FEBS Letters

FEBS PRESS science publishing by scientists

N-degron specificity of chloroplast ClpS1 in plants

Cyrille Montandon[1], David A. Dougan[2] and Klaas J. van Wijk[1]

1 Plant Biology Section, School of Integrative Plant Sciences (SIPS), Cornell University, Ithaca, NY, USA
2 Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Australia

Correspondence
K. J. van Wijk, Plant Biology Section, School of Integrative Plant Sciences (SIPS), Emerson Hall, Cornell University, Ithaca, NY 14853, USA
Tel: +1-607-255-3664
E-mail: kv35@cornell.edu

(Received 8 March 2019, revised 29 March 2019, accepted 1 April 2019, available online 20 April 2019)

doi:10.1002/1873-3468.13378

Edited by Ulf-Ingo Flügge

The prokaryotic N-degron pathway depends on the Clp chaperone-protease system and the ClpS adaptor for recognition of N-degron bearing substrates. Plant chloroplasts contain a diversified Clp protease, including the ClpS homolog ClpS1. Several candidate ClpS1 substrates have been identified, but the N-degron specificity is unclear. Here, we employed *in vitro* ClpS1 affinity assays using eight N-degron green fluorescence protein reporters containing either F, Y, L, W, I, or R in the N-terminal position. This demonstrated that 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. This affinity is dependent on two conserved residues in the ClpS1 binding pocket and is sensitive to FR dipeptide competition, suggesting a unique chloroplast N-degron pathway.

Keywords: adaptor; *Arabidopsis thaliana*; chloroplast; ClpS1; N-degron; protease

The N-terminal (Nt) residue of a protein is prone to several modifications and is a major determinant of protein stability in the cytosol of prokaryotes and eukaryotes, and perhaps also in chloroplasts and non-photosynthetic plastids in apicomplexan, algae, and plants [1–4]. The role of the N-terminus in protein stability was conceptualized in the N-end rule, which stated that certain amino acids, when exposed at the N-terminus of a protein, act as triggers (N-degrons) for degradation. However, our current understanding of protein degradation via the N-terminus suggests that, at least in some organisms, all residues can act as N-degrons therefore the N-end rule pathway was recently renamed the N-degron pathway [2].

In prokaryotes, only a limited number of residues can act as N-degrons and as such Nt residues can be defined as "destabilizing" residues which "tag" the protein for degradation and "stabilizing" residues which are found on stable proteins. In prokaryotes, N-degrons are specifically recognized by a unique N-recognin (or adaptor protein) named ClpS, which delivers the substrate to the ClpA/C chaperone of the ClpP proteolytic system [5–7]. ClpS is a 10–15 kDa protein with an unstructured Nt extension (NTE) and a folded core bearing an hydrophobic N-degron binding pocket. Substrate recognition by ClpS is mediated via the N-degron binding pocket in ClpS specifically recognizing the α-amino group and side chain of an Nt primary destabilizing residue (L, F, Y, and W) together with carbonyl oxygen of the first peptide bond. In *Escherichia coli*, Nt primary destabilizing residues recognized by ClpS are generated either through proteolytic processing or attachment of a primary destabilizing residue (L or F) by L/F-tRNA-protein transferases (LTFRs) onto a secondary destabilizing residue (in particular such as R, K, and in one case M) [8,9].

The sequence of events and molecular details for N-degron recognition and degradation by the ClpAPS system in non-photosynthetic bacteria have been resolved in detail [6,10,11]. The first step involves N-degron recognition by ClpS, followed by docking of the adaptor-substrate complex to the Nt domain of

Abbreviations
BB, binding buffer; FT, flow-through; GluTR1, glutamyl tRNA reductase 1; GSH, glutathione; LFTR, L/F-tRNA-protein transferase; Nt, N-terminal; NTE, N-terminal extension; WB, washing buffer.

ClpA to form a low-affinity complex. Next, the NTE of ClpS engages with the pore of ClpA to form what has been termed a high-affinity delivery complex. The NTE of ClpS is then "pulled" into the ClpA pore in an ATP-dependent fashion and the resulting distortion of the ClpS structure allows release of the substrate inside the ClpA pore. Finally, ClpS is released from ClpA, facilitating the unfolding and degradation of the substrate by the ClpAP complex. Notably, ClpS is necessary and sufficient for the recognition and delivery of N-degron substrates to the bacterial Clp system. In addition, the ClpS adaptor can repress substrate degradation by preventing ClpA to engage in substrate interactions with non-ClpS targets [12–14].

Clp proteases are also present in photosynthetic bacteria such as *Synechococcus elongatus*, in non-photosynthetic apicoplasts of the malaria pathogen *Plasmodium falciparum* [15–17], in plant chloroplasts and mitochondria [18,19], as well as mitochondria of non-plant eukaryotes [20–24]. The chloroplast Clp proteolytic system plays an essential role in chloroplast biogenesis, proteostasis, and metabolism [25,26]. Many α-proteobacteria such as *Agrobacterium tumefaciens* and cyanobacteria such as *S. elongatus* contain two ClpS homologs, ClpS1 similar to bacterial ClpS, and the more divergent ClpS2 [27–29]. Plants contain at least one ClpS1 and in some plant species also a ClpS1-like homolog [30].

The Clp machinery greatly diversified during evolution, with plant chloroplasts harboring the most complex Clp system. In the higher plant *Arabidopsis thaliana,* the chloroplast Clp system consists of a hetero-oligomeric protease core comprising five proteolytically active subunits (ClpP1, ClpP3–6) and four proteolytically inactive proteins (ClpR1–4), as well as two stabilizing/activating factors (ClpT1–2), three AAA+ (ATPases associated with a variety of cellular activities) chaperones (ClpC1, ClpC2, and ClpD), the adaptor ClpS1 [30], and its putative co-adaptor ClpF [31]. The stability of bacterial ClpA depends on the presence of ClpS [32], but ClpC1 stability in chloroplasts is independent of ClpS1 [30]. ClpS1 physically interacts with ClpC, and ClpS1 affinity studies showed that it recognizes a subset of chloroplast proteins, including glutamyl tRNA reductase 1 (GluTR1) [30]. Follow-up experiments showed that dark-induced degradation of GluTR1 requires the Clp system [33]. The interaction between ClpS1 and these candidate substrates, but not ClpF, is dependent on the conserved substrate binding residues (D89/N90) in the ClpS1 core domain, but N-degrons in these substrates have not been identified. However, systematic mass spectrometry analysis of the N-termini of 165 chloroplast stromal proteins using an Nt tagging (TAILS) strategy showed that A, V, T (often in N-α-acetylated form), and S were by far the most observed Nt residues, followed by M and G, even after normalization for their frequency in the plastid proteome, whereas other residues were absent or highly underrepresented [34]. This is compatible with N-degron mechanisms operating in the chloroplast.

In this paper, we examined the specificity of chloroplast ClpS1 via *in vitro* affinity assays using a series of fusion proteins bearing either a classical bacterial ClpS dependent type 2 Nt-destabilizing residues (L, F, Y, and W), an eukaryotic-specific type 2 Nt-destabilizing residue (I) or a type 1 destabilizing residue (R) as a control. The specificity of the interaction was validated through competitive elution of the substrate from wild-type ClpS1 using an FR dipeptide, in addition to the use of a ClpS1 N-degron binding mutant ClpS1-D89A/N90A (further referred to as ClpS1-DN/AA) as negative control. We compare these results to the specificity of ClpS homologs in bacteria and the apicoplast of *P. falciparum*.

Materials and methods

Generation of N-degron-green fluorescence protein (GFP) fusion proteins

Eight different purified recombinant N-degron-GFP fusion proteins were generated as detailed in [9]. These consisted of the Nt residues FR, YR, LR, WR, IR, or R, followed by the 10-amino acid linker SKGEELFTGV (10) fused to GFP. Two additional N-degron-GFP fusion proteins consisted of WLTMITDSLAGV (WLβgal) or ILFVQEL (IL5), followed by GFP. These eight N-degrons were expressed as His6-Ubiquitin fusion proteins in *E. coli*, purified, followed by removal of the ubiquitin moiety using deubiquitinating enzyme to uncover the desired Nt amino acid [9].

Expression and purification of the ClpS1, ClpS1-DN/AA, and GST baits

Recombinant *A. thaliana* ClpS1 and the N-degron binding pocket mutant ClpS1-DN/AA were expressed in *E. coli* as GST-fusion proteins, alongside GST alone and purified as described in [30]. Only the mature portion of ClpS1 was used (A45-C159). Purified recombinant proteins were stored in TBS (50 mM Tris-HCl, pH 7.5, 150 mM NaCl) with 10% glycerol at −80 °C.

In vitro binding assays

The storage buffer of recombinant ClpS1, ClpS1-DN/AA, or GST was exchanged for binding buffer [BB; 10 mM HEPES KOH pH8, 50 mM KOAc, 5 mM MgOAc, 10% (v/w) glycerol, 0.5% (v/v) TritonX-100] using 4 mL Amicon Ultra

concentration columns (Millipore, Millipore Sigma Headquarters, Burlington, MA). Protein concentrations were determined using a BCA assay kit (Thermo Fisher Scientific, Waltham, MA). Equimolar amounts of each bait protein (15 μg of GST and 20 μg of ClpS1, or ClpS1-DN/AA per assay) were batch incubated with GSTrap beads (15 μL per assay; GE Healthcare, Chicagio, IL) rotating end-over-end for 30 min in BB. The beads were washed twice with five column volumes of BB, and 50 μL aliquots of each bait-GSTrap beads were dispatched in 8 × 0.5 mL spin columns (Thermo Fisher Scientific, Waltham, MA). Each N-degron (15 μg) was incubated (for 30 min on an end-over-end shaker) in five column volumes of BB with each of the three GST columns. Beads were washed (2 × 10 min) in five column volumes of washing buffer [WB; 10 mM HEPES KOH pH8, 50 mM KOAc, 5 mM MgOAc, 10% (v/v) glycerol, 0.25% (v/v) TritonX-100].

Step-wise elution with FR dipeptide and reduced glutathione (GSH)

Bound N-degron reporter proteins were eluted following incubation of the beads (30 min on an end-over-end shaker) in four column volumes of 1 mg·mL^{-1} FR dipeptide in WB. After an additional washing step (~ 30 min), the bait and the remaining bound N-degron were eluted in five column volumes of 10 mM GSH in WB.

Detection of N-degron reporters and the baits

Fractions of the N-degron reporter input, flow-through (FT), and FR dipeptide elution were separated by SDS/PAGE, and visualized by staining with Coomassie Brilliant Blue. The remaining amount of the FR dipeptide eluates were precipitated with 1 : 3 : 4 v/v chloroform : methanol : water, and precipitates were washed with methanol, dried, and resuspended in 1 × Laemmli SDS sample buffer for separation by SDS/PAGE, transferred to poly(vinylidene difluoride) membrane followed by immunodetection using an anti-GFP antibody (GenScript, Piscataway, NJ) and detection by chemiluminescence using standard procedures.

Results

Overview of the experimental setup

The general overview of the ClpS1 affinity experiments for candidate N-degrons and controls is illustrated in Fig. 1 (panels A and B). A comparison of the experimental ClpS structures from *E. coli* and *P. falciparum* with a model of Arabidopsis ClpS1 (panel C) and a sequence alignment of all three ClpS homologs (panel D) is also shown in Fig. 1. The structural comparison indicates a general similarity between folds, that is, three α-helices (α1–α3) connected to three antiparallel β-strands with a β1α1α2β2α3β3 topology with the N-degron binding pocket/cavity formed by α1α2. As indicated by the surface color coding, the N-degron binding pocket in ClpS1 is bordered by positive charges, potentially impacting N-degron affinity (see further below).

We selected eight N-degron GFP reporters for the affinity experiments as in [9] (Fig. 1A). These reporters consisted of an Nt residue, a residue in the second position, and a linker followed by GFP. For the Nt residue we selected the classic bacterial type 2 destabilizing residues (L, F, Y, and W), R as control, and I (assigned "expanded" N-degron) based on its affinity to the ClpS homolog in the apicoplast of *P. falciparum* [16]. The influence of the residue in the second position was tested by replacing the R (observed to enhance N-degron affinity in bacterial ClpS) with the non-charged, apolar residue L. Based on the structural and functional characterisation of N-degron binding, it is generally understood that the identification of residues downstream of the second position do not contribute to ClpS interaction; however, these downstream residues do play a crucial role in substrate delivery to the cognate unfoldase [6,9,10].

Each of the N-degron reporters were then tested for affinity to chloroplast ClpS1 using ClpS1 affinity beads; ClpS1 was in its mature form with the Nt chloroplast targeting signal removed, as we previously established [30]. To bind ClpS1 to these affinity beads, ClpS1 was expressed as GST-ClpS1 fusion protein in *E. coli* and purified. As a negative control, we used GST affinity beads containing immobilised GST (alone). To test the specific recognition of ClpS1 for N-degrons, we used ClpS1-DN/AA, in which the critical N-degron binding residues are replaced with alanine [30]. Previously, we showed that GST-ClpS1 interacts with several (candidate) endogenous chloroplast stromal substrates, as well as with the co-adaptor ClpF [30,31]. These stromal substrates had little-to-no affinity for GST-ClpS1-DN/AA, whereas the interaction of ClpF with ClpS1 was unaffected by these mutations. Each of the three bait proteins (GST, GST-ClpS1, and GST-ClpS1-DN/AA) were bound to GSTrap affinity beads and incubated with eight different N-degron reporter proteins (total of 24 incubations per replicate). Unbound N-degron reporters were washed from each affinity column, and bound N-degrons were eluted using a FR dipeptide solution (FR eluate). Following another wash to remove residual FR dipeptide and free N-degron reporters, the GST-fusion proteins and remaining N-degron reporters were eluted from the affinity beads using 10 mM GSH (GSH eluate). These GSH eluates served as another control for the amount of bait and also show to what extent each N-degron reporter remained associated with the bait. The reason for this retention is currently unclear;

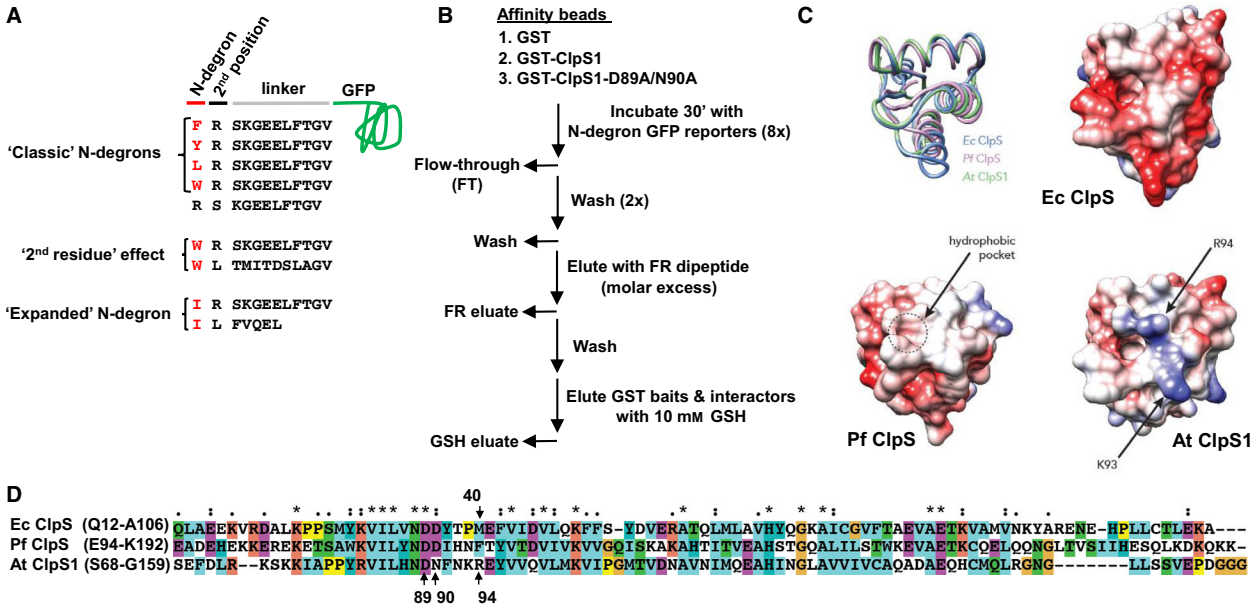

Fig. 1. Overview of the experimental setup and comparison of *Escherichia coli* ClpS and *Arabidopsis thaliana* ClpS1. (A) Eight N-degron GFP reporters were expressed in *E. coli* and purified to test for affinity to At ClpS1. These reporters consisted of the Nt residue, that is, the classic bacterial type 2 Nt-destabilizing residues (L, F, Y, and W), control R, and the expanded type 2 Nt-destabilizing residue I, based on its affinity to the apicoplast ClpS homolog in *Plasmodium falciparum* [16]. The influence of the residue in the second position was tested by replacing the "typical/model" residue R with the uncharged, nonpolar residue L. As indicated, three different linkers were used, namely a 10 aa sequence of SKGEELFTGV [6], a 10 aa sequence TMITDSLAGV (βgal, derived from [8,43]) in case of WL, and in case of IL a 5 aa sequence FVQEL (see further explanation for the choice of these linkers in [9–11]), (B) Three different bait proteins were expressed in *E. coli* and purified. These were the GST protein as a negative control, and ClpS1 or ClpS1 D89A/N90A (referred to here as ClpS1-DN/AA) each fused to the C-terminus of GST. Each of the substrate proteins were incubated with each bait (8 × 3 incubations per replicate). Unbound N-degron reporters were washed from each affinity column, and bound N-degron reporters were eluted using a FR dipeptide solution (FR eluate). Following another wash, bait proteins with remaining bound N-degron reporters were eluted using 10 mM GSH (GSH eluate). FT samples, wash samples, FR eluates, and GSH eluates were evaluated for the amounts of bait and N-degron reporters after SDS/PAGE, Coomassie staining, transfer to blots followed by Ponceau S staining and detection with anti-GFP anti-serum. (C) Structural comparison of the core domains of *E. coli* ClpS, *P. falciparum* ClpS, and *A. thaliana* ClpS1 Top left panel: String representation of the C backbone of *E. coli* ClpS (PDB: 2W9R), *P. falciparum* ClpS (PDB: 4O2X) and *A. thaliana* ClpS1 (modeled on 2W9R, using Swiss Model) in blue, pink, and green, respectively. Top right panel: Electrostatic surface representation of Ec ClpS. Bottom left panel: Electrostatic surface representation of Pf ClpS, highlighting the proposed hydrophobic pocket. Bottom right panel: Electrostatic surface representation of ClpS1. Residues K93 and R94 are indicated. Coulombic surface +10, −10 – blue is positively charged, red is negatively charged [10]. (D) Sequence alignment of the core regions of the three ClpS homologs shown in panel C. Key residue numbers for Ec CpS and At ClpS1 are indicated.

however, it appears to be nonspecific, as all reporters, including R_{10}GFP were retained to ClpS1. FT samples, wash samples, FR eluates, and GSH eluates were evaluated for the amounts of bait and N-degron reporters using SDS/PAGE, Coomassie Brilliant Blue staining, transferred to blots followed by Ponceau S staining and detection with anti-GFP anti-serum.

ClpS1 specifically interacts with two of the four classic type 2 bacterial Nt-destabilizing residues, F and W and weak recognition of L

Initially, we focused on ClpS1 affinity for the classic type 2 bacterial degrons, LR, FR, YR and WR, and R as control (Fig. 2). Figure 2A shows that the amount of each N-degron reporter, loaded on the three bait-affinity columns (GST, GST-ClpS1, and GST-ClpS1-DN/AA) was fully saturating. Following washes to remove unbound N-degron reporters, columns were incubated with molar excess of FR peptide and the eluates were collected. Separation of proteins by SDS/PAGE, from each FR eluate, followed by Coomassie staining showed weak bands for FR_{10}GFP and WR_{10}GFP were eluted from the ClpS1 column, but not the GST or ClpS1-DN/AA controls (Fig. 2A). Immunodetection with anti-GFP serum of FR eluate confirmed a strong and specific recovery of proteins bearing either a FR or WR degron with the ClpS1 bait, but not for GST or ClpS1-DN/AA (Fig. 2B). Additionally, a weak, yet specific recovery of proteins bearing the LR degron was also observed for

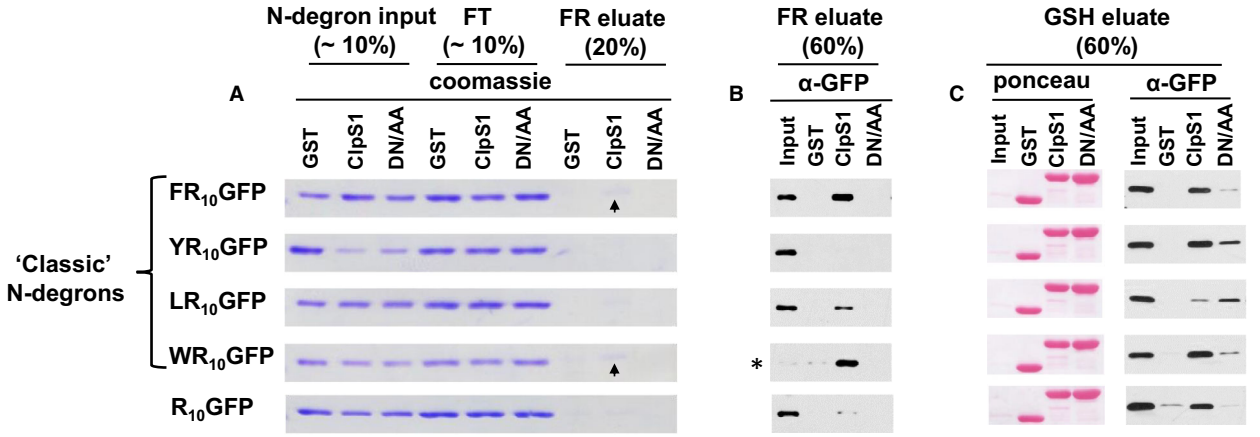

Fig. 2. Probing ClpS1 interactions with bacterial type 2 N-degron, L, F, Y, and W. Equimolar amounts (ca. the equivalent of 15 μg of GST) of recombinant GST (GST), GST-ClpS1 (ClpS1), or mutant GST-ClpS1 (DN/AA) were immobilized on 100 μL of GSTrap beads. Fifteen micrograms of five different recombinant N-degron-GFP reporters were incubated with the different affinity columns. After washing, bound N-degron-GFP reporters were specifically eluted with four column volumes of a 1 mg·mL^{-1} FR dipeptide solution. Subsequently, each bait protein (GST, wild-type and mutant GST-ClpS1) was eluted with five column volumes 10 mM GSH. (A) Ten percent of the N-degron input, 10% of the flow-through, and 20% of the FR eluted fractions were separated by SDS/PAGE and stained with Coomassie brilliant blue. (B) Sixty percent of each FR eluate was analyzed by immunoblotting using an anti-GFP antibody. One percent of the respective N-degron input was loaded as a reference. *The signal for "input" in this blot for WR_{10}GFP was relatively weak, likely due to technical reasons; an aliquot of the same input was detected with "normal signal strength" in the blot of panel C. (C) Sixty percent of the GSH elution from each column was separated by SDS/PAGE. One percent of the respective N-degron input was loaded as a reference. The presence of bound N-degron reporters (not eluted by the FR dipeptide) was analyzed by immunoblotting using an anti-GFP antibody. Ponceau S staining is shown for the bait. N-degron-GFP = 27 kDa; GST = 26 kDa; GST-ClpS1-DN/AA and GST-ClpS1 = 39 kDa.

ClpS1, with an even weaker background signal for proteins bearing the R-degron (Fig. 2B).

To determine the full extent of N-degron binding to the bait, each bait protein was eluted from the beads (together with any bound interacting proteins) using GSH. GSH eluates were separated by SDS/PAGE, followed by blotting and detection by Ponceau S staining. Ponceau S staining of the bait proteins showed consistent loading of the bait proteins across all columns (Fig. 2C). The blot was then probed with anti-GFP serum to detect the N-degron reporter proteins; an N-degron input control was included. This showed the strongest signals for WR, FR, and YR bound to ClpS1, with considerable binding of YR to ClpS1-DN/AA and weak interaction of LR_{10}-GFP to both ClpS1 and ClpS1-DN/AA. Finally, signals in the control GST column and ClpS1 column were also observed for the R_{10}-GFP control (Fig. 2C). The Ponceau stain of the GSH eluates shows equal distribution of the baits (GST, ClpS1, and ClpS1-DN/AA) for each affinity column. The immunoblots (with anti-GFP) of the GSH eluates show that (a) very little of each N-degron reporter interacts with the beads (alone), with the exception of a minor amount of the R N-degron reporter (see Fig. 2C, GST lane); (b) the FR dipeptide is unable to compete with all of the bound N-degron reporters; and (c) that a small amount of N-degron reporters, in particular YR and LR, interact with ClpS independent of the conserved N-degron binding residues (see Fig. 2C, DN/AA lane).

The ClpS1 affinity for the W N-degron is not significantly influenced by the residue in the second position, comparison of the charged R with the apolar, hydrophobic L

Next, we examined the role of the penultimate residue within the N-degron for ClpS1 recognition. It has been shown for *E. coli* ClpS that the best binding was observed for N-degrons in which the penultimate residue was basic, namely R or K [6]. In the case of N-degrons generated by LFTR the penultimate residue is typically R or K, as was also observed in *E. coli* substrates trapped on ClpS affinity columns in [35]. This preference for basic residues in the second position is complemented by the acidic nature of the surface surrounding the hydrophobic pocket of *E. coli* ClpS [10] (see Fig. 1C). In contrast, the surface surrounding the hydrophobic pocket of ClpS1 is more positively charged, due in part to the presence of two basic residues (K93 and R94) which replace P39 and M40 (the "specificity" residue in *E. coli* ClpS; see Fig. 1D).

Because of the presence of the positively charged residue (R) within the binding pocket of ClpS1 and the potential for charge repulsion of the substrate, we also examined the effect of the second residue in the substrate. Therefore, we compared the WR and WL N-degron reporter for binding to ClpS1 and controls (Fig. 3A–C). Consistently, we observed that the affinity of ClpS1 for WL, was comparable to or better than, the affinity of ClpS1 for WR, as determined by FR dipeptide elution (both in the Coomassie stain and immunoblot; Fig. 3B,C). It should be noted that, as described in [10,11], we have used different linkers for the WR and WL reporters; however based on the structure of ClpS in complex with several N-degron peptides [10,11], neither linker is expected to influence substrate affinity for ClpS (see Fig. 1A). Upon elution of the bait using GSH, somewhat variable signals were observed for WR and WL with both ClpS1 and ClpS1-DN/AA, however very little non-specific binding to the beads alone was observed (Fig. 3C, GST lane). As observed also in Fig. 2, the FR dipeptide was unable to outcompete all of the bound N-degron reporters. In contrast to most reporter proteins, the WL-reporter showed slightly more binding to ClpS1-DN/AA when compared to wild-type ClpS1.

ClpS1 has little specific affinity for I as N-degron, irrespective of the having R or L in the second position

In the case of the ClpS homolog from the non-photosynthetic apicoplast of *P. falciparum*, a specific affinity for an Nt I reporter (IL_5-GFP) was observed [16]. We note that the *E. coli* ClpS-M40A variant recognizes β-branched (V and I) residues in addition to the type 2 N-degrons [11]. Figure 4 shows a direct comparison for this N-degron and the variant IR_{10}-GFP, for binding to ClpS1 and controls. We did not observe affinity for IL or IR to ClpS1 or the negative controls when eluting with the FR dipeptides (Fig. 4A,B). In contrast, clear signals were again observed both through Coomassie staining and immunoblot with anti-GFP for the FR and WL degrons (Fig. 4A,B), similar to Fig. 2. Upon elution of the bait with GSH, weak and somewhat variable signals were observed for IR and IL with both ClpS1 and ClpS1-DN/AA, while strong signals were again observed for both FR and WL, indicating that the FR dipeptide could not fully outcompete the bound N-degron reporters (Fig. 4C). We do note that the linkers for IL and IR were different (see also Fig. 1A), but as explained above, neither are expected to impact the specific affinity for the N-degron binding pocket.

Discussion

Chloroplast ClpS1 from Arabidopsis is likely to be a functional homolog of bacterial ClpS. However, chloroplast ClpS1 exhibits a low overall sequence identify (< 30%) with its bacterial counterparts and contains a number of unique features, including a unique composition of substrate "specificity" residues, a longer NTE (~ 10 residues longer, excluding the chloroplast transit peptide of 44 residues), and interaction with a protein that is unique to higher plants – the co-adaptor ClpF [31]. The NTE of bacterial ClpS enters the ClpA chaperone pore for substrate transfer

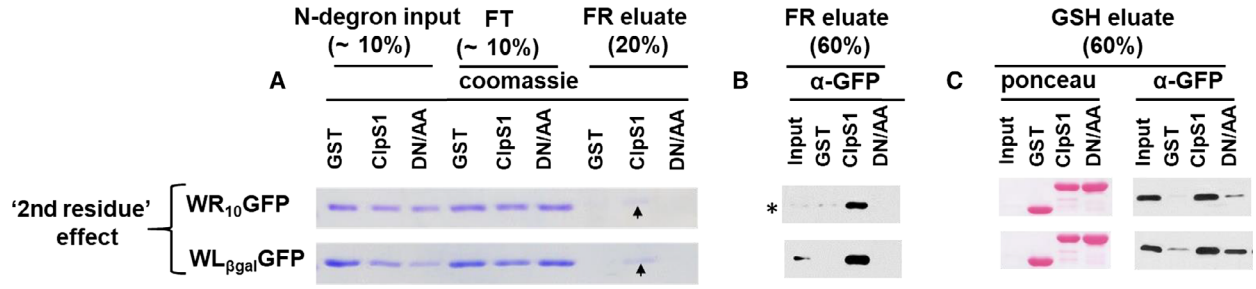

Fig. 3. Probing the W N-degron interactions with ClpS1 for the effect of the second position. The experimental design was as described in Fig. 2 but only N-degron reporters WR_{10}-GFP and $WL_{βgal}$-GFP were used. (A) Ten percent of the N-degron input, 10% of the flow-through, and 20% of the FR eluted fractions were separated by SDS/PAGE and stained with Coomassie Brilliant Blue. (B) Sixty percent of each FR eluate was analyzed by immunoblotting using an anti-GFP antibody. One percent of the respective N-degron input was loaded as a reference. The signal for "input" in this blot for WR_{10}GFP was relatively weak, likely due to technical reasons; an aliquot of the same input was detected with "normal signal strength" in the blot of panel C. (C) Sixty percent of the GSH elution of the different bait proteins were separated by SDS/PAGE. The presence of bound N-degron reporters (not eluted by the FR dipeptide) was analyzed by immunoblotting using an anti-GFP antibody. Ponceau S staining shows the bait proteins. N-degron-GFP = 27 kDa; GST = 26 kDa; GST-ClpS1-DN/AA and GST-ClpS1 = 39 kDa.

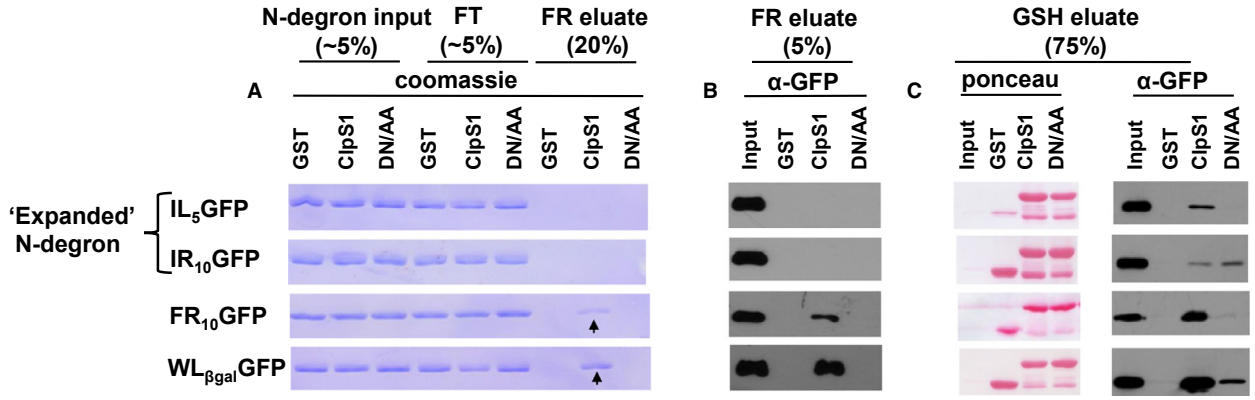

Fig. 4. ClpS1 has no specific affinity for I as an N-degron. The experimental design was as described for Fig. 2; however, in this case only N-degrons-reporters IR_{10}-GFP and IL_5-GFP were used with FR and WL as positive controls. (A) Approximately five percent of the N-degron input, ~ 5% of the flow-through, and 20% of the FR eluted fractions were separated by SDS/PAGE and stained with Coomassie Brilliant Blue. (B) Five percent of each FR eluate was analyzed by immunoblotting using an anti-GFP antibody. One percent of the respective N-degron inputs was loaded as a reference. (C) Seventy-five percent of the GSH elution of the different baits was separated by SDS/PAGE. The presence of bound N-degron reporters (not eluted by the FR dipeptide) was analyzed by immunoblotting using an anti-GFP antibody. Ponceau S staining is shown for the bait. N-degron GFP = 27 kDa; GST = 26 kDa; GST-ClpS1-DN/AA and GST-ClpS1 = 39 kDa.

[7] and this longer NTE in ClpS1 likely reflects structural differences between bacterial ClpA and chloroplast ClpC. Furthermore, N-degrons for chloroplast ClpS1 have yet to be established *in vivo* or *in vitro* [30]. A recent study used recombinant ClpS1 to indirectly examine the affinity for several classic N-degrons through competition in degradation assays with the *E. coli* ClpAPS system [36]. This suggested that ClpS1 has much lower affinity (~ 20 fold) for the model degrons FR, YL, YK, WF, and LL, and surprisingly, with the lowest affinity to WF. Furthermore, crosslinkers were needed to visualize any interaction of the N-degron model substrates with ClpS1, in contrast to our current study, where the interaction between ClpS1 and the WR and FR N-degron substrates is readily visible. The poor affinity reported by Ceccarelli and colleagues was attributed to the presence of a so-called "gatekeeper" residue (termed R50, which is the equivalent to what we refer to here as ClpS1 R94 (Fig. 1C,D) based on the amino acid sequence of the full-length protein before processing) at the position equivalent to M40 in *E. coli* ClpS [36]. As previously described in more detail [30], this arginine is conserved in ClpS1 homologs in angiosperms, but replaced by Glu in ClpS1-like plant and algal species (and Phe in *P. falciparum* ClpS), reinforcing the question to what extent N-degrons in chloroplasts are the same as in *E. coli*. This also relates to the question if chloroplasts have secondary N-degrons that require L/F-tRNA-protein transferase (LFTR) or even Arg-D/E amino transferase activity. So far, no known L/F-transferase homologs (similar to either Bpt or LFTR) have been identified in the chloroplast of Arabidopsis. However, two cytosolic Arg-D/E transferase (Ate) homologs have been identified in Arabidopsis (ATE1 and ATE2) [37–39] and the transferase activity of ATE1 has been demonstrated *in vitro* [40]. However, neither ATE1 nor ATE2, nor other types of R-transferases, are currently known to localize to chloroplasts. The lack of known Nt aminoacyl-transferases in plastids raises the question if secondary N-degrons operate in plastids. It is important to note that type 2 residues (W, F, Y, and L) and type 1 residues (R, K, and H) are rarely observed in Nt positions of the chloroplast stromal proteome [34] which could suggest that these residues confer instability when placed in the Nt position of stromal exposed proteins; for discussion see also [4,41,42].

Our N-degron ClpS1 binding and FR dipeptide elution assays demonstrate that chloroplast ClpS1 binds specifically and consistently to the N-degron reporters with the N-termini FR, WR, and WL. This interaction is strongly dependent on the presence of the conserved ClpS1 binding pocket residue D89 and N90, suggesting that the role of the equivalent residues in *E. coli* (D35 and D36) in the binding of the N-degron is conserved in ClpS1. Following FR dipeptide elution, ClpS1 also showed very weak interaction with LR, but no interaction was observed for the Nt residues YR, IR, IL, or R. The ability of the FR dipeptide to elute the bound FR, WR, and WL reporters confirms the specificity of the interaction. However, a portion of the N-degron reporters could not be eluted with the FR dipeptides, as evidenced when each bait was released from the

affinity matrix using GSH. These FR-insensitive interactions reflect either that the interaction with the ClpS1 N-degron binding pocket cannot be outcompeted effectively with FR dipeptide and/or that N-degron reporters also exhibit affinity to other regions of ClpS1, (but not to the GSH beads). In conclusions, our *in vitro* results suggest an N-degron pathway exists in chloroplasts, which is compatible with the N-terminome data for the chloroplast stromal proteome [34]. More direct *in vivo* (*in planta*) degradation assays are needed to test the N-degron pathway, including the role of the plant chloroplast-specific ClpF.

Acknowledgements

This study was supported by a grant from the National Science Foundation, MCB #1614629 to KVW and a postdoctoral fellowship from the Swiss National Science Foundation (SNF) to CM. DAD was supported by an Research Fellowship (DP110103936) from the Australian Research Council (ARC). We thank Mr. Adnan Alhuwaider for his technical assistance in the purification of N-degron-GFP substrates.

Author contributions

CM carried out all the experiments and wrote an initial draft of the manuscript. KVW and DAD analyzed the data and wrote the paper. KVW oversaw the project and supervised CM.

References

1 Gibbs DJ, Bailey M, Tedds HM and Holdsworth MJ (2016) From start to finish: amino-terminal protein modifications as degradation signals in plants. *New Phytol* **211**, 1188–1194.
2 Varshavsky A (2019) N-degron and C-degron pathways of protein degradation. *Proc Natl Acad Sci USA* **116**, 358–366.
3 Dissmeyer N, Rivas S and Graciet E (2018) Life and death of proteins after protease cleavage: protein degradation by the N-end rule pathway. *New Phytol* **218**, 929–935.
4 Dissmeyer N (2019) Conditional protein function via N-degron pathway-mediated proteostasis in stress physiology. *Annu Rev Plant Biol* **70**. https://doi.org/10.1146/annurev-arplant-050718-095937.
5 Zeth K, Ravelli RB, Paal K, Cusack S, Bukau B and Dougan DA (2002) Structural analysis of the adaptor protein ClpS in complex with the N- terminal domain of ClpA. *Nat Struct Biol* **9**, 906–911.
6 Erbse A, Schmidt R, Bornemann T, Schneider-Mergener J, Mogk A, Zahn R, Dougan DA and Bukau B (2006) ClpS is an essential component of the N-end rule pathway in Escherichia coli. *Nature* **439**, 753–756.
7 Rivera-Rivera I, Roman-Hernandez G, Sauer RT and Baker TA (2014) Remodeling of a delivery complex allows ClpS-mediated degradation of N-degron substrates. *Proc Natl Acad Sci USA* **111**, E3853–E3859.
8 Tobias JW, Shrader TE, Rocap G and Varshavsky A (1991) The N-end rule in bacteria. *Science* **254**, 1374–1377.
9 Ninnis RL, Spall SK, Talbo GH, Truscott KN and Dougan DA (2009) Modification of PATase by L/F-transferase generates a ClpS-dependent N-end rule substrate in *Escherichia coli*. *EMBO J* **28**, 1732–1744.
10 Schuenemann VJ, Kralik SM, Albrecht R, Spall SK, Truscott KN, Dougan DA and Zeth K (2009) Structural basis of N-end rule substrate recognition in *Escherichia coli* by the ClpAP adaptor protein ClpS. *EMBO Rep* **10**, 508–514.
11 Wang KH, Roman-Hernandez G, Grant RA, Sauer RT and Baker TA (2008) The molecular basis of N-end rule recognition. *Mol Cell* **32**, 406–414.
12 Liu J, Francis LI, Jonas K, Laub MT and Chien P (2016) ClpAP is an auxiliary protease for DnaA degradation in *Caulobacter crescentus*. *Mol Microbiol* **102**, 1075–1085.
13 Wang KH, Sauer RT and Baker TA (2007) ClpS modulates but is not essential for bacterial N-end rule degradation. *Genes Dev* **21**, 403–408.
14 Dougan DA, Mogk A, Zeth K, Turgay K and Bukau B (2002) AAA+ proteins and substrate recognition, it all depends on their partner in crime. *FEBS Lett* **529**, 6–10.
15 El Bakkouri M, Pow A, Mulichak A, Cheung KL, Artz JD, Amani M, Fell S, de Koning-Ward TF, Goodman CD, McFadden GI *et al.* (2010) The Clp chaperones and proteases of the human malaria parasite *Plasmodium falciparum*. *J Mol Biol* **404**, 456–477.
16 Tan JL, Ward L, Truscott KN and Dougan DA (2016) The N-end rule adaptor protein ClpS from *Plasmodium falciparum* exhibits broad substrate specificity. *FEBS Lett* **590**, 3397–3406.
17 AhYoung AP, Koehl A, Vizcarra CL, Cascio D and Egea PF (2016) Structure of a putative ClpS N-end rule adaptor protein from the malaria pathogen *Plasmodium falciparum*. *Protein Sci* **25**, 689–701.
18 van Wijk KJ (2015) Protein maturation and proteolysis in plant plastids, mitochondria, and peroxisomes. *Annu Rev Plant Biol* **66**, 75–111.
19 Peltier JB, Ripoll DR, Friso G, Rudella A, Cai Y, Ytterberg J, Giacomelli L, Pillardy J and Van Wijk KJ (2004) Clp protease complexes from photosynthetic and non-photosynthetic plastids and mitochondria of plants. Their predicted three-dimensional structures, and functional implications. *J Biol Chem* **279**, 4768–4781.

20 Olivares AO, Baker TA and Sauer RT (2016) Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. *Nat Rev Microbiol* **14**, 33–44.
21 Liu K, Ologbenla A and Houry WA (2014) Dynamics of the ClpP serine protease: a model for self-compartmentalized proteases. *Crit Rev Biochem Mol Biol* **49**, 400–412.
22 Sauer RT and Baker TA (2011) AAA+ proteases: ATP-fueled machines of protein destruction. *Annu Rev Biochem* **80**, 587–612.
23 Alexopoulos JA, Guarne A and Ortega J (2012) ClpP: a structurally dynamic protease regulated by AAA+ proteins. *J Struct Biol* **179**, 202–210.
24 Truscott KN, Bezawork-Geleta A and Dougan DA (2011) Unfolded protein responses in bacteria and mitochondria: a central role for the ClpXP machine. *IUBMB Life* **63**, 955–963.
25 Nishimura K and van Wijk KJ (2015) Organization, function and substrates of the essential Clp protease system in plastids. *Biochim Biophys Acta* **1847**, 915–930.
26 Rodriguez-Concepcion M, D'Andrea L and Pulido P (2018) Control of plastidial metabolism by the Clp protease complex. *J Exp Bot*, https://doi.org/10.1093/jxb/ery441.
27 Stein BJ, Grant RA, Sauer RT and Baker TA (2016) Structural basis of an N-degron adaptor with more stringent specificity. *Structure* **24**, 232–242.
28 Andersson FI, Blakytny R, Kirstein J, Turgay K, Bukau B, Mogk A and Clarke AK (2006) Cyanobacterial ClpC/HSP100 protein displays intrinsic chaperone activity. *J Biol Chem* **281**, 5468–5475.
29 Tryggvesson A, Stahlberg FM, Topel M, Tanabe N, Mogk A and Clarke AK (2015) Characterization of ClpS2, an essential adaptor protein for the cyanobacterium *Synechococcus elongatus*. *FEBS Lett* **589**, 4039–4046.
30 Nishimura K, Asakura Y, Friso G, Kim J, Oh SH, Rutschow H, Ponnala L and van Wijk KJ (2013) ClpS1 is a conserved substrate selector for the chloroplast Clp protease system in *Arabidopsis*. *Plant Cell* **25**, 2276–2301.
31 Nishimura K, Apitz J, Friso G, Kim J, Ponnala L, Grimm B and van Wijk KJ (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.
32 Dougan DA, Reid BG, Horwich AL and Bukau B (2002) ClpS, a substrate modulator of the ClpAP machine. *Mol Cell* **9**, 673–683.
33 Apitz J, Nishimura K, Schmied J, Wolf A, Hedtke B, van Wijk KJ 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.
34 Rowland E, Kim J, Bhuiyan NH and van Wijk KJ (2015) The *Arabidopsis* chloroplast stromal N-terminome: complexities of amino-terminal protein maturation and stability. *Plant Physiol* **169**, 1881–1896.
35 Humbard MA, Surkov S, De Donatis GM, Jenkins LM and Maurizi MR (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.
36 Colombo CV, Rosano GL, Mogk A and Ceccarelli EA (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.
37 Yoshida S, Ito M, Callis J, Nishida I and Watanabe A (2002) A delayed leaf senescence mutant is defective in arginyl-tRNA:protein arginyltransferase, a component of the N-end rule pathway in *Arabidopsis*. *Plant J* **32**, 129–137.
38 Graciet E, Walter F, Maoileidigh DO, Pollmann S, Meyerowitz EM, Varshavsky A and Wellmer F (2009) The N-end rule pathway controls multiple functions during *Arabidopsis* shoot and leaf development. *Proc Natl Acad Sci USA* **106**, 13618–13623.
39 Holman TJ, Jones PD, Russell L, Medhurst A, Ubeda Tomas S, Talloji P, Marquez J, Schmuths H, Tung SA, Taylor I *et al.* (2009) The N-end rule pathway promotes seed germination and establishment through removal of ABA sensitivity in *Arabidopsis*. *Proc Natl Acad Sci USA* **106**, 4549–4554.
40 White MD, Klecker M, Hopkinson RJ, Weits DA, Mueller C, Naumann C, O'Neill R, Wickens J, Yang J, Brooks-Bartlett JC *et al.* (2017) Plant cysteine oxidases are dioxygenases that directly enable arginyl transferase-catalysed arginylation of N-end rule targets. *Nat Commun* **8**, 14690.
41 Millar AH, Heazlewood JL, Giglione C, Holdsworth MJ, Bachmair A and Schulze WX (2019) The scope, functions, and dynamics of posttranslational protein modifications. *Annu Rev Plant Biol* **70**, https://doi.org/10.1146/annurev-arplant-050718-100211.
42 Apel W, Schulze WX and Bock R (2010) Identification of protein stability determinants in chloroplasts. *Plant J* **63**, 636–650.
43 Tobias JW, Shrader TE, Rocap G and Varshavsky A (1991) The N-end rule in bacteria. *Science* **254**(5036), 1374–1377.