (~6–9 aa) are released from the Clp protease core through dynamic lateral pores and, once the substrate degradation is complete, the Clp chaperone–protease complex disassembles (Fig. 2). It should be noted that the Clp chaperones may accumulate as dimers (e.g., in chloroplasts) when not engaged in the degradation cycle and that formation of the chaperone hexamer requires priming of the chaperone by adaptors and/or ATP, leading to the formation of the activate hexamer in the ATP-bound state (9, 29, 30).

### Organization, expression, coevolution, structures, and functions of Clp systems

The basic structure of the Clp machinery is generally conserved across cyanobacteria, plastids, and apicomplexes and



**Figure 2. Working model of the Clp proteolytic cycle and the Clp components.** Proteins become substrates after undergoing a post-translational modification (PTM) that must result in the generation of a substrate recognition signal, a degron. The substrate is recruited to the Clp chaperone complex, possibly in dependence of an adaptor (or an adaptor complex). ATP-dependent unfolding and threading of the substrate through the central pore of the Clp hexamer follows, and the Clp protease core complex docks onto the hexameric Clp chaperone. Small substrate fragments (~6–9 aa) are released from the Clp protease core through dynamic lateral pores, and once the substrate degradation is completed, the Clp chaperone–protease complex disassembles. Clp chaperones may accumulate as dimers when not engaged in the degradation cycle and the formation of the chaperone hexamer requires priming of the chaperone by adaptors and/or ATP, leading to the formation of the activate hexamer in ATP-bound state. Clp, caseinolytic protease.

consists of a cylinder-like protease core and an AAA+ chaperone ring complex (10, 21, 31) (Figs. 1 and 2; Table 1). The chaperone complex serves as a molecular gate that controls substrate access and recognizes, unfolds, and translocates protein substrates into the core cavity in an ATP-dependent manner. Recent cryoelectron microscopy studies of bacterial ClpAP, ClpXP1P2, and ClpXP assemblies have provided an unparalleled insight into the organization and dynamics of these supercomplexes from *Escherichia coli*, *Listeria monocytogenes*, and *Neisseria meningitidis*, respectively (32–34). So far, no cryoelectron microscopy structures of Clp chaperone–protease assemblies have been resolved for cyanobacteria, plastids, or apicoplasts, but high-resolution structures for some of the individual Clp components have been determined by X-ray crystallography for these systems, as will be discussed below. Key differences across cyanobacteria, plastids, and apicoplasts are the number of different proteins that make up the tetradecameric protease core, the number of ClpC and ClpS homologs, and the presence of ClpX (only in cyanobacteria) or ClpD (only in higher plant plastids) or ClpT homologs (only in plastids of both algae and higher plants). In addition to ClpS homologs observed in all three ‘systems,’ just two additional substrate adaptors have been identified so far, namely ClpF for plastids and NbIA for cyanobacteria. Based on identification of additional Clp adaptors in various non-photosynthetic bacteria, as reviewed (9, 35), we speculate that

cyanobacteria, plastids, and apicoplasts have additional adaptors that are likely to be diverse, as they probably evolved independently in each of the lineages. In the remainder of this section, we briefly summarize the organization, expression, coevolution, and structures of Clp systems across these systems and point out which components are essential.

### The Clp system in cyanobacteria

The cyanobacterial Clp system in *Synechocystis* species has mostly been studied in rod-shaped *S. elongatus* PCC 7942 and coccoid *Synechocystis* sp. PCC803. The Clp system in both species is comprised of a heterotetradecameric ClpP1P2 core (36) and an essential heterotetradecameric ClpP3R core, as reviewed (21). ClpX and ClpC are chaperones for the ClpP1P2 and ClpP3R cores, respectively (Fig. 1; Table 1). ClpP2 is on an operon with ClpX (37), and ClpP2 shows regulated (diel) polar cellular localization impacting its activity (36). ClpS1 and ClpS2 are adaptor proteins interacting with ClpC, but cyanobacteria contain an additional ClpC adaptor, NbIA (some species have two NbIA homologs), which is specialized to deliver phycobilisome (PBS) proteins for degradation (38, 39). The structures of NbIA from *Anabaena* PCC 7120, *Thermosynechococcus vulcanus*, and *Synechococcus elongatus* sp. PCC 794 have been determined by X-ray crystallography and were shown to form four-helix bundles of dimers (40, 41). The DNA damage repair protein UmuD was recently identified as an

*in vivo* substrate for both ClpCP3R and ClpXP1P2. UmuD was found to have an unusual dual degron that is recognized by the ClpS homologs or by ClpXP (42) (ClpS homologs and N-degrons across cyanobacteria, plastids, and apicoplasts). ClpXP1P2 also plays an important function in regulating the circadian rhythm (36). The loss of ClpXP1P2 protease results in long-period circadian rhythms. Deletion of the ClpX chaperone, but not ClpP1/2, results in cell division defects in a manner that is dependent on the expression of a dusk-peaking factor (36). However, the Clp substrates that impact division and circadian rhythm are not yet known (36) (Discovery of Clp substrates and functions based on phenotypes, comparative proteomics and *in vivo* and *in vitro* affinity enrichment).

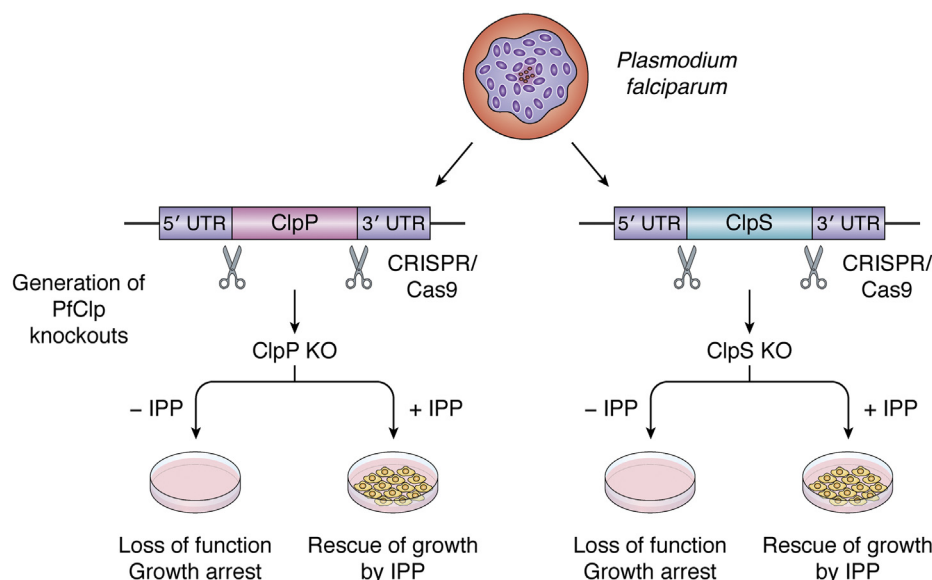
### The Clp system in plastids of algae and plants

The Clp protease in plastids of higher plants originates from its ancient cyanobacterial endosymbiont, but it increased in complexity since endosymbiosis (Fig. 1). Indeed, higher plant plastids and chloroplasts contain the most diversified Clp protease core compared with all other species. Forward and/or reverse genetics in *Arabidopsis*, maize, rice, and tobacco demonstrated the essential nature of the plastid Clp system, as reviewed (20, 21). Complete loss of ClpC chaperone or ClpPR protease capacity results in embryo lethality, whereas partial loss results in delayed growth and development and virescent leaves. The plastid Clp system in *Arabidopsis* consists of a hetero-oligomeric protease core comprising five proteolytically active subunits (ClpP1 and ClpP3-ClpP6) and four proteolytically inactive proteins (ClpR1-4), as well as two plant-specific accessory proteins (ClpT1,2), three AAA+ chaperones (ClpC1, ClpC2, and ClpD), and two adaptors (ClpS1 and ClpF) (43) (Fig. 1; Table 1). Plastids do not contain any ClpX homologs but instead are present in mitochondria. The ClpPR core is constructed from the heptameric P-ring containing the four ClpP subunits (P3: P4: P5: P6 = 1: 2: 3: 1) and the heptameric R-ring containing the four ClpR proteins and ClpP1 (P1: R1: R2: R3: R4 = 3: 1: 1: 1) (21). A recent study showed that there is a tight correlation between amino acid substitution rates in the plastid-encoded ClpP1 and the nuclear-encoded Clp subunits across a broad sampling of angiosperms, suggesting continuing selection on interactions within this complex (44). This is also consistent with the finding that ClpP5 (3 copies per Clp core) is crucial for Clp catalysis, whereas ClpP3 (one copy per core) plays an essential role in Clp structure even if its catalytic activity is dispensable (45). The ClpC1,2 and ClpD chaperones carry two ATPase domains and an IGF motif that is essential for binding to the Clp protease core complex. ClpC1,2 and likely also ClpD (but not experimentally demonstrated) are involved in gating the ClpPRT complexes. Whereas ClpC chaperones are found in cyanobacteria and a variety of bacteria (46), the ClpD chaperone is unique to plants and ClpD has only ~45% sequence identity to ClpC proteins in plastids and cyanobacteria (21). Similar to the ClpC homologs, ClpD is a so-called class-1 HSP100 chaperone containing two conserved ATPase domains with all typical features (Walker A and B domains and the pore loops) per monomer, as well as the

conserved IGF motif (21). However, ClpD neither has the R-motif nor the uvrB/C motif present in ClpC1,2 and ClpF. uvrB/C domains are generally believed to support protein-protein interactions <https://prosite.expasy.org/PDOC50151>. Removal of the R motif in cyanobacterial ClpC results in loss of interaction with the ClpP3/R core but enables binding of *E. coli* ClpP and subsequent proteolytic activity (47). It remains to be determined if ClpD (lacking this R motif) can interact with the plastid ClpPR complex. As we pointed out previously, the uvrB/C of ClpF was suggested to interact with the uvrB/C domain in the ClpC homolog (43), perhaps indicating that ClpF does not interact with ClpD. A recent X-ray structure showed that the N-domain of ClpD is structurally divergent from ClpC1, and on that basis, it was suggested that ClpD may have different substrate recognition mechanisms than ClpC homologs (48). Nuclear genes encoding for Clp subunits coexpress with each other, except ClpD which clearly coexpresses with proteins involved in drought stress and senescence (25). This sets ClpD apart as a chaperone with specific stress and developmental function (ClpD has also been named ERD1 for Early Drought Induced Protein 1), but substrates of ClpD are so far unknown. Identification of ClpD interactors and substrates should help identify ClpD function and perhaps explain its diversification from ClpC. Plants have one or two ClpS1 homologs that likely function to recruit chloroplast proteins with N-degrons to ClpC1,2 and perhaps also to ClpD, for degradation (4). ClpS1 interacts with the plant-specific ClpF protein that contains uvrB/C and YccV domains. We postulated that ClpS1 and ClpF form a binary complex for delivery of glutamyl t-RNA reductase (GluTR) and perhaps other substrates. The core domain of ClpS1 and the N-terminal/UVR domains of ClpF are important for ClpC1 interaction (43). Finally, plastids contain a fourth Clp chaperone, ClpB3 (49, 50), which also has two ATPase domains but lacks an IGF motif and is, therefore, unable to dock onto the Clp protease core (21). ClpB3 is a member of the well-studied group of ClpB unfoldases, which rescue aggregated proteins for reactivation rather than degradation in collaboration with HSP70 proteins (51, 52) (Other substrate adaptors and delivery systems).

### The Clp system in apicoplasts

*T. gondii* and *P. falciparum* are both well studied members of the apicomplexan, but the apicoplast Clp system has only been studied in *P. falciparum*. *P. falciparum* apicoplasts contain single ClpP, ClpR, ClpC, and ClpS proteins (53). Functional apicoplasts are essential for the viability of the *P. falciparum* because apicoplasts synthesize and supply the cell with the isoprenoid precursor, isopentenyl pyrophosphate (IPP). Indeed, apicoplasts are dispensable for growth and viability of the organism if IPP is supplemented in the culture medium. A recent study reported on the engineering of a *P. falciparum* strain with a cytosolic mevalonate bypass system that produces isoprenoid precursors and therefore replicated normally after the loss of the apicoplast organelle (54). This approach provides an alternative to chemical



**Figure 3. Schematic representation summarizing the essential role of Clp components for apicoplast biogenesis and survival of *Plasmodium falciparum*.** ClpPKO and ClpSKO lines were generated by CRISPR-cas9. The absence of either ClpP or ClpS proteins led to the loss of function and growth arrest, suggesting that ClpP and ClpS are essential for apicoplast biogenesis and survival of the parasite. This growth inhibition was completely rescued by addition of the isoprenoid precursor IPP (isopentenyl pyrophosphate) (53). Clp, caseinolytic protease.

complementation with IPP to probe the essential apicoplast function. ClpP, ClpR, ClpC, and surprisingly also ClpS (ClpS is not essential for viability in plants, algae, cyanobacteria, and nonphotosynthetic bacteria) are each essential for apicoplast viability because they are required for apicoplast biogenesis, and the absence of Clp components led to growth arrest and apicoplast loss (53, 55) (Fig. 3). Loss-of-function mutants for ClpP, ClpC, and ClpS could be obtained when grown in presence of IPP, allowing the introduction of modified ClpP and ClpC transgenes to probe Clp processing, assembly, protein–protein interactions, and trap substrates (53, 55). ClpP and ClpR form each homoheptameric rings, but they do not appear to form a stable tetradecameric complex together (56, 57). However, *in vivo* affinity tagging with ClpC, ClpP, or ClpS as bait, identified ClpR in all screens, suggesting that ClpR does form a complex with the rest of the Clp subunits. It has been observed for *E. coli* and plastid Clp protease core complexes that the intraheptameric ring interactions are far stronger than the interactions between the heptameric rings: an increase in ionic strength is sufficient to destabilize these ring–ring interactions, but not intraring interactions (58). This may explain why it has been difficult to purify an apicoplast ClpPR complex. It seems therefore quite likely that homoheptameric rings of apicoplast ClpR and ClpP do interact to form a complex. The structure of *P. falciparum* ClpR was determined by X-ray crystallography and showed that the ClpR monomer adopts a similar fold as bacterial ClpPs (57). *In vivo* experiments using *P. falciparum* expressing a catalytically inactive ClpP subunit and its endogenous functional ClpP subunit showed that ClpP undergoes processing by neighboring ClpP subunits to remove its prodomain. In contrast, ClpR has no prodomain and hence does not require any processing after its N-terminal signal peptide is cleaved (53).

The structure of ClpS was determined by X-ray crystallography (59) and interaction with N-degrons investigated (59, 60), as will be discussed in more detail below (ClpS homologs and N-degrons across cyanobacteria, plastids, and apicoplasts).

### Discovery of Clp substrates and functions based on phenotypes, comparative proteomics, and *in vivo* and *in vitro* affinity enrichment

Clp adaptors and chaperones must selectively recognize the degradation signal in substrates through their substrate-interacting domains and deliver the substrates to the proteolytic chamber (21). Confirmed and candidate Clp substrates have been identified in cyanobacteria, chloroplasts, and apicoplasts by (i) a combination of comparative proteomics and follow-up studies, as reviewed (20, 21), (ii) discovery through direct interaction assays with adaptors ClpS1 and/or ClpF (43, 61) or NblA (62), or (iii) *in vivo* (*in planta*) trapping with the plastid ClpC chaperone (29) or ClpP, ClpR, ClpC, and ClpS proteins in apicoplasts (53), or (iv) targeted analysis of specific proteins to probe for the role of Clp instability, for example, for UmuD in cyanobacteria (42).

So far, no comparative proteomic screens or *in vivo* trapping approaches (as will be explained below) have been conducted for loss-of-function Clp mutants in cyanobacteria. It would be of great interest to carry out such studies for different cyanobacterial species. However, as mentioned above, there are two confirmed cyanobacterial Clp substrates, namely UmuD involved in DNA damage repair (42) and PBS proteins such as  $\beta$ -phycocyanin (62). The proteolysis adaptor, NblA, binds to the N-terminus of  $\beta$ -phycocyanin and delivers it to the ClpC chaperone. It should be noted that NblA is also involved in the disassembly of the protein pigment complex (62–64).

Proteolytic ClpP1P2 complex activity is needed to support a normal circadian rhythm. It was suggested that the core oscillator protein KaiC is a target for the ClpXP protease because the KaiC level increased in the *clpxp2* null background under specific conditions (36). In the same study, it has also been proposed that a factor, defined as resetting factor, is targeted for degradation by the ClpXP1P2 protease system. It is important to note that *in vitro*, the Kai-based oscillator can function without the need for protein synthesis and degradation and relying on reversible phosphorylation and protein–protein interaction, but the *in vivo* scenario is more complicated, as discussed (36). Finally, no Clp protease activity besides ClpX function is needed to control cell division (even if the cell division protein FtsZ is a substrate of ClpXP in *E. coli* (65)).

The impact of the loss of chloroplast Clp capacity on plant phenotypes, through reduced accumulation of ClpC/D chaperones, ClpPRT subunits, or ClpS1 and ClpF adaptors, has been studied in *Arabidopsis* and to a lesser degree also in maize, tomato, rice, tobacco, and *Chlamydomonas*. Collectively, these studies suggest that Clp degrades a wide range of substrates and molecular plastid functions and that insufficient Clp protease or chaperone activity results in protein folding stress and probably protein aggregation (Clp in the proteostasis network, retrograde signaling and the cpUPR). Importantly, this has shown that the Clp system in *Arabidopsis* targets specific metabolic enzymes to help fine-tune metabolic activity, as exemplified by GluTR, a key enzyme in the tetrapyrrole pathway, deoxyxylulose 5-phosphate synthase (DXS) in the methylerythritol 4-phosphate pathway, phytoene synthase in the carotenoid pathway, and the thylakoid P-type ATPase of *Arabidopsis* 2 (PAAM2/HMA8) involved in copper transport, as summarized in an excellent review on control of plastid metabolism by the Clp system (20). Reduced activity of the Clp system also strongly affects plastid mRNA metabolism and plastid ribosomes, but it is still not clear whether these are pleiotropic effects or because the Clp system is essential to degrade or cleave specific proteins involved in plastid gene expression. Finally, other studies show the impact of the Clp system on tomato fruit ripening (66), and temporal proteomics using inducible RNAi lines of Clp subunits in *Nicotiana benthamiana* showed a very similar impact on the chloroplast proteome as in *Arabidopsis*, suggesting a conserved role of the Clp system across plant species (67). Two recent studies (68, 69) examined various phenotypic effects of the cosuppression of ClpC1 and ClpC2 in *N. benthamiana*. Similar as in *Arabidopsis*, loss of ClpC capacity resulted in reduced growth, virescent leaves, and a range of developmental defects. Not surprisingly, steady-state concentrations of a range of soluble primary metabolites were affected. These results underscore the essential nature of the Clp system not only in *Arabidopsis* but also in other plant species such as tobacco.

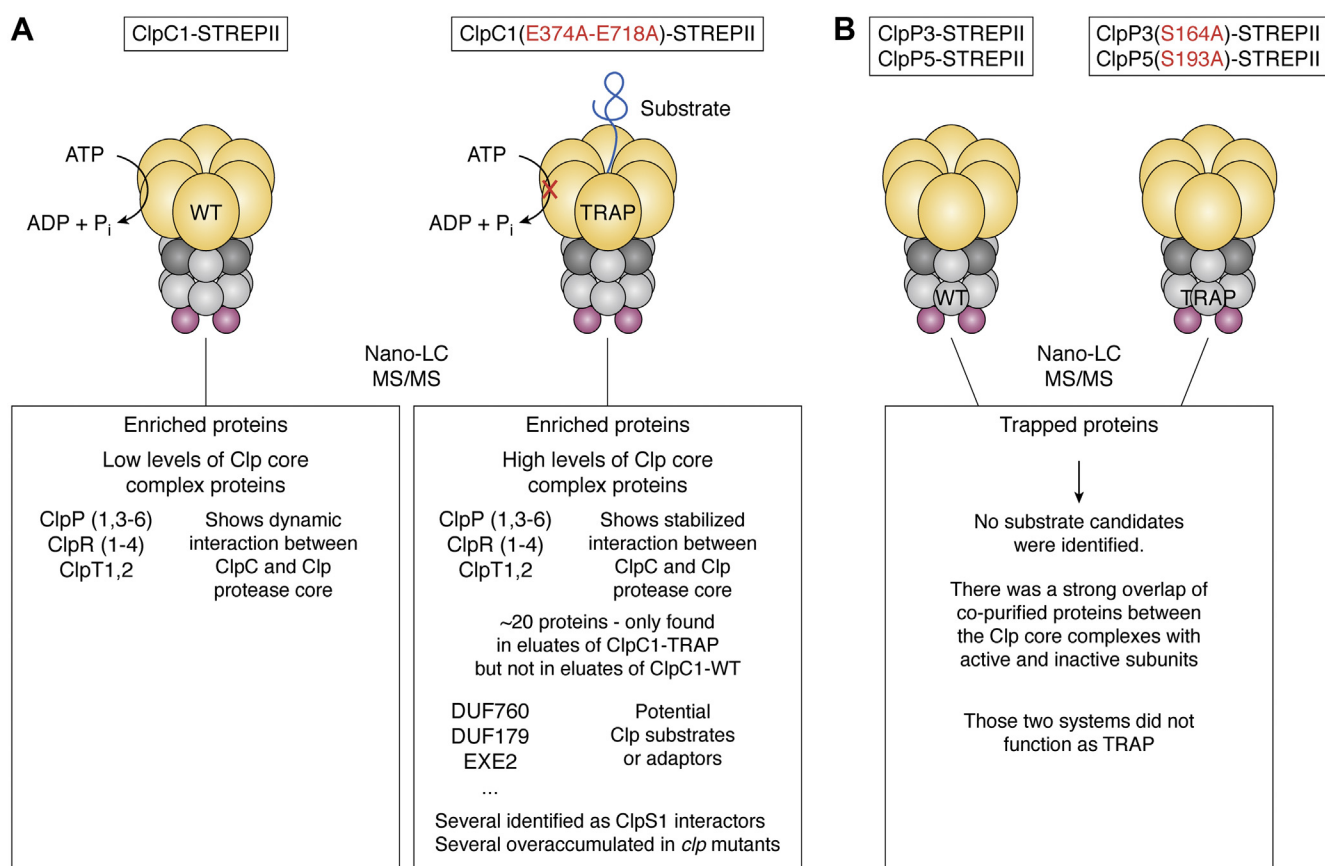
An *in vivo* trapping approach in *Arabidopsis* identified a dozen potential substrates interacting with *Arabidopsis* chloroplast ClpC1 (29), following strategies successfully used for substrates trapping of other AAA+ proteins in bacterial systems, as reviewed (11). The *in vivo* trap was generated by

expressing ClpC1 mutated in two critical glutamate residues in the two Walker B domains of ClpC1 required for the hydrolysis of ATP and with a C-terminal STREPII affinity tag for purification (29) (Fig. 4). When expressed in the WT, this ClpC1–TRAP induced a dominant visible phenotype, and no viable mutants that express ClpC1–TRAP in the *clpc1-1* null mutant could be recovered. This complete loss of viability and the dominant negative phenotype, together with MS evidence, suggests that (i) the ClpC1–TRAP forms inactive mixed oligomers with endogenous ClpC1 (in the WT background) and ClpC2 (in both WT and *clpc1-1* backgrounds) and (ii) the ClpC1–TRAP recruits and stably associates with Clp protease cores forming nonfunctional (jammed) complexes. Affinity purification of the ClpC1–TRAP followed by MS analysis resulted in a dozen proteins highly enriched compared with affinity-purified ClpC1 with a C-terminal STREPII affinity tag. These enriched proteins likely represent Clp protease substrates and/or new adaptors (Fig. 4). Several of these trapped proteins overaccumulated in *clp* mutants and/or were found as interactions for the adaptor ClpS1, supporting their functional relationship to Clp. Importantly, all ClpP (1, 3–6), ClpR (1–4), and ClpT1,2 subunits that make up the plastid protease core complex were strongly enriched in the ClpC1–TRAP complexes, providing the first robust support for ClpC and Clp core physical and functional interactions (29). *In vivo* traps were generated by expressing ClpP3 or ClpP5 mutated in one critical serine residue and with a C-terminal STREPII affinity tag for purification (ClpP3(S164A)–STREPII, ClpP5(S193A)–STREPII). Those two systems did not function well as a TRAP compared with the ClpC1 TRAP. No substrate candidates were identified, suggesting that the bottle neck for degradation is likely substrate recognition and unfolding by Clp adaptors and chaperones, upstream of the Clp core (45) (Fig. 3).

*In vivo* Clp protein interactome studies for the apicoplast Clp system in *P. falciparum* were carried out through affinity enrichments with ClpC, ClpR, ClpP, ClpP-inactive, and ClpS followed by MS to identify the enriched proteins (53) (Fig. 3). Enriched proteins were found for each of these affinity analyses. This supported the interaction between the four Clp subunits and identified multiple candidate substrates. Several proteins were identified in all five pull-downs and showed a diverse array of functions (*e.g.*, DNA replication, protein trafficking, RNA methyltransferase), whereas others were identified in four or less pull-downs. The pool of candidate Clp substrates is an excellent resource to probe apicoplast Clp function in more detail and discover degrons for the Clp system. We compared these Clp interactors with proteins identified in the *in vivo* ClpC1 trap for *Arabidopsis* (29) and confirmed plastid substrates, but no obvious overlapping homologs were identified.

### Clp in the proteostasis network, retrograde signaling, and the cpUPR

Cyanobacteria, plastids, and apicoplasts all contain multiple protease systems that collectively control protein processing and degradation. Based on the strong phenotypes of *clp*



**Figure 4. Schema summarizing *in vivo* trapping in *Arabidopsis*.** A, the *in vivo* trap was generated by expressing ClpC1 mutated in two critical glutamate residues in the two Walker B domains of ClpC1 required for the hydrolysis of ATP and with a C-terminal STREPII affinity tag for purification (ClpC1(E374A-E718A)-STREPII). ClpP (1, 3–6), ClpR (1–4) and ClpT1,2 subunits were strongly enriched in the ClpC1–TRAP, providing the first robust support for ClpC and Clp protease physical and functional interaction. A dozen potential substrates were only detected in eluates of ClpC1–TRAP but not in eluates of ClpC1–WT. Several of these trapped proteins also overaccumulated in Clp mutants and were previously identified as ClpS1 interactors, supporting the role of Clp proteases in the degradation of these targets (29). B, *in vivo* traps were generated by expressing ClpP3 or ClpP5 mutated in one critical serine residue and with a C-terminal STREPII affinity tag for purification (ClpP3(S164A)-STREPII and ClpP5(S193A)-STREPII). Those two systems did not function well as a TRAP compared with the ClpC1 TRAP. No substrate candidates were identified, suggesting that the bottle neck for degradation is likely substrate recognition and unfolding by Clp adaptors and chaperones, upstream of the Clp core (45).

mutants with reduced Clp capacity, the Clp system clearly plays a central role in proteostasis in all three systems. Loss of Clp capacity not only affects the accumulation of specific proteins, as discussed in the previous sections, but also can result in protein folding stress and protein aggregation, leading to loss of viability, often referred to as proteotoxicity (70, 71). A typical response to proteotoxicity is the upregulation of chaperone capacity in an effort by the cell (or subcellular compartment) to remove protein aggregates and avoid more misfolding. Such upregulation involves signaling from the affected subcellular compartment to the nucleus to activate transcription of chaperones and possibly proteases and/or protein repair enzymes, summarized under the term unfolded protein response (UPR) (70–73).

To the best of our knowledge, UPR pathways in response to loss of Clp capacity have not been studied in cyanobacteria. However, upregulation of chaperones in cyanobacteria is clearly associated with a variety of stresses including heat, cold, and reactive oxygen species stress (74, 75). A UPR signaling pathway has not been identified for apicoplasts, but it seems likely that such a pathway must exist to alleviate apicoplast

proteome folding stress and aggregation. It will be important to evaluate a possible UPR to a loss of Clp capacity in cyanobacteria and apicoplasts and compare this to recent evidence for plastid UPR pathways in plants and algae.

Loss of chloroplast Clp capacity results in an increase of plastid chaperones CPN60/HSP70 and the unfoldase ClpB3 (52) as first observed in comparative proteomics of the *Arabidopsis clpr2-1* mutant (with only ~20% Clp protease capacity) (76) and subsequently demonstrated in many additional comparative proteomics studies in *Arabidopsis*, as reviewed (20, 21). The upregulation of these chaperones suggests a retrograde signaling pathway that relays proteotoxic/folding stress in the chloroplast to the nucleus (20, 21, 77). The unfoldase activity of *Arabidopsis* ClpB3 was recently studied using recombinant ClpB3. It was shown to form a hexameric complex in the presence of MgATP and display intrinsic ATPase activity across a wide range of pH and was resistant to heat. This *in vitro* unfoldase activity was observed in absence of HSP70 (or other chloroplast stromal chaperones) and supports the role of ClpB3 as unfoldase/disaggregase in chloroplasts (52) and complements an earlier *Arabidopsis in vivo*



study which showed that ClpB3 is critical for plastid biogenesis and possibly heat stress tolerance (50). A recent study in *Arabidopsis* identified the transcription factor HsfA2 (AT2G26150) as a component of this retrograde signaling pathway (77). Expression of HsfA2 was independent of the pentapeptide repeat protein Genomes Uncoupled 1 (GUN1; AT2G31400), a central integrator of chloroplast to nucleus retrograde signaling pathways. Interestingly, whereas GUN1 mRNA levels and the translation rate remain high across development, the half-life of the GUN1 protein is very short (~4 h) because of ClpC-dependent degradation, resulting in low GUN1 protein levels under nonstress conditions (78). Thus, GUN1 is a deliberately unstable protein and is likely a genuine substrate of the Clp protease system. There is a clear parallel to the Ethylene Response Factor VII family of transcription factors involved in oxygen signaling and the anoxia response, as reviewed (79). These Ethylene Response Factor VII transcription factors are constantly degraded by the Cys/Arg branch of the N-degron pathway involving an N-recognin E3 ligase and the proteasome; however, at low cellular oxygen concentrations, degradation slows down, allowing these transcription factors to accumulate and activate the anoxia response (79). Testing a series of truncated GUN1 variants for *in vivo* stability suggested that the pentapeptide repeat domains play a role in stabilization, perhaps indicating that they function as (part of) a degron (78). However, it remains to be determined how GUN1 is recognized as a substrate and how it can be stabilized: many scenarios can be envisioned including a conformational change that could expose a degron, specific PTMs at selected amino acid residues generating or removing a degron, protection of a degron through binding by other proteins (e.g., an antiadaptor), or even dynamic regulation of substrates' adaptors.

A chloroplast Protein Unfolding Response (cpUPR) and retrograde signaling pathway was also proposed for the green algae *Chlamydomonas reinhardtii* (80). Recently, a cytosolic kinase MARS1 was identified as a chloroplast retrograde signaling component in *Chlamydomonas* (81). Loss of MARS1 (Cre16.g692228) reduced the ability of cells to cope with excessive light stress, whereas overexpression of MARS1 improved the cellular response to light stress. Furthermore, mutation of the catalytic triad of the kinase impaired MARS1 function in this signaling pathway. MARS1 was identified in a suppressor screen of cpUPR activation based on the expression of VIPP2 and HSP22 in response to downregulation of plastid-encoded ClpP1. These reporters are quite different in their functions than the stromal chaperones ClpB3/HSP70 that are used as markers for folding stress and aggregation in the *Arabidopsis* cpUPR studies. A recent study showed that VIPP2 forms a complex with its homolog VIPP1 and HSP22 to alleviate membrane stress and protein membranes from oxidative damage and to aid in the removal of dysfunctional proteins (82). BLAST searches identified AT1G67890 and AT5G49470 as the closest *Arabidopsis* MARS1 homologs in plants. Both *Arabidopsis* homologs have a PYP-like sensor domain (IPR035965) and a protein kinase domain. PYP-like sensor domain proteins have been identified in a wide range of species

and often act as oxygen or redox sensors (83). Extensive work for higher plant chloroplasts identified multiple retrograde signaling pathways involving, for example, apocarotenoids, oxylipins, or 3'-phosphoadenosine 5'-phosphosulfate (84–87). It is likely that multiple retrograde signaling pathways also operate in green algae such as *Chlamydomonas* (88). To further test if the MARS1-dependent pathway is truly involved in chloroplast protein folding and aggregation stress (rather than oxidative stress), it would be beneficial to test the genetic interaction between the Clp system and MARS1 and monitor the accumulation of plastid chaperones such as ClpB3.

In addition to being recognized for their function in unfolding and delivering substrates for degradation, chloroplast ClpC homologs have also been reported to participate in chloroplast protein import (89, 90). This was suggested because (i) a small fraction of total chloroplast ClpC1 has consistently been identified in chloroplast inner envelope fractions (most is in the stroma), (ii) Clp mutants have reduced import efficiency, and (iii) ClpC homologs have an affinity to chloroplast transit peptides. ClpC has also been suggested to aid in import quality control, that is, removal of precursor proteins that maybe dysfunctional. The precise functions and mechanistic details of ClpC contributions at the chloroplast inner envelope are controversial, as debated in recent letters published in the *Plant Cell* (89, 90). Probably for historical reasons, the chloroplast import literature refers to the ClpC homologs often as HSP93-III (ClpC2) and HSP93-V (ClpC1), with the roman numerical referring to the *Arabidopsis* chromosome number. This does add to the confusion because this family of AAA+ chaperones is referred to in the broader literature as Clp chaperones.

### ClpS homologs and N-degrons across cyanobacteria, plastids, and apicoplasts

Very little is known about the degrons for the Clp systems in cyanobacteria, plastid, or apicoplasts. Recognizing and understanding degrons is key to unraveling the role of Clp systems in proteostasis. Even in the most studied Clp system (i.e., *E. coli*), understanding of substrate selection for degradation by Clp is far from complete. It has been demonstrated that substrate selection by the *E. coli* ClpXP system can involve the SsrA tag and the SsrB adaptor for protein nascent chains stalled on ribosomes, or a variety of C-terminal tags (10). The recognin/adaptor ClpS recognizes substrates based on degrons in the N-terminal region for delivery to ClpAP (e.g., in *E. coli*) or ClpCP (e.g., in *Mycobacterium tuberculosis*) (91, 92). Hence, ClpS fulfills a similar function as N-recognin E3 ligases for substrate selection and degradation by the proteasome (e.g., PRT1 and PRT6 in *Arabidopsis thaliana*). The ClpA/C/X chaperones also recognize disordered regions (10, 93) as substrates. Adding to the complexity of substrate recognition and selection is the notion that adaptor proteins not only bind and deliver substrates but also they can also modulate substrate selection of proteases, as exemplified by, for example, HspQ and ClpS, influencing the activity of both ClpAP and the Lon protease in *Salmonella enterica* (94). In this section, we review the current understanding of ClpS

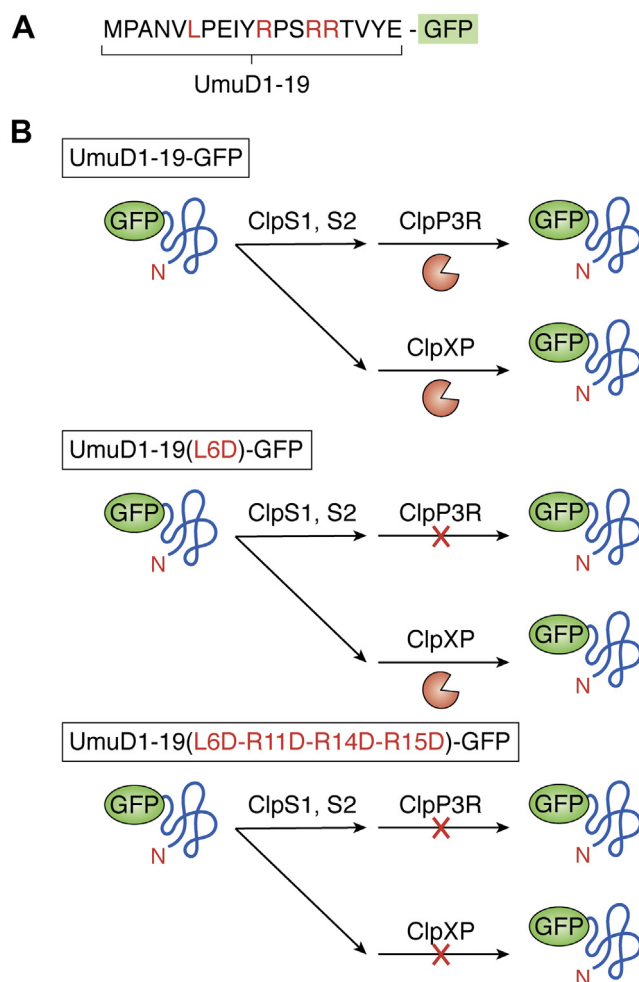
homologs found in cyanobacteria, plastids, and apicoplasts and insight in N-degrons.

### Cyanobacteria

Not much is known about how Clp-driven proteolysis pathways are regulated in cyanobacteria. *Synechococcus elongates* contains ClpS1 and ClpS2 that dynamically associate with ClpC (95). Using peptide affinity arrays and recombinant ClpS1 and ClpS2, a differential affinity for N-terminal residues was observed (96). ClpS1 showed an affinity for F, Y, and W, whereas ClpS2 showed an affinity for L, E, and Y as well as V and I. In all cases, R or K at the P2 position (the residue immediately downstream of the N-terminal residue) increased binding, whereas D or E at the P2 position decreased binding (96). These *in vitro* results support ClpS-dependent N-degron pathway in cyanobacteria; however, N-degron pathways remain to be demonstrated *in vivo*. A recent breakthrough in the discovery of degrons in cyanobacteria came from a study in *Synechocystis* sp. PCC6803 focused on *in vivo* degradation of UmuD involved in DNA damage repair. The UmuD homolog in *E. coli* was shown to be a substrate for both Clp and Lon, and degrons were located in the N-terminal region (97, 98) (Fig. 5). We note that, unlike *E. coli* and plastids in algae and plants, most members of cyanobacteria including, *Synechocystis* sp. PCC6803, lack Lon proteases (42). UmUD was degraded by ClpCP3R and required the presence of either ClpS1 or ClpS2, and in absence of both ClpS1 and ClpS2, UmUD was not degraded (42). Mutagenesis within the N-terminal region of UmuD (MPANVLPEIERPSRRTVYE) fused to GFP showed that the leucine residue was essential for ClpS-triggered degradation by ClpCP3R and that the arginine residues contributed to degradation by ClpXP1P2 (Fig. 5). *In vivo* N-degron substrates for *E. coli* ClpS were all found to be cleaved in the N-terminal region by unknown proteases, thus exposing the N-degron, and indicating that the ClpS-ClpAP pathway operated downstream of other proteases (99). A similar scenario may occur in *Synechocystis*, but despite significant efforts, the authors were unsuccessful to determine if leucine residue was the most N-terminal residue in the UmuD protein that ClpS recognized. Although this recent study (42) was a breakthrough in deciphering Clp degradation pathways in cyanobacteria, much remains to be done.

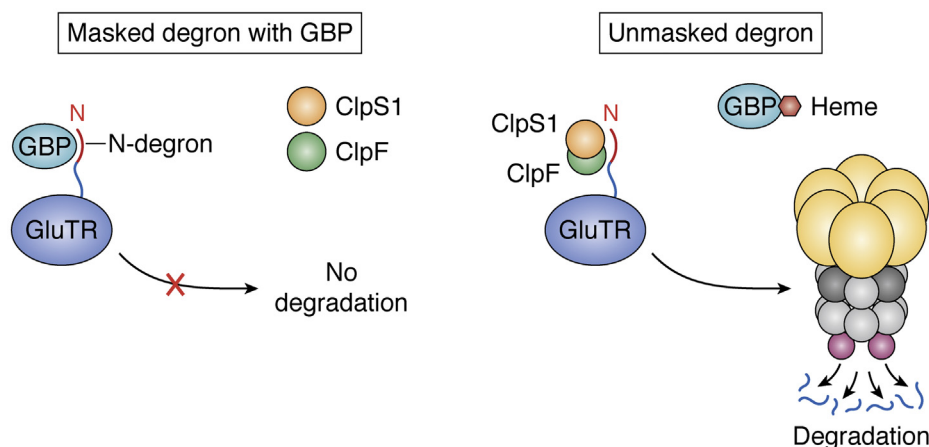
### Plastids

Phylogenetic analysis 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) (61, 100). Recently, we conducted *in vitro* *A. thaliana* ClpS1 affinity assays using eight N-degron protein reporters containing F, Y, L, W, I, or R in the N-terminal position and a positive residue (R) in the P2 position (101). Elution of the bound reporters with an FR dipeptide suggested that compared with *E. coli*, ClpS1 has a restricted N-degron specificity, recognizing proteins bearing an N-terminal F or W, only weakly recognizing L, but not recognizing Y or I. These



**Figure 5. Schematic representation of the dual degron of UmuD substrate of ClpCPR and ClpXP in cyanobacteria.** A, amino acid sequence of the first 19 residues of UmuD in *Synechocystis* sp. PCC6803 (42). N-terminal region ((MPANVLPEIERPSRRTVYE) of UmuD protein (1–19 AA) fused to GFP (UmuD1-19AA-GFP). B, the N-terminus leucine residue (L6) is mutated to aspartate (D) (UmuD1-19AA-GFP(L6D)). The N-terminus leucine residue (L6) and the arginine residues (R11, R14, and R15) are mutated to aspartate (UmuD1-19AA-GFP (L6DR11DR14–15D)). L6 is essential for ClpS triggered degradation by ClpCPR and arginine residues in the N-terminal region contributed to degradation by ClpXP1P2 (42).

findings could have been impacted by the choice of the dipeptide FR used to competitively elute the degron reporters. Our findings differed from a similar *in vitro* study (102), which suggested that *A. thaliana* ClpS1 affinity for F, Y, L, and W, with the lowest affinity for W. The 2.0 Å structure of *Arabidopsis* ClpS1 in the presence of a bound N-degron was recently determined by X-ray crystallography (103). Whereas many features of *Arabidopsis* ClpS1 are conserved, the structure showed a somewhat enlarged substrate binding pocket and reduced hydrophobic interactions with the ‘classical’ L N-degron. The hydrophobic residues in the pocket and the basic gatekeeper (R instead of M) at the entrance of the pocket control the N-degron selectivity of *Arabidopsis* ClpS1 (102, 103). *In vivo* testing of ClpS1 affinity for N-termini will be critical to truly assess ClpS1 substrate affinity and test the efficacy of ClpS delivery to ClpC. Large-scale *in vivo* chloroplast



**Figure 6. Schema illustrating how GluTR N-degrons can be masked and unmasked by the GluTR-binding protein (GBP).** In the absence of heme, GBP binds to the N-degron of GluTR protein and prevents its degradation by the Clp system (masked N-degron). When heme levels increase, the release of GBP from GluTR1 enables the binding of ClpS1 and ClpF adaptors to the latter and its concomitant proteolytic degradation (unmasked N-degron) (106).

N-terminome analysis showed that A, V, T, and S, and to a lesser extent G and M, were the most frequent N-terminal residues of the steady-state stromal proteome, whereas other residues were very infrequent or absent as N-terminal residues. This suggested that N-degron pathways do operate in the stroma, but the rules remain to be determined (104). Moreover, we showed that *A. thaliana* ClpS1 interacts with ClpF, a protein unique to higher plant plastids (43). ClpF is likely to be a coadaptor of ClpS1, together involved in substrate (such as GluTR) delivery to the ClpC chaperones (22).

GluTR was identified as a direct interactor of ClpS1 using a combination of *in vitro* trapping and affinity purification with ClpS1 and mutated ClpS1 carrying a mutation in the ClpS1 substrate binding site (61). The interacting domains between ClpC1, ClpS1, ClpF, and GluTR were then mapped through *in vitro* biochemical studies using recombinant proteins, but there was no obvious degron in GluTR (43). Follow-up studies showed that GluTR degradation is delayed in Clp mutants and that expression of a truncated GluTR protein lacking the N-terminal region remained in the stroma (rather than associating with the thylakoid) and was more stable during prolonged darkness (105) (Fig. 6). Consistent with the role of the N-terminal domain as a degron, the rate of GluTR proteolysis in the dark increased in the absence of the GluTR-binding protein (GBP) but decreased in mutants impaired in ClpC1 or the proteolytic core subunit ClpR2 (105). Although GluTR accumulates at higher levels in ClpC1 and ClpS1 mutants, the degradation rate of the enzyme in the dark appears not to be altered when ClpS1 is missing. Accumulation of heme, induced by feeding with 5-aminolevulinic acid, stimulated Clp-protease-dependent degradation of *Arabidopsis* GluTR1. Binding of heme to the GBP inhibited the interaction of GBP with the N-terminal regulatory domain of GluTR1, thus making it accessible to the Clp protease (Fig. 6). These results show a functional link between the heme content and the post-translational control of GluTR stability by the Clp system, which helps ensure adequate availability of chlorophyll and heme (106).

### Apicoplasts

The structure of the apicoplast ClpS homolog has been determined by X-ray crystallography (59). *In vitro* ClpS affinity assays suggested that N-degrons included not only classic residues such as F, W, Y, L but also I (59, 60). Isoleucine is not a typical N-degron in other bacteria, which indicates that apicoplast ClpS exhibits an expanded substrate specificity. Despite this progress in identifying ClpS N-degrons *in vitro*, physiological substrates and *in vivo* demonstration of N-degron pathways apicoplasts have not yet been shown. The ClpS interactors identified using affinity purification screens discussed earlier (53) need to be further evaluated for their degrons and affinity to ClpS.

### Other substrate adaptors and delivery systems

Cyanobacteria contain PBSs that are large light-harvesting complexes that capture and funnel light energy to the photosystems. These complexes are degraded during nutrient deprivation leading to chlorosis (62). Degradation of PBS complexes is specifically triggered by proteolysis adaptors called nonbleaching proteins A and B (NblA and NblB) that act directly on the PBS (38, 62, 63, 107, 108). NblA is a small proteolysis adaptor of about 60 amino acids and its helix-loop-helix structure mimics fragments of the common PBP structure (41). This protein facilitates the interaction between ClpC and PBS. In fact, NblA interacts with PBS rods (109),  $\alpha$ - and  $\beta$ -core subunits (41, 62), and ClpC (38). NblA can thereby induce degradation of both the PBS rod and core (39, 64). NblA expression is adjusted by a response regulator NblR and a sensor NblS (110). NblB is distantly related to the CpcE subunit of E/F-type lyases (108) and is implicated in the degradation of the core component allophycocyanin in the presence of NblA (63, 111).

In chloroplasts, HSP70 and the J20 cochaperone At4g13830 are involved as ClpC chaperone partners for the removal of aggregated and misfolded proteins (112–115). In particular, the nonfunctional and aggregated forms of the chloroplast

isoprenoid pathway enzyme (DXS) are delivered to Hsp70 by J20 that thus functions as an Hsp70 adaptor. HSP70 then delivers these nonfunctional DXS proteins to ClpC for degradation by the ClpPR protease or delivers them to the unfoldase ClpB3 for disaggregation (52). Indeed, mutants defective in J20 activity accumulated increased levels of DXS protein (but not transcripts) and displayed reduced levels of DXS enzyme activity, indicating that loss of J20 function causes post-transcriptional accumulation of DXS in an inactive form. This J20 protein is part of a larger family of J-domain containing proteins, also named J-proteins or DnaJ-like proteins, that can be divided in subfamilies based on the presence and absence of specific domains (116, 117). J-domain-containing proteins have been recognized as cochaperones for HSP proteins in bacteria (118), *Arabidopsis* (119, 120), yeast (121), and *Chlamydomonas* (122). Early on, J-domain proteins have also been recognized as important regulators for plastid biogenesis and light responses (123–125) and received new attention in recent years (116, 117). It is possible that in addition to the J20 protein At4g13830, other J-domain (like) proteins also could participate directly or indirectly in substrate recognition and delivery of plastid proteins for degradation.

Substrate adaptors have been identified for a variety of protease systems in both prokaryotes and eukaryotes. This includes E3-ligase adaptors and F-box proteins in the ubiquitin-proteasome degradation system in *Arabidopsis*, broadening the substrate selection capacity for cullin E3 ligases (126–128). The adaptors for these E3 ligases thus further enhance the capacity to select specific proteins for degradation. Similarly, diverse adaptors have been identified for Clp systems in bacteria such as RcdA, CpdR, and popA in *Caulobacter crescentus* or SsrB in *E. coli*, as reviewed (35). We, therefore, speculate that, in addition to ClpS1 and ClpF, there could be additional adaptor proteins for the Clp system in chloroplasts. We suggest that there might be novel adaptors among the proteins specifically captured by the *in vivo* ClpC1 Walker B domain trap lines in *Arabidopsis* (29). However, further experimentation will be needed to test this hypothesis.

Finally, another mode to control protein degradation is through proteins that either directly protect substrates from targeting by adaptors or by proteins that modify the substrate selection by a protease. Both cases have been identified for the Clp system in nonphotosynthetic bacteria. For example, in *Salmonella enterica*, MgtC protects PhoP from degradation by the Clp system by outcompeting ClpS for binding to PhoP (129), whereas acetylated HspQ inhibits ClpS substrate selection by direct binding ClpS (94). Both MgtC and acetylated HspQ thus function as antiadaptors in the Clp system. It is quite striking that chloroplast ClpF also has an YccV domain similar as HspQ, perhaps suggesting that it can modify protease substrate selectivity and/or affect ClpS substrate selection. Future studies will be required to test these ClpF functions and antiadaptors and other protease specificity modulators in chloroplasts.

## Is the relationship between substrates and Clp conserved across lineages?

There are several specific cases of conserved substrates across diverse lineages, and these could be useful to explore to understand principles of substrate selection and degrons. Key examples are GluTR homologs involved in tetrapyrrole biosynthesis (for heme and chlorophyll) observed as Clp substrate in *Salmonella typhimurium* (130, 131) and chloroplasts, and UmuD involved in DNA repair in *E. coli* and cyanobacteria (42, 98). Systematic comparison across diverse lineages for these and other Clp substrates could help understand evolutionary adaptation in proteostasis.

## Conclusions and outlook

Clp systems are central in proteostasis in cyanobacteria, plastids, and apicoplasts. The basic components of these Clp machineries have now been established and several tools and techniques are available to better understand substrate degrons as well as recognition and delivery of substrates by known and unknown adaptors. In particular, *in vivo* substrate trapping techniques (11) under different abiotic conditions, developmental stages, and in loss-of-function mutants in genes that affect biogenesis or central plastid/proplastid functions, could provide a better understanding of the Clp substrate pool and the relevant degrons. Deliberate experimental and *in silico* strategies should be used to search and test for additional candidate Clp substrate adaptors.

Recent developments in CryoEM technologies and the establishment of many CryoEM facilities across universities and research institutions allow for structural determination of Clp complex assemblies, possibly together with stabilized and trapped substrates, as has been demonstrated in the last 2 years for several nonphotosynthetic bacterial systems (32, 33, 132, 133). Such structural information for plastid, apicoplast, and cyanobacterial Clp systems will be needed to truly understand mechanisms of substrate recognition, unfolding, and degradation.

Cyanobacteria, plastids, and apicoplasts require the activity of processing peptidases, proteases, and amino-peptidases (cleaving just one, two, or three amino acids) for a broad range of activities, including (i) removal of presequences of nuclear-encoded plastid and apicoplast proteins, (ii) N-terminal methionine removal of newly synthesized proteins, (iii) additional N- or C-terminal cleavages for maturation, stabilization, and possibly activation of proteins, (iv) removal of misfolded, damaged, or aggregated proteins, and (v) removal of unwanted proteins in response to environmental or developmental transitions. In particular, plastids are highly dynamic as they can undergo specialization during developmental transitions (*e.g.*, from etioplast to chloroplast, or from chloroplast to chromoplast). Therefore, coordination between the Clp protease systems and other proteases localized in the same sub-cellular compartments will be important and should be explored. A large-scale MS analysis of protein affinity enriched using a ClpS affinity column in *E. coli* suggested widespread

proteolytic processing to generate N-end rule substrates for ClpS. Many of these affinity enriched substrates had also been modified by Aat (an amino transferase adding leucine or phenylalanine to proteins with N-terminal lysine or arginine) (99). This suggests that the N-degron pathway for the Clp system in *E. coli* operates mostly downstream of other proteases. Experimental tools such as comparative and quantitative peptidomics and proteomics in single and higher order protease mutant backgrounds are now sufficiently developed such that they can be effectively used to begin unraveling these protease networks. Finally, the adaptor ClpS (and possibly other adaptors) can be modified and tuned to deliberately target other proteins; this provides exciting opportunities for biotechnology and synthetic biology (134, 135).

**Author contributions**—I. B. and K. J. V. W. wrote the paper together.

**Funding and additional information**—This study was supported by funding from the National Science Foundation through awards #1834636 and #1940961.

**Conflict of interest**—The authors declare that they have no conflicts of interest with the contents of this article.

**Abbreviations**—The abbreviations used are: AAA+, ATPases Associated with diverse cellular Activities; Clp, caseinolytic protease; cpUPR, chloroplast Protein Unfolding Response; DXS, deoxyxylulose 5-phosphate synthase; GBP, GluTR-binding protein; GluTR, glutamyl t-RNA reductase; GUN, Genomes Uncoupled 1; IPP, isopentenyl pyrophosphate; PTM, post-translational modification; UPR, unfolded protein response.

## References

- Dohmen, R. J., Wu, P., and Varshavsky, A. (1994) Heat-inducible degron: A method for constructing temperature-sensitive mutants. *Science* **263**, 1273–1276
- Bachmair, A., Finley, D., and Varshavsky, A. (1986) *In vivo* half-life of a protein is a function of its amino-terminal residue. *Science* **234**, 179–186
- Varshavsky, A. (2019) N-degron and C-degron pathways of protein degradation. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 358–366
- Bouchnak, I., and van Wijk, K. J. (2019) N-degron pathways in plastids. *Trends Plant Sci.* **24**, 917–926
- Rawlings, N. D., Barrett, A. J., and Finn, R. (2016) Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* **44**, D343–D350
- Seraphim, T. V., and Houry, W. A. (2020) AAA+ proteins. *Curr. Biol.* **30**, R251–R257
- Puchades, C., Sandate, C. R., and Lander, G. C. (2020) The molecular principles governing the activity and functional diversity of AAA+ proteins. *Nat. Rev. Mol. Cell Biol.* **21**, 43–58
- Marshall, R. S., and Vierstra, R. D. (2019) Dynamic regulation of the 26S proteasome: From synthesis to degradation. *Front. Mol. Biosci.* **6**, 40
- Mahmoud, S. A., and Chien, P. (2018) Regulated proteolysis in bacteria. *Annu. Rev. Biochem.* **87**, 677–696
- Olivares, A. O., Baker, T. A., and Sauer, R. T. (2018) Mechanical protein unfolding and degradation. *Annu. Rev. Physiol.* **80**, 413–429
- Rei Liao, J. Y., and van Wijk, K. J. (2019) Discovery of AAA+ protease substrates through trapping approaches. *Trends Biochem. Sci.* **44**, 528–545
- Nonoyama, T., Kazamia, E., Nawaly, H., Gao, X., Tsuji, Y., Matsuda, Y., Bowler, C., Tanaka, T., and Dorrell, R. G. (2019) Metabolic innovations underpinning the origin and diversification of the diatom chloroplast. *Biomolecules* **9**, 322
- Jarvis, P., and Lopez-Juez, E. (2013) Biogenesis and homeostasis of chloroplasts and other plastids. *Nat. Rev. Mol. Cell Biol.* **14**, 787–802
- van Wijk, K. J., and Kessler, F. (2017) Plastoglobuli: Plastid micro-compartments with integrated functions in metabolism, plastid developmental transitions, and environmental adaptation. *Annu. Rev. Plant Biol.* **68**, 253–289
- van Wijk, K. J. (2015) Protein maturation and proteolysis in plant plastids, mitochondria, and peroxisomes. *Annu. Rev. Plant Biol.* **66**, 75–111
- Nishimura, K., Kato, Y., and Sakamoto, W. (2017) Essentials of proteolytic machineries in chloroplasts. *Mol. Plant* **10**, 4–19
- Izumi, M., and Nakamura, S. (2018) Chloroplast protein turnover: The influence of extraplastidic processes, including autophagy. *Int. J. Mol. Sci.* **19**, 828
- van Dooren, G. G., and Striepen, B. (2013) The algal past and parasite present of the apicoplast. *Annu. Rev. Microbiol.* **67**, 271–289
- Nishimura, K., Kato, Y., and Sakamoto, W. (2016) Chloroplast proteases: Updates on proteolysis within and across suborganellar compartments. *Plant Physiol.* **171**, 2280–2293
- Rodriguez-Concepcion, M., D'Andrea, L., and Pulido, P. (2019) Control of plastidial metabolism by the Clp protease complex. *J. Exp. Bot.* **70**, 2049–2058
- 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
- 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
- Cheregi, O., Wagner, R., and Funk, C. (2016) Insights into the cyanobacterial Deg/HtrA proteases. *Front. Plant Sci.* **7**, 694
- Rigas, S., Daras, G., Tsitsekian, D., Alatzas, A., and Hatzopoulos, P. (2014) Evolution and significance of the Lon gene family in Arabidopsis organelle biogenesis and energy metabolism. *Front. Plant Sci.* **5**, 145
- Majsec, K., Bhuiyan, N. H., Sun, Q., Kumari, S., Kumar, V., Ware, D., and van Wijk, K. J. (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
- Ferrall-Fairbanks, M. C., Kieslich, C. A., and Platt, M. O. (2020) Reassessing enzyme kinetics: Considering protease-as-substrate interactions in proteolytic networks. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 3307–3318
- Margulis, B., Tsimokha, A., Zubova, S., and Guzhova, I. (2020) Molecular chaperones and proteolytic machineries regulate protein homeostasis in aging cells. *Cells* **9**, 1308
- Yu, A. Y., and Houry, W. A. (2007) ClpP: A distinctive family of cylindrical energy-dependent serine proteases. *FEBS Lett.* **581**, 3749–3757
- Montandon, C., Friso, G., Liao, J. R., Choi, J., and van Wijk, K. J. (2019) *In vivo* trapping of proteins interacting with the chloroplast CLPC1 chaperone; potential substrates and adaptors. *J. Proteome Res.* **18**, 2585–2600
- Carroni, M., Franke, K. B., Maurer, M., Jager, J., Hantke, I., Gloge, F., Linder, D., Gremer, S., Turgay, K., Bukau, B., and Mogk, A. (2017) Regulatory coiled-coil domains promote head-to-head assemblies of AAA+ chaperones essential for tunable activity control. *Elife* **6**, e30120
- Liu, K., Ologbenla, A., and Houry, W. A. (2014) Dynamics of the ClpP serine protease: A model for self-compartmentalized proteases. *Crit. Rev. Biochem. Mol. Biol.* **49**, 400–412
- Lopez, K. E., Rizo, A. N., Tse, E., Lin, J., Scull, N. W., Thwin, A. C., Lucius, A. L., Shorter, J., and Southworth, D. R. (2020) Conformational plasticity of the ClpAP AAA+ protease couples protein unfolding and proteolysis. *Nat. Struct. Mol. Biol.* **27**, 406–416
- Gatsogiannis, C., Balogh, D., Merino, F., Sieber, S. A., and Raunser, S. (2019) Cryo-EM structure of the ClpXP protein degradation machinery. *Nat. Struct. Mol. Biol.* **26**, 946–954
- Ripstein, Z. A., Vahidi, S., Houry, W. A., Rubinstein, J. L., and Kay, L. E. (2020) A processive rotary mechanism couples substrate unfolding and proteolysis in the ClpXP degradation machinery. *Elife* **9**, e52158

35. Kuhlmann, N. J., and Chien, P. (2017) Selective adaptor dependent protein degradation in bacteria. *Curr. Opin. Microbiol.* **36**, 118–127
36. Cohen, S. E., McKnight, B. M., and Golden, S. S. (2018) Roles for ClpXP in regulating the circadian clock in *Synechococcus elongatus*. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E7805–E7813
37. Vijayan, V., Jain, I. H., and O’Shea, E. K. (2011) A high resolution map of a cyanobacterial transcriptome. *Genome Biol.* **12**, R47
38. Baier, A., Winkler, W., Korte, T., Lockau, W., and Karradt, A. (2014) Degradation of phycobilisomes in *Synechocystis* sp. PCC6803: Evidence for essential formation of an NblA1/NblA2 heterodimer and its codegradation by a Clp protease complex. *J. Biol. Chem.* **289**, 11755–11766
39. Sendersky, E., Kozer, N., Levi, M., Garini, Y., Shav-Tal, Y., and Schwarz, R. (2014) The proteolysis adaptor, NblA, initiates protein pigment degradation by interacting with the cyanobacterial light-harvesting complexes. *Plant J.* **79**, 118–126
40. Bienert, R., Baier, K., Volkmer, R., Lockau, W., and Heinemann, U. (2006) Crystal structure of NblA from *Anabaena* sp. PCC 7120, a small protein playing a key role in phycobilisome degradation. *J. Biol. Chem.* **281**, 5216–5223
41. Dines, M., Sendersky, E., David, L., Schwarz, R., and Adir, N. (2008) Structural, functional, and mutational analysis of the NblA protein provides insight into possible modes of interaction with the phycobilisome. *J. Biol. Chem.* **283**, 30330–30340
42. Jin, H., Kim, R., and Bhaya, D. (2020) Deciphering proteolysis pathways for the error-prone DNA polymerase in cyanobacteria. *Environ. Microbiol.* <https://doi.org/10.1111/1462-2920.14911>
43. Nishimura, K., Apitz, J., Friso, G., Kim, J., Ponnala, L., Grimm, B., and van Wijk, K. J. (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
44. Williams, A. M., Friso, G., van Wijk, K. J., and Sloan, D. B. (2019) Extreme variation in rates of evolution in the plastid Clp protease complex. *Plant J.* **98**, 243–259
45. Liao, J.-Y. R., Friso, G., Kim, J., and van Wijk, K. J. (2018) Consequences of the loss of catalytic triads in chloroplast CLPPR protease core complexes *in vivo*. *Plant Direct* **2**, 1–10
46. Miller, J. M., Chaudhary, H., and Marsee, J. D. (2018) Phylogenetic analysis predicts structural divergence for proteobacterial ClpC proteins. *J. Struct. Biol.* **201**, 52–62
47. Tryggvesson, A., Stahlberg, F. M., Mogk, A., Zeth, K., and Clarke, A. K. (2012) Interaction specificity between the chaperone and proteolytic components of the cyanobacterial Clp protease. *Biochem. J.* **446**, 311–320
48. Mohapatra, C., Kumar Jagdev, M., and Vasudevan, D. (2017) Crystal structures reveal N-terminal domain of *Arabidopsis thaliana* ClpD to be highly divergent from that of ClpC1. *Sci. Rep.* **7**, 44366
49. Lee, U., Rioflorida, I., Hong, S. W., Larkindale, J., Waters, E. R., and Vierling, E. (2007) The *Arabidopsis* ClpB/Hsp100 family of proteins: Chaperones for stress and chloroplast development. *Plant J.* **49**, 115–127
50. Myounga, F., Motohashi, R., Kuromori, T., Nagata, N., and Shinozaki, K. (2006) An *Arabidopsis* chloroplast-targeted Hsp101 homologue, APG6, has an essential role in chloroplast development as well as heat-stress response. *Plant J.* **48**, 249–260
51. Avellaneda, M. J., Franke, K. B., Sunderlikova, V., Bukau, B., Mogk, A., and Tans, S. J. (2020) Processive extrusion of polypeptide loops by a Hsp100 disaggregase. *Nature* **578**, 317–320
52. Parcerisa, I. L., Rosano, G. L., and Ceccarelli, E. A. (2020) Biochemical characterization of ClpB3, a chloroplasmic disaggregase from *Arabidopsis thaliana*. *Plant Mol. Biol.* **104**, 451–465
53. Florentin, A., Stephens, D. R., Brooks, C. F., Baptista, R. P., and Muralidharan, V. (2020) Plastid biogenesis in malaria parasites requires the interactions and catalytic activity of the Clp proteolytic system. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 13719–13729
54. Swift, R. P., Rajaram, K., Liu, H. B., Dziedzic, A., Jedlicka, A. E., Roberts, A. D., Matthews, K. A., Jhun, H., Bumpus, N. N., Tewari, S. G., Wallqvist, A., and Prigge, S. T. (2020) A mevalonate bypass system facilitates elucidation of plastid biology in malaria parasites. *PLoS Pathog.* **16**, e1008316
55. Florentin, A., Cobb, D. W., Fishburn, J. D., Cipriano, M. J., Kim, P. S., Fierro, M. A., Striepen, B., and Muralidharan, V. (2017) PfClpC is an essential Clp chaperone required for plastid integrity and Clp protease stability in *Plasmodium falciparum*. *Cell Rep.* **21**, 1746–1756
56. El Bakkouri, M., Pow, A., Mulichak, A., Cheung, K. L., Artz, J. D., Amani, M., Fell, S., de Koning-Ward, T. F., Goodman, C. D., McFadden, G. I., Ortega, J., Hui, R., and Houry, W. A. (2010) The Clp chaperones and proteases of the human malaria parasite *Plasmodium falciparum*. *J. Mol. Biol.* **404**, 456–477
57. El Bakkouri, M., Rathore, S., Calmettes, C., Wernimont, A. K., Liu, K., Sinha, D., Asad, M., Jung, P., Hui, R., Mohammed, A., and Houry, W. A. (2013) Structural insights into the inactive subunit of the apicoplast-localized caseinolytic protease complex of *Plasmodium falciparum*. *J. Biol. Chem.* **288**, 1022–1031
58. Olinares, P. D., Kim, J., Davis, J. I., and van Wijk, K. J. (2011) Subunit stoichiometry, evolution, and functional implications of an asymmetric plant plastid ClpP/R protease complex in *Arabidopsis*. *Plant Cell* **23**, 2348–2361
59. AhYoung, A. P., Koehl, A., Vizcarra, C. L., Cascio, D., and Egea, P. F. (2016) Structure of a putative ClpS N-end rule adaptor protein from the malaria pathogen *Plasmodium falciparum*. *Protein Sci.* **25**, 689–701
60. Tan, J. L., Ward, L., Truscott, K. N., and Dougan, D. A. (2016) The N-end rule adaptor protein ClpS from *Plasmodium falciparum* exhibits broad substrate specificity. *FEBS Lett.* **590**, 3397–3406
61. Nishimura, K., Asakura, Y., Friso, G., Kim, J., Oh, S. H., Rutschow, H., Ponnala, L., and van Wijk, K. J. (2013) ClpS1 is a conserved substrate selector for the chloroplast Clp protease system in *Arabidopsis*. *Plant Cell* **25**, 2276–2301
62. Nguyen, A. Y., Bricker, W. P., Zhang, H., Weisz, D. A., Gross, M. L., and Pakrasi, H. B. (2017) The proteolysis adaptor, NblA, binds to the N-terminus of beta-phycoerythrin: Implications for the mechanism of phycobilisome degradation. *Photosynth. Res.* **132**, 95–106
63. Hu, P. P., Hou, J. Y., Xu, Y. L., Niu, N. N., Zhao, C., Lu, L., Zhou, M., Scheer, H., and Zhao, K. H. (2020) The role of lyases, NblA and NblB proteins and bilin chromophore transfer in restructuring the cyanobacterial light-harvesting complex(double dagger). *Plant J.* **102**, 529–540
64. Sendersky, E., Kozer, N., Levi, M., Moizik, M., Garini, Y., Shav-Tal, Y., and Schwarz, R. (2015) The proteolysis adaptor, NblA, is essential for degradation of the core pigment of the cyanobacterial light-harvesting complex. *Plant J.* **83**, 845–852
65. Camberg, J. L., Hoskins, J. R., and Wickner, S. (2009) ClpXP protease degrades the cytoskeletal protein, FtsZ, and modulates FtsZ polymer dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 10614–10619
66. D’Andrea, L., Simon-Moya, M., Llorente, B., Llamas, E., Marro, M., Loza-Alvarez, P., Li, L., and Rodriguez-Concepcion, M. (2018) Interference with Clp protease impairs carotenoid accumulation during tomato fruit ripening. *J. Exp. Bot.* **69**, 1557–1568
67. Moreno, J. C., Martinez-Jaime, S., Schwartzmann, J., Karcher, D., Tillich, M., Graf, A., and Bock, R. (2018) Temporal proteomics of inducible RNAi lines of Clp protease subunits identifies putative protease substrates. *Plant Physiol.* **176**, 1485–1508
68. Ali, M. S., and Baek, K. H. (2020) Co-suppression of NbClpC1 and NbClpC2, encoding Clp protease chaperons, elicits significant changes in the metabolic profile of *Nicotiana benthamiana*. *Plants (Basel)* **9**, 259
69. Ali, M. S., and Baek, K. H. (2019) Co-suppression of NbClpC1 and NbClpC2 alters plant morphology with changed hormone levels in *Nicotiana benthamiana*. *Plant Cell Rep.* **38**, 1317–1328
70. Moehle, E. A., Shen, K., and Dillin, A. (2019) Mitochondrial proteostasis in the context of cellular and organismal health and aging. *J. Biol. Chem.* **294**, 5396–5407
71. Lam, I., Hallaceli, E., and Khurana, V. (2020) Proteome-scale mapping of perturbed proteostasis in living cells. *Cold Spring Harb. Perspect. Biol.* **12**, a034124
72. Dogra, V., and Kim, C. (2019) Chloroplast protein homeostasis is coupled with retrograde signaling. *Plant Signal. Behav.* **14**, 1656037

73. Hetz, C., Zhang, K., and Kaufman, R. J. (2020) Mechanisms, regulation and functions of the unfolded protein response. *Nat. Rev. Mol. Cell Biol.* **21**, 421–438
74. Rajaram, H., Chaurasia, A. K., and Apte, S. K. (2014) Cyanobacterial heat-shock response: Role and regulation of molecular chaperones. *Microbiology (Reading)* **160**, 647–658
75. Sinetova, M. A., and Los, D. A. (2016) Systemic analysis of stress transcriptomics of *Synechocystis* reveals common stress genes and their universal triggers. *Mol. Biosyst.* **12**, 3254–3258
76. Rudella, A., Friso, G., Alonso, J. M., Ecker, J. R., and van Wijk, K. J. (2006) Downregulation of ClpR2 leads to reduced accumulation of the ClpPRS protease complex and defects in chloroplast biogenesis in *Arabidopsis*. *Plant Cell* **18**, 1704–1721
77. Llamas, E., Pulido, P., and Rodriguez-Concepcion, M. (2017) Interference with plastome gene expression and Clp protease activity in *Arabidopsis* triggers a chloroplast unfolded protein response to restore protein homeostasis. *PLoS Genet.* **13**, e1007022
78. Wu, G. Z., Chalvin, C., Hoelscher, M., Meyer, E. H., Wu, X. N., and Bock, R. (2018) Control of retrograde signaling by rapid turnover of GENOMES UNCOUPLED1. *Plant Physiol.* **176**, 2472–2495
79. Weits, D. A., van Dongen, J. T., and Licausi, F. (2021) Molecular oxygen as a signaling component in plant development. *New Phytol.* **229**, 24–35
80. Ramundo, S., Casero, D., Muhlhaus, T., Hemme, D., Sommer, F., Crevecoeur, M., Rahire, M., Schroda, M., Rusch, J., Goodenough, U., Pellegrini, M., Perez-Perez, M. E., Crespo, J. L., Schaad, O., Civic, N., et al. (2014) Conditional depletion of the *Chlamydomonas* chloroplast ClpP protease activates nuclear genes involved in autophagy and plastid protein quality control. *Plant Cell* **26**, 2201–2222
81. Perlaza, K., Toutkoushian, H., Boone, M., Lam, M., Iwai, M., Jonikas, M. C., Walter, P., and Ramundo, S. (2019) The Mars1 kinase confers photoprotection through signaling in the chloroplast unfolded protein response. *Elife* **8**, e49577
82. Theis, J., Niemeyer, J., Schmollinger, S., Ries, F., Rutgers, M., Gupta, T. K., Sommer, F., Muranaka, L. S., Venn, B., Schulz-Raffelt, M., Willmund, F., Engel, B. D., and Schroda, M. (2020) VIPP2 interacts with VIPP1 and HSP22E/F at chloroplast membranes and modulates a retrograde signal for HSP22E/F gene expression. *Plant Cell Environ.* **43**, 1212–1229
83. Taylor, B. L., and Zhulin, I. B. (1999) PAS domains: Internal sensors of oxygen, redox potential, and light. *Microbiol. Mol. Biol. Rev.* **63**, 479–506
84. Munoz, P., and Munne-Bosch, S. (2020) Oxylipins in plastidial retrograde signaling. *Redox Biol.* **37**, 101717
85. D'Alessandro, S., and Havaux, M. (2019) Sensing beta-carotene oxidation in photosystem II to master plant stress tolerance. *New Phytol.* **223**, 1776–1783
86. Kleine, T., and Leister, D. (2016) Retrograde signaling: Organelles go networking. *Biochim. Biophys. Acta* **1857**, 1313–1325
87. Chan, K. X., Phua, S. Y., Crisp, P., McQuinn, R., and Pogson, B. J. (2016) Learning the languages of the chloroplast: Retrograde signaling and beyond. *Annu. Rev. Plant Biol.* **67**, 25–53
88. Rea, G., Antonacci, A., Lambrevia, M. D., and Mattoo, A. K. (2018) Features of cues and processes during chloroplast-mediated retrograde signaling in the alga *Chlamydomonas*. *Plant Sci.* **272**, 193–206
89. Li, H. M., Schnell, D., and Theg, S. M. (2020) Protein import motors in chloroplasts: On the role of chaperones. *Plant Cell* **32**, 536–542
90. Nakai, M. (2020) Reply: The revised model for chloroplast protein import. *Plant Cell* **32**, 543–546
91. Gao, X., Yeom, J., and Groisman, E. A. (2019) The expanded specificity and physiological role of a widespread N-degron recognin. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 18629–18637
92. Dougan, D. A., Truscott, K. N., and Zeth, K. (2010) The bacterial N-end rule pathway: Expect the unexpected. *Mol. Microbiol.* **76**, 545–558
93. Lunge, A., Gupta, R., Choudhary, E., and Agarwal, N. (2020) The unfoldase ClpC1 of *Mycobacterium tuberculosis* regulates the expression of a distinct subset of proteins having intrinsically disordered termini. *J. Biol. Chem.* **295**, 9455–9473
94. Yeom, J., and Groisman, E. A. (2019) Activator of one protease transforms into inhibitor of another in response to nutritional signals. *Genes Dev.* **33**, 1280–1292
95. Stanne, T. M., Pojidaeva, E., Andersson, F. I., and Clarke, A. K. (2007) Distinctive types of ATP-dependent Clp proteases in cyanobacteria. *J. Biol. Chem.* **282**, 14394–14402
96. Stein, B. J., Grant, R. A., Sauer, R. T., and Baker, T. A. (2016) Structural basis of an N-degron adaptor with more stringent specificity. *Structure* **24**, 232–242
97. Smith, C. K., Baker, T. A., and Sauer, R. T. (1999) Lon and Clp family proteases and chaperones share homologous substrate-recognition domains. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 6678–6682
98. Neher, S. B., Sauer, R. T., and Baker, T. A. (2003) Distinct peptide signals in the UmuD and UmuD' subunits of UmuD/D' mediate tethering and substrate processing by the ClpXP protease. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 13219–13224
99. Humbard, M. A., Surkov, S., De Donatis, G. M., Jenkins, L. M., and Maurizi, M. R. (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
100. 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
101. Montandon, C., Dougan, D. A., and van Wijk, K. J. (2019) N-degron specificity of chloroplast ClpS1 in plants. *FEBS Lett.* **593**, 962–970
102. Colombo, C. V., Rosano, G. L., Mogk, A., and Ceccarelli, E. A. (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
103. Kim, L., Heo, J., Kwon, D. H., Shin, J. S., Jang, S. H., Park, Z. Y., and Song, H. K. (2021) Structural basis for the N-degron specificity of ClpS1 from *Arabidopsis thaliana*. *Protein Sci.* **30**, 700–708
104. Rowland, E., Kim, J., Bhuiyan, N. H., and van Wijk, K. J. (2015) The *Arabidopsis* chloroplast stromal N-terminome: Complexities of amino-terminal protein maturation and stability. *Plant Physiol.* **169**, 1881–1896
105. Apitz, J., Nishimura, K., Schmied, J., Wolf, A., Hedtke, B., van Wijk, K. J., and Grimm, B. (2016) Posttranslational control of ALA synthesis includes GluTR degradation by Clp protease and stabilization by GluTR-binding protein. *Plant Physiol.* **170**, 2040–2051
106. Richter, A. S., Banse, C., and Grimm, B. (2019) The GluTR-binding protein is the heme-binding factor for feedback control of glutamyl-tRNA reductase. *Elife* **8**, e46300
107. Collier, J. L., and Grossman, A. R. (1994) A small polypeptide triggers complete degradation of light-harvesting phycobiliproteins in nutrient-deprived cyanobacteria. *EMBO J.* **13**, 1039–1047
108. Dolganov, N., and Grossman, A. R. (1999) A polypeptide with similarity to phycocyanin alpha-subunit phycocyanobilin lyase involved in degradation of phycobilisomes. *J. Bacteriol.* **181**, 610–617
109. Luque, I., Ochoa De Alda, J. A., Richaud, C., Zabulon, G., Thomas, J. C., and Houmard, J. (2003) The NblAI protein from the filamentous cyanobacterium *Tolypothrix PCC 7601*: Regulation of its expression and interactions with phycobilisome components. *Mol. Microbiol.* **50**, 1043–1054
110. van Waasbergen, L. G., Dolganov, N., and Grossman, A. R. (2002) nblS, a gene involved in controlling photosynthesis-related gene expression during high light and nutrient stress in *Synechococcus elongatus PCC 7942*. *J. Bacteriol.* **184**, 2481–2490
111. Levi, M., Sendersky, E., and Schwarz, R. (2018) Decomposition of cyanobacterial light harvesting complexes: NblA-dependent role of the bilin lyase homolog NblB. *Plant J.* **94**, 813–821
112. Pulido, P., Llamas, E., and Rodriguez-Concepcion, M. (2017) Both Hsp70 chaperone and Clp protease plastidial systems are required for protection against oxidative stress. *Plant Signal. Behav.* **12**, e1290039
113. Pulido, P., Llamas, E., Llorente, B., Ventura, S., Wright, L. P., and Rodriguez-Concepcion, M. (2016) Specific Hsp100 chaperones determine the fate of the first enzyme of the plastidial isoprenoid pathway for either refolding or degradation by the stromal Clp protease in *Arabidopsis*. *PLoS Genet.* **12**, e1005824
114. D'Andrea, L., and Rodriguez-Concepcion, M. (2019) Manipulation of plastidial protein quality control components as a new strategy to improve carotenoid contents in tomato fruit. *Front. Plant Sci.* **10**, 1071

115. Pulido, P., Toledo-Ortiz, G., Phillips, M. A., Wright, L. P., and Rodriguez-Concepcion, M. (2013) Arabidopsis J-protein J20 delivers the first enzyme of the plastidial isoprenoid pathway to protein quality control. *Plant Cell* **25**, 4183–4194
116. Pulido, P., and Leister, D. (2018) Novel DnaJ-related proteins in Arabidopsis thaliana. *New Phytol.* **217**, 480–490
117. Zhang, B., Qiu, H. L., Qu, D. H., Ruan, Y., and Chen, D. H. (2018) Phylogeny-dominant classification of J-proteins in Arabidopsis thaliana and Brassica oleracea. *Genome* **61**, 405–415
118. Hennessy, F., Cheetham, M. E., Dirr, H. W., and Blatch, G. L. (2000) Analysis of the levels of conservation of the J domain among the various types of DnaJ-like proteins. *Cell Stress Chaperones* **5**, 347–358
119. Miernyk, J. A. (2001) The J-domain proteins of Arabidopsis thaliana: An unexpectedly large and diverse family of chaperones. *Cell Stress Chaperones* **6**, 209–218
120. Rajan, V. B., and D'Silva, P. (2009) Arabidopsis thaliana J-class heat shock proteins: Cellular stress sensors. *Funct. Integr. Genomics* **9**, 433–446
121. Walsh, P., Bursac, D., Law, Y. C., Cyr, D., and Lithgow, T. (2004) The J-protein family: Modulating protein assembly, disassembly and translocation. *EMBO Rep.* **5**, 567–571
122. Willmund, F., Dorn, K. V., Schulz-Raffelt, M., and Schroda, M. (2008) The chloroplast DnaJ homolog CDJ1 of Chlamydomonas reinhardtii is part of a multichaperone complex containing HSP70B, CGE1, and HSP90C. *Plant Physiol.* **148**, 2070–2082
123. Orme, W., Walker, A. R., Gupta, R., and Gray, J. C. (2001) A novel plastid-targeted J-domain protein in Arabidopsis thaliana. *Plant Mol. Biol.* **46**, 615–626
124. Chen, K. M., Piippo, M., Holmstrom, M., Nurmi, M., Pakula, E., Suorsa, M., and Aro, E. M. (2011) A chloroplast-targeted DnaJ protein AtJ8 is negatively regulated by light and has rapid turnover in darkness. *J. Plant Physiol.* **168**, 1780–1783
125. Chen, K. M., Holmstrom, M., Raksajit, W., Suorsa, M., Piippo, M., and Aro, E. M. (2010) Small chloroplast-targeted DnaJ proteins are involved in optimization of photosynthetic reactions in Arabidopsis thaliana. *BMC Plant Biol.* **10**, 43
126. Chico, J. M., Lechner, E., Fernandez-Barbero, G., Canibano, E., Garcia-Casado, G., Franco-Zorrilla, J. M., Hammann, P., Zamarrero, A. M., Garcia-Mina, J. M., Rubio, V., Genschik, P., and Solano, R. (2020) CUL3(BPM) E3 ubiquitin ligases regulate MYC2, MYC3, and MYC4 stability and JA responses. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 6205–6215
127. Julian, J., Coego, A., Lozano-Juste, J., Lechner, E., Wu, Q., Zhang, X., Merilo, E., Belda-Palazon, B., Park, S. Y., Cutler, S. R., An, C., Genschik, P., and Rodriguez, P. L. (2019) The MATH-BTB BPM3 and BPM5 subunits of Cullin3-RING E3 ubiquitin ligases target PP2CA and other clade A PP2Cs for degradation. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 15725–15734
128. Williams, C., Fernandez-Calvo, P., Colinas, M., Pauwels, L., and Goossens, A. (2019) Jasmonate and auxin perception: How plants keep F-boxes in check. *J. Exp. Bot.* **70**, 3401–3414
129. Yeom, J., Wayne, K. J., and Groisman, E. A. (2017) Sequestration from protease adaptor confers differential stability to protease substrate. *Mol. Cell* **66**, 234–246.e235
130. Wang, L., Elliott, M., and Elliott, T. (1999) Conditional stability of the HemA protein (glutamyl-tRNA reductase) regulates heme biosynthesis in Salmonella typhimurium. *J. Bacteriol.* **181**, 1211–1219
131. Wang, L., Wilson, S., and Elliott, T. (1999) A mutant HemA protein with positive charge close to the N terminus is stabilized against heme-regulated proteolysis in Salmonella typhimurium. *J. Bacteriol.* **181**, 6033–6041
132. Fei, X., Bell, T. A., Barkow, S. R., Baker, T. A., and Sauer, R. T. (2020) Structural basis of ClpXP recognition and unfolding of ssrA-tagged substrates. *Elife* **9**, e61496
133. Vahidi, S., Ripstein, Z. A., Juravsky, J. B., Rennella, E., Goldberg, A. L., Mittermaier, A. K., Rubinstein, J. L., and Kay, L. E. (2020) An allosteric switch regulates Mycobacterium tuberculosis ClpP1P2 protease function as established by cryo-EM and methyl-TROSY NMR. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 5895–5906
134. Tullman, J., Callahan, N., Ellington, B., Kelman, Z., and Marino, J. P. (2019) Engineering ClpS for selective and enhanced N-terminal amino acid binding. *Appl. Microbiol. Biotechnol.* **103**, 2621–2633
135. Kunjapur, A. M., Stork, D. A., Kuru, E., Vargas-Rodriguez, O., Landon, M., Soll, D., and Church, G. M. (2018) Engineering posttranslational proofreading to discriminate nonstandard amino acids. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 619–624