Sequence Motifs in Transit Peptides Act as Independent Functional Units and Can Be Transferred to New Sequence Contexts[1][OPEN]

Dong Wook Lee, Seungjin Woo, Kyoung Rok Geem, and Inhwan Hwang*

Division of Integrative Biosciences and Biotechnology (D.W.L., S.W., I.H.) and Department of Life Sciences (K.R.G., I.H.), Pohang University of Science and Technology, Pohang 790–784, Korea

A large number of nuclear-encoded proteins are imported into chloroplasts after they are translated in the cytosol. Import is mediated by transit peptides (TPs) at the N termini of these proteins. TPs contain many small motifs, each of which is critical for a specific step in the process of chloroplast protein import; however, it remains unknown how these motifs are organized to give rise to TPs with diverse sequences. In this study, we generated various hybrid TPs by swapping domains between Rubisco small subunit (RbcS) and chlorophyll *a/b*-binding protein, which have highly divergent sequences, and examined the abilities of the resultant TPs to deliver proteins into chloroplasts. Subsequently, we compared the functionality of sequence motifs in the hybrid TPs with those of wild-type TPs. The sequence motifs in the hybrid TPs exhibited three different modes of functionality, depending on their domain composition, as follows: active in both wild-type and hybrid TPs, active in wild-type TPs but inactive in hybrid TPs, and inactive in wild-type TPs but active in hybrid TPs. Moreover, synthetic TPs, in which only three critical motifs from RbcS or chlorophyll *a/b*-binding protein TPs were incorporated into an unrelated sequence, were able to deliver clients to chloroplasts with a comparable efficiency to RbcS TP. Based on these results, we propose that diverse sequence motifs in TPs are independent functional units that interact with specific translocon components at various steps during protein import and can be transferred to new sequence contexts.

The chloroplasts of plant cells have more than 3,000 different types of proteins involved in their functions (Leister, 2003; Li and Chiu, 2010), and more than 90% of these proteins are encoded in the nucleus and translated by cytosolic ribosomes (Li and Chiu, 2010; Lee et al., 2013a). Consequently, one of the most critical processes in chloroplast proteome biogenesis is the specific, posttranslational delivery of these nuclear-encoded proteins to chloroplasts (Jarvis, 2008; Li and Chiu, 2010; Lee et al., 2013a, 2014). Delivery to chloroplasts requires a specific targeting signal whose form depends on the type of protein and its location in the chloroplast. Most proteins imported into the chloroplast contain an N-terminal transit peptide (TP) as a targeting signal (Lee et al., 2006, 2008, 2013a; Chotewutmontri et al., 2012; Li and Teng, 2013). The TP is cleaved off after import into the chloroplast; thus, the proteins that still contain the TP are called preproteins. Despite progress made in previous studies (Lee et al., 2008; Chotewutmontri et al., 2012; Li and Teng, 2013), the types of information encoded by the long TPs, as well as how this information determines translocation through the import channel, remain to be elucidated.

One long-lasting question regarding the mechanism of TP-mediated protein import is how TPs can specifically deliver proteins into chloroplasts. In striking contrast to endoplasmic reticulum (ER)-targeting signals, TPs are highly diverse at the primary sequence level and do not converge toward a consensus sequence. The leader sequence, which contains the N-terminal ER-targeting signal, is composed of a stretch of hydrophobic amino acids ranging from 15 to 20 residues. Although the exact sequence is highly variable, the residues tend to be hydrophobic, making a high degree of hydrophobicity a common characteristic feature for both luminal and membrane proteins. Despite their diversity in primary sequence, TPs also share certain characteristics that serve as the basis for the software prediction of chloroplast proteins; these features include an amino acid composition with a high concentration of hydroxylated residues and a lack of acidic residues (Bruce, 2000; Bhushan et al., 2006), an unfolded and extended structure, an α-helix-containing secondary structure that may be induced by binding to the lipids of chloroplasts (Wienk et al., 1999; Bruce, 2000), and an abundance of Pro residues that may contribute to the unstructured nature of TPs (Pilon et al., 1995; Bruce, 2000; Zybailov et al., 2008).

[1] This work was supported by the National Research Foundation, Ministry of Science, Technology and Future Planning (grant no. 2013070270 and NRF–2013R1A1A2060635 to D.W.L.) and the Ministry of Ocean and Fisheries, Korea (grant no. D11413814H480000121).

* Address correspondence to ihhwang@postech.ac.kr.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Inhwan Hwang (ihhwang@postech.ac.kr).

D.W.L. and I.H. conceived the project; D.W.L. and I.H. designed the research; D.W.L., S.W., and K.R.G. performed the experiments; D.W.L. and I.H. wrote the article.

[OPEN] Articles can be viewed without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.15.00842

These features provide insight into the sequence information carried by TPs. However, we are still far from fully understanding how TPs function in the mechanism of protein import into chloroplasts. Recent studies have identified sequence motifs by analyzing various deletion and substitution mutants (Pilon et al., 1995; Lee et al., 2006, 2008, 2013a; Chotewutmontri et al., 2012). These motifs, or domains, are thought to be involved in the interaction with components of the translocon (Chotewutmontri et al., 2012; Li and Teng, 2013). Moreover, multiple sequence motifs function individually, or in a combinatorial manner, during specific steps of the import process (Lee et al., 2006, 2008, 2009a). In addition, certain motifs share functional redundancy, or are additive or synergistic. However, despite the progress in identifying sequence motifs from different TPs, it remains unknown how the large number of diverse TPs, as a whole, can deliver proteins to chloroplasts. In ER targeting, the targeting machinery recognizes hydrophobicity, a common feature of the leader sequences, but not the primary sequence (Hessa et al., 2005). Therefore, leader sequences with different primary sequences can be recognized by the same molecular machinery. However, in contrast to the leader sequences, the TPs of chloroplast preproteins contain different sets of sequence motifs (Lee et al., 2006, 2008). These observations raise several questions, including (1) how the large number of TPs with different sets of sequence motifs can be recognized by only a few import receptors (Li and Chiu, 2010; Lee et al., 2013a; Li and Teng, 2013), and (2) how TPs can have such diverse sequences while still retaining their function.

In this study, we investigated the design principles of TPs with diverse primary sequences. Using TPs of the Rubisco small subunit (RbcS) and chlorophyll *a/b*-binding protein (Cab) proteins, which have completely different primary sequences and functional motifs (Lee et al., 2006, 2008), we generated hybrid TPs and examined their activities in chloroplast protein import within protoplasts. We provide evidence that sequence motifs are independent functional units that interact with various components of the translocon during import into chloroplasts and can be transferred to new sequence contexts. However, the functionalities as well as the activities of these motifs are greatly dependent on the overall sequence context of, and their positions in, TPs. In addition, we demonstrated that functional synthetic transit peptides (SynTPs) can be generated by incorporating only a few sequence motifs from RbcS and Cab TPs into an unrelated sequence.

RESULTS

TPs of RbcS and Cab Share the Same General Import Pathway

To gain insight into the functionality of sequence motifs in TPs, we first examined whether TPs with different sequence motifs are imported by the same molecular machinery. To this end, we fused the TPs of Rubisco small subunit (RbcS-nt [N-terminal region]) and Cab (Cab-nt), which contain completely different sequence motifs (Fig. 1A), to GFP and examined whether they are imported into chloroplasts by the same general import pathway. In this study, we used the protoplast system among many different approaches that had been used in the study of protein import into chloroplasts. Compared with the in vitro import system, the protein import process in protoplasts more closely resembles that in plants, although it is more difficult to elucidate the exact action mechanism at each step of the import process in detail (Lee et al., 2006, 2008, 2009a; Lee and Hwang, 2011). Therefore, we took advantage of the protoplast system to examine the behavior of various TPs in delivering the C-terminal GFP moiety into chloroplasts after translation. Initially, we introduced the two constructs, *RbcS-nt*:*GFP* and *Cab-nt*:*GFP*, into protoplasts from wild-type, *plastid protein import2* (*ppi2*), or *heat shock protein93-V* (*hsp93-V*) plants, and their import into chloroplasts was compared. *ppi2* and *hsp93-V* plants have a mutation in Arabidopsis *translocon at the outer envelope of chloroplasts159* (*atToc159*) and *Hsp93-V*, respectively (Bauer et al., 2000; Su and Li, 2010). Toc159 and Hsp93-V are outer membrane-localized import receptors and members of the chloroplast-localized heat shock protein Hsp93 family that play crucial roles at early and late steps, respectively, during protein import into chloroplasts (Smith et al., 2004; Kovacheva et al., 2005; Su and Li, 2010). In wild-type protoplasts, both fusion proteins localized to chloroplasts; moreover, they were properly processed, as revealed by western-blot analysis with an anti-GFP antibody (Fig. 1, B and C). By contrast, in *ppi2* and *hsp93-V* mutant protoplasts, both RbcS-nt:GFP and Cab-nt:GFP localized in the cytosol as well as in chloroplasts (Fig. 1Bb, Bc, Be, and Bf), indicating that these mutant protoplasts show a defect in import of these reporter proteins. Consistent with this, both constructs yielded considerable amounts of unprocessed precursors in addition to the processed form, indicating that Toc159 and Hsp93-V play crucial roles in chloroplast import of both RbcS-nt:GFP and Cab-nt:GFP (Fig. 1, B–D). These results suggest that, despite the differences in their primary sequences and sequence motifs, RbcS-nt:GFP and Cab-nt:GFP are imported into chloroplasts by the same molecular machinery.

Functionality of Sequence Motifs Is Critically Dependent on Their Locations within TPs

Next, we examined how different sequence motifs are used in the same import machinery. The entire import process of chloroplast proteins can be divided into multiple steps, each of which involves one or more interactions between a selected translocon component and its corresponding sequence motifs in the TP (Li and Teng, 2013). Consistent with this idea, TPs contain specific binding motifs for Toc159, chloroplast heat shock protein70 (cpHsp70), or 14-3-3 (May and Soll, 2000; Smith et al., 2004; Lee et al., 2009a, 2013a; Chotewutmontri et al., 2012; Li and Teng, 2013). In addition,

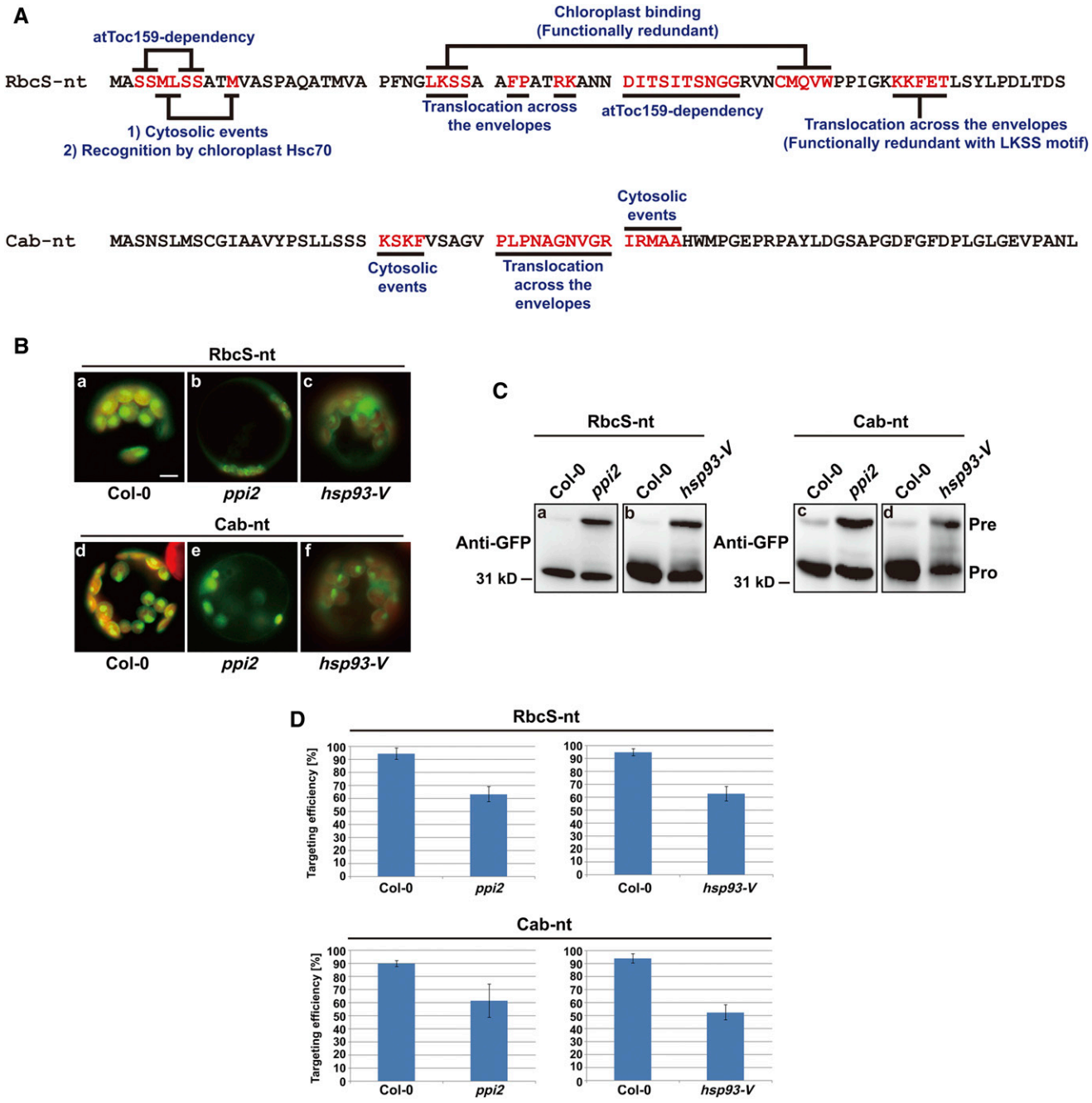

Figure 1. The TPs of RbcS and Cab share the import machinery for preprotein import into chloroplasts. A, Sequences of wild-type RbcS-nt and Cab-nt. Functional motifs are indicated in red and underlined. The activities of sequence motifs are described in blue. B, Localization of reporter proteins. Protoplasts from wild-type (Columbia-0 [Col-0]), *ppi2*, and *hsp93-V* plants were transformed with the indicated constructs, and GFP patterns were observed 12 h after transformation. Green, red, and yellow signals represent GFP, autofluorescence of chlorophyll, and the overlap between green and red signals, respectively. Scale bar = 20 μm. C, Western-blot analysis of the reporter proteins. Total protein extracts from transformed protoplasts were analyzed by western blotting with anti-GFP antibody. Pre, Precursor form; Pro, processed form. D, Targeting efficiency of RbcS-nt and Cab-nt in wild-type, *ppi2*, and *hsp93-V* plants, obtained by quantitation of western-blot data shown in (C). Signal intensity of protein bands was measured using software installed on the LAS3000 imager (FUJIFILM), and targeting efficiency was defined as the percentage of the processed form relative to the total amount of expressed protein. Three independent transformation experiments were performed; the data represent means with SDs.

a specific component of the translocon may interact with multiple motifs in different sequences (Li and Teng, 2013). Therefore, we hypothesized that a particular TP should contain one or more motifs corresponding to each of these translocon components, and that for each component, the interacting motif can be selected from among multiple potential motifs. This hypothesis introduces the concept that sequence motifs act as independent functional units, and is therefore similar to the multiselection and multiorder (M&M) model for generation of TPs proposed by Li and Teng (2013). To test this hypothesis, we generated various hybrid TPs by swapping regions of sequence between RbcS-nt and Cab-nt. In our previous studies, we identified critical sequence motifs in both RbcS-nt and Cab-nt (Fig. 1A). Based on this information, we divided the TPs into four blocks (namely, amino acid positions 1–22, 23–31, 32–41, and 42–79 or 42–80; Fig. 2A) and exchanged these blocks between RbcS-nt (R) and Cab-nt (C; Fig. 2A). The resulting hybrid TPs were fused to GFP and introduced into protoplasts to examine their ability to deliver proteins into chloroplasts (Fig. 2, B–F). Of these six reporters, C(1–22)R(23–79):GFP, C(1–31)R(32–79):GFP, C(1–41)R(42–79):GFP, and R(1–41)C(42–80):GFP localized primarily to chloroplasts and mainly produced the processed forms, indicating that they were efficiently imported into chloroplasts (Fig. 2, Bc and E). By contrast, R(1–22)C(23–80):GFP and R(1–31)C(32–80):GFP exhibited defects in protein import: R(1–22)C(23–80):GFP was partially imported into chloroplasts, with a significant proportion remaining in the cytosol (Fig. 2, Ba, C, and D), whereas R(1–31)C(32–80):GFP was present mainly in the cytosol as the unprocessed form (Fig. 2, Bb, C, and D). In general, hybrid TPs consisting of the N-terminal segment from Cab-nt and the C-terminal segment from RbcS-nt were more efficient in mediating protein import into chloroplasts than those with the inverse arrangement (Fig. 2, E and F). These results indicate that sequence motifs in TPs are largely, but not always, functional in different sequence contexts and thus raise the possibility that the motifs act as functional units independent of TPs where they originated.

The Function of Motifs Depends on Their Position and Sequence Context of TPs

Of the six hybrid TPs, two exhibited significant or severe defects in protein import efficiency despite the presence of motifs that are functional in their original TPs. One conceivable explanation is that artificial swapping of domains between RbcS-nt and Cab-nt resulted in improper placement of the sequence motifs in the hybrid TPs. Consistent with this idea, internal deletions in TPs inhibit protein import into chloroplasts (Rensink et al., 2000; Lee et al., 2006). To test this idea, we modified R(1–22)C(23–80) by inserting three residues (SSS, amino acids 20–22 of Cab-nt) to give R(1–22)C(20–80) (Fig. 3A). The import efficiency of R(1–22)C(20–80) was significantly improved relative to R(1–22)C(23–80), but it was still not as efficient as Cab-nt (Fig. 3, Ba, Bc, and C). We then further extended the Cab-nt domain by two additional residues (LL, amino acids 18–19 of Cab-nt) to generate R(1–22)C(18–80), which was as efficient as Cab-nt (Fig. 3, Bb, Bc, and C); thus, insertion of five residues fully restored import efficiency to the wild-type level (Fig. 3, B and C). Next, to determine whether the increase in import efficiency was caused by repositioning of sequence motifs contained in the C-terminal segment of Cab-nt, we replaced LLSSS, the residues inserted in R(1–22)C(18–80), with five alanines (Fig. 3D); as a control, we introduced the same mutation into Cab-nt. Ala substitution in R(1–22)C(18–80) significantly reduced import efficiency, but substitution of the same sequence in wild-type Cab-nt had no effect (Fig. 3, D–F), indicating that the specific amino acid sequence LLSSS, but not the five-residue insertion, was critical for the import efficiency of R(1–22)C(18–80). These results suggest that the motif LLSSS acts as a critical sequence motif for protein import into chloroplasts in the hybrid TP R(1–22)C(18–80), but not the wild-type Cab-nt, and therefore raise the possibility that the functionality of sequence motifs is dependent on local sequence context. Next, we attempted to rescue the import efficiency of R(1–31)C(32–80) (Fig. 4). As in the approach used for R(1–22)C(23–80), we inserted nine residues (KSKFVSAGV, corresponding to amino acids 23–31 of Cab-nt) to yield R(1–31)C(23–80) (Fig. 4A). R(1–31)C(23–80) was imported as efficiently as wild-type Cab-nt (Fig. 4, B–D). We then assessed the contribution of individual sequence motifs to import efficiency. The inserted segment, KSKFVSAGV, contains the essential sequence motif KSKF from Cab-nt (Lee et al., 2008), and another critical sequence motif, LKSS, from RbcS-nt is N terminal to that sequence (Fig. 1A; Lee et al., 2006). To investigate the role of the LKSS motif in this hybrid TP, R(1–31)C(23–80), during protein import, we generated two mutants by deleting three residues (SSA, amino acids 29–31 of RbcS-nt) or five residues (LKSSA, amino acids 27–31 of RbcS-nt) to yield R(1–28)C(23–80) and R(1–26)C(23–80), respectively (Fig. 4A). The import efficiency of both deletion-mutant hybrid TPs was similar to that of R(1–31)C(23–80), indicating that the motif LKSS of RbcS-nt no longer acted as a critical sequence motif in this context (Fig. 4, B–D). This observation raises the possibility that a new sequence motif might have substituted functionally for the LKSS motif in R(1–31)C(23–80). A previous report proposed that FNGLK residues (amino acids 24–28) in the RbcS TP, located just upstream of the LKSS motif, are semiconserved and postulated to be a potential binding site for Toc33/34, which is essential for chloroplast biogenesis and viability (Constan et al., 2004; Chotewutmontri et al., 2012; Li and Teng, 2013). To test whether the FNGLK motif is critical for protein import into chloroplasts in the hybrid TP, we replaced it with alanines (Fig. 4E); as a control, the same mutation was introduced into wild-type RbcS-nt. Ala substitution of this motif in R(1–28)C(23–80) caused a significant defect in protein import into chloroplasts, therefore, GFP

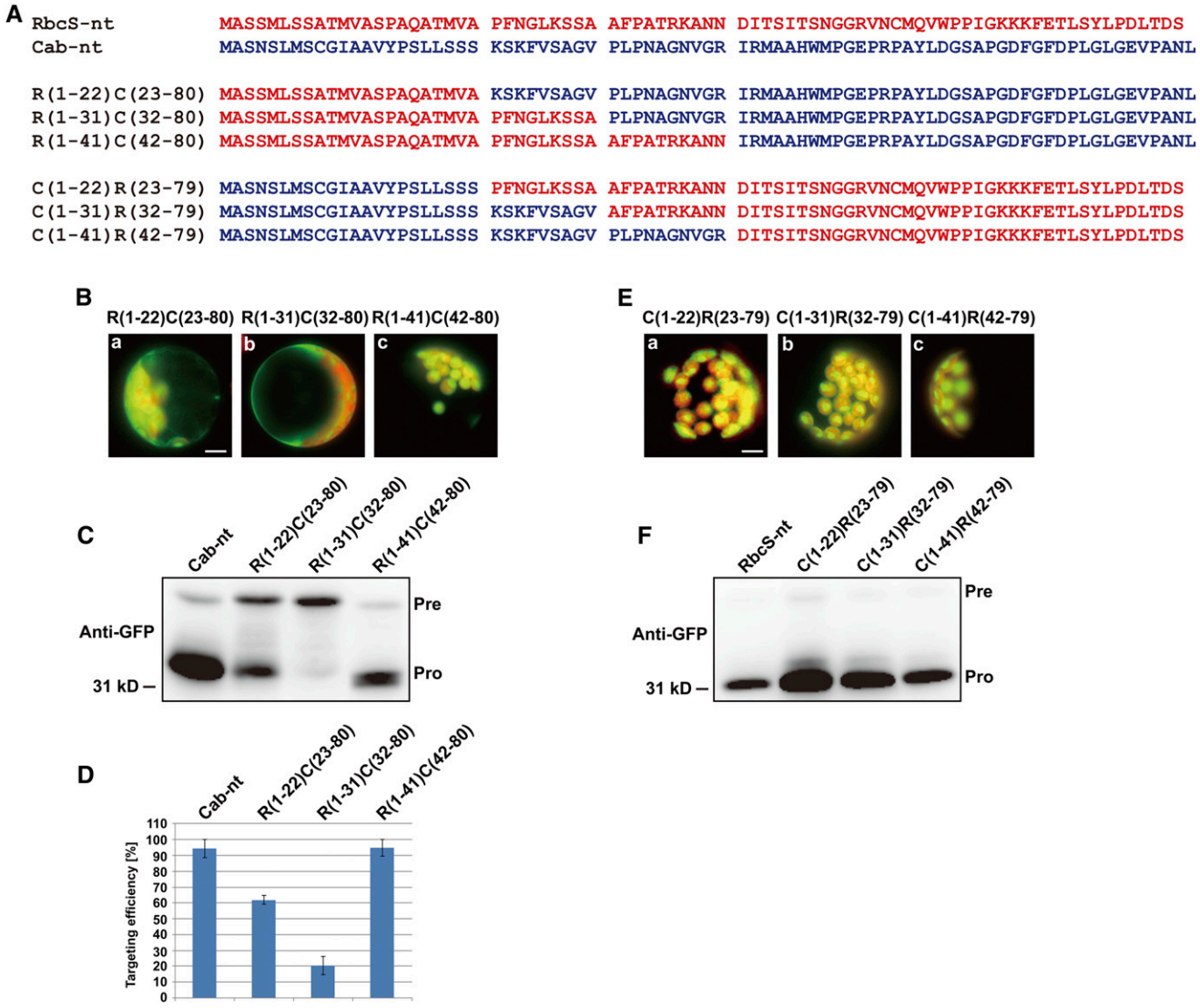

Figure 2. The sequence motifs of RbcS and Cab TPs are only partially interchangeable. A, Sequences of RbcS-nt, Cab-nt, and hybrid TPs. B and E, Localization of reporter proteins. Protoplasts from wild-type plants were transformed with the indicated constructs, and the GFP patterns were observed 12 h after transformation. Green, red, and yellow signals represent GFP, autofluorescence of chlorophyll, and the overlap between green and red signals, respectively. Scale bar = 20 μm. C and F, Western-blot analysis of reporter proteins. Protein extracts from protoplasts transformed with indicated constructs were analyzed by western blotting with anti-GFP antibody. Pre, Precursor form; Pro, processed form. D, Targeting efficiency of various domain-swapping mutants. Western-blot data shown in C were quantitated as described in Figure 1D.

signals were detected in both the cytosol and chloroplasts (Fig. 4, Fc, Gc, and H). However, the same mutation in wild-type RbcS-nt did not affect chloroplast targeting (Fig. 4Fa, G, and H), as reported previously (Lee et al., 2006). Next, we examined the role of the KSKF motif in the hybrid TP R(1–28)C(23–80) (Fig. 4E). In Cab-nt, the KSKF motif is essential for an early step in protein import (Lee et al., 2008). Consistent with this, the majority of Cab-nt[KSKF/4A] was present in the cytosol as the unprocessed form (Fig. 4, Fb and Gb). The Ala substitution of the KSKF motif in R(1–28)C(23–80) also caused a severe defect in protein import, but with improved targeting efficiency compared with Cab-nt [KSKF/4A], suggesting that the KSKF motif in the R (1–28)C(23–80) context may have partially redundant function with the FNGLK motif (Fig. 4, F–H). Taken together, these results support the idea that the functionality and activity of sequence motifs in TPs are critically dependent on their position in and sequence context of TPs. These motifs are thought to be involved in interactions with various translocon components either simultaneously or sequentially during protein import into chloroplasts. The motif-interacting translocon components are located at the specific locations along the route of protein import into chloroplasts, such as the surface of chloroplasts, outer and inner envelope

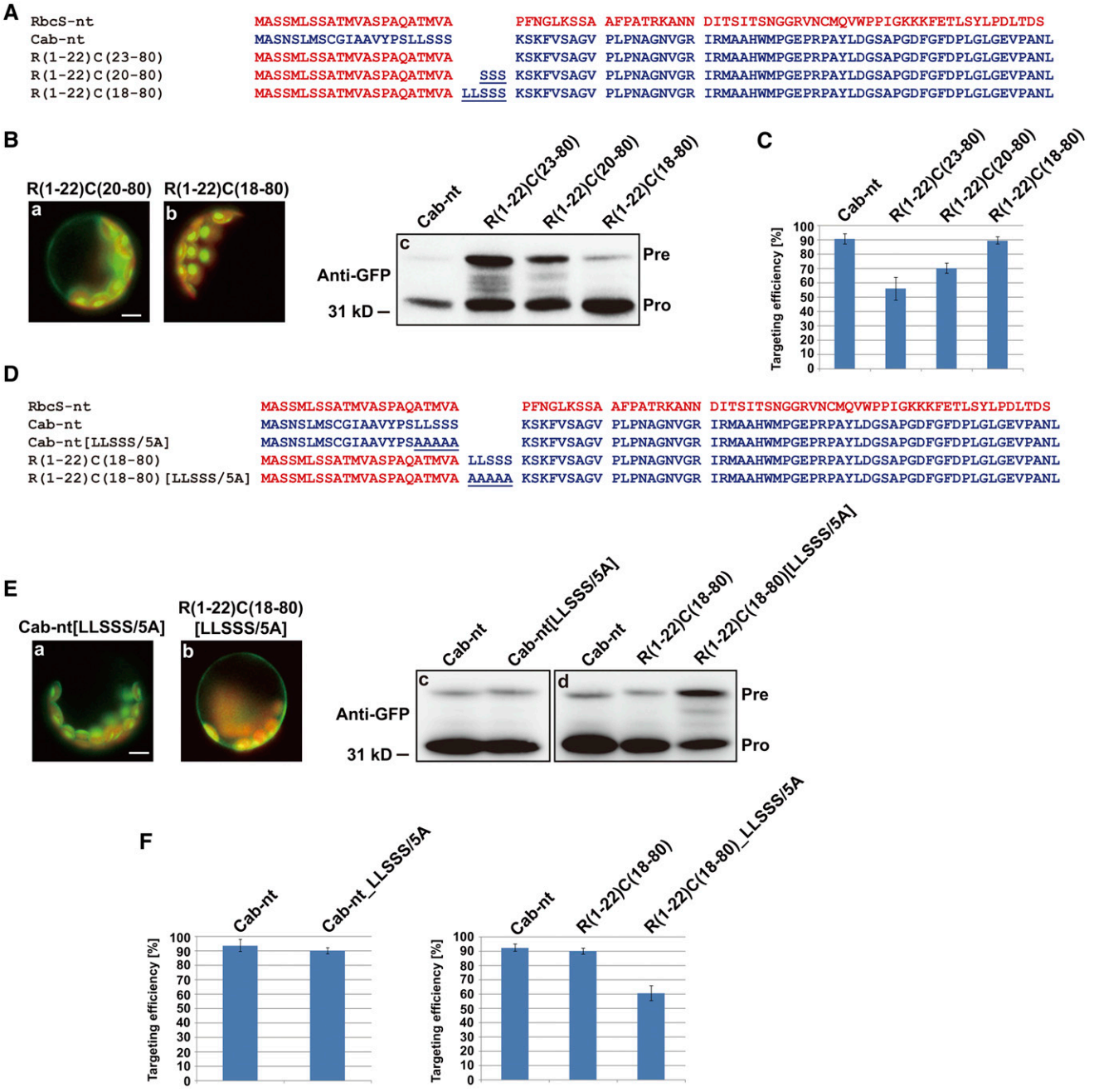

Figure 3. Functions of sequence motifs are critically dependent on their locations within transit peptides. A and D, Sequences of RbcS-nt, Cab-nt, and hybrid TPs. B and E, Localization of reporter proteins. Protoplasts were transformed with the indicated constructs, and the GFP patterns were observed 12 h after transformation. Green, red, and yellow signals represent GFP, autofluorescence of chlorophyll, and the overlap between green and red signals, respectively. Scale bar = 20 μm. Protein extracts from protoplasts transformed with indicated constructs were analyzed by western blotting with anti-GFP antibody. Pre, Precursor form; Pro, processed form. C and F, Western-blot data shown in B and E were quantitated as described in Figure 1D.

membranes, and stroma, and may have little mobility. Therefore, the position of these translocon components may dictate the position of sequence motifs in TPs for successful interaction between translocon components and their interacting motifs during protein import.

New Sequence Contexts Lead to Changes in Both Composition and Function of Active Sequence Motifs

The results described in Figure 4 revealed that active sequence motifs became inactive, and vice versa, in the new sequence context of hybrid TPs. This observation

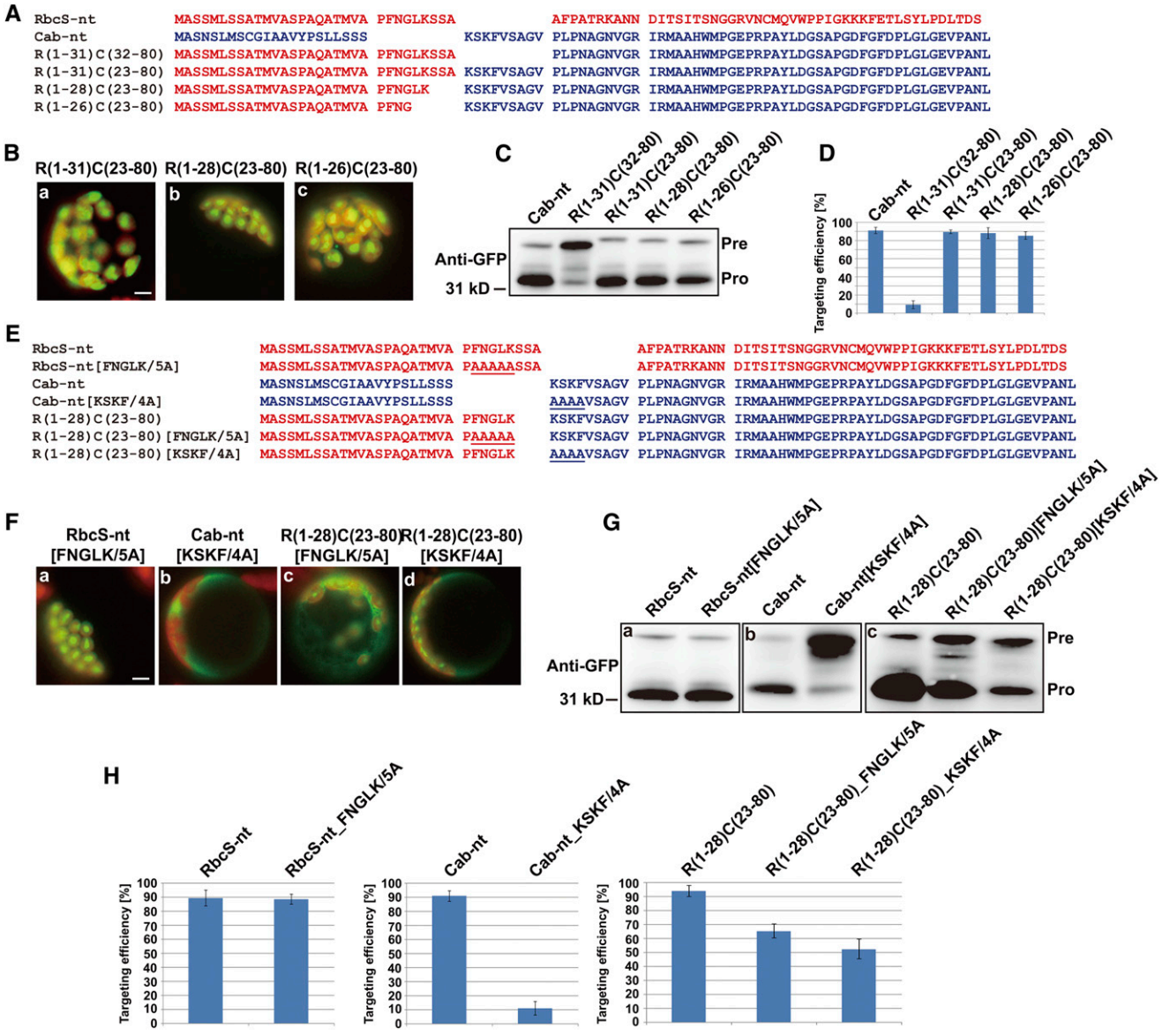

Figure 4. Functions of sequence motifs are critically dependent on the overall context of transit peptides. A and E, Sequences of RbcS-nt, Cab-nt, and hybrid TPs. B and F, Localization of reporter proteins. Protoplasts were transformed with the indicated constructs, and the GFP patterns were observed 12 h after transformation. Green, red, and yellow signals represent GFP, autofluorescence of chlorophyll, and the overlap between green and red signals, respectively. Scale bar = 20 μm. C and G, Western-blot analysis of reporter proteins. Protein extracts from protoplasts transformed with the indicated constructs were analyzed by western blotting with anti-GFP antibody. Pre, Precursor form; Pro, processed form. D and H, Western-blot data shown in C and G were quantitated as described in Figure 1D.

prompted an investigation of whether critical sequence motifs originally identified in RbcS-nt and Cab-nt are still active in the hybrid TPs. In two hybrid TPs, C(1–22)R(23–79) and C(1–31)R(32–79), we replaced sequence motifs identified in wild-type Cab-nt and RbcS-nt with Ala (Fig. 5A).

We included two motifs, FP and RK, from RbcS-nt and motif KSKF from Cab-nt, as well as two newly identified motifs, LLSSS from Cab-nt and FNGLK from RbcS-nt, in the two hybrid TPs (Figs. 1A and 5A). The resultant constructs were fused to GFP and introduced into protoplasts. In both hybrid TPs, C(1–22)R(23–79) and C(1–31)R(32–79), simultaneous Ala substitution of FP and RK caused a severe defect in protein import, indicating that these motifs were still functional in the hybrid TPs (Fig. 5, B–E). However, in the context of the hybrid TP C(1–22)R(23–79), Ala substitution of the newly identified sequence motifs LLSSS and FNGLK did not cause a significant defect in protein import. Similarly, in the hybrid TP C(1–31)R(32–79), Ala substitution of the motif KSKF, which is critical in wild-type Cab-nt, did not cause any significant defect in

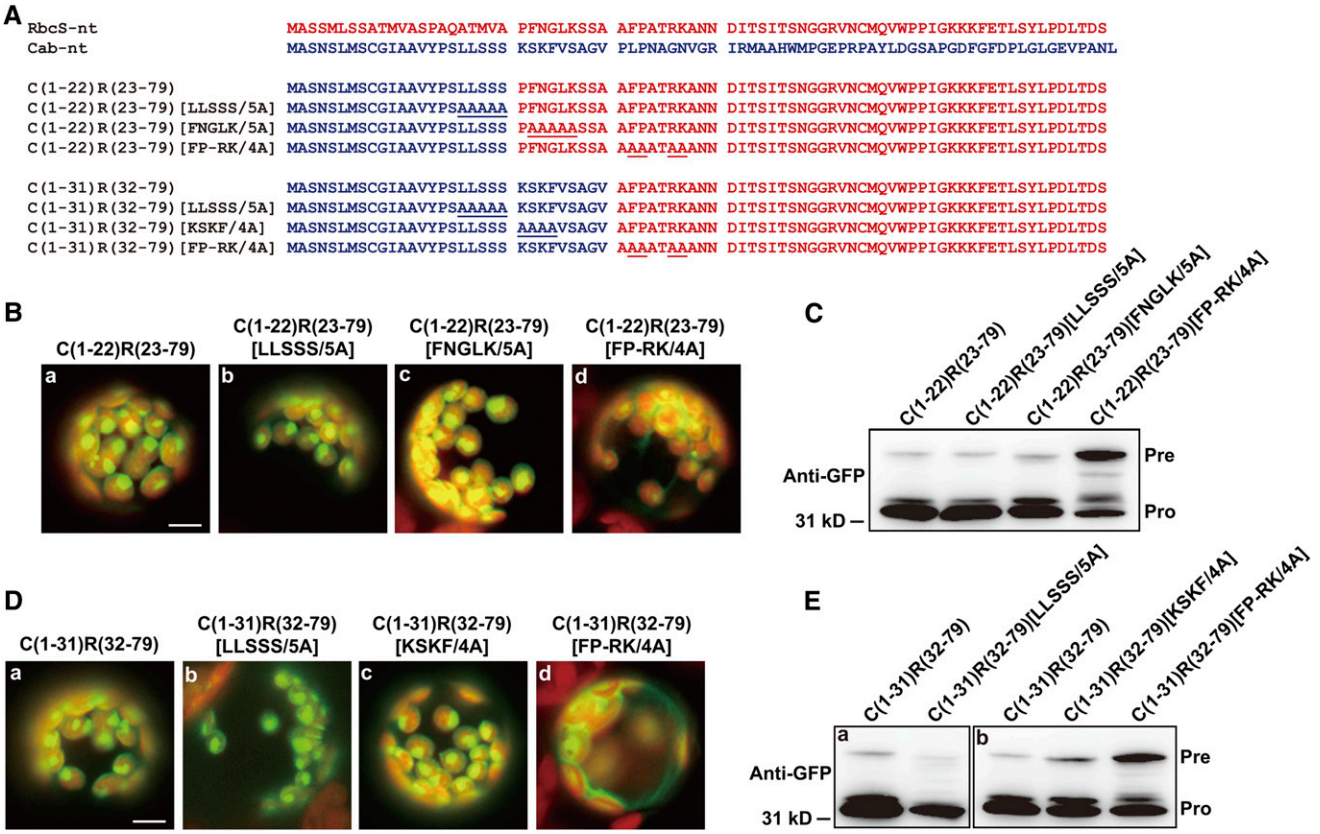

Figure 5. New sequence contexts of TPs lead to inactivation of critical sequence motifs in the original or other contexts. A, Sequences of RbcS-nt, Cab-nt, and Ala substitution mutants of hybrid TPs. B and D, Localization of reporter proteins. Protoplasts were transformed with indicated constructs and analyzed as described in Figure 1. C and E, Protein extracts from protoplasts transformed with indicated constructs were analyzed by western blotting with anti-GFP antibody. Pre, Precursor form; Pro, processed form.

protein import (Fig. 5, B–E), indicating that these motifs were not functional in the hybrid sequences C(1–22)R(23–79) and C(1–31)R(32–79). These results suggest that, in the context of C(1–22)R(23–79) and C(1–31)R(32–79), different sequence motifs, yet unidentified, may play crucial roles in protein import.

Finally, we focused on sequence motifs in the N-terminal regions of TPs. Previous studies have shown that the moderate hydrophobicity of the N-terminal region in TPs is important for an early cytosolic step, as well as for recognition by cpHsc70 in the stroma (Lee et al., 2006, 2008; Chotewutmontri et al., 2012; Chotewutmontri and Bruce, 2015). In RbcS-nt, residues ML (amino acids 5–6) and MV (amino acids 11–12) contribute to the moderate hydrophobicity of the N-terminal region. By contrast, Ala substitution of four hydrophobic residues (NLMC, amino acids 4, 6, 7, and 9) in Cab-nt did not cause any defect in chloroplast import (Lee et al., 2008). Instead of NLMC, we performed Ala substitution of the residues LMCI (amino acids 6, 7, 9, and 11) in wild-type Cab-nt and two hybrid TPs, C(1–22)R(23–79) and C(1–31)R(32–79), and examined the import efficiency of the resultant mutants (Fig. 6A). The LMCI/4A mutation in Cab-nt abrogated chloroplast import almost completely, indicating that the LMCI residues are responsible for the moderate hydrophobicity of the N-terminal region. However, the same mutation in the two hybrid TPs, C(1–22)R(23–79) and C(1–31)R(32–79), had almost no adverse effect on import efficiency (Fig. 6, B and C), indicating that the sequence motif responsible for the moderate hydrophobicity of the N-terminal region also depends on the sequence context. It is possible that other hydrophobic residues may substitute functionally for LMCI to provide the required hydrophobicity in this region.

Next, we examined whether the lower import efficiency of some of the hybrid sequence is caused by proteasome-mediated proteolytic degradation of precursors in the cytosol. Previous studies have shown that unimported precursors are subject to the proteasome-mediated proteolytic degradation in the cytosol to prevent cytotoxicity to the cell (Lee et al., 2009b; Bischof et al., 2011; Grimmer et al., 2014; Tillmann et al., 2015). Moreover, many chloroplast precursor proteins are known to be modified by N-terminal acetylation in the cytosol, which may possibly mediate the proteasomal degradation in the cytosol (Hwang et al., 2010; Bischof et al., 2011; Grimmer et al., 2014). To test the possibility of proteasomal degradation of unimported hybrid preproteins, four defective hybrid TPs, *R*(*1–22*)*C*(*23–80*),

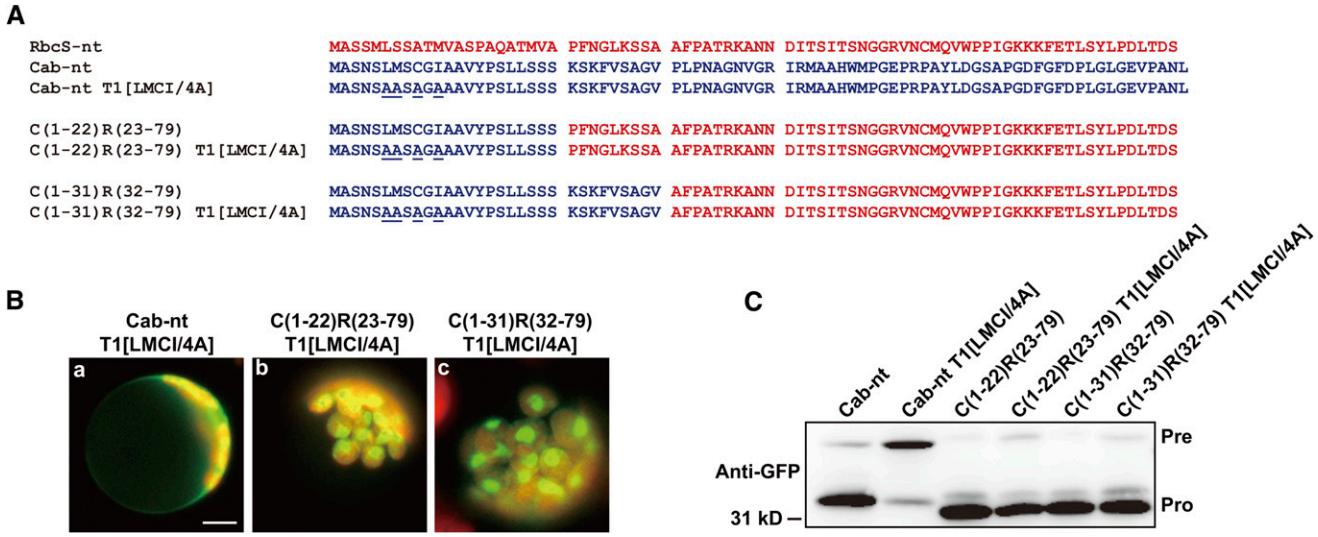

Figure 6. The functionality of the N-terminal hydrophobic residues in Cab TP can be dramatically changed depending on the contexts. A, Sequences of RbcS-nt, Cab-nt, and Ala substitution mutants of hybrid TPs. B, Localization of reporter proteins. Protoplasts were transformed with the indicated constructs and analyzed as described in Figure 1. C, Protein extracts from protoplasts transformed with indicated constructs were analyzed by western blotting using anti-GFP antibody. Pre, Precursor form; Pro, processed form.

R(*1–31*)*C*(*32–80*), *R*(*1–28*)*C*(*23–80*)[*FNGLK/5A*], and *R*(*1–28*)*C*(*23–80*)[*KSKF/4A*], and an efficient TP, *R*(*1–28*)*C*(*23–80*), were transformed into protoplasts (Supplemental Fig. S1). At 8 h after transformation, protoplasts were treated with dimethyl sulfoxide (DMSO) or MG132 and incubated for an additional 6 h. In this experiment, *RbcS-nt*[*T4A/T7A*] was included as a positive control for proteasome-mediated proteolytic degradation (Lee et al., 2006), and *Cab-nt*:*monomeric red fluorescent protein* was cotransformed as a negative control for the MG132 treatment. All four preproteins containing defective TPs were degraded at varying degrees, which is in agreement with the previous report showing that the stability of chloroplast preproteins in the cytosol can vary depending on the precursors (Supplemental Fig. S1; Grimmer et al., 2014). However, their targeting efficiencies were not improved in the presence of MG132, indicating that the low targeting efficiencies of these four hybrid TPs are not caused by the cytosolic quality control (Supplemental Fig. S1).

Hybrid TPs Have Different Targeting Efficiencies in *ppi2* and *hsp93-V* Protoplasts Depending on the Sequence Context

Protein import into chloroplasts is mediated by the interaction between sequence motifs in TPs and components located in the cytosol and outer/inner envelopes of chloroplasts, namely, Toc/Tic (translocon at the inner envelope of chloroplasts) translocons (Li and Chiu, 2010; Li and Teng, 2013). Thus, we examined the import efficiency of hybrid TPs with different sequence contexts in *ppi2* and *hsp93-V* protoplasts (Fig. 7). In *ppi2* protoplasts, the import efficiency of all of the hybrid constructs was decreased, similar to their parental TPs RbcS-nt and Cab-nt:GFP (Fig. 7, A and B), indicating that these hybrid TPs are also imported through an atToc159-containing translocon, similar to their parental TPs. However, the degree of the decrease varied among these hybrid TPs. In *ppi2* protoplasts, the import efficiency of C(1–22)R(23–79) and C(1–31)R(32–79) was higher than that of wild-type RbcS-nt and Cab-nt at both early and late time points (Fig. 7, A and B), whereas in the import efficiency of R(1–22)C(18–80) and R(1–28)C(23–80) was higher than that of Cab-nt and RbcS-nt at early and late time points, respectively (Fig. 7, A and B). This indicates that the import efficiency of TPs depends on the sequence context.

Next, we examined the import efficiency of these hybrid TPs in *hsp93-V* protoplasts. Unlike in *ppi2* protoplasts, in *hsp93-V* protoplasts, the import efficiency of R(1–22)C(18–80) and R(1–28)C(23–80) was nearly identical to that of Cab-nt but lower than that of RbcS-nt (Fig. 7, C and D). On the other hand, the import efficiency of C(1–22)R(23–79) and C(1–31)R(32–79) was higher than that of RbcS-nt and Cab-nt in *hsp93-V* protoplasts (Fig. 7, E and F), again confirming that the import efficiency depends on the sequence contexts of TPs. It is possible that the sequence contexts of TPs as well as the availability of translocon components influence the mode of interaction between motifs and translocon components.

An Efficient SynTP Can Be Generated from an Unrelated Sequence by Incorporating Critical Sequence Motifs from RbcS or Cab TPs

The fact that various sequence motifs in TPs are independent functional units capable of interacting with

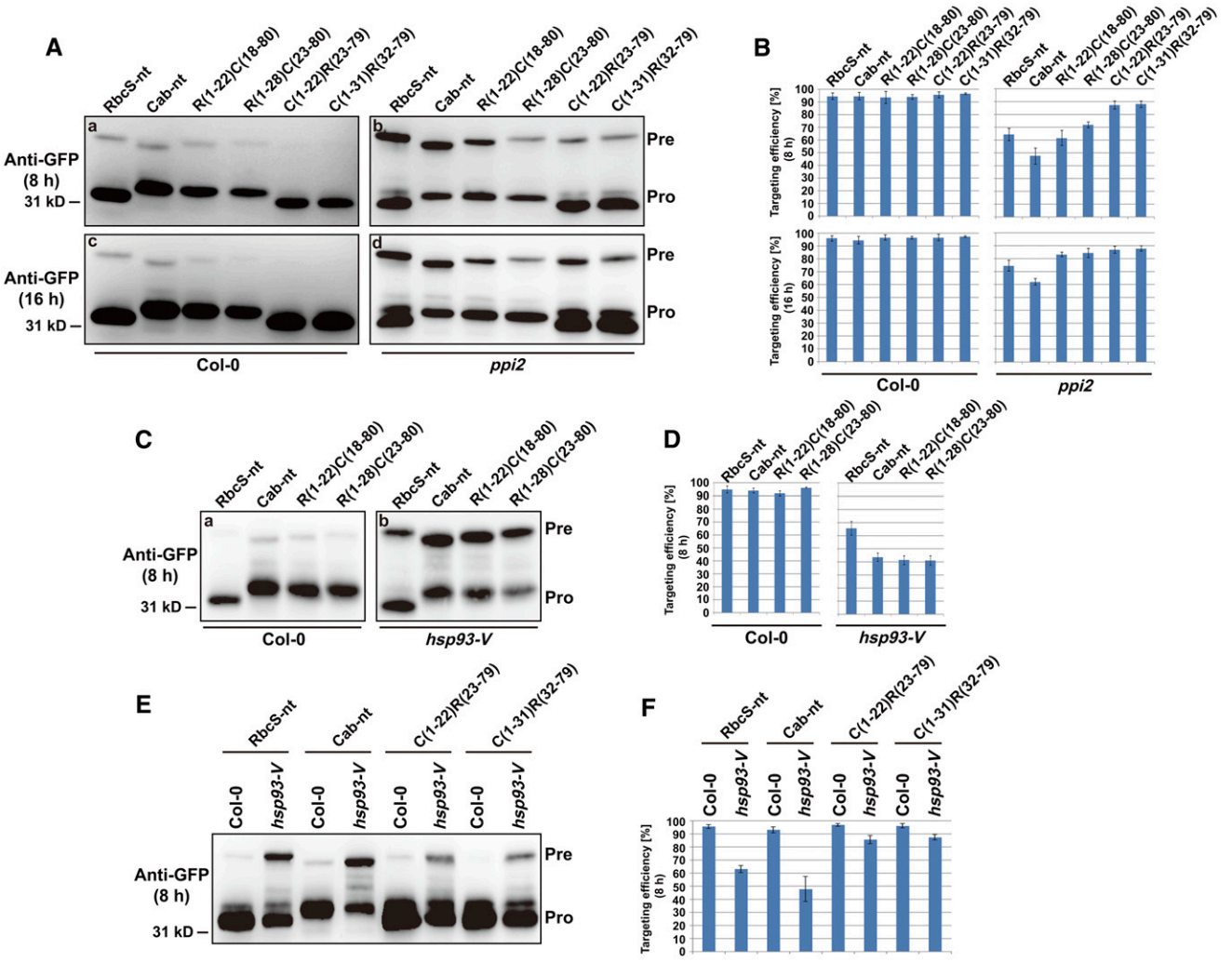

Figure 7. TPs with different sequence contexts exhibit differences in their import into *ppi2* and *hsp93-V* protoplasts. A, C, and E, Western-blot analyses of reporter proteins. Protein extracts from wild-type (Columbia-0 [Col-0]), *ppi2*, and *hsp93-V* mutant protoplasts transformed with the indicated constructs were analyzed by western blotting using an anti-GFP antibody. Pre, Precursor form; Pro, processed form. B, D, and F, Western-blot data shown in A, C, and E were quantitated as described in Figure 1D.

their cognate translocon components (Figs. 3–6) raises the intriguing possibility that diverse TPs evolved such that each TP contains a unique set of translocon component-interacting motifs. This idea is similar to the concept proposed by the M&M model (Li and Teng, 2013). To test this idea, we first introduced two critical sequence motifs, FNGLK and FP/RK, of the RbcS TP into the N-terminal 80-amino acid segment (CPY80) of carboxypeptidase Y (CPY), a vacuolar protein that is targeted to the ER via a signal recognition particle-dependent pathway and finally transported to the vacuole via vesicle trafficking through the Golgi apparatus and prevacuolar compartment (Lee et al., 2013b). As expected, CPY[1–95] delivered GFP to the central vacuole, but not to the chloroplast (Fig. 8, A, Ba, and Bb). Next, we removed the outermost N-terminal charged residues EK by substituting them with AS (commonly present in RbcS and Cab TPs), and we substituted four acidic residues (E or D) with alanines (indicated in blue characters) because TPs do not favor charged residues in the N-terminal region or negatively charged residues (Bruce, 2000; Bhushan et al., 2006). Finally, we substituted 10 hydrophobic residues in the leader sequence (underlined) of CPY with alanines to prevent ER targeting. However, the SynTP was not able to efficiently deliver proteins into chloroplasts (Fig. 8, A–D). Next, we incorporated the motif LLSSS of the Cab TP or the corresponding region of the RbcS TP to generate SynTP[LLSSS] and SynTP[ATMVASPA], respectively. These two SynTPs were able to deliver proteins into chloroplasts (Fig. 8, A–D). These results clearly show that specific sequence motifs from functional TPs can be transferred to a completely unrelated sequence to confer the ability to deliver proteins into chloroplasts.

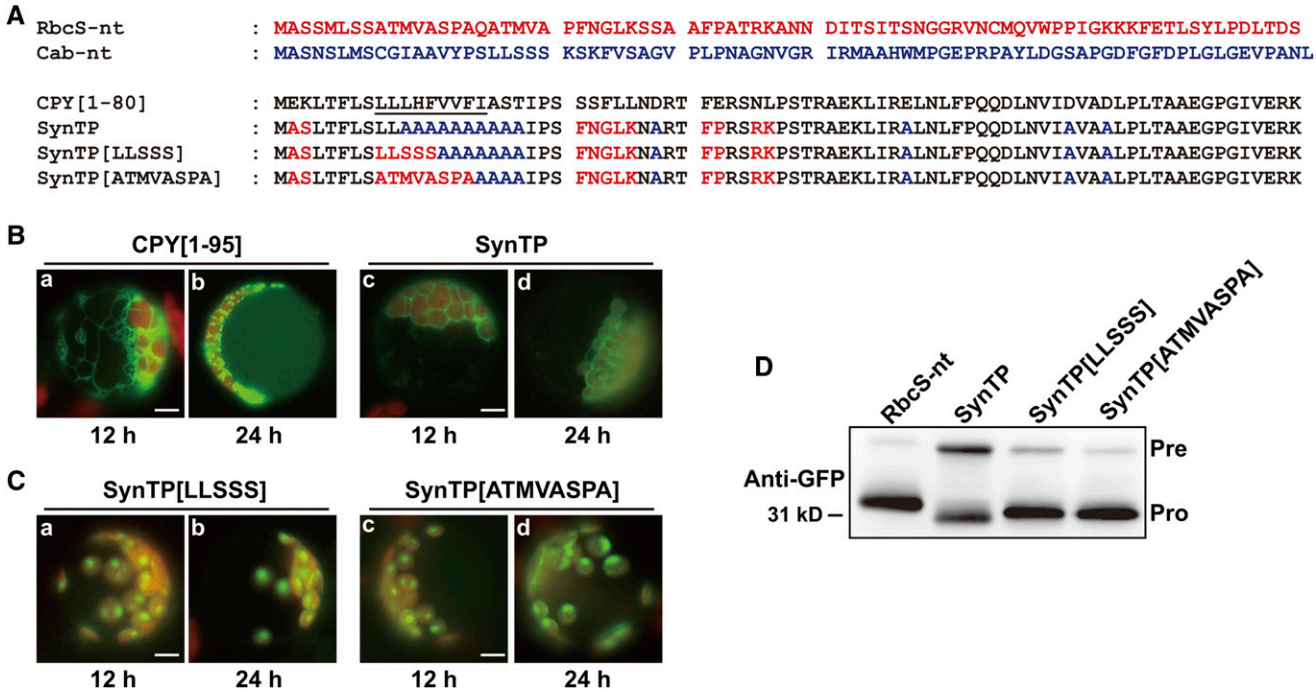

Figure 8. SynTPs generated using critical sequence motifs from RbcS and Cab TPs efficiently deliver proteins into chloroplasts. A, Sequences of RbcS-nt, Cab-nt, and SynTPs. B and C, Localization of reporter proteins. Protoplasts were transformed with the indicated constructs and analyzed as described in Figure 1. D, Protein extracts from protoplasts transformed with the indicated constructs were analyzed by western blotting using an anti-GFP antibody. Pre, Precursor form; Pro, processed form.

DISCUSSION

In this study, we provide compelling evidence that the sequence motifs are functional units independent of any specific TPs from which they originated. This conclusion is based on the finding that the majority of hybrid TPs, generated by domain swapping at three different arbitrarily selected positions between two different TPs (RbcS-nt and Cab-nt), were as efficient in chloroplast protein import as their parental wild-type TPs (Fig. 2). The process of protein import into chloroplasts can be divided into multiple independent steps: navigation through the cytosol (May and Soll, 2000; Qbadou et al., 2006; Li and Chiu, 2010; Fellerer et al., 2011; Lee et al., 2013a), recognition by receptors at the surface of chloroplasts (Becker et al., 2004; Smith et al., 2004; Li and Chiu, 2010), translocation through the outer and inner membranes (Rensink et al., 2000; Kikuchi et al., 2009, 2013; Li and Chiu, 2010), and possibly the pulling step by chaperones at the stroma (Chou et al., 2006; Li and Chiu, 2010; Shi and Theg, 2010; Su and Li, 2010; Chotewutmontri et al., 2012; Inoue et al., 2013; Liu et al., 2014). Each of these steps requires one or more specific factors to mediate the import process. These factors include cytosolic factors, import receptors at the surface of chloroplasts, channel proteins at both chloroplastic inner and outer membranes, and stromal proteins (Jarvis, 2008; Li and Chiu, 2010). Translocon components may bind to TPs using their cognate sequence motifs. Indeed, this notion is consistent with previous studies showing that cpHsp70 and the import receptors Toc159 and Toc33 bind to specific motifs in the TPs (Becker et al., 2004; Chotewutmontri et al., 2012). At the inner membrane, the TPs that emerge from the Tic translocon bind directly to Tic110 (Chou et al., 2006; Tsai et al., 2013). The TPs released from Tic110 by the action of Tic40 interact with several stromal chaperones, such as Heat shock cognate70 (Hsc70), Hsp93, and possibly Hsp90C, during or after translocation (Shi and Theg, 2010; Su and Li, 2010; Rosano et al., 2011; Chotewutmontri et al., 2012; Inoue et al., 2013; Liu et al., 2014). Cytosolic factors, such as 14-3-3 and Hsp70, also bind to specific sites in the TP and facilitate import into chloroplasts (May and Soll, 2000). Thus, TPs may contain at least one motif for each of these translocon components for efficient import. Together with the notion that sequence motifs are independent functional units, this idea may constitute a key concept to explain how such diverse TPs evolved as the targeting signal of plastid proteins. Based on these ideas, we were able to design SynTPs from a completely unrelated sequence using a few sequence motifs from RbcS and Cab TPs (Fig. 8).

Another important finding of our study is that the activity and/or functionality of sequence motifs are dependent on the sequence context of TPs where they are located. One example is the semiconserved FNGLK motif, a potential Toc34-binding motif in RbcS-nt

(Chotewutmontri et al., 2012; Li and Teng, 2013); this motif is not critical for import into chloroplasts in its original context (Fig. 4), but in the context of R(1–28)C (23–80), this motif was essential for import (Fig. 4). Another example is the KSKF motif in Cab-nt; this motif is essential for an early cytosolic step in chloroplast targeting (Lee et al., 2008), but it became less critical in the context of R(1–28)C(23–80) (Fig. 4) and became inactive in the context of C(1–31)R(32–79) (Fig. 5). RbcS-nt and Cab-nt have differences not only in the primary sequence of their sequence motifs, but also in the organization of, and the relationship among, these motifs: RbcS-nt sequence motifs have a high degree of functional redundancy and are often involved in additive and synergistic relationships, whereas the sequence motifs in Cab-nt are more independent of one another (Fig. 1A; Lee et al., 2006, 2008). Because of the complexities in the relationships among sequence motifs in TPs, the hybrid TPs, which were generated by domain swapping at arbitrary positions between RbcS-nt (which has motifs that interact with one another) and Cab-nt (which has more independent motifs), may not easily recapitulate the relationship the motifs had in their native context. Instead, in the hybrid TPs, the motifs originating from the two different TPs may engage in unique relationships. Thus, one possibility is that the sequence motifs in the hybrid TPs may be reassembled or reorganized using whatever sequence motifs are available in the two domains. Reanalysis of the hybrid TPs confirmed this idea and revealed that the reassembly of a set of sequence motifs in the hybrid TPs appears not to be straightforward. In the hybrid TPs, the wild-type activity (or functionality) of a sequence motif had one of three different fates in the new sequence context: sequence motifs active in the original TPs remained active, sequence motifs originally active became inactive, or latent sequence motifs in the original TPs became active. Therefore, an important conclusion of this study is that the activity and functionality of the sequence motifs we investigated are not inherent, but can change depending on the sequence context of the TP.

What is the underlying cause of changes in the activity or functionality of sequence motifs in new hybrid TPs? In the process of reassembling a new set of sequence motifs in hybrid TPs, whether a particular sequence motif is included should depend on whether it can make a successful interaction with a specific component of the translocon as part of the entire import process, thereby contributing to the protein import. Thus, the import process should be considered as a series of sequential interactions between TP sequence motifs and translocon components (or cytosolic factors) that continue until the precursor has been completely translocated into the chloroplast. Moreover, at any given time during translocation, multiple interactions may occur. Consequently, an important factor is the spacing between potential sequence motifs, especially when the two domains are randomly fused to each other. Indeed, it has been proposed that sequence motifs must be arranged with proper spacing in the TPs to facilitate interaction(s) with translocon components (located at various sites within the chloroplast), as well as with cytosolic factors (Bruce, 2000). Consistent with this, an internal deletion in TPs causes severe defects in the import efficiency (Pilon et al., 1995; Lee et al., 2002, 2006). The fact that the defective hybrid can be rescued by insertion of a few amino acid residues at the junction is also consistent with this notion. The spacing between sequence motifs may be determined by the distance between translocon components, and their corresponding cognate sequence motifs, within the chloroplast. Moreover, it has been proposed that the length of TPs is determined by the thickness of the two chloroplast envelope membranes (Bionda et al., 2010; Chotewutmontri et al., 2012). The translocon components are located at particular locations within the chloroplasts (Jarvis, 2008; Li and Chiu, 2010) and can be considered fixed, with a limited range of mobility. These translocon components may engage in simultaneous or sequential interactions using their cognate sequence motifs en route to the stroma. Therefore, the distance between translocon components may be a key determining factor in selecting the interacting motifs in the TPs. For example, if a sequence motif is located too close or too far from another motif, it may not be selected due to the difficulty of simultaneous or sequential interactions.

TPs exhibit a great deal of flexibility and diversity in sequence motif composition. Accordingly, one fundamental question is how this can occur. One possible explanation is that a particular translocon component may bind to more than one sequence motif. Consistent with this notion, Toc159 binds to many different TPs with very diverse primary sequences (Smith et al., 2004). In ER-to-Golgi trafficking, Sec24, a component of COPII vesicles acting as a cargo receptor, contains three separate cargo-binding sites (Miller et al., 2003). Another possibility is that a translocon component interacts with the main characteristics of amino acids in a certain region of TPs, but does not bind to the motifs with a specific sequence, as in the case of the signal recognition particle that binds to the hydrophobic leader sequences. For example, cpHsc70 interacts with the N-terminal motif of TPs, which has a moderate hydrophobicity conferred by a few hydrophobic amino acids (Lee et al., 2008; Chotewutmontri et al., 2012; Chotewutmontri and Bruce, 2015). A recent study by Teng et al. (2012) also supports this notion; they showed that preproteins with two consecutive basic residues are actively imported into older chloroplasts, whereas those with two consecutive acidic residues are not. The two mechanisms may not be mutually exclusive, but are utilized depending on the situation or the specific translocon component(s) involved. If a particular translocon component can interact with more than one motif, diverse TPs can be generated by selectively assembling a TP-specific set of translocon component-interacting motifs from a large potential pool of such motifs. The flexibility and diversity may also be increased by the existence of multiple pathways that

involve different components during protein delivery into chloroplasts (Ivanova et al., 2004; Kessler and Schnell, 2009; Lee et al., 2009a; Andrès et al., 2010; Li and Chiu, 2010; Su and Li, 2010; Bischof et al., 2011; Li and Teng, 2013). In protein import mediated by RbcS and Cab TPs, translocon components such as Toc159 and Hsp93-V play crucial roles in various steps. Protein import was severely impaired in *ppi2* and *hsp93-V* mutant plants (Figs. 1 and 7). However, in *ppi2* and *hsp93-V* mutant protoplasts, the import efficiency of some preproteins with hybrid TPs was significantly higher than that of wild-type RbcS and Cab TPs, despite the fact that these hybrid TPs were generated using domains of RbcS and Cab TPs (Fig. 7). One possibility is that RbcS and Cab TPs contain motifs that are active in at Toc159- or Hsp93-independent pathways. Indeed, RbcS and Cab TPs can support protein import into chloroplasts even in *ppi2* protoplasts, albeit at a lower efficiency (Lee et al., 2009a; Su and Li, 2010; Bischof et al., 2011; Inoue et al., 2013). These results suggest that TPs with different sequence contexts use different pathways and translocon components during protein import.

Finally, we demonstrated that SynTPs with import efficiencies comparable with that of the RbcS TP can be generated with a completely unrelated sequence using a few sequence motifs from RbcS and/or Cab TPs (Fig. 8). One possible scenario for TP evolution is that a sequence was converted into a functional TP by gradually incorporating sequence motifs until it included at least one motif randomly selected from multiple potential motifs for each type of translocon component to attain the maximal import efficiency. As a result, diverse TPs can be generated such that each contains a unique set of motifs. This idea is similar to the M&M model recently proposed by Li and Teng (2013) and may also be consistent with the exon shuffling model, in which TPs originated from the acquisition of several preexisting exons (Quigley et al., 1988; Gregerson et al., 1994; Long et al., 1996; Bruce, 2001), and the design principle of eukaryotic gene promoters, which use small sequence elements as binding sites for transcription factors.

MATERIALS AND METHODS

Growth of Plants

Arabidopsis (*Arabidopsis thaliana*; Columbia-0) was grown in a growth chamber at 22°C to 23°C with a 16-h-light/8-h-dark cycle on Gamborg B5 (Duchefa) plates. Leaf tissues were harvested from 2- to 3-week-old plants and used immediately for protoplast preparation.

Construction of Reporter Constructs

The swapping and Ala substitution mutant constructs were generated by PCR-based mutagenesis as described previously (Lee et al., 2006, 2008, 2009a). The sequences of the primers used to construct reporter constructs are shown in Supplemental Table S1. To generate *CPY*[*1–95*]:*GFP*, *CPY:GFP* (Lee et al., 2013b) was digested with *Xba*I/*Sal*I, and the digested DNA fragment corresponding to *CPY*[*1–95*] was ligated to an expression vector digested with the same restriction endonucleases. To generate SynTPs, we performed four successive PCR reactions with four forward primers (in the order of Syn-1, Syn-2, Syn-3, and Syn-4; Supplemental Table S1) and nosT-B primer (reverse primer) using *RbcS-nt*:*GFP* as template. After the fourth round of PCR, PCR products were digested with *Xba*I/*Xho*I and ligated to an expression vector digested with the same restriction endonucleases.

Polyethylene Glycol-Mediated Transformation and in Vivo Targeting of Reporter Proteins

Plasmid DNA used for polyethylene glycol-mediated transformation was purified using Qiagen columns. DNA was introduced into Arabidopsis protoplasts by the polyethylene glycol-mediated transformation method, as described previously (Jin et al., 2001; Lee and Hwang, 2011). Images of transformed protoplasts were acquired, as described previously (Jin et al., 2001; Lee and Hwang, 2011).

DMSO or MG132 Treatment of Transformed Protoplasts

Protoplasts transformed with plasmid DNA were incubated for 8 h. Subsequently, the transformed protoplasts were treated with DMSO or 20 μM MG132 for 6 h. Protein extracts from the transformed protoplasts were analyzed by western blotting with anti-GFP and red fluorescent protein antibodies.

Sequence data from this article can be found in the Arabidopsis Genome Initiative or GenBank/EMBL data libraries under the following accession numbers: *RbcS*, At1g67090; *Cab*, At3g54890.

Supplemental Data

The following supplemental materials are available.

Supplemental Figure S1. Proteasome-mediated degradation of preproteins.

Supplemental Table S1. Primer sequences used to generate the swapping and Ala substitution mutant constructs.

ACKNOWLEDGMENTS

The *hsp93-V* knockout mutant seeds were provided by Hsou-min Li (Institute of Molecular Biology, Academia Sinica, Taipei, Taiwan).

Received June 4, 2015; accepted July 2, 2015; published July 6, 2015.

LITERATURE CITED

Andrès C, Agne B, Kessler F (2010) The TOC complex: preprotein gateway to the chloroplast. Biochim Biophys Acta **1803:** 715–723

Bauer J, Chen K, Hiltbunner A, Wehrli E, Eugster M, Schnell D, Kessler F (2000) The major protein import receptor of plastids is essential for chloroplast biogenesis. Nature **403:** 203–207

Becker T, Jelic M, Vojta A, Radunz A, Soll J, Schleiff E (2004) Preprotein recognition by the Toc complex. EMBO J **23:** 520–530

Bhushan S, Kuhn C, Berglund AK, Roth C, Glaser E (2006) The role of the N-terminal domain of chloroplast targeting peptides in organellar protein import and miss-sorting. FEBS Lett **580:** 3966–3972

Bionda T, Tillmann B, Simm S, Beilstein K, Ruprecht M, Schleiff E (2010) Chloroplast import signals: the length requirement for translocation in vitro and in vivo. J Mol Biol **402:** 510–523

Bischof S, Baerenfaller K, Wildhaber T, Troesch R, Vidi PA, Roschitzki B, Hirsch-Hoffmann M, Hennig L, Kessler F, Gruissem W, et al. (2011) Plastid proteome assembly without Toc159: photosynthetic protein import and accumulation of *N*-acetylated plastid precursor proteins. Plant Cell **23:** 3911–3928

Bruce BD (2000) Chloroplast transit peptides: structure, function and evolution. Trends Cell Biol **10:** 440–447

Bruce BD (2001) The paradox of plastid transit peptides: conservation of function despite divergence in primary structure. Biochim Biophys Acta **1541:** 2–21

Chotewutmontri P, Bruce BD (2015) Non-native, N-terminal Hsp70 molecular motor recognition elements in transit peptides support plastid protein translocation. J Biol Chem **290:** 7602–7621

Chotewutmontri P, Reddick LE, McWilliams DR, Campbell IM, Bruce BD (2012) Differential transit peptide recognition during preprotein binding and translocation into flowering plant plastids. Plant Cell **24:** 3040–3059

Chou ML, Chu CC, Chen LJ, Akita M, Li HM (2006) Stimulation of transit-peptide release and ATP hydrolysis by a cochaperone during protein import into chloroplasts. J Cell Biol **175:** 893–900

Constan D, Patel R, Keegstra K, Jarvis P (2004) An outer envelope membrane component of the plastid protein import apparatus plays an essential role in Arabidopsis. Plant J **38:** 93–106

Fellerer C, Schweiger R, Schöngruber K, Soll J, Schwenkert S (2011) Cytosolic HSP90 cochaperones HOP and FKBP interact with freshly synthesized chloroplast preproteins of Arabidopsis. Mol Plant **4:** 1133–1145

Gregerson RG, Miller SS, Petrowski M, Gantt JS, Vance CP (1994) Genomic structure, expression and evolution of the alfalfa aspartate aminotransferase genes. Plant Mol Biol **25:** 387–399

Grimmer J, Rödiger A, Hoehenwarter W, Helm S, Baginsky S (2014) The RNA-binding protein RNP29 is an unusual Toc159 transport substrate. Front Plant Sci **5:** 258

Hessa T, Kim H, Bihlmaier K, Lundin C, Boekel J, Andersson H, Nilsson I, White SH, von Heijne G (2005) Recognition of transmembrane helices by the endoplasmic reticulum translocon. Nature **433:** 377–381

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

Inoue H, Li M, Schnell DJ (2013) An essential role for chloroplast heat shock protein 90 (Hsp90C) in protein import into chloroplasts. Proc Natl Acad Sci USA **110:** 3173–3178

Ivanova Y, Smith MD, Chen K, Schnell DJ (2004) Members of the Toc159 import receptor family represent distinct pathways for protein targeting to plastids. Mol Biol Cell **15:** 3379–3392

Jarvis P (2008) Targeting of nucleus-encoded proteins to chloroplasts in plants. New Phytol **179:** 257–285

Jin JB, Kim YA, Kim SJ, Lee SH, Kim DH, Cheong GW, Hwang I (2001) A new dynamin-like protein, ADL6, is involved in trafficking from the *trans*-Golgi network to the central vacuole in Arabidopsis. Plant Cell **13:** 1511–1526

Kessler F, Schnell D (2009) Chloroplast biogenesis: diversity and regulation of the protein import apparatus. Curr Opin Cell Biol **21:** 494–500

Kikuchi S, Bédard J, Hirano M, Hirabayashi Y, Oishi M, Imai M, Takase M, Ide T, Nakai M (2013) Uncovering the protein translocon at the chloroplast inner envelope membrane. Science **339:** 571–574

Kikuchi S, Oishi M, Hirabayashi Y, Lee DW, Hwang I, Nakai M (2009) A 1-megadalton translocation complex containing Tic20 and Tic21 mediates chloroplast protein import at the inner envelope membrane. Plant Cell **21:** 1781–1797

Kovacheva S, Bédard J, Patel R, Dudley P, Twell D, Ríos G, Koncz C, Jarvis P (2005) In vivo studies on the roles of Tic110, Tic40 and Hsp93 during chloroplast protein import. Plant J **41:** 412–428

Lee DW, Hwang I (2011) Transient expression and analysis of chloroplast proteins in Arabidopsis protoplasts. Methods Mol Biol **774:** 59–71

Lee DW, Jung C, Hwang I (2013a) Cytosolic events involved in chloroplast protein targeting. Biochim Biophys Acta **1833:** 245–252

Lee DW, Kim JK, Lee S, Choi S, Kim S, Hwang I (2008) Arabidopsis nuclear-encoded plastid transit peptides contain multiple sequence subgroups with distinctive chloroplast-targeting sequence motifs. Plant Cell **20:** 1603–1622

Lee DW, Lee S, Lee GJ, Lee KH, Kim S, Cheong GW, Hwang I (2006) Functional characterization of sequence motifs in the transit peptide of Arabidopsis small subunit of Rubisco. Plant Physiol **140:** 466–483

Lee DW, Lee S, Oh YJ, Hwang I (2009a) Multiple sequence motifs in the Rubisco small subunit transit peptide independently contribute to Toc159-dependent import of proteins into chloroplasts. Plant Physiol **151:** 129–141

Lee J, Kim DH, Hwang I (2014) Specific targeting of proteins to outer envelope membranes of endosymbiotic organelles, chloroplasts, and mitochondria. Front Plant Sci **5:** 173

Lee KH, Kim DH, Lee SW, Kim ZH, Hwang I (2002) In vivo import experiments in protoplasts reveal the importance of the overall context but not specific amino acid residues of the transit peptide during import into chloroplasts. Mol Cells **14:** 388–397

Lee S, Lee DW, Lee Y, Mayer U, Stierhof YD, Lee S, Jürgens G, Hwang I (2009b) Heat shock protein cognate 70-4 and an E3 ubiquitin ligase, CHIP, mediate plastid-destined precursor degradation through the ubiquitin-26S proteasome system in *Arabidopsis*. Plant Cell **21:** 3984–4001

Lee Y, Jang M, Song K, Kang H, Lee MH, Lee DW, Zouhar J, Rojo E, Sohn EJ, Hwang I (2013b) Functional identification of sorting receptors involved in trafficking of soluble lytic vacuolar proteins in vegetative cells of Arabidopsis. Plant Physiol **161:** 121–133

Leister D (2003) Chloroplast research in the genomic age. Trends Genet **19:** 47–56

Li HM, Chiu CC (2010) Protein transport into chloroplasts. Annu Rev Plant Biol **61:** 157–180

Li HM, Teng YS (2013) Transit peptide design and plastid import regulation. Trends Plant Sci **18:** 360–366

Liu L, McNeilage RT, Shi LX, Theg SM (2014) ATP requirement for chloroplast protein import is set by the K_m for ATP hydrolysis of stromal Hsp70 in *Physcomitrella patens*. Plant Cell **26:** 1246–1255

Long M, de Souza SJ, Rosenberg C, Gilbert W (1996) Exon shuffling and the origin of the mitochondrial targeting function in plant cytochrome c1 precursor. Proc Natl Acad Sci USA **93:** 7727–7731

May T, Soll J (2000) 14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants. Plant Cell **12:** 53–64

Miller EA, Beilharz TH, Malkus PN, Lee MC, Hamamoto S, Orci L, Schekman R (2003) Multiple cargo binding sites on the COPII subunit Sec24p ensure capture of diverse membrane proteins into transport vesicles. Cell **114:** 497–509

Pilon M, Wienk H, Sips W, de Swaaf M, Talboom I, van 't Hof R, de Korte-Kool G, Demel R, Weisbeek P, de Kruijff B (1995) Functional domains of the ferredoxin transit sequence involved in chloroplast import. J Biol Chem **270:** 3882–3893

Rensink WA, Schnell DJ, Weisbeek PJ (2000) The transit sequence of ferredoxin contains different domains for translocation across the outer and inner membrane of the chloroplast envelope. J Biol Chem **275:** 10265–10271

Rosano GL, Bruch EM, Ceccarelli EA (2011) Insights into the Clp/HSP100 chaperone system from chloroplasts of Arabidopsis thaliana. J Biol Chem **286:** 29671–29680

Shi LX, Theg SM (2010) A stromal heat shock protein 70 system functions in protein import into chloroplasts in the moss *Physcomitrella patens*. Plant Cell **22:** 205–220

Smith MD, Rounds CM, Wang F, Chen K, Afitlhile M, Schnell DJ (2004) atToc159 is a selective transit peptide receptor for the import of nucleus-encoded chloroplast proteins. J Cell Biol **165:** 323–334

Su PH, Li HM (2010) Stromal Hsp70 is important for protein translocation into pea and *Arabidopsis* chloroplasts. Plant Cell **22:** 1516–1531

Teng YS, Chan PT, Li HM (2012) Differential age-dependent import regulation by signal peptides. PLoS Biol **10:** e1001416

Tillmann B, Röth S, Bublak D, Sommer M, Stelzer EH, Scharf KD, Schleiff E (2015) Hsp90 is involved in the regulation of cytosolic precursor protein abundance in tomato. Mol Plant **8:** 228–241

Tsai JY, Chu CC, Yeh YH, Chen LJ, Li HM, Hsiao CD (2013) Structural characterizations of the chloroplast translocon protein Tic110. Plant J **75:** 847–857

Qbadou S, Becker T, Mirus O, Tews I, Soll J, Schleiff E (2006) The molecular chaperone Hsp90 delivers precursor proteins to the chloroplast import receptor Toc64. EMBO J **25:** 1836–1847

Quigley F, Martin WF, Cerff R (1988) Intron conservation across the prokaryote-eukaryote boundary: structure of the nuclear gene for chloroplast glyceraldehyde-3-phosphate dehydrogenase from maize. Proc Natl Acad Sci USA **85:** 2672–2676

Wienk HL, Czisch M, de Kruijff B (1999) The structural flexibility of the preferredoxin transit peptide. FEBS Lett **453:** 318–326

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