# Different lumen-targeting pathways for nuclear-encoded versus cyanobacterial/plastid-encoded Hcf136 proteins

Peter J. Hynds[a], Henning Plücken[b], Peter Westhoff[b], Colin Robinson[a,*]

[a] *Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK*
[b] *Institut für Entwicklungs und Molekularbiologie der Pflanzen, Heinrich-Heine-Universität, Universitätsstrasse 1, 40225 Düsseldorf, Germany*

Received 13 December 1999; received in revised form 12 January 2000

Edited by Gunnar von Heijne

**Abstract Lumenal proteins are transported across the thylakoid membrane by two very different pathways: Sec-dependent or twin-arginine translocase (Tat)-dependent, where the substrate protein can be transported in a folded state. We present the first evidence that a given protein can be targeted by different pathways in different organisms. *Arabidopsis* Hcf136 is targeted exclusively by the Tat pathway in pea chloroplasts and no Sec-dependent transport is evident even when the twin-arginine is replaced by twin-lysine. However, twin-arginine motifs are absent from the presequences of Hcf136 proteins encoded by plastid or cyanobacterial genomes, strongly implying translocation by another pathway (presumably Sec). We suggest that the Hcf136 protein was transferred to the Tat pathway when the gene became incorporated into the nuclear genome, possibly due to the tighter folding associated with the more involved, post-translational targeting pathway.**



*Key words:* Chloroplast; Protein transport; Twin-arginine translocase; Signal peptide; Sec system

## 1. Introduction

Two very different mechanisms are used for the translocation of lumenal proteins across the thylakoid membrane of plant chloroplasts, and the properties of the protein translocation systems have been the focus of considerable attention in recent years. Most lumenal proteins are encoded in the nucleus and hence imported across the chloroplast envelope, apparently by a similar mechanism that is also used by most imported proteins (reviewed in [1]). Once inside the chloroplast, however, these proteins are translocated across the thylakoid membrane by either a Sec- or ΔpH-dependent mechanism (reviewed in [2,3]). The Sec-dependent mechanism appears to be similar to bacterial Sec-type protein export systems, involving stromal SecA activity, ATP hydrolysis and a membrane-bound SecYEG translocon [4–7]. The Sec-independent translocation system has completely different properties. Proteins such as the 23 and 16 kDa photosystem II proteins (23K, 16K) are likewise synthesised with a cleavable lumen-targeting peptide, but these precursors are translocated across the thylakoid membrane by a mechanism that requires a ΔpH but not nucleoside triphosphates [8,9]. The choice of translocation pathway is specified largely by the type of lumen-targeting signal present [10,11, and a key determinant for ΔpH-dependent targeting is the presence of an essential twin-arginine motif in the lumen-targeting signal [12].

Recent studies have confirmed the existence of a related pathway in bacteria (reviewed in [13]). The cloning of the Hcf106 component of the ΔpH-dependent translocase [14,15] revealed the presence of related genes in numerous bacteria and several have now been shown to be required for the Sec-independent export of a subset of periplasmic proteins [16–19]. Substrates for the bacterial system are similarly synthesised with twin-arginine signal peptides [20] and the bacterial translocase has been termed the twin-arginine translocation (Tat) system. Both the thylakoidal and bacterial systems appear to be primarily used for the translocation of fully-folded proteins [20–22]. This ability to transport folded protein domains is unusual among known protein transporters and is unique, to date, among those that operate in tightly sealed energy-transducing membranes.

It is essentially accepted that chloroplasts evolved from endosymbiotic cyanobacterial-type organisms, in which case the chloroplastic Tat system was also inherited from such an organism. This appears to be the case, since the known higher plant *tat* genes are highly homologous to open reading frames (ORFs) in the *Synechocystis* genome (not shown). Some cyanobacterial Tat substrates have probably also been maintained in chloroplasts, and we would predict that these would continue to utilise the Tat pathway. However, none has been identified to date and all of the known chloroplastic Tat substrates (23K, 16K, PSI-N, PSII-T and P16) are absent from the *Synechocystis* genome. Here, we report a situation that is, to date, unique: the apparent transfer of a lumenal protein to a different pathway during the course of evolution. Hcf136 is a lumenal protein in chloroplasts that is required for photosystem II stability or assembly [23]. We show that it is targeted exclusively by the Tat pathway in chloroplasts, yet the plastid-encoded homologue in the alga *Cyanophora paradoxa* and the homologous cyanobacterial proteins all lack a twin-arginine and are therefore predicted to be Sec-dependent. This is the first example of a lumenal protein following different targeting pathways in different organisms, and we speculate that the higher plant protein may have been forced onto the Tat pathway as a consequence of the more complex import pathway.

## 2. Materials and methods

### 2.1. Import assays

Precursor proteins were synthesised in vitro by transcription of cDNA clones encoding wheat pre-23K or *Arabidopsis* pre-Hcf136 [23] followed by translation in a wheat germ lysate in the presence of [$^{35}$S]methionine. Assays for the import of proteins by intact chlor-

*Corresponding author. Fax: (44)-1203-523701.
E-mail: cg@dna.bio.warwick.ac.uk


PII: S0014-5793(00)01129-7

oplasts and isolated pea thylakoids were as in [24]. Inhibitor experiments involved the use of nigericin at 4 μM (+10 mM KCl), Na azide at 10 mM or wheat pre-23K competitor as detailed in [25].

*2.2. Constructs*

A twin-lysine mutant of *Arabidopsis* pre-Hcf136 was generated by site-specific mutagenesis of the cDNAs in which 4 bp mutations were introduced by inverse PCR mutagenesis. Back-to-back primers were designed complementary to the target region, with a single bp gap between the 5′ ends of the forward and reverse primers (necessary due to the insertion of an extra, random base by *Taq* polymerase, and therefore the gap between the primers was positioned at the third base of a wobbly codon, to ensure the fidelity of amino acid at this codon). The desired mismatch was placed close to the 5′ end of the forward primer. For mutagenesis of the *hcf136* gene (accession Y15628), forward and reverse primers were 5′-TTCAGTAAGAAGGAGC-TTCTGTACCAATCG-3′ and 5′-GATAAGGATGAAGAGGACG-ATG-3′, respectively. Prior to iPCR, the primers were phosphorylated with T4 kinase (Gibco). Each iPCR reaction (10 μl final volume) contained 0.5 μM each phosphorylated primer, 200 μM each dATP, dCTP, dGTP and dTTP, 1.5 mM $MgCl_2$, 1×PCR buffer (Gibco) and 0.5 U *Taq* polymerase (Gibco). The iPCR reaction was carried out for 35 cycles, each cycle consisting of a denaturation step (20 s, 94°C), an annealing step (20 s, 60°C) and a polymerisation step (20 s, 72°C). After 35 cycles, samples were heated for 10 min at 72°C, and electrophoresed through a 1% agarose/TBE gel, before being visualised with ethidium bromide. The band was extracted and self-ligated using T4 ligase (Gibco) and transformed into *Escherichia coli*. Positive clones were sequenced.

## 3. Results and discussion

*3.1. The signal peptide of imported pre-Hcf136, but not of plastid-encoded/cyanobacterial pre-Hcf136 homologs, contains a twin-arginine*

*Arabidopsis* Hcf136 is known to be localised in the thylakoid lumen [23] and the protein is synthesised with an apparently typical bipartite lumen-targeting presequence in which two distinct regions can be discerned: an N-terminal 'envelope transit peptide', which specifies targeting into the chloroplast, followed by a thylakoid lumen-targeting signal peptide. As shown in Fig. 1, the latter contains three characteristic domains [2,3]: a charged N-terminal domain, a hydrophobic core domain (underlined) and a more polar C-terminal domain ending with Ala-Xaa-Ala, the consensus signal for the thylakoidal processing peptidase in chloroplasts [2]. Homologous ORFs are present in the plastid genome of the unicellular alga *C. paradoxa*, and in two cyanobacterial genomes: those of *Synechocystis* PCC6803 and *Anabaena variabilis*. The proteins encoded by these ORFs each contain an N-terminal hydrophobic region at the position corresponding to the *Arabidopsis* signal peptide, and we therefore assume that each is similarly synthesised with a signal peptide containing the three domains described above. Putative cleavage sites are denoted with asterisks.

It is notable that the signal peptide of the *Arabidopsis* protein contains a twin-arginine motif (italicised) just prior to the hydrophobic region, suggesting that it might be a substrate for the Tat system. This was considered to be of significant potential interest because none of the other Hcf136 signal peptides contain this motif, almost certainly precluding transport by this pathway (see later).

*3.2. Arabidopsis pre-Hcf136 is targeted exclusively by the Tat pathway*

The mere presence of a twin-arginine is not a definite indicator of targeting pathway because Sec-type signal peptides can also contain twin-arginine at this position, at least in bacteria. However, lumen-targeting pathways can be dissected with precision using in vitro translocation assays because several diagnostic inhibitors of the Sec and Tat pathways are available. Transport across the thylakoid membrane by the Tat pathway is totally dependent on the transthylakoidal ΔpH and hence sensitive to the proton ionophore nigericin, while Sec-dependent transport is inhibited (though not completely) by azide, which inhibits SecA action [11]. Most importantly, competition assays have proved to be very powerful in this type of study because over-expressed precursors of Tat substrates compete very effectively with other Tat substrates while having no apparent effect on the transport of Sec substrates [25]. Fig. 2 shows experiments in which these inhibitors have been used in assays for the import and sorting of pre-Hcf136 by intact pea chloroplasts.

In the control panel, the import of *Arabidopsis* pre-Hcf136 has been analysed in the absence of inhibitors. The data show that the protein is synthesised as a 44 kDa precursor which is imported into the organelles and converted to two polypep-

```
A.thal      MASLQLCDGYLLFKPSVSPRFLSQRISHRLIPKASSSPPPSPSPSSSSSSLSFSRRELLY
C.para      ------------------------------------------------------MILNWRKV
Synech      ----------------------------------------------MPVKFPSLKFEQLKQ
Anabaena    ----------------------------------------------------MKSWQK

A.thal      QSAAVSLSLSSIVGPARA♦DEQLSEWERVFLPIDPGVVLLDIAFVPDEPSRGFLLGTRQTL
C.para      IVSFLVIILTNFYNISFTHA♦ESYKWEQIPLNTDE--ILLDIGFVPDQPQRGWLLGTRSTL
Synech      LVLVAAIAVFCVSCSHVPDLA♦FNPWQEIALETDS--TFADIAFTED-PNHGWLVGTKETI
Anabaena    IFALLVVLLLCIGCSKVPSTS♦YNPWAVVSLPTEA--KLLDIAFTEN-PQHGFLVGSNATL
                   . .               *  . *  .      ** *    * .* *.*.  *.

A.thal      LLETKDGGSTWNPRSIPSAEEEDFNYRFNSISFKGKEGWIIGKPAILLYTADAGENWDRIP
C.para      LFETTDKGKTWELRSLNLEDDK---YRLNSISFSGKEGWVTGKPAILLHTTDGGSSWSRIP
Synech.     IFETTDGGDTWEQKLIDLGEEK---ASFSAVSFSGNEGWITGKPSILLHTTDGGQTWARIP
Anabaena    LLETNDGGNNWQPLNLALDDDR---YRFDSVSFAGKEGWIAGEPSLLLHTTDEGRSWSRIP
            . ** * * .*    .   ..        ..** * ***. * *..** *.* *  * ***
```

Fig. 1. Primary sequences of plastid and cyanobacterial Hcf136 proteins and their putative signal peptides. The figure shows the N-terminal region of *Arabidopsis* pre-Hcf136 (A. thal) aligned against the corresponding regions of ORF333 from the plastid genome of *C. paradoxa* (C. para; accession U30821), ORF341 from *Synechocystis* sp. PCC6803 (Synech; D90903) and ORF1 from *A. variabilis* (AJ012181). Identical residues are denoted by asterisks and conserved residues by dots. Hydrophobic regions of the predicted signal peptides are underlined and putative terminal cleavage sites indicated by ♦. A twin-arginine in the signal peptide of *Arabidopsis* Hcf136 is italicised.

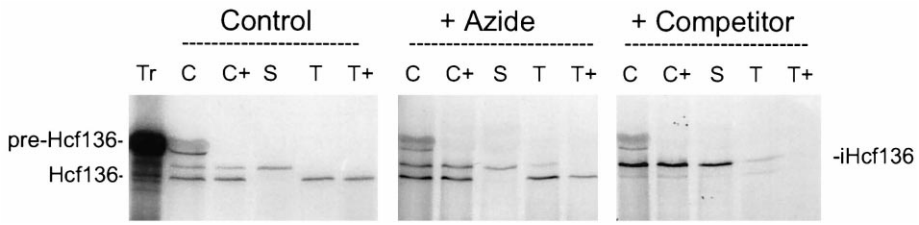


Fig. 2. Targeting of *Arabidopsis* Hcf136 by the ΔpH/Tat pathway in chloroplasts. Pre-Hcf136 was synthesised by transcription-translation of a cDNA clone and incubated with intact pea chloroplasts. Import incubations contained no added inhibitors (control panel), 10 mM Na azide (+azide panel) or 3 μM over-expressed pea pre-23K (+competitor panel). After incubation, samples were analysed of the chloroplasts (lane C), protease-treated chloroplasts (C+) and the stromal (S) and thylakoid (T) fractions after lysis of the organelles in 10 mM HEPES-KOH, pH 8.0, 5 mM $MgCl_2$ and centrifugation for 10 min. Lane T+: thylakoid fraction was incubated with 0.2 mg/ml thermolysin for 30 min on ice. Lane Tr: translation product. iHcf136, intermediate form.

tides: a 37 kDa thylakoid-associated protein (lane T) which is believed to be the mature protein, together with a larger intermediate that is found predominantly in the stroma (lane S), presumably en route to the thylakoids. The mature protein is resistant to proteolysis (lane T+) and is thus located in the lumen. These import data closely resemble those observed with other lumenal proteins, where intermediate forms are usually generated by removal of the envelope transit peptide by the stromal processing peptidase [2,3], and they confirm that *Arabidopsis* Hcf136 is indeed synthesised with a bipartite presequence.

Azide has no detectable effect although substantial inhibition was observed in parallel tests with a Sec substrate, pre-33K (data not shown). Significantly, the 'competitor' panel shows that the presence of relatively high concentrations of unlabeled, over-expressed pea pre-23K results in a complete block in transport across the thylakoid membrane, with the result that the intermediate-size protein accumulates in the stroma. Taken together, these data clearly demonstrate that pre-Hcf136 is translocated primarily, if not exclusively, by the Tat pathway.

As predicted for Tat-dependent translocation, translocation across the thylakoid membrane depends totally on a trans-thylakoidal ΔpH. This can be observed in chloroplast import assays, where accumulation of the stromal intermediate is again observed (not shown) or in assays for the import of proteins into isolated thylakoids as shown in Fig. 3. Incubation of the in vitro synthesised precursor with washed pea thylakoids results in the appearance of mature-size protein which is resistant to proteolysis of the membranes and hence imported. Import is completely blocked by nigericin, consistent with Tat-specific translocation.

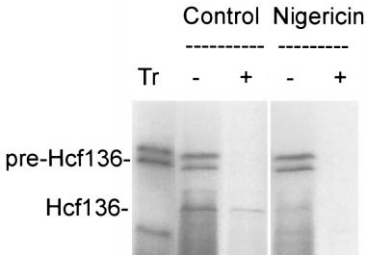


Fig. 3. Import of pre-Hcf136 into isolated thylakoids requires a ΔpH. Pre-Hcf136 (lane Tr) was incubated with isolated pea thylakoids in the absence of inhibitors (control) or the presence of 4 μM nigericin. After incubation, the thylakoids were pelleted and analysed directly (−) or after incubation with 0.2 mg/ml thermolysin for 30 min on ice (+).

These data are are of interest because the *Synechocystis* and *Anabaena* pre-Hcf136 proteins lack arginine residues in the N-terminus of the signal peptide, and are thus almost certainly not transported by this pathway. Translocation is probably by the Sec pathway, although this remains to be confirmed. We therefore considered it possible that Hcf136 might be transported by both pathways in chloroplasts, although primarily by the Tat pathway given the competition data shown above. This possibility was probed by removing the twin-arginine motif known to be important for Tat-dependent translocation [12]; any residual translocation would strongly suggest that the protein could utilise the Sec pathway, especially in the absence of any competition for substrate by the Tat machinery. Fig. 4 shows import experiments conducted using this twin-Lys mutant. The data show that the wild-type pre-protein is again imported and sorted as expected, but the twin-Lys mutant accumulates quantitatively in the stroma and no translocation into the lumen is apparent. The presence of twin-Lys at this position is no barrier to the thylakoidal Sec system [26] and we conclude that pre-Hcf136 is totally unable to use the Sec pathway.

### 3.3. Implications for the evolution of the Tat pathway

Hcf136 is the sixth confirmed substrate for the Tat system in chloroplasts and the first for which homologous proteins can be identified in cyanobacteria or the plastid genomes of algae. However, none of the Hcf136 homologues in *Cyanophora*/*Synechocystis*/*Anabaena* contains a twin-arginine motif, and we propose that this protein must utilise alternative pathways in these organisms. This scenario assumes that Tat-specific signal peptides do contain an important and conserved twin-arginine in cyanobacteria and *Cyanophora*, but this is extremely probable. Tat-type targeting peptides contain this motif in a very wide range of bacterial species [20] and it

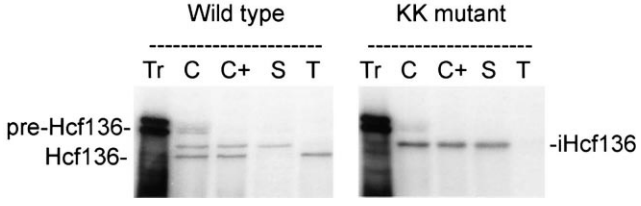


Fig. 4. Mutagenesis of the twin-arginine in pre-Hcf136 blocks translocation across the thylakoid membrane. *Arabidopsis* pre-Hcf136 and a mutant containing twin-Lys in place of the twin-arginine motif (KK-mutant) were imported into pea chloroplasts and the organelles fractionated as in Fig. 2. Other symbols as in Figs. 2 and 3.

would be totally unexpected if twin-arginine motifs were important in chloroplasts but not in cyanobacteria, especially since these are the bacteria that are most closely related to chloroplasts. The Tat system has not been studied in cyanobacteria but analysis of one predicted substrate supports the above premise. The Rieske protein binds FeS centres in the cytoplasm, is firmly predicted to be transported by the Tat pathway and is accordingly synthesised with an RR-signal peptide in a wide range of bacteria [20]. We note that this protein is likewise synthesised with apparently typical RR-signal peptides in various cyanobacteria, such as *Synechocystis*, *Agmenellum quadruplicatum*, *Synechococcus elongatus* and *C. nostoc* (not shown).

It is nevertheless possible that Hcf136 may be targeted by the Tat pathway in other cyanobacteria, or in other eukaryotic algae when plastid-encoded. In other words, targeting by the Tat pathway may not be strictly correlated with transfer of the *hcf136* gene to the nucleus. However, no other known bacterial proteins are believed to be Sec-dependent in some species but Tat-dependent in others [20]. Rather, the most likely explanation is that Hcf136 transferred to the Tat pathway when the *hcf136* gene moved to the nucleus after the initial endosymbiotic events.

Why might Hcf136 have become 'Sec-incompatible' at this point? One possibility is that Hcf136 is translocated either during or shortly after synthesis in cyanobacteria (or when encoded by the plastid genome), and has little opportunity to fold under these circumstances. In contrast, nuclear-encoded chloroplast proteins are transported across the thylakoid membrane in a strictly post-translational manner and probably have ample opportunity to fold in the stroma. Transport by the Sec pathway may be precluded if the protein folds at all tightly, and bacterial Sec systems are indeed unable to export a large proportion (in fact the majority) of cytoplasmic or heterologous proteins tested (e.g. [27]). The same problem may apply to those proteins that have arisen since the initial endosymbiotic events, such as 23K, 16K, PSI-N, PSII-T and P16. All are transported by the Tat pathway and yet none is believed to bind cofactors. We suggest that higher plants have made full use of the unique properties of the Tat pathway in order to evolve and efficiently target a range of novel, nuclear-encoded proteins whose folding properties precluded Sec-dependent translocation in chloroplasts.

*Acknowledgements:* This research was supported by Biotechnology and Biological Sciences Research Council Grant C08676 to C.R.

## References

[1] Heins, L., Collinson, I. and Soll, J. (1998) Trends Plant Sci. 3, 56–61.
[2] Robinson, C., Hynds, P.J., Robinson, D. and Mant, A. (1998) Plant Mol. Biol. 38, 209–221.
[3] Cline, K. and Henry, R. (1996) Annu. Rev. Cell Dev. Biol. 12, 1–26.
[4] Yuan, J., Henry, R., McCaffery, M. and Cline, K. (1994) Science 266, 796–798.
[5] Nakai, M., Goto, A., Nohara, T., Sugita, D. and Endo, T. (1994) J. Biol. Chem. 269, 31338–31341.
[6] Schuenemann, D., Amin, P., Hartmann, E. and Hoffman, N.E. (1999) J. Biol. Chem. 274, 12177–12182.
[7] Mori, H., Summer, E.J., Ma, X. and Cline, K. (1999) J. Cell Biol. 146, 45–55.
[8] Mould, R.M. and Robinson, C. (1991) J. Biol. Chem. 266, 12189–12193.
[9] Cline, K., Ettinger, W.F. and Theg, S.M. (1992) J. Biol. Chem. 267, 2688–2696.
[10] Robinson, C., Cai, D., Hulford, A., Brock, I.W., Michl, D., Hazell, L., Schmidt, I., Herrmann, R.G. and Klösgen, R.B. (1994) EMBO J. 13, 279–285.
[11] Henry, R., Kapazoglou, A., McCaffery, M. and Cline, K. (1994) J. Biol. Chem. 269, 10189–10192.
[12] Chaddock, A.M., Mant, A., Karnauchov, I., Brink, S., Herrmann, R.G., Klösgen, R.B. and Robinson, C. (1995) EMBO J. 14, 2715–2722.
[13] Dalbey, R.E. and Robinson, C. (1999) Trends Biochem. Sci. 24, 17–22.
[14] Voelker, R. and Barkan, A. (1995) EMBO J. 14, 3905–3914.
[15] Settles, M.A., Yonetani, A., Baron, A., Bush, D.R., Cline, K. and Martienssen, R. (1997) Science 278, 1467–1470.
[16] Santini, C.L., Ize, B., Chanal, A., Müller, M., Giordano, G. and Wu, L.F. (1998) EMBO J. 17, 101–112.
[17] Weiner, J.H., Bilous, P.T., Shaw, G.M., Lubitz, S.P., Frost, L., Thomas, G.H., Cole, J.A. and Turner, R.J. (1998) Cell 93, 93–101.
[18] Sargent, F., Bogsch, E.G., Stanley, N.R., Wexler, M., Robinson, C., Berks, B.C. and Palmer, T. (1998) EMBO J. 17, 3640–3650.
[19] Bogsch, E.G., Sargent, F., Stanley, N.R., Berks, B.C., Robinson, C. and Palmer, T. (1998) J. Biol. Chem. 273, 18003–18006.
[20] Berks, B.C. (1996) Mol. Microbiol. 22, 393–404.
[21] Clark, S.A. and Theg, S.M. (1997) Mol. Biol. Cell 8, 923–934.
[22] Hynds, P.J., Robinson, D. and Robinson, C. (1998) J. Biol. Chem. 273, 34868–34874.
[23] Meurer, J., Plücken, H., Kowallik, K.V. and Westhoff, P. (1998) EMBO J. 17, 5286–5297.
[24] Brock, I.W., Hazell, L., Michl, D., Nielsen, V.S., Møller, B.L., Herrmann, R.G., Klösgen, R.B. and Robinson, C. (1993) Plant Mol. Biol. 23, 717–725.
[25] Cline, K., Henry, R., Li, C. and Yuan, J. (1993) EMBO J. 12, 1405–4114.
[26] Brink, S., Bogsch, E., Mant, A. and Robinson, C. (1997) Eur. J. Biochem. 245, 340–348.
[27] Simonen, M. and Palva, I. (1993) Microbiol. Rev. 57, 109–137.