Chloroplast transit peptides often require downstream unstructured sequence in *Chlamydomonas reinhardtii*

Caspari, OD[1,2]

[1]Affiliation: Chloroplast Biology and Light Sensing in Microalgae research group (UMR7141), Institut de Biologie Physico-Chimique (CNRS/Sorbonne Université), 13 Rue Pierre et Marie Curie, 75005 Paris, France

[2]Current address: Stress Adaptation and Metabolism in Enterobacteria research group, Department of Microbiology, Pasteur Institute, 28 Rue du Docteur Roux, 75015 Paris, France

*** Correspondence:**
Caspari, OD
odc20@cantab.net

Keywords: disordered protein, protein import, chloroplast, targeting peptide, transit peptide, presequence, *Chlamydomonas reinhardtii*

1 Abstract

The N-terminal sequence stretch that defines subcellular targeting for most nuclear encoded chloroplast proteins is usually considered identical to the sequence that is cleaved upon import. Yet here this study shows that for nine out of ten tested Chlamydomonas chloroplast transit peptides, additional sequence past the cleavage site is required to enable chloroplast targeting. Using replacements of native post-cleavage residues with alternative sequences points to a role for unstructured sequence at mature protein N-termini.

2 Introduction

Chloroplasts derive from once free-living cyanobacteria and retain genomes of their own, yet the vast majority of genes coding for plastid proteins are located in the nucleus (Shi and Theg, 2013). Proteins are expressed in the cytosol as preproteins containing N-terminal chloroplast targeting peptides (cTP), which direct import into the chloroplast (Chotewutmontri et al., 2017). Upon import, the N-terminus of the preprotein is cut off at a defined cleavage site and degraded (Teixeira and Glaser, 2013; Kmiec et al., 2014).

The cTP, i.e. the sequence both necessary and sufficient to generate chloroplast targeting, is generally assumed to coincide with the sequence stretch that is cleaved off. It is this cleaved presequence that has been studied for decades, with a well-documented tripartite structure and multiple sequence elements now recognized to contribute to targeting (von Heijne et al., 1989; Chotewutmontri et al., 2017; Caspari and Lafontaine, 2021). Yet the cleavage site may not always be a reliable guide for defining the cTP. Parts of the mature protein N-terminus have been recognized before as a requirement for efficient targeting in higher plants (Bionda et al., 2010; Rolland et al., 2016; Shen et al., 2017). Bionda et al. (2010) described a length requirement of ~55-60 amino acids for *Arabidopsis thaliana*, and proposed that in the ~50% of chloroplast targeted proteins where the cleavable part of the cTP is shorter, residues C-terminal of the cleavage site contribute unstructured sequence towards targeting. There are even examples of import through the canonical TOC/TIC system of proteins

lacking a cleavable N-terminal presequence altogether (Chang et al., 2014), implying the entire targeting information is present in the mature protein.

It was noted decades ago that cleavable cTPs are generally shorter in the model green alga *Chlamydomonas reinhardtii* compared to higher plants based on the placement of cleavage sites (Franzén et al., 1990), suggesting there may be a role for mature protein N-termini. In Chlamydomonas, the Rubisco activase (RBCA) cTP has recently been found to require post-cleavage site residues to target a fluorescent reporter protein to the chloroplast (Garrido et al., 2020), in line with earlier work regarding Rubisco small subunit (RBCS) cTP from Chlamydomonas (Razzak et al., 2017) and higher plants (Comai et al., 1988). This present study revisits the role of post-cleavage site residues, and finds that substantial stretches of unstructured sequence appear to be required to enable targeting by most Chlamydomonas cTPs.

3 Materials and Methods

Constructs were Gibson assembled as described in (Garrido et al., 2020) based on pMO611, a variant of published pMO449 (Onishi and Pringle, 2016) where an upstream start codon was mutated to CTG. Sequences inserted upstream of Venus were amplified from Chlamydomonas genomic DNA extracted from wild-type strain T222+, with the following exceptions: construct 'RBCS-cTP+23' was amplified from cDNA generated from T222+ to avoid introducing an intron not present in the other RBCS-cTP constructs; the G-rich artificial linker sequence was amplified from plasmid CrRaf1-pJHL encoding Raf1-strep (Wietrzynski et al., 2021). Corresponding amino acid sequences are given in Table 1. Constructs were verified by sequencing (Eurofins). Transformants were obtained and microscopy performed as described before (Garrido et al., 2020).

4 Results

Native cTPs were tested by inserting the respective genomic sequences upstream of a Venus fluorescent reporter in a bicistronic expression vector (Onishi and Pringle, 2016; Caspari, 2020). Neither RBCA nor RBCS cTPs enabled import into the chloroplast (Fig. 1A,B), instead showing Venus fluorescence from across the cytosol or from within globular structures outside the chloroplast respectively. Given that successful targeting using the RBCS-cTP including 23 residues downstream of the cleavage site had previously been reported in Chlamydomonas (Razzak et al., 2017) and higher plants (Comai et al., 1988) and extending the RBCA-cTP by 23 residues had equally restored targeting (Garrido et al., 2020), a series of successive extensions by 3 (Fig. 1C,D), 5 (Fig. 1E,F), 10 (Fig. 1G,H) and 23 residues (Fig. 1I,J) was prepared. The +3, +5 and +10 constructs of both RBCA-cTP and RBCS-cTP all show significant Venus fluorescence from outside the chloroplast (Fig. 1C-H), and only the RBCA-cTP+10 construct also shows some partial localization of Venus to the chloroplast (Fig. 1G). By contrast, the +23 constructs show clear accumulation of Venus inside the chloroplast (Fig. 1I,J), strikingly visible as fluorescence from inside the pyrenoid, a proteinaceous structure within the chloroplast that generates a visible dip in chlorophyll fluorescence (Mackinder et al., 2017). Note that the RBCA-cTP+23 construct contains a pyrenoid-localisation motif (Meyer et al., 2020). By contrast, mitochondrial targeting peptides (mTPs) from carbonic anhydrase 4 (CAH4) and γ-carbonic anhydrase 2 (CAG2) show mitochondrial targeting, evident as co-localisation of Venus and MitoTracker fluorescence, in the absence of any post-cleavage site residues (Fig. 1K,L). CAG2-mTP shows some remaining signal from the cytoplasm, which persists when the construct is extended by 23 residues past the cleavage site (Garrido et al., 2020).

This is a provisional file, not the final typeset article

To test whether the presence of a relatively long stretch of mature protein is a general requirement rather than an exception concerning specifically RBCA and RBCS, a further eight cTPs were assayed for import (Fig. 2). These cTPs were each extended by 5 residues beyond the cleavage site to ensure any lack of import was not an artifact arising from a missing cleavage site, for which several residues past the cleavage site seem to be important (Tardif et al., 2012). Only one of the tested cTP+5 constructs, that of PSAD, generates chloroplast localized Venus signal (Fig. 2A-H). Chloroplast targeting with the PSAD-cTP also works in the absence of any post-cleavage site residues (Fig. 2I).

To follow up the hypothesis that post-cleavage site residues contribute unstructured sequence towards chloroplast targeting, four constructs were prepared that each extend the RBCA-cTP+5 by different sequences (Fig. 3). Five mature RBCA residues were kept to ensure the cleavage site would be maintained intact. Three constructs contained unstructured sequence downstream (A-C) and were expected to generate chloroplast targeting, whereas a fourth (D) contained a α-helical stretch that was expected to impede chloroplast import. Two of the unstructured sequences as well as the helical stretch were copied from Chlamydomonas nuclear genes encoding chloroplast-localized proteins EPYC1 (also known as LCI5), VIPP1 and RBCXa respectively. In each of these cases, sequence was chosen from the C-terminal half of the protein, to avoid any sequence that might contribute to the cTPs of these proteins. In addition, a standard G-rich linker sequence, such as is commonly used to join elements when generating fusion proteins in molecular biology, was used to test whether an artificial flexible sequence could substitute for a native unstructured one.
Of these, natively unstructured sequences from the third repeat of intrinsically disordered EPYC1 and the disordered C-terminus of VIPP1 generated chloroplast localization of the Venus reporter (Fig. 3A,B). Note that the VIPP1-derived sequence also shows additional Venus signal emanating from cytosolic structures (Fig. 3B). By contrast, both the artificial flexible linker and the C-terminal α-helix of RBCXa resulted in only a small amount of Venus localizing to the chloroplast, with a majority of signal emanating from structures outside the chloroplast (Fig. 3C,D).

5 Discussion

By definition, targeting peptides comprise the sequence that is both necessary for targeting a native cargo protein to its subcellular destination, and sufficient to redirect a heterologously expressed cargo to the same location (Chotewutmontri et al., 2017). The present study demonstrates that in Chlamydomonas, the amino-terminal sequence required to specify chloroplast targeting exceeds the cleavable part for nine out of ten tested cTPs. Care was taken to make the set of studied cTPs as representative as possible within the limited choice of cTPs with a well-characterized cleavage site (Terashima et al., 2011; Tardif et al., 2012). In particular, studied cTPs were chosen to include proteins related to photosynthesis (e.g. RBCA, RBCS, PSAD) as well as to plastid housekeeping (e.g. ribosomal proteins PRPL4, PSRP3), as import of these groups may be mediated by different receptors at the translocon (Chotewutmontri et al., 2017). The cleavage site should thus no longer be considered a reliable guide for identifying the N-terminal stretch containing all chloroplast targeting information.

The reduced length of algal cTPs (Franzén et al., 1990) may partly be accounted for by specific sequence elements required for targeting in plant cTPs being reduced in Chlamydomonas, such as a TOC-interacting FGLK motif (Chotewutmontri et al., 2017; Razzak et al., 2017; Caspari et al., 2021). Yet the requirement for post-cleavage site residues implies that the amino-terminal sequences contributing to targeting in Chlamydomonas are in fact longer than recognized to date, with sequence elements that play a role in import present past the cleavage site just as in short plant cTPs (Caspari et al., 2021). With lengths of 56 and 60 residues respectively, the import competent RBCA-cTP+23

124 and RBCS-cTP+23 constructs mirror the ~55-60 amino acid requirement Arabidopsis cTPs (Bionda
125 et al., 2010). In this sense, algal cTPs may not be shorter than their plant counterparts after all.

126 Previous studies suggested that the contribution of mature protein N-termini towards targeting may
127 simply be a supply of unstructured sequence (Bionda et al., 2010; Shen et al., 2017). Bionda et al.
128 (2010) found that using tightly folded titin as cargo protein prevented chloroplast import by short
129 cTPs unless first denatured. Shen et al., (2017) found that ~20 residues of unfolded sequence
130 downstream of the cTP cleavage site were required to target two *E. coli* proteins into rice
131 chloroplasts. While less tightly folded than titin, the Venus fluorescent reporter used as cargo protein
132 in the present study contains a short, tightly folded α-helix at the N-terminus (Rekas et al., 2002),
133 which may conceivably impair import in the absence of an unfolded spacer sequence. In principle,
134 the chloroplast import machinery is capable of unfolding any protein for post-translational import,
135 driven by an ATP-powered motor complex in the chloroplast stroma (Chotewutmontri et al., 2017).
136 However, initiation of import requires the cTP to reach across both TOC and TIC passively before
137 being able to contact stromal motor proteins. It is this initiation step that is most likely disrupted by
138 the presence of structured elements too close to the N-terminus (Bionda et al., 2010). Mitochondrial
139 import lacks similar constraints, as mitochondrial targeting peptides are guided across the outer
140 membrane translocase using interactions with increasing affinity (Schatz, 1997; Fukasawa et al.,
141 2017), and then make use of the proton motive force across the inner membrane to power import
142 even before contacting the matrix-localized PAM complex that uses ATP to energize further import
143 (Martin et al., 1991; Wiedemann and Pfanner, 2017). On the other hand, Razzak et al. (2017) report
144 the presence of a specific motif (Q|MMVW) required for targeting in the Chlamydomonas RBCS-
145 cTP that spans the cleavage site (Schmidt et al., 1979), suggesting post-cleavage site sequence
146 elements may also play a more active role in specifying targeting.

147 Here, we found that chloroplast import by RBCA-cTP could indeed be restored using unstructured
148 sequences taken from chloroplast proteins EPYC1 and VIPP1 as post-cleavage site spacers. EPYC1
149 is an intrinsically disordered protein comprised of four near-identical internal repeats that act as
150 linker protein within the chloroplast pyrenoid (Mackinder et al., 2016), a proteinaceous structure
151 within the chloroplast that is central to the algal carbon concentrating mechanism (Caspari et al.,
152 2017). VIPP1 plays a role in the biogenesis of thylakoid membranes and has an unstructured C-
153 terminus (Zhang et al., 2016; Theis et al., 2019). C-terminal sequences from these proteins are not
154 expected to contain any specific targeting motifs, thus restoration of targeting lends further support to
155 the importance of unstructured sequence at the N-terminus. In the same vain, copying an inherently
156 structured sequence from the most C-terminal helix of RBCXa (Bracher et al., 2015) largely
157 prevented chloroplast targeting (Fig. 3D).

158 An overall picture thus emerges of short algal cTPs up to the cleavage site containing the information
159 for chloroplast targeting, but relying on additional unstructured sequence at mature protein N-termini
160 to allow import to be successfully initiated. The importance of sufficient length and the presence of
161 unstructured sequence elements for chloroplast targeting was equally evident in recent work aimed at
162 retracing likely steps in the evolution of targeting peptides from putative antimicrobial origins
163 (Caspari et al., 2021). Additional complexities likely exist, as several observations point to further
164 factors being at play: A G-rich linker failed to restore chloroplast import beyond what had been seen
165 for the RBCXa helix (Fig. 3C), suggesting that this artificial flexible sequence differs in important
166 ways from native unstructured sequences. Furthermore, the very long PSRP3-cTP+5 construct (89
167 residues) failed to target whereas the relatively short PSAD-cTP (35 residues) achieved targeting,
168 suggesting additional factors are in play beyond cTP length. The case of PSAD is particularly
169 intriguing, as this cTP should be too short to reach into the stroma passively, yet appears to

This is a provisional file, not the final typeset article

circumvent the impedance of a folded Venus cargo. Presumably, the cargo protein is kept in an unfolded state prior to import, e.g. through recruitment of chaperones interacting with the cTP, or perhaps through co-translational import following polysome recruitment to the chloroplast envelope by this cTP. Either way, for the purposes of sending heterologous proteins to the chloroplast (Crozet et al., 2018), the PSAD-cTP is likely the best choice among Chlamydomonas cTPs studied so far to ensure targeting without a need for additional sequence at the mature protein N-terminus.

6 Conflict of Interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

7 Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

8 Author Contributions

Conceptualisation, Methodology, Investigation, Analysis, Interpretation, Visualisation and Writing: ODC

9 Funding

Financial support from the Rothschild Foundation, LabEx Dynamo (ANR-LABX-011) and the ChloroMitoRAMP grant (ANR-19-CE13-0009) to ODC, as well as annual funding from CNRS and Sorbonne University to UMR7141, is gratefully acknowledged.

10 Acknowledgments

I would like to thank Francis-André Wollman for critical reading of the manuscript, as well as Richard Kuras and Yves Choquet for helpful discussions and advice. Katia Wostrikoff kindly provided plasmid CrRaf1-pJHL from which the G-rich linker sequence was amplified. Thank you also to Stéphane Lemaire and Pierre Crozet for suggesting that I should try PSAD-cTP.

11 References

Bionda, T., Tillmann, B., Simm, S., Beilstein, K., Ruprecht, M., and Schleiff, E. (2010). Chloroplast import signals: The length requirement for translocation in vitro and in vivo. *J. Mol. Biol.* 402, 510–523. doi:10.1016/j.jmb.2010.07.052.

Bracher, A., Hauser, T., Liu, C., Hartl, F. U., and Hayer-Hartl, M. (2015). Structural analysis of the rubisco-assembly chaperone RbcX-II from Chlamydomonas reinhardtii. *PLoS One* 10, 1–17. doi:10.1371/journal.pone.0135448.

Caspari, O. D. (2020). Introduction of a leaky stop codon as molecular tool in Chlamydomonas reinhardtii. *PLoS One* 15, e0237405. doi:10.1371/journal.pone.0237405.

Caspari, O. D., Garrido, C., Choquet, Y., Wollman, F.-A., and Lafontaine, I. (2021). Retracing the

206 evolution from antimicrobial to targeting peptides.

207 Caspari, O. D., and Lafontaine, I. (2021). The role of antimicrobial peptides in the evolution of
208 endosymbiotic protein import. *PLoS Pathog.* 17, e1009466. doi:10.1371/journal.ppat.1009466.

209 Caspari, O. D., Meyer, M. T., Tolleter, D., Wittkopp, T. M., Cunniffe, N. J., Lawson, T., et al.
210 (2017). Pyrenoid loss in *Chlamydomonas reinhardtii* causes limitations in CO2 supply, but not
211 thylakoid operating efficiency. *J. Exp. Bot.* 68, 3903–3913. doi:10.1093/jxb/erx197.

212 Chang, W. L., Soll, J., and Bölter, B. (2014). A new member of the psToc159 family contributes to
213 distinct protein targeting pathways in pea chloroplasts. *Front. Plant Sci.* 5, 1–14.
214 doi:10.3389/fpls.2014.00239.

215 Chotewutmontri, P., Holbrook, K., and Bruce, B. D. (2017). Plastid Protein Targeting: Preprotein
216 Recognition and Translocation. *Int. Rev. Cell Mol. Biol.* 330, 227–294.
217 doi:10.1016/bs.ircmb.2016.09.006.

218 Comai, L., Larson-Kelly, N., Kiser, J., Mau, C. J., Pokalsky, A. R., Shewmaker, C. K., et al. (1988).
219 Chloroplast transport of a ribulose bisphosphate carboxylase small subunit-5-enolpyruvyl 3-
220 phosphoshikimate synthase chimeric protein requires part of the mature small subunit in
221 addition to the transit peptide. *J. Biol. Chem.* 263, 15104–15109. doi:10.1016/s0021-
222 9258(18)68151-7.

223 Crozet, P., Navarro, F., Willmund, F., Mehrshahi, P., Bakowski, K., Lauersen, K., et al. (2018). Birth
224 of a Photosynthetic Chassis: a MoClo Toolkit Enabling Synthetic Biology in the Microalga
225 *Chlamydomonas reinhardtii*. *ACS Synth. Biol.* 7, 2074–2086. doi:10.1021/acssynbio.8b00251.

226 Franzén, L.-G. G., Rochaix, J.-D. D., and von Heijne, G. (1990). Chloroplast transit peptides from
227 the green alga Chlamydomonas reinhardtii share features with both mitochondrial and higher
228 plant chloroplast presequences. *FEBS Lett.* 260, 169–172. doi:10.1016/0014-5793(90)80094-Y.

229 Fukasawa, Y., Oda, T., Tomii, K., and Imai, K. (2017). Origin and Evolutionary Alteration of the
230 Mitochondrial Import System in Eukaryotic Lineages. *Mol. Biol. Evol.* 34, 1574–1586.
231 doi:10.1093/molbev/msx096.

232 Garrido, C., Caspari, O. D., Choquet, Y., Wollman, F. A., and Lafontaine, I. (2020). Evidence
233 Supporting an Antimicrobial Origin of Targeting Peptides to Endosymbiotic Organelles. *Cells* 9,
234 1795. doi:10.3390/cells9081795.

235 Kmiec, B., Teixeira, P. F., and Glaser, E. (2014). Shredding the signal : targeting peptide degradation
236 in mitochondria and chloroplasts. *Trends Plant Sci.* 19, 771–778.
237 doi:10.1016/j.tplants.2014.09.004.

238 Mackinder, L. C. M., Chen, C., Leib, R. D., Patena, W., Blum, S. R., Rodman, M., et al. (2017). A
239 Spatial Interactome Reveals the Protein Organization of the Algal CO2-Concentrating
240 Mechanism. *Cell* 171, 133–147. doi:10.1016/j.cell.2017.08.044.

241 Mackinder, L. C. M., Meyer, M. T., Mettler-Altmann, T., Chen, V. K., Mitchell, M. C., Caspari, O.
242 D., et al. (2016). A repeat protein links Rubisco to form the eukaryotic carbon concentrating
243 organelle. *Proc. Natl. Acad. Sci. USA* 113, 5958–5963. doi:10.1073/pnas.1522866113.

This is a provisional file, not the final typeset article

244 Martin, J., Mahlke, K., and Pfanner, N. (1991). Role of an energized inner membrane in
245 mitochondrial protein import: ΔΨ drives the movement of presequences. *J. Biol. Chem.* 266,
246 18051–18057. doi:10.1016/s0021-9258(18)55235-2.

247 Meyer, M., Itakura, A., Patena, W., Wang, L., He, S., Emrich-Mills, T., et al. (2020). Assembly of
248 the algal CO2-fixing organelle, the pyrenoid, is guided by a Rubisco-binding motif. *Sci. Adv.* 6,
249 eabd2408. doi:10.1126/sciadv.abd2408.

250 Onishi, M., and Pringle, J. R. (2016). Robust Transgene Expression from Bicistronic mRNA in the
251 Green Alga *Chlamydomonas reinhardtii*. *G3 (Bethesda).* 6, 4115–4125.
252 doi:10.1534/g3.116.033035.

253 Razzak, A., Lee, D. W., Yoo, Y., and Hwang, I. (2017). Evolution of rubisco complex small subunit
254 transit peptides from algae to plants. *Sci. Rep.* 7, 9279. doi:10.1038/s41598-017-09473-x.

255 Rekas, A., Alattia, J. R., Nagai, T., Miyawaki, A., and Ikura, M. (2002). Crystal structure of venus, a
256 yellow fluorescent protein with improved maturation and reduced environmental sensitivity. *J.*
257 *Biol. Chem.* 277, 50573–50578. doi:10.1074/jbc.M209524200.

258 Rolland, V., Badger, M. R., and Price, G. D. (2016). Redirecting the Cyanobacterial Bicarbonate
259 Transporters BicA and SbtA to the Chloroplast Envelope: Soluble and Membrane Cargos Need
260 Different Chloroplast Targeting Signals in Plants. *Front. Plant Sci.* 7, 185.
261 doi:10.3389/fpls.2016.00185.

262 Schatz, G. (1997). Just follow the acid chain. *Nature* 388, 121. doi:10.1038/40510.

263 Schmidt, G. W., Devillers-thiery, A., Desruisseaux, H., Blobel, G., and Chua, N. (1979). NH2 -
264 Terminal Amino Acid Sequences of Precursor and Mature Forms of the Small Subunit From
265 Chlamydomonas Reinhardtii. 83.

266 Shen, B. R., Zhu, C. H., Yao, Z., Cui, L. L., Zhang, J. J., Yang, C. W., et al. (2017). An optimized
267 transit peptide for effective targeting of diverse foreign proteins into chloroplasts in rice. *Sci.*
268 *Rep.* 7, 1–12. doi:10.1038/srep46231.

269 Shi, L. X., and Theg, S. M. (2013). The chloroplast protein import system: From algae to trees.
270 *Biochim. Biophys. Acta - Mol. Cell Res.* 1833, 314–331. doi:10.1016/j.bbamcr.2012.10.002.

271 Tardif, M., Atteia, A., Specht, M., Cogne, G., Rolland, N., Brugière, S., et al. (2012). Predalgo: A
272 new subcellular localization prediction tool dedicated to green algae. *Mol. Biol. Evol.* 29, 3625–
273 3639. doi:10.1093/molbev/mss178.

274 Teixeira, P. F., and Glaser, E. (2013). Processing peptidases in mitochondria and chloroplasts.
275 *Biochim. Biophys. Acta - Mol. Cell Res.* 1833, 360–370. doi:10.1016/j.bbamcr.2012.03.012.

276 Terashima, M., Specht, M., and Hippler, M. (2011). The chloroplast proteome: A survey from the
277 Chlamydomonas reinhardtii perspective with a focus on distinctive features. *Curr. Genet.* 57,
278 151–168. doi:10.1007/s00294-011-0339-1.

279 Theis, J., Gupta, T. K., Klingler, J., Wan, W., Albert, S., Keller, S., et al. (2019). VIPP1 rods engulf
280 membranes containing phosphatidylinositol phosphates. *Sci. Rep.* 9, 1–11. doi:10.1038/s41598-

281 019-44259-3.

282 von Heijne, G., Steppuhn, J., and Herrmann, R. G. (1989). Domain structure of mitochondrial and
283 chloroplastic targeting peptides. *Eur. J. Biochem.* 180, 535–545.

284 Wiedemann, N., and Pfanner, N. (2017). Mitochondrial Machineries for Protein Import and
285 Assembly. *Annu. Rev. Biochem.* 86, 685–714. doi:10.1146/annurev-biochem-060815-014352.

286 Wietrzynski, W., Traverso, E., Wollman, F.-A., and Wostrikoff, K. (2021). The state of
287 oligomerization of Rubisco controls the rate of synthesis of the Rubisco large subunit in
288 Chlamydomonas reinhardtii . *Plant Cell*, 1–22. doi:10.1093/plcell/koab061.

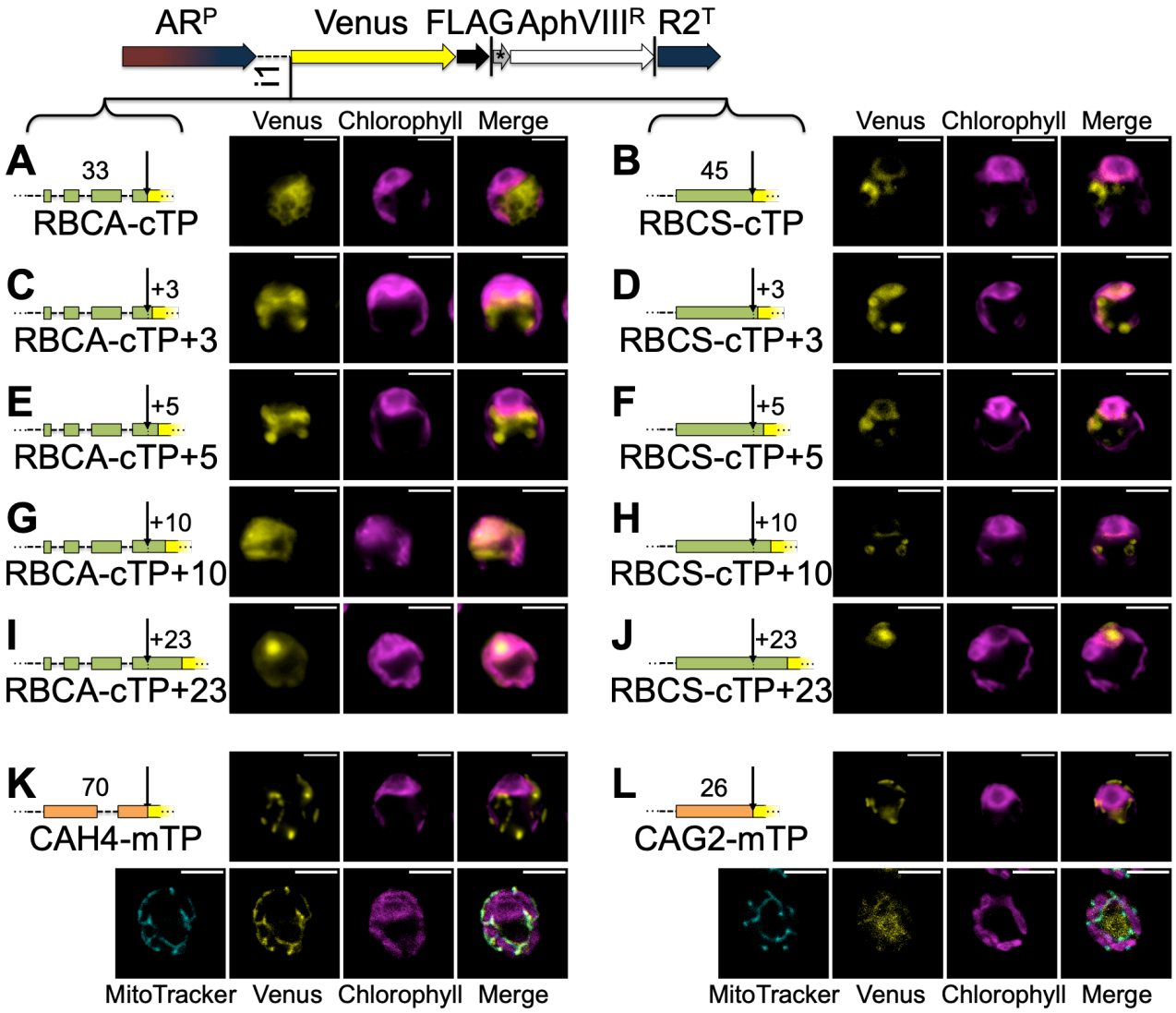
289 Zhang, L., Kondo, H., Kamikubo, H., Kataoka, M., and Sakamoto, W. (2016). VIPP1 has a
290 disordered C-terminal tail necessary for protecting photosynthetic membranes against stress.
291 *Plant Physiol.* 171, 1983–1995. doi:10.1104/pp.16.00532.

292

293

This is a provisional file, not the final typeset article

294 **12 Figures**



295

296 **Fig. 1 – A stretch of post-cleavage site residues is necessary for chloroplast targeting peptides.**
297 Constructs were generated by inserting candidate peptides upstream of Venus in a bicistronic
298 expression vector made of the following components: AR^P: Hybrid *HSP70a-RBCS2* promoter, i1:
299 *RBCS2* intron 1, Venus: Venus fluorescent protein (a YFP variant), FLAG: triple-FLAG tag, |: stop
300 codon, *: bicistronic intergenic sequence (tagcat), $AphVIII^R$: Paromomycin resistance gene, $R2^T$:
301 *RBCS2* terminator. Candidate peptides are Rubisco activase (RBCA, A-E) and Rubisco small subunit
302 (RBCS, B-J) chloroplast targeting peptides (cTP) extended by 0, 3, 5, 10 or 23 residues beyond the
303 cleavage site into the mature protein, as well as carbonic anhydrase 4 (CAH4, K) and γ-carbonic
304 anhydrase 2 (CAG2, L) mitochondrial targeting peptides (mTP). Lengths in amino acid residues
305 are indicated above each peptide. Introns within peptides are symbolized by dotted lines. Typical
306 epifluorescence images of resulting transformants in strain $T222^+$ show fluorescence from the Venus
307 channel in yellow and chlorophyll autofluorescence in magenta. Confocal microscopy images include
308 an additional MitoTracker channel in cyan for mTPs. All scale bars are 5µm in length.

309

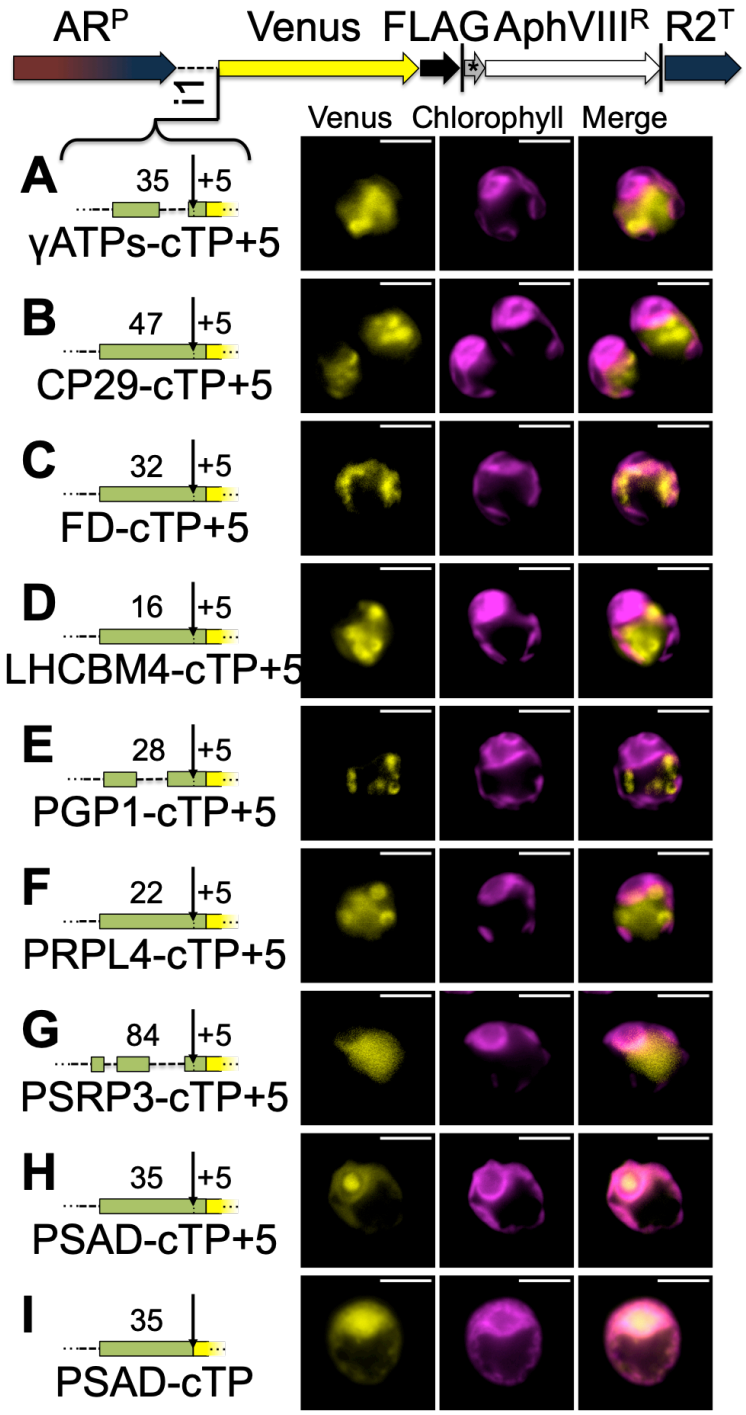


Fig. 2 – Without a stretch of post-cleavage site residues, Chlamydomonas chloroplast targeting peptides fail to import. Candidate chloroplast targeting peptides, each including 5 post-cleavage site residues, from (A) chloroplast ATP-synthase gamma chain (γATPs), (B) chlorophyll a/b binding protein of photosystem II (CP29), (C) ferredoxin (FD), (D) chlorophyll a/b binding protein of light harvesting complex II (LHCBM4), (E) phosphoglycolate phosphatase (PGP1), (F) plastid ribosomal protein L4 (PRPL4), (G) plastid-specific ribosomal protein 3 (PSRP3) and (H) Photosystem 1 subunit D (PSAD), were inserted in-frame upstream of Venus. (I) A second PSAD construct lacking post-cleavage site residues was probed. Lengths in amino acid residues are indicated above each peptide. Representation as in Fig. 1.

This is a provisional file, not the final typeset article

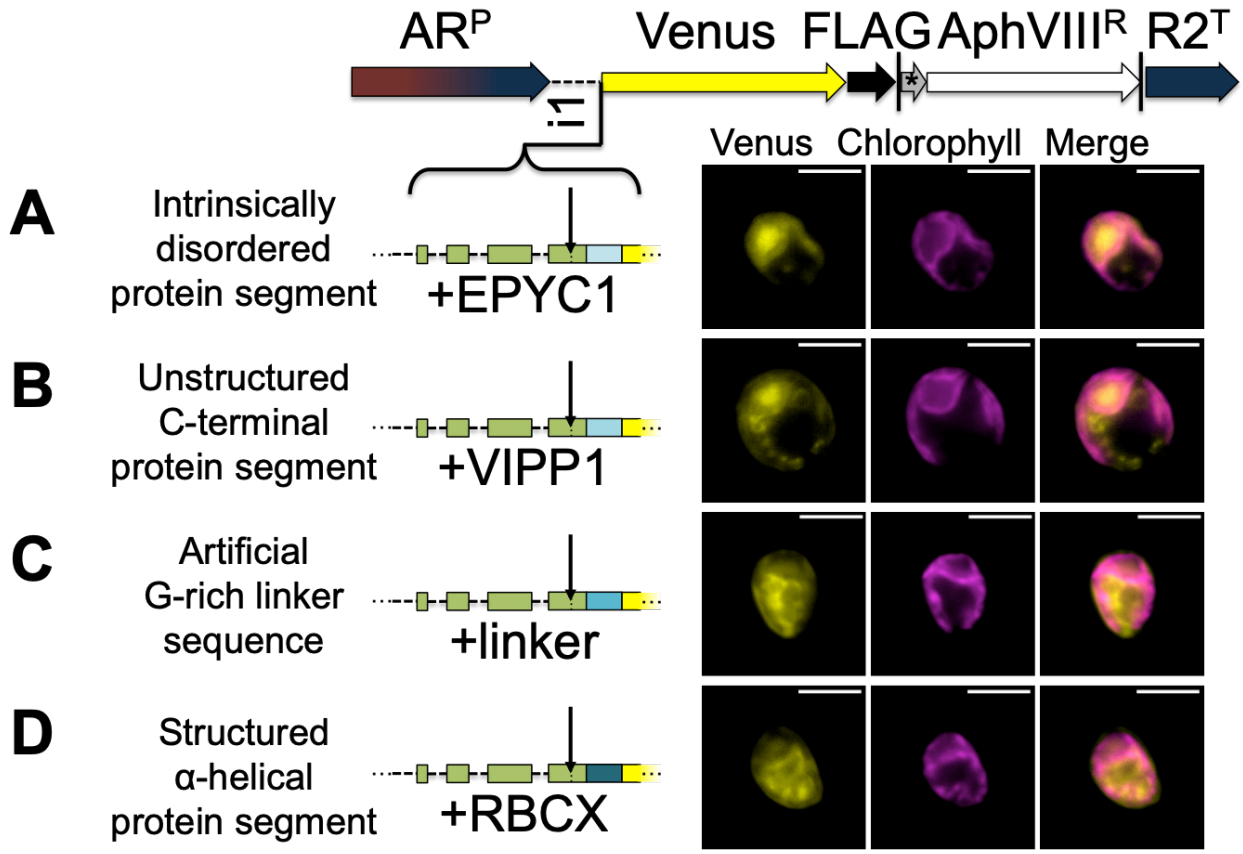


Fig. 3 – Unstructured downstream sequence aids cTP targeting. The following candidate sequences were inserted downstream of RBCA-cTP+5 and upstream of Venus: (A) residues 195-213 of intrinsically disordered protein EPYC1, which are part of the 3rd out of 4 repeats; (B) residues 260-278 of VIPP1, which form part of an unstructured C-terminus; (C) an artificial linker sequence mainly consisting of quintuple-G interspersed with S; (D) residues 135-154 of RBCXa, which form part of the most C-terminal α-helix. Refer to Table 1 for amino acid sequences. Representation as in Fig. 1.

13 Tables

Table 1 – Amino acid sequences. Cleavage sites are indicated with a full stop between residues.

Element	Amino acid sequence
RBCS-cTP (+23)	MAAVIAKSSVSAAVARPARSSVRPMAALKPAVKAAPV AAPAQANQ.MMVWTPVNNKMFETFSYLPPLSD
RBCA-cTP (+23)	MQVTMKSSAVSGQRVGGARVATRSVRRAQLQVV.ASS RKQMGRWRSIDAGVDASDDQ
CAH4-mTP	MSSRNVATALRMFATLGRSQAGEASAMMGTGSALLA QRAAALGGASAVNKGCSCRCGRVACMGACMPMRH.

CAG2-mTP	MLKRVGQSLVPFARAGLTQTAESFRG.
γATPs-cTP (+5)	MAAMLASKQGAFMGRSSFAPAPKGVASRGSLQVVA.GLKEV
CP29-cTP (+5)	MVFKFPTPPGTQKKAGTTATKPAPKATTKKVATSTGTRSGGVGYRKY.QGDAL
FD-cTP (+5)	MAMAMRSTFAARVGAKPAVRGARPASRMSCMA.YKVTL
LHCBM4-cTP (+5)	MAFALATSRKALQVTC.KATGK
PGP1-cTP (+5)	MVAAQASARPIATNEQKLELLKKVECFI.FDCDG
PRPL4-cTP (+5)	MQTMRVAFRPAATSRSTVVTRA.SAVAA
PSRP3-cTP (+5)	MMLVRQVPCLAQRASARSAVRSRPAPFTPARRVFARAQADAALLEDDEVAYEVTNEDFELDENDPDLLLGLEAVELMGSDVDAM.TAAAV
PSAD-cTP (+5)	MAVMMRTQAPAATRASSRVAVAARPAARRAVVVRA.EAEAA
+EPYC1	SSSPAPASSAPAPARSSSA
+VIPP1	SAVGSLPEPRAVDALDLEL
+linker	SGGGGSGGGGSMGGGGSN
+RBCX	KTMEMASRETREANTRLM

330

This is a provisional file, not the final typeset article