Volume 260, number 2, 165-168 FEB 08028 January 1990

Chloroplast transit peptides from the green alga *Chlamydomonas reinhardtii* share features with both mitochondrial and higher plant chloroplast presequences

Lars-Gunnar Franzén*, Jean-David Rochaix° and Gunnar von Heijne+

*Department of Biochemistry, Arrhenius laboratories, University of Stockholm, Sweden,
°Departments of Molecular Biology and Plant Biology, University of Geneva, Switzerland and
+Department of Molecular Biology, Karolinska Institute Center for Biotechnology, NOVUM, Huddinge, Sweden

Received 2 november 1989

Chloroplast transit peptides from the green alga *Chlamydomonas reinhardtii* have been analyzed and compared with chloroplast transit peptides from higher plants and mitochondrial targeting peptides from yeast, Neurospora and higher eukaryotes. In terms of length and amino acid composition, chloroplast transit peptides from *C. reinhardtii* are more similar to mitochondrial targetting peptides than to chloroplast transit peptides from higher plants. They also contain the potential amphiphilic α-helix characteristic of mitochondrial presequences. However, in similarity with chloroplast transit peptides from higher plants, they contain a C-terminal region with the potential to form an amphiphilic β-strand. As in higher plants, transit peptides that route proteins to the thylakoid lumen consist of an N-terminal domain similar to stroma-targeting transit peptides attached to a C-terminal apolar domain that share many characteristics with secretory signal peptides.

Chloroplast; Protein import; Transit peptide; (*Chlamydononas reinhardtii*)

1. INTRODUCTION

The majority of proteins in mitochondria and chloroplasts are encoded by the nuclear genome. They are synthesized in the nucleocytoplasm as preproteins with N-terminal extensions (transit peptides, targeting peptides or presequences) that are required for protein import into the organelles. In most cases, the targeting peptides are removed by intraorganellar proteases during or shortly after import. Although the targeting peptides should contain the information for correct sorting and import into the organelle, they do not contain any conserved amino acid sequences; rather, secondary structure appears to be important for targeting proteins to the mitochondrion or the chloroplast [1]. mTPs contain an N-terminal domain that appears to form an amphiphilic α-helix and a C-terminal domain with different amphiphilic properties. cTPs from higher plants contain an uncharged N-terminal part, a central non-amphiphilic part and a C-terminal part that may form an amphiphilic β-strand [1].

Correspondence address: L.-G. Franzén, Department of Biochemistry, Arrhenius laboratories, University of Stockholm, S-10691 Stockholm, Sweden

Abbreviations: cTP, chloroplast transit peptide; mTP, mitochondrial targeting peptide; OEE1, 2 and 3, oxygen evolution enhancer proteins 1, 2 and 3; Rubisco, ribulose bisphosphate carboxylase/oxygenase; PS, photosystem

To test the specificity of organellar targeting, chimeric proteins containing an mTP or a cTP and a 'passenger' protein have been constructed. Hurt et al. [2] found that a cTP from *Chlamydomonas reinhardtii* can direct proteins into yeast mitochondria. In contrast, Boutry et al. [3] found that organelle targeting is highly specific in *Nicotiana*; an mTP and a cTP from *Nicotiana* directed a passenger protein into *Nicotiana* mitochondria and chloroplasts, respectively. One obvious difference between the two experiments is that in the former, a heterologous system was used (*C. reinhardtii* cTP/yeast mitochondria) while in the latter, a homologous system was used (targeting peptides and organelles from *Nicotiana*). Since some preliminary observations suggest that *C. reinhardtii* cTPs are different from higher plant cTPs [1,4], it may be that the signals for organelle targeting vary between organisms.

In this paper, we present an analysis of 12 cTPs from *C. reinhardtii*. Both in length and amino acid composition, these cTPs are more similar to mTPs from yeast and higher eukaryotes than to cTPs from higher plants. Their structure appears to be a hybrid between the structures for mTPs and cTPs suggested in [1]. *C. reinhardtii* cTPs have a short uncharged N-terminal region, a central region rich in arginine, alanine, valine and serine with a high propensity for forming an amphiphilic α-helix similar to mTPs, and a C-terminal region that may form an amphiphilic β-strand as in cTPs from higher plants.

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

```
PSI P28                              MALVARPVLSARVAASRPRVAARKAVRVSA^KYGEN
PSI P30                                  MQALSSRVNIAAKPQRAQRLVVRA^EEVKA
PSI P35                              MQTLASRPSLRASARVAPRRAPRVAVVTKA^ALDPQ
PSI P37                                MQALATRPSAIRPTKAARRSSVVVRA^DGFIG
Rubisco SS          MAAVIAKSSVSAAVARPARSSVRPMAALKPAVKAAPVAAPAQANQ^MMVWT
ATPase γ                         MAAMLASKQGAFMGRSSFAPAPKGVASRGSLQVVA^GLKEV
cabII-1             MAFALASRKALQVTCKATGKKTAAKAAAPKSSGVEFYGPNRAKWLGPYSEN

PSI P21     MALTMRNPAVKASSRVAPSSRRALRVACQAQKNETASKVGTALAASALAAAVSLSAPSAAMA^DIAGL
PSII OEE1                MALRAAQSAKAGVRAARPNRATAVVCKAQKVGQAAAAAALATAMVAGSANA^LTFDE
PSII OEE2          MATALCNKAFAAAPVARPASRRSAVVVRASGSDVSRRAALAGFAGAAALVSSSPANA^AYGDS
PSII OEE3                MALASKVATRPAVASRRGAVVVRASGESRRAVLGGLLASAVAAVAPKAALA^LTPVD
cyt c-552           MLQLANRSVRAKAARASQSARSVSCAAAKRGADVAPLTSALAVTASILLTTGAASASA^ADLAL
```

Fig.1. Chloroplast transit peptides from *C. reinhardtii*. Stroma-targeting cTPs: photosystem I (PSI) subunits P28, P30, P35 and P37 (genes *psa*H, *psa*E, *psa*G and *psa*K, respectively) [4,5], the small subunit of Rubisco [6], the gamma subunit of ATPase [7] and cabII-1 protein (chlorophyll a/b protein) [8]. Lumen-targeting proteins: photosystem I subunit P21 (gene *psa*F) [5], the OEE proteins of photosystem II [9,10] and the Cu(II)-repressible plastidic cytochrome *c*-552 [11]. Positively charged amino acids are underlined, and the processing sites between transit peptides and mature proteins are indicated. For the cabII-1 protein, the processing site is not known. For the OEE3 protein, the processing site has not been determined experimentally but is deduced from comparisons with transient peptides of other lumenal proteins [10].

2. MATERIALS AND METHODS

2.1. *Sequence samples*

A total of 12 cTPs from *C. reinhardtii* were collected from the literature (fig.1). Seven of the cTPs are of the stroma-targeting type and five are of the lumen-targeting type. For one of the stroma-targeting cTPs (cabII-1 protein), the cleavage site (stromal protease) is not known.

2.2. *Hydrophobic moment analysis*

Hydrophobic moment analysis was carried out according to [12]. Window sizes of 8 and 11 residues were used for the analysis of potential amphiphilic β-strands and α-helices, respectively. The statistical significance of peaks in the hydrophobic moment vs ∂ plots was assessed by comparison with the hydrophobic moment profiles for control samples of 100 scrambled copies of each wild-type sequence.

3. RESULTS

3.1. C. reinhardtii *cTPs are distinct from cTPs from higher plants but similar to mTPs in terms of length and amino acid composition*

The mean length of the stroma-targeting *C. reinhardtii* cTPs in our sample is 29 residues. This is close to the mean length for mTPs (30 residues) but clearly much shorter than the mean length for cTPs from higher plants (60 residues) [1].

The overall amino acid composition of the *C. reinhardtii* cTPs (excluding the apolar regions in the lumen-targeting cTPs) is shown in fig.2, together with the amino acid profiles for mTPs and higher plant cTPs (from [1]). It is readily apparent that *C. reinhardtii* cTPs are quite similar to mTPs, with a high content of Arg (14%), a moderate Ser content (10%), and virtually no acidic residues. However, the *C. reinhardtii* cTP typically has much more Ala (29%) than either mTPs or higher plant cTPs.

3.2. *The* C. reinhardtii *cTP has an uncharged N-terminal region*

cTPs from higher plants lack charged amino acids together with turn-promoting Pro and Gly residues among their ~10 N-terminal residues, whereas mTPs do not, as a rule, lack charged residues in the immediate N-terminal region [1]. The *C. reinhardtii* cTPs also have an uncharged region next to the N-terminus (Fig.1), although this region seems shorter than in cTPs from higher plants (~5 residues ~10 residues)

3.3. *The central region of* C. reinhardtii *cTPs has a high potential for forming an amphiphilic α-helix*

In order to search for amphiphilic structures in the *C. reinhardtii* cTPs, we employed hydrophobic moment

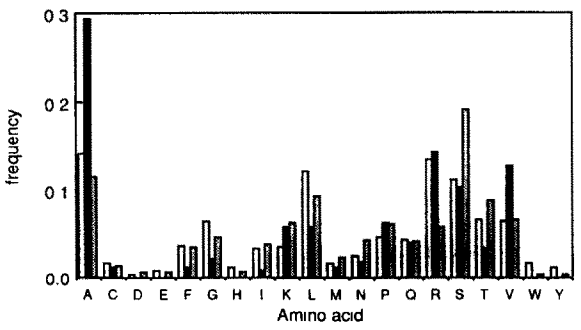

Fig.2. Amino acid composition of cTPs from *C. reinhardtii* (black bars), cTPs from higher plants (shaded bars) and mTPs from yeast, Neurospora and higher eukaryotes (white bars).

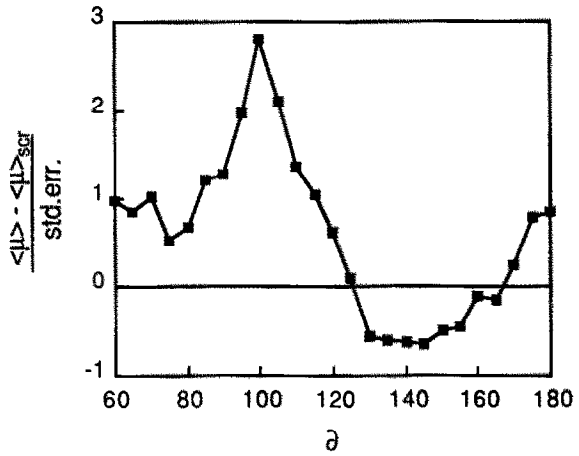

Fig.3. Mean normalized hydrophobic moment of *C. reinhardtii* cTPs as a function of ∂. Window length = 11 residues. The peak at ∂ = 100° is indicative of an amphiphilic α-helix.

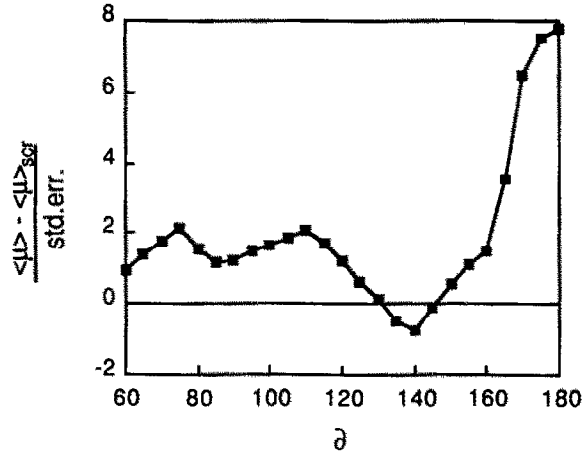

Fig.4. Mean normalized hydrophobic moment as a function of ∂ for the −8 to −1 region of five *C. reinhardtii* cTPs with known stromal cleavage sites. Window length = 8 residues. The peak at ∂ = 180° is indicative of an amphiphilic β-strand.

analysis [12,13]. Any periodic polypeptide structure can be characterized by an angle ∂ at which the side chains of successive residues project from the backbone in a helical wheel plot. The regular α-helix is caracterized by ∂ = 100°, and an ideal β-strand has ∂ = 180°. We have scanned the collection of *C. reinhardtii* cTPs in fig.1 with a window of 11 residues and with ∂ varying between 60° and 180°. For each sequence the highest hydrophobic moment found for each value of ∂ was recorded, and the mean values for the whole sample were calculated. Finally, the mean values were normalized by subtracting the mean values similarly calculated for a sample of 100 randomized copies of each sequence and dividing by the standard error of the mean hydrophobic moment for the original sequences (fig.3).

Surprisingly, a strong signal was found at ∂ = 100°, suggesting that *C. reinhardtii* cTPs can form highly amphiphilic, positively charged α-helices similar to those found in mTPs.

3.4. *The region next to the cleavage site for stromal protease may form an amphiphilic β-strand*

Since cTPs from higher plants, but not mTPs, have a potential to form an amphiphilic β-strand close to the cleavage site [1], we calculated the mean hydrophobic moment in region −8 to −1 for the five stroma-targeting *C. reinhardtii* cTPs with known cleavage sites in fig.1. (The cTP of the Rubisco small subunit was excluded for reasons detailed below.) From the results presented in fig.4, it is clear that a strong amphiphilic β-strand signal is present.

An inspection of the amino acids surrounding the stromal cleavage sites also indicates that a fairly well-conserved sequence Val-X-Ala is present in positions −3 to −1 (fig.1). A similar, though less well-conserved, (Val/Ile)-X-(Ala/Cys)↓Ala pattern can be discerned in cTPs from higher plants (Von Heijne, in preparation). We note that this pattern is absent from the cleavage site of the Rubisco small subunit cTP from *C. reinhardtii*, which suggests that this transit peptide may undergo multiple cleavages upon import.

3.5. *Lumen-targeting cTPs often contain one negative charge*

As in higher plants, *C. reinhardtii* cTPs from lumenal proteins are mosaic structures with an N-terminal stroma-targeting cTP attached to a C-terminal apolar domain (fig.1). The N-terminal domains show all the characteristics of *C. reinhardtii* stroma-targeting cTPs. The C-terminal domains are very similar to the corresponding domains of lumen-targeting cTPs from higher plants and to bacterial signal peptides [1]. In four of the five *C. reinhardtii* lumen-targeting cTPs, one negatively charged aspartate or glutamate is found in the N-terminal part of the lumen-targeting domain. One or more acidic residues is present in similar positions in most lumen-targeting cTPs from higher plants (not shown). The significance of this observation is unclear, but it may be related to the targeting to the thylakoid membrane. Interestingly, pea cytochrome *f*, a plastid-encoded protein located in the thylakoid membrane, also has an acidic residue in its positively charged N-terminal region [14].

4. DISCUSSION

From the analysis presented above, the *C. reinhardtii* cTP appears to be an interesting 'hybrid' between the mTPs and the higher plant cTPs. It contains both the potential amphiphilic α-helix of mTPs and the potential amphiphilic β-strand of higher plant cTPs. The presence of an amphiphilic α-helix may explain why a *C. reinhardtii* cTP can direct proteins into yeast mitochondria [2]. The differences between *C. reinhardtii* and higher plant cTPs may explain why the precursor of

the Rubisco small subunit from *C. reinhardtii* is incorrectly processed by chloroplasts from spinach and pea [15].

The structure of *C. reinhardtii* cTPs poses an interesting problem for mitochondrial import. Does *C. reinhardtii* use similar signals for both mTPs and cTPs? If that is the case, is *C. reinhardtii* more prone to mistargeting between mitochondrion and chloroplast? To answer these questions, we need sequences of mTPs from *C. reinhardtii*. A cDNA sequence for mitochondrial cytochrome *c* has been published [16] but, unfortunately, this protein does not have a targeting peptide. Experiments to isolate other nuclear-encoded mitochondrial proteins are currently in progress.

According to the endosymbiont hypothesis, chloroplasts originate from prokaryotes similar to cyanobacteria; a prokaryotic intruder gradually lost its autonomy, developed into an organelle and became more and more dependent on the nuclear gene activity of its host [17,18]. Since there are indications that the mitochondria of *C. reinhardtii* and higher plants are of different origin [19], it is tempting to speculate that their chloroplasts are also the result of two different events of endosymbiosis, in which two similar photosynthetic prokaryotes invaded two different eukaryotic hosts. If that is the case, the mechanisms for chloroplast protein import could have developed independently, resulting in different structures of the cTPs. In contrast, the signals for protein translocation through the thylakoid membrane (the C-terminal domain of lumen-targeting cTPs) in *C. reinhardtii* and higher plants are very similar, indicating a common origin of these signals. In fact, these signals are probably much older than the signals for chloroplast protein import. Similar structures can be found in signal peptides involved in routing proteins into the secretory pathway in prokaryotic cells [20]. In particular, proteins located in the thylakoid lumen in cyanobacteria have signal peptides of this type [21–23]. Thus, according to the endosymbiont hypothesis, these signal peptides were already present in the prokaryote(s) that developed into chloroplast(s). When genes were transferred from the endosymbiont/organelle to the nuclear genome, different signals for chloroplast protein import may have been added in *C. reinhardtii* and higher plants.

An open question is what determines specificity of protein import into chloroplasts and mitochondria in plant and algal cells. *C. reinhardtii* provides interesting possibilities for studying this subject. Mutants lacking chloroplastic or mitochondrial proteins are available [24] and a reliable nuclear transformation system has recently been established for this alga [25,26]. Thus, *C. reinhardtii* will be an interesting system for studies of protein sorting and import, in which mutant and chimeric proteins can be tested in vivo. In addition, the short length of *C. reinhardtii* cTPs facilitates in vitro studies on synthetic peptides corresponding to cTPs. Studies on synthetic peptides corresponding to mTPs have revealed that mTPs are surface active and insert spontaneously into membranes [27]. Our analysis here suggests that *C. reinhardtii* cTPs will be found to have similar properties.

Acknowledgement: This work was supported by a grant from the Swedish Natural Sciences Research Council to GvH.

REFERENCES

[1] Von Heijne, G., Steppuhn, J. and Herrmann, R.G. (1989) Eur. J. Biochem. 180, 535–545.
[2] Hurt, E.C., Soltanifar, N., Goldschmidt-Clermont, M., Rochaix, J.D. and Schatz, G. (1986) EMBO J. 5, 1343–1350.
[3] Boutry, M., Nagy, F., Poulsen, C., Aoyagi, K. and Chua, N.-H. (1987) Nature 328, 340–342.
[4] Franzén, L.-G., Frank, G., Zuber, H. and Rochaix, J.-D. Mol. Gen. Genet. 219, 137–144.
[5] Franzén, L.-G., Frank, G., Zuber, H. and Rochaix, J.-D. (1989) Plant Mol. Biol. 12, 463–474.
[6] Goldschmidt-Clermont, M. and Rahire, M. (1986) J. Mol. Biol. 191, 421–432.
[7] Yu, L.M. and Selman, B.R. (1988) J. Biol. Chem. 263, 19342–19345.
[8] Imbault, P., Wittemer, C., Johanningmeier, U., Jacobs, J.D. and Howell, S.H. (1988) Gene 73, 397–407.
[9] Mayfield, S.P., Rahire, M., Frank, G., Zuber, H. and Rochaix, J.-D. (1987) Proc. Natl. Acad. Sci. USA 84, 749–753.
[10] Mayfield, S.P., Schirmer-Rahire, M., Frank, G., Zuber, H. and Rochaix, J.-D. (1989) Plant Mol. Biol. 12, 683–693.
[11] Merchant, S. and Bogorad, L. (1987) J. Biol. Chem. 262, 9062–9067.
[12] Eisenberg, D. (1984) Annu. Rev. Biochem. 53, 595–623.
[13] Von Heijne, G. (1986) EMBO J. 5, 1335–1342.
[14] Willey, D.L., Auffret, A.D. and Gray, J.C. (1984) Cell 36, 555–562.
[15] Mishkind, M.L., Wessler, S.R. and Schmidt, G.W. (1985) J. Cell Biol. 100, 226–234.
[16] Amati, B.B., Goldschmidt-Clermont, M., Wallace, C.J.A. and Rochaix, J.-D. (1988) J. Mol. Evol. 28, 151–160.
[17] Raven, P.H. (1970) Science 169, 641–646.
[18] Margulis, L. (1981) Symbiosis in Cell Evolution, W.H. Freeman, San Fransisco.
[19] Gray, M.W. and Boer, P.H. (1988) Philos. Trans. R. Soc. Lond. Ser. B 319, 135–147.
[20] Von Heijne, G. (1985) J. Mol. Biol. 184, 99–105.
[21] Wallace, T.P., Stewart, A.C., Pappin, D. and Howe, C.J. (1989) Mol. Gen. Genet. 216, 334–339.
[22] Philbrick, J.B. and Zilinskas, B.A. (1988) Mol. Gen. Genet. 212, 418–425.
[23] Kuwabara, T., Reddy, K.J. and Sherman, L.A. (1987) Proc. Natl. Acad. Sci. USA 84, 8230–8324.
[24] Rochaix, J.-D. (1987) FEMS Microbiol. Rev. 46, 13–34.
[25] Debuchy, R., Purton, S. and Rochaix, J.-D. (1989) EMBO J. 8, 2803–2809.
[26] Fernández, E., Schnell, R., Ranum, L.P.W., Hussey, S.C., Silflow, C.D. and Lefebvre, P.A. (1989) Proc. Natl. Acad. Sci. USA 86, 6449–6453.
[27] Tamm, L.K. (1986) Biochemistry 25, 7470–7476.