The EMBO Journal vol.14 no.12 pp.2715–2722, 1995

A new type of signal peptide: central role of a twin-arginine motif in transfer signals for the ΔpH-dependent thylakoidal protein translocase

Alison M.Chaddock, Alexandra Mant, Ivan Karnauchov[1], Susanne Brink, Reinhold G.Herrmann[1], Ralf Bernd Klösgen[1] and Colin Robinson[3]

Department of Biological Sciences, University of Warwick, Coventry, CV4 7AL, UK and [1]Botanisches Institut der Ludwig-Maximilians-Universität, Menzinger Strasse 67, 80638 München, Germany

[3]Corresponding author

The ΔpH-driven and Sec-related thylakoidal protein translocases recognise distinct types of thylakoid transfer signal, yet all transfer signals resemble bacterial signal peptides in structural terms. Comparison of known transfer signals reveals a single concrete difference: signals for the ΔpH-dependent system contain a common twin-arginine motif immediately before the hydrophobic region. We show that this motif is critical for the ΔpH-driven translocation process; substitution of the arg–arg by gln–gln or even arg–lys totally blocks translocation across the thylakoid membrane, and replacement by lys–arg reduces the rate of translocation by >100-fold. The targeting information in this type of signal thus differs fundamentally from that of bacterial signal peptides, where the required positive charge can be supplied by any basic amino acid. Insertion of a twin–arg motif into a Sec-dependent substrate does not alter the pathway followed but reduces translocation efficiency, suggesting that the motif may also repel the Sec-type system. Other information must help to specify the choice of translocation mechanism, but this information is unlikely to reside in the hydrophobic region because substitution by a hydrophobic section from an integral membrane protein does not affect the translocation pathway.

Key words: chloroplast/protein transport/signal peptide/thylakoid biogenesis

Introduction

Many proteins are targeted to extracytosolic compartments in both eukaryotic and prokaryotic cells, and a variety of translocation mechanisms have evolved to ensure the efficient and accurate delivery of proteins to these compartments. In simple terms, the translocation mechanisms usually involve the synthesis of a given protein with a specific targeting signal, and the recognition of this signal by the appropriate translocation system. In many cases, including most of the proteins which are translocated across the mitochondrial and chloroplast envelopes, the bacterial plasma membrane and the endoplasmic reticulum (ER), the targeting signals take the form of cleavable N-terminal extensions which are removed following translocation across the target membrane. These targeting signals fall into two broad categories. The presequences of imported chloroplast and mitochondrial proteins are hydrophilic and basic, whereas proteins destined for transport across the ER and bacterial plasma membranes are synthesized with presequences, termed signal peptides, which lack primary structure homology but contain three distinct regions: a hydrophilic, basic N-terminal region (N-region), a hydrophobic core region (H-region) and a more polar C-terminal region (C-region) which contains short-chain residues at the –3 and –1 positions (reviewed by von Heijne *et al.*, 1989 and Izard and Kendall, 1994). The 'consensus' signal peptides differ slightly between eukaryotes and prokaryotes, but the two types of signal are on the whole remarkably similar.

Some imported chloroplast proteins are synthesized with two types of targeting signal. Cytosolically synthesized thylakoid lumen proteins, such as plastocyanin (PC) and the 33, 23 and 16 kDa proteins (33K, 23K, 16K) of the oxygen-evolving complex, are imported by a two-phase pathway which requires the operation of two distinct translocation systems. Accordingly, they are synthesized with bipartite presequences containing two signals in tandem. The first 'envelope transit' signal ensures delivery into the stromal phase by means of a translocase located in the envelope membranes; this signal is usually (though not always) removed by the stromal processing peptidase (Hageman *et al.*, 1986; James *et al.*, 1989; Michl *et al.*, 1994; Nielsen *et al.*, 1994). The second 'thylakoid transfer' signal then directs translocation across the thylakoid membrane, after which it is removed by a thylakoidal processing peptidase. Intriguingly, thylakoid transfer signals share key features with signal peptides, namely the possession of H-regions and short-chain –3 and –1 residues; the N-regions are likewise charged although equal numbers of basic and acidic residues are found in those thylakoid transfer signals which have been delineated to date (von Heijne *et al.*, 1989; Bassham *et al.*, 1991). These similarities are not simply coincidental: the presequences of 33K and PC have been shown to direct export of proteins in *Escherichia coli*, and the reaction specificities of bacterial and thylakoidal processing peptidases are virtually identical (Seidler and Michel, 1990; Meadows and Robinson, 1991; Shackleton and Robinson, 1991; Haehnel *et al.*, 1994).

Since all thylakoid transfer signals contain superficially similar N-, H- and C-regions, it was originally considered likely that they were transported across the thylakoid membrane by a common, Sec-type mechanism inherited from the cyanobacterial type progenitor of higher plant chloroplasts. In fact, the situation is considerably more complex and there is now good evidence that lumenal proteins are transported by two completely different

mechanisms (reviewed in Robinson and Klösgen, 1994). A subset of proteins, including 33K, PC and PSI-F, are transported by a mechanism which requires the presence of a stromal protein factor and ATP (Hulford *et al.*, 1994; Karnauchov *et al.*, 1994; Mant *et al.*, 1994; Robinson *et al.*, 1994). SecA is involved in this mechanism, strongly suggesting that the translocation mechanism was indeed inherited from a prokaryotic ancestor (Nakai *et al.*, 1994; Yuan *et al.*, 1994). The rest of the known lumenal proteins (including 23K, 16K and PSI-N) are transported by a mechanism which requires neither stromal factors nor ATP, but which is instead reliant on the thylakoidal ΔpH (Mould and Robinson, 1991; Cline *et al.*, 1992; Klösgen *et al.*, 1992; Nielsen *et al.*, 1994). These proteins are believed to be absent from cyanobacteria, prompting speculation that their emergence in chloroplasts has been accompanied by the acquisition of a new system for their transport across the thylakoid membrane. Remarkably (in view of the similarities among thylakoid transfer signals), the two translocases recognize only their cognate thylakoid transfer signals; chimeric proteins containing the pre-sequence from one group and the mature protein from another are transported quantitatively by the mechanism specified by the presequence (Henry *et al.*, 1994; Robinson *et al.*, 1994). In view of these findings, we have examined the thylakoid transfer signals from all of the known lumenal proteins, and we have noted that all of the substrates for the ΔpH-dependent translocase possess a twin-arginine motif immediately before the H-region. In this report, we show that the presence of both arginines is crucial for the ΔpH-driven translocation process, and that this motif may also function as a negative signal for the Sec-dependent system. The composition of the H-region, on the other hand, does not appear to influence the import pathway into the thylakoid lumen.

Results

Thylakoid transfer signals for the ΔpH-dependent pathway possess a common twin-arg motif

A total of nine proteins have been shown to be targeted into or across the thylakoid membrane by means of cleavable bipartite presequences. Of these, 33K, PC and PSI-F are translocated across the thylakoid membrane by the ATP-dependent mechanism (Hulford *et al.*, 1994; Karnauchov *et al.*, 1994; Mant *et al.*, 1994; Robinson *et al.*, 1994). 23K, 16K, PSII-T and PSI-N are translocated by the ΔpH-dependent pathway (Mould *et al.*, 1991; Cline *et al.*, 1992; Klösgen *et al.*, 1992; Henry *et al.*, 1994; Mant *et al.*, 1994) and it appears likely that polyphenol oxidase is also translocated by this mechanism because dissipation of the ΔpH blocks thylakoid transport in intact chloroplasts (Sommer *et al.*, 1994). The ninth protein, CF_0II, is probably integrated into the thylakoid membrane by a different, spontaneous mechanism (Michl *et al.*, 1994). Figure 1 illustrates the thylakoid transfer signals of the substrates for the two types of protein translocase in higher plants. In two cases (wheat 23K and 33K), the start sites of the thylakoid transfer signals have been delineated by identification of the SPP cleavage sites within the presequences (Bassham *et al.*, 1991). The start site of *Silene pratensis* PC was also identified in this study and, because the transfer signals are very similar, the start

A. Chloroplast proteins

	N-region	H-region	C-region
Wh 33K	AFGVDAGARITCSLQSDIREVASKCADAAKMAGFALATSALLVSGATA		
Sp PC	ASLKNVGAAVVATAAAGLLAGNAMA		
Sp PSI-F	--QENDQQQPKKLELAKVGANAAAALALSSVLLSSWSVAPDAAMA		
Wh 23K	AQKNDEAASDAAVVTSRRAALSLLAGAAAIAVKVSPAAA		
Sp 16K	--NIRAQQVSAEAETSRRAMLGFVAAGLASGSFVKAVLA		
Bar PSI-N	--AAAKRVQVAPAKDRRSALLGLAAVFAATAASAGSARA		
Cot PSII-T	--AAKGAQVESVQMSGERKTEGNNGRREMMFAAAAAAICSVAGVATA		
Tom PPO	--NGNQDETNSVDRRNVLLGLGGLYGVANAIPLAASA		

B. Cyanobacterial proteins.

Synechocystis	33K	MRFRPSIVALLSVCFGLLTFLYSGSAFA
	PC	MSKKFLTILAGLLLVVSSFFLSVSPAAAA
	c553	MFKLFNQASRIFFGIALPCLIFLGGIFSLGNTALA
	PSI-F	MKHLLALLLAFTLWFNFAPSASA
	CtpA	MGKRTRRFWALAFSLLMGALIYLGNTPSALA
Synechococcus	33K	MRYRAFLAAFLAVCLGVLTACSSGPTAA
	c553	MKRILGTAIAALVVLLAFIAPAQA
Anabaena	c553	MKKIFSLVLLGIALFTFAFSSPALA

Fig. 1. Thylakoid transfer signals for the ΔpH-driven thylakoidal protein translocase contain a conserved twin-arg motif. (**A**) The figure shows the primary structures of the thylakoid transfer signals of wheat 33K (Meadows *et al.*, 1991), spinach PC (Rother *et al.*, 1986) and spinach PSI-F (Steppuhn *et al.*, 1988), all of which are targeted by the Sec-dependent pathway, and wheat 23K (James and Robinson, 1991), spinach 16K (Jansen *et al.*, 1987), barley PSI-N (Knoetzel and Simpson, 1993), cotton PSII-T (A.Kapazoglou and L.Dure, unpublished data) and tomato PPO (Newman *et al.*, 1993) which are transported by the ΔpH-driven mechanism. (**B**) The lower panel shows the transfer signals of cyanobacterial thylakoid lumen proteins elucidated by (from top to bottom) Philbrick and Zilinskas (1988), Briggs *et al.* (1990), Zhang *et al.* (1994), Chitnis *et al.* (1991), Shestakov *et al.* (1994), Kuwabara *et al.* (1987), Laudenbach *et al.* (1990) and Bovy *et al.* (1992). H-regions are underlined, charged residues are given in bold, and twin-arg motifs common to substrates for the ΔpH-driven system are italicized.

of the spinach PC signal is assumed to be at an analogous position within the presequence. All of the transfer signals contain H-regions and apparently similar C-regions, with short-chain residues at the –3 and –1 positions which are known to be important for the terminal processing reaction (Shackleton and Robinson, 1991). The N-regions are more variable and there are few clear-cut differences between the signals for the ΔpH- and ATP-dependent mechanisms. However, one feature common to all of the substrates for the ΔpH-dependent system is the presence of a twin-arg motif in the N-region immediately prior to the H-region (we have termed this position the N–H junction). This feature is conserved within all of the published structures for these substrates, which in total amount to at least 14 23K/16K/PSI-N/PSII-T/PPO sequences in higher plants (not shown). Equally striking is the observation that arginine is totally absent from this position in the published 33K/PC/PSI-F sequences from higher plants (which again total 14). A single lysine is present in every case with the exception of barley pre-PC, which contains lys–lys at the N–H junction. This is unexpected because, although signal peptides usually contain a basic residue at this position, the identity of the residue is deemed to be totally irrelevant; lysine and arginine are interchangeable in bacterial signal peptides, and even histidine is tolerated at the appropriate pH (Sasaki *et al.*, 1990). This point is borne out by analysis of prokaryotic thylakoid transfer peptides. As shown in Figure 1B, both lysine and arginine are found at the N–H junction of these signals in various species of

```
                    48  49  50  51  52  53  54
pre-23K     --------Val-Thr-Ser-Arg-Arg-Ala-Ala---
R51/52Q     --------Val-Thr-Ser-Gln-Gln-Ala-Ala---
R51K        --------Val-Thr-Ser-Lys-Arg-Ala-Ala---
R52K        --------Val-Thr-Ser-Arg-Lys-Ala-Ala---

                    45  46  47  48  49  50
pre-PC              Ala-Ser-Leu-Lys-Asn-Val---
PC-RR               Ala-Ser-Arg-Arg-Asn-Val---
```

Fig. 2. Mutations at the N–H junction in wheat pre-23K and spinach pre-PC. The figure illustrates the site-specific mutants generated at the N–H junctions of the two precursors; the residues of the adjacent H-regions are underlined.

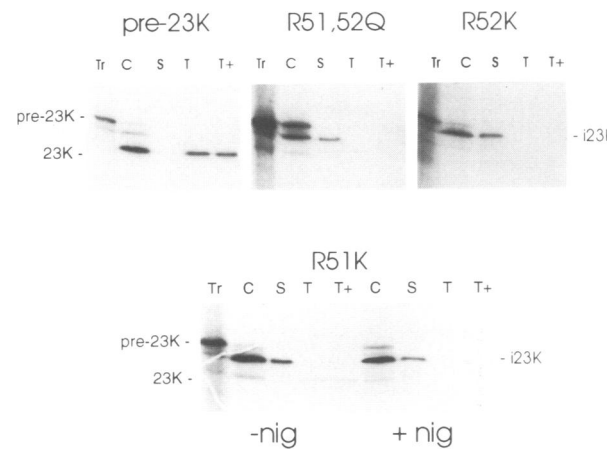

Fig. 3. Mutations within the 23K twin-arg motif either abolish, or drastically reduce, the efficiency of translocation across the thylakoid membrane in pea chloroplasts. Wild-type pre-23K and the mutant precursors shown in Figure 2 were synthesized *in vitro* by transcription–translation (lanes Tr) and incubated with intact pea chloroplasts as detailed in Materials and methods. After incubation, samples were analysed of the non-protease-treated chloroplasts (lanes C) and the stromal (lanes S) and thylakoid (lanes T) fractions after lysis of protease-treated chloroplasts. Lanes T+; the thylakoids were protease treated after lysis of the chloroplasts. In the case of the R51K mutant (lower panel), a parallel incubation was carried out in the presence of 2 μM nigericin (panel '+nig'). i23K, stromal intermediate form of 23K.

cyanobacteria, with arginine in fact the more common of the two.

These analyses give rise to some interesting possibilities. First, the twin-arg motif is a good candidate for the elusive targeting signal for the ΔpH-dependent translocase and, second, there is a clear possibility that the same twin-arg motif may act to repel the Sec machinery. These possibilities were tested by site-specific mutagenesis of the N–H junctions in pre-23K (targeted by the ΔpH-dependent pathway) and pre-PC (a substrate for the Sec-dependent mechanism).

The twin-arg motif in wheat pre-23K is critical for translocation across the thylakoid membrane

The importance of the twin-arg motif was tested directly by substituting these residues in wheat pre-23K using oligonucleotide-directed site-specific mutagenesis. In the first mutant (R51/52Q), both arginines were replaced by glutamine in order to simultaneously test the importance of the arginine side chains and the positive charges at these positions. In the other mutants (R51K and R52K), the arginines were replaced individually by lysine, hence the charge characteristics of these precursor proteins are unchanged. Finally, we inserted a twin-arg motif in place of the single lysine at the corresponding position in spinach pre-PC, to test whether this motif alters the import pathway for Sec-dependent substrates. These alterations are illustrated in Figure 2, and Figure 3 shows the results obtained when the mutant pre-23K proteins were imported into intact pea chloroplasts. In the control incubation, wild-type pre-23K is efficiently imported into the organelles, targeted into the thylakoid lumen and processed to the mature size (lanes T+ represent the thylakoid lumen contents). The absence of a detectable stromal intermediate form is confirmation that the intraorganellar sorting of this protein takes place with high efficiency under these conditions. In contrast, the three mutant 23K forms are all imported into the chloroplasts with high efficiency, but in each case the subsequent thylakoid-translocation step is drastically affected. Replacement of the arg–arg motif with either gln–gln or arg–lys (R51/52Q and R52K mutants, respectively) completely blocks translocation across the thylakoid membrane (top panel), and the only detectable imported protein is the stromal intermediate form (lanes S). The effects of introducing lys-arg are almost as drastic; in this case, some mature-size protein is found in the thylakoid lumen, but the predominant form is again the stromal intermediate form. Nigericin completely inhibits this low-level translocation of the R51K mutant (lower panel), indicating that the ΔpH-dependent pathway is being followed. The import characteristics of the R51K mutant are shown more effectively in a time-course analysis of import (Figure 4A), in which it is observed that the mature-size protein accumulates only at late time points in the incubation (note that the non-protease-treated samples contain a polypeptide which migrates slightly faster than mature-size 23K; this polypeptide is present in the translation mix and is not generated in the import incubation). At earlier time points (for example 5 min), the mature form is barely detectable.

The translocation of the R51K mutant was further examined (Figure 4B) using an assay for the import of proteins by isolated thylakoids (Brock *et al.*, 1993) in which the thylakoid transfer processes can be compared more quantitatively. Pre-23K is imported with high efficiency in this reconstituted system, and after only 5 min a substantial proportion of available wild-type precursor is internalized. However, no import whatsoever of the R51K mutant is observed, confirming the pivotal role of the twin-arg motif in this translocation mechanism. We estimate that the import efficiency of the R51K mutant is <1% of the control value, since import efficiencies of this order can be detected in this assay system. No import of the R51/52Q or R52K mutants into thylakoids has been observed, in agreement with the results of the chloroplast import assays (not shown).

The presence of the twin-arg motif does not divert a Sec-dependent substrate onto the ΔpH-dependent pathway

Is the presence or absence of this motif the sole means by which the two thylakoidal protein translocases recog-

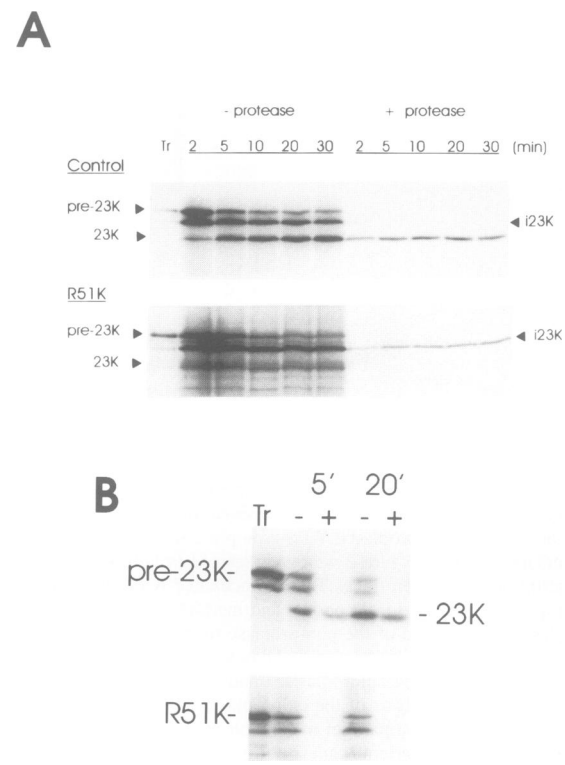

Fig. 4. Import and sorting kinetics of the R51K pre-23K mutant. (**A**) Pre-23K (top panel) and the R51K mutant (lower panel) were incubated with intact chloroplasts and samples of the incubation were removed at the indicated time points (in min). The chloroplasts were then pelleted and analysed directly (– protease) or after protease treatment of the organelles (+ protease). Symbols as in Figure 3. (**B**) Pre-23K and the R51K mutant were incubated under illumination with washed pea thylakoids for 5 or 20 min as indicated. After incubation, samples of the import mixtures were analysed directly (–) or after protease treatment of the thylakoids (+).

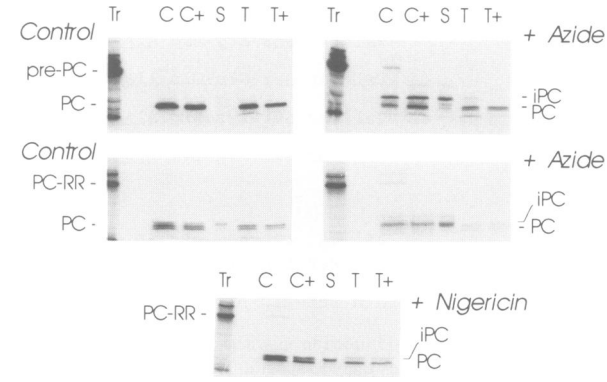

Fig. 5. The PC-RR mutant is transported across the thylakoid membrane by the Sec-dependent pathway. Pre-PC (top two panels) and the PC-RR mutant (lower three panels) were incubated with intact chloroplasts in the absence of any inhibitors, or in the presence of azide or nigericin as indicated. After incubation, samples were analysed of the chloroplasts (C) protease-treated chloroplasts (C+), and of the stromal and thylakoid fractions from protease-treated chloroplasts (lanes S and T, respectively). Lanes T+, protease-treated thylakoids. PC, mature-size PC; iPC, stromal intermediate forms of PC or PC-RR.

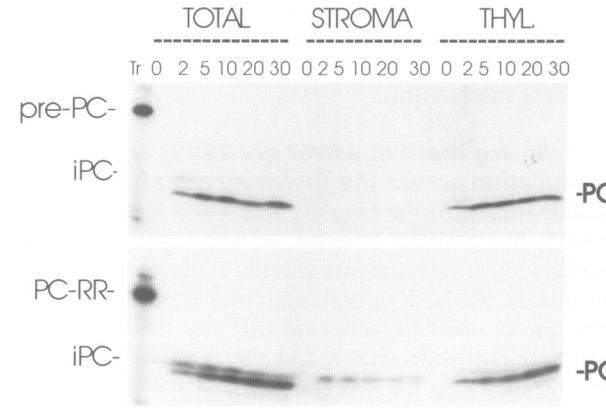

Fig. 6. The PC-RR mutant is transported across the thylakoid membrane at a much reduced rate. Pre-PC and PC-RR were incubated with intact pea chloroplasts for the times (in min) indicated above the lanes, after which they were diluted in 1 ml import buffer containing 10 mM Na azide and protease treated. Samples were analysed of the chloroplasts ('total') panel, and of the stromal and thylakoid (thyl.) fractions following lysis of the organelles and centrifugation. The thylakoid fraction was washed once before analysis.

nize their cognate substrates? We tested this possibility by introducing a twin-arg into the corresponding position within pre-PC and importing the mutant precursor (PC-RR) into chloroplasts in the presence of nigericin (an ionophore which inhibits translocation by the ΔpH-dependent pathway) and azide, which inhibits the Sec-dependent pathway. The data are shown in Figure 5. In the control incubations, wild-type PC is efficiently targeted into the thylakoid lumen and processed to the mature size, which occasionally runs as a close doublet in this gel system. No stromal intermediate is apparent, but this form becomes prominent when import takes place in the presence of the SecA inhibitor azide, in accordance with previous findings (Knott and Robinson, 1994). As expected from previous studies (Theg *et al.*, 1989; Knott and Robinson, 1994; Robinson *et al.*, 1994), dissipation of the thylakoidal ΔpH with nigericin has no detectable effect on import (data not shown). The PC-RR mutant is also imported into the lumen and processed to the mature size, although with somewhat diminished efficiency. In the absence of any inhibitors, most of the imported protein is found as the mature size form, but some stromal form is also found (denoted iPC). This stromal form has a greater mobility than the normal stromal intermediate, suggesting that the stromal intermediate either runs aberrantly in the gel due to the additional positive charge, or is smaller due to abnormal cleavage by the stromal processing peptidase or another stromal protease. Tests with the diagnostic inhibitors clearly indicate that this mutant is targeted by the normal Sec-dependent pathway; the import characteristics are largely unaffected by nigericin (the stromal form is only slightly more pronounced), but azide again leads to inhibition of thylakoid transfer, and the stromal intermediate-size form accumulates.

The data shown in Figure 5 appear to suggest that the PC-RR mutant is transported across the thylakoid membrane with almost wild-type efficiency, but a more accurate picture of the import kinetics is obtained from time-course studies. In fact, Figure 6 shows that at early time points the stromal form of PC-RR is as abundant as the mature size protein, whereas wild-type PC is again present almost exclusively as the mature form. Moreover,

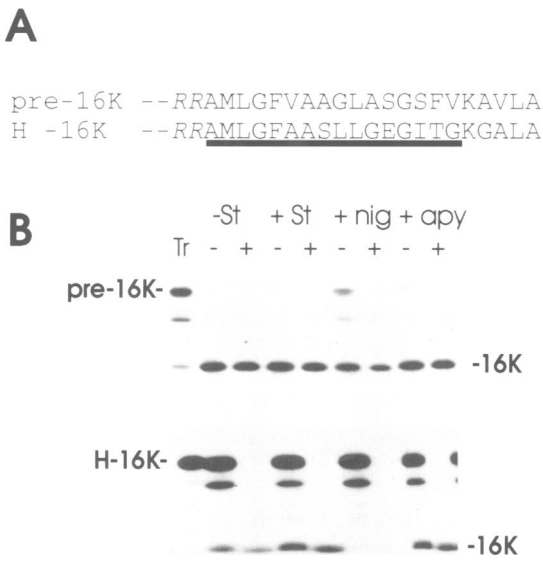

Fig. 7. Energetics of thylakoid transfer of a pre-16K mutant, H-16K, containing a foreign H-region. (**A**) Primary structure of the C-terminal section of the spinach 16K thylakoid transfer signal, and of the H-16K mutant containing 14 residues of the 22 kDa psbS gene product in place of the 16K H-region. The H-regions are underlined and the twin-arg motifs are italicized. (**B**) H-16K was incubated with pea thylakoids in the absence or presence of stromal extract (–St, +St) and in the presence of stroma together with 2 μM nigericin (+nig). Panel '+apy': the import mixture was pre-treated with 1 U of apyrase for 5 min on ice to hydrolyse NTPs. Samples were analysed directly or after protease treatment of the thylakoids (+ and –, respectively).

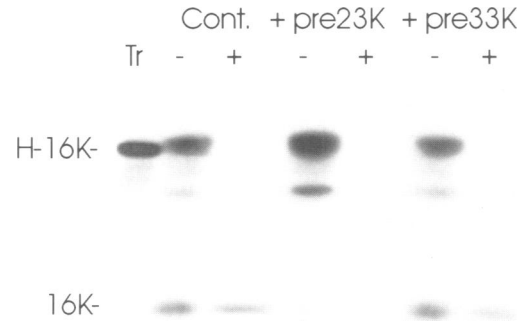

Fig. 8. Pre-23K, but not pre-33K, competes with the H-16K mutant for translocation into isolated thylakoids. H-16K was incubated with pea thylakoids in the absence of competitors (control panel) or in the presence of 4 μM spinach pre-23K or pre-33K as indicated. After incubation, samples were analysed directly (–) or after protease treatment of the thylakoids (+).

the data show that the stromal form of PC-RR declines in abundance after the early peak, consistent with this form being a bona fide intermediate on the correct transport pathway. The alternative explanation is that the stromal polypeptide is degraded, but this is unlikely for two reasons. First, we have found that the stromal form is not degraded during long import incubations performed in the presence of azide and, secondly, partially purified stromal processing peptidase generates a polypeptide of the same mobility (data not shown). It is unlikely that the processing peptidase cleaves aberrantly (the actual processing site is unaltered in the PC-RR mutant) and we therefore believe that the stromal iPC polypeptide is of the correct size. Whatever the explanation, the translocation of PC is significantly inhibited by the presence of the twin-arg motif, and we therefore propose that this motif is to some extent inhibitory in the context of a Sec-dependent precursor protein. It is unlikely that translocation is affected by the presence of the additional positive charge at the N–H junction, since barley pre-PC contains lys–lys at this position (Nielsen and Gausing, 1987).

The hydrophobic core region is unlikely to contain specific targeting information for the ΔpH-dependent pathway

The inability of the twin-arg motif to divert PC onto the ΔpH-dependent pathway implies that additional information in the 23K-type transfer signal is required for the specific interaction with its translocase. One obvious candidate for this targeting information is the H-region which is located immediately downstream of the twin-arg. The H-region of a ΔpH-driven transfer signal was therefore replaced, and in this experiment we used a hydrophobic segment derived from a transmembrane anchor of the 22 kDa PsbS protein, a member of the Cab-protein family. Integration of another member of this family, LHCPII, has been shown to proceed by a nucleoside triphosphate-dependent process which is not strictly dependent on the thylakoidal ΔpH (Cline *et al.*, 1992). The transfer signal of pre-16K, another substrate for the ΔpH-driven translocase, was chosen for the mutagenesis and the structure of this mutant (H-16K) is depicted in Figure 7A. The H-region is altered in 10 out of 14 residues, and additionally carries a negative charge in H-16K. Figure 7B shows thylakoid import assays carried out to determine the effects of the introduced H-region. The control import assays with pre-16K show that this substrate is imported almost quantitatively in both the absence and presence of stromal extract, in accordance with previous findings (Klösgen *et al.*, 1992). Nigericin causes a marked, though by no means complete, inhibition of import (lanes +nig). It remains unclear why import is not totally blocked, but one contributory factor is probably the unprecedented efficiency with which this precursor is imported. Apyrase does not inhibit import at doses which completely block import by the Sec-dependent mechanism and integration of LHCPII (Cline *et al.*, 1992; Hulford *et al.*, 1994; Robinson *et al.*, 1994). The H-16K mutant is also imported into thylakoids, although there is little doubt that the rate of import is reduced when compared with authentic pre-16K. The precise composition of the H-region is thus important for translocation to proceed with maximal efficiency. Importantly, however, import is completely inhibited by nigericin, but unaffected by apyrase, indicating that the introduced H-region has not altered the targeting requirements. These data strongly suggest that the ΔpH-driven mechanism is being exclusively utilized by the mutant, and supporting data are presented in Figure 8, in which saturating concentrations of pre-23K or pre-33K were present in the incubations to specifically block the ΔpH- and Sec-dependent mechanisms, respectively; import of H-16K is competed out by pre-23K but unaffected by pre-33K.

Discussion

Recent studies on thylakoid protein translocation/integration have revealed a surprising variety of mechanisms, with two completely different systems involved in the translocation of lumenal proteins. There is no evidence for any cross-talk between these mechanisms, and we have therefore addressed the all-important question: how do the two thylakoidal protein transport systems discriminate so successfully between two types of thylakoid transfer signal which are so similar in overall structural terms? In particular, we have sought to understand two issues: the molecular signals which are recognized by the ΔpH-dependent system (because this system is so unusual in terms of operating mechanism) and the means by which the two translocases avoid transporting the 'wrong' substrates. Our approach was to mutagenize putative targeting signals in transfer signals and to analyse the translocation requirements for the mutant proteins. First, we focused on the single concrete difference (in terms of primary structure) between the two types of transfer signal: the presence of a twin-arg motif at the N–H junction in the substrates for the ΔpH-driven translocase. In a second set of experiments, the influence of the H-region on targeting specificity was examined.

The data clearly show that the twin-arg motif plays a critical role in the mechanism by which the ΔpH-driven system recognizes its cognate precursor protein substrates. The presence of both arginine residues is essential for efficient translocation across the thylakoid membrane, and neither residue can be substituted by even lysine. The second of the arginine residues is perhaps the more critical, since substitution by lysine completely blocks transport across the thylakoid membrane, but the effects of substituting the first arginine are almost as dramatic. It is therefore almost certain that this is a specific recognition signal for the translocase, although this is in fact difficult to prove because no binding assay is available for this translocation mechanism.

Despite the importance of the twin-arg motif, the 23K transfer signal must contain additional information which is required for translocation by the ΔpH-dependent translocase, because the PC-RR mutant is not diverted onto this pathway to any detectable extent. We considered the H-region to be a good candidate for this 'additional information', because the twin-arg is always located just prior to the H-region (with the exception of cotton PSII-T, where a single glu residue separates the two). However, this is unlikely to be the case since replacement of the H-region in pre-16K by the transmembrane anchor of the PsbS protein affects the efficiency of thylakoid transfer, but does not appear to shift it to another pathway. It should be noted that we cannot exclude the possibility that substitution of the H-region enables the Sec system to recognize the precursor protein to some extent, because low-level Sec-mediated transfer may have been inhibited by the presence of the twin-arg motif (see below). Nevertheless, we feel that the function of the H-region in this type of thylakoid transfer signal is unlikely to be an important factor in enabling the two translocases to identify their correct substrates, although it is highly likely that the H-region is important for translocation by both types of mechanism.

At the outset of this work, we considered it possible that the twin-arg motif might play a dual role as a positive signal for the ΔpH-driven system, and as a negative signal masking precursors from the Sec system. The available sequence data are consistent with such a dual role, because Sec-dependent substrates contain no arginine residues at all (let alone twin-arg motifs) at the N–H junction, at least in higher plants. We believe that the twin-arg does indeed play such a masking role because its presence in the PC-RR mutant does inhibit the Sec mechanism to a significant extent. In other words, the transport characteristics of the PC-RR mutant may well explain the absence of even single arginine residues at the N–H junction of known higher plant transfer signals. Other factors must be involved in the discrimination process, however, because the PC-RR mutant is still imported into the lumen at an appreciable rate.

Even taking into account the possible masking effect of the twin-arg, it remains unclear why 23K-type transfer signals fail to be recognized by the Sec system. If the presence of either twin-arg or a single lysine at the N–H junction in pre-PC is tolerated, at least to an extent, by the Sec system, why does this system not recognize either pre-23K or the mutants generated in this study? The available evidence suggests that wild-type 23K does not bind to any translocation factors in the stroma (Creighton *et al.*, 1995), implying that the 23K transfer signal is specifically recognized only at the thylakoid membrane. There are thus ample opportunities for the soluble components of the Sec apparatus (SecA and SecB, if present) to interact with 23K-type transfer signals, yet these signals appear to be completely ignored. In the absence of translocation by the ΔpH-dependent translocase, the mutants analysed in this study are certainly not transported by the Sec system to any great extent, yet there are no other clear-cut differences between the two types of thylakoid transfer signal. In general, transfer signals for the ΔpH-driven translocase contain negatively charged areas in the N-regions which are notably absent in bacterial signal peptides. However, the Sec-dependent substrates also differ from their cyanobacterial counterparts in containing acidic residues in this region, casting doubt on the ability of acidic residues to repel the Sec machinery. Nevertheless, it is interesting that the ΔpH-dependent substrates tend to contain a 'patch' of acidic residues in the N-regions, and these residues are certainly worthy of further study. Given that the H-region is unlikely to contain specific targeting information, and the C-region is relatively small, the N-region is the most likely location for further specific targeting signals, and further studies should indicate whether this is in fact the case.

Finally, it is interesting to speculate that these results suggest a fundamental difference between the 23K-type transfer signal and the signal peptides which target proteins across the ER and bacterial plasma membranes. The latter types of peptide also contain a critical basic residue(s) at the N–H junction, but in these cases it is clear that the identity of the side chain(s) is irrelevant (Sasaki *et al.*, 1990); it is simply necessary for one or more positive charges to be positioned before the H-region. Lysine, arginine and histidine are interchangeable in bacterial signal peptides, whereas the arginine guanidyl side chains are a specific requirement in thylakoid transfer signals for

the ΔpH-driven translocase. This finding implies an equally fundamental difference between the mechanisms by which the two types of signal interact with translocation systems and promote membrane transfer. An overall positive charge in the N-region is undoubtedly important for the functioning of the Sec-dependent mechanism in bacteria, and it has been suggested that the positive charge is required for interaction with the negatively charged phospholipids in the membrane bilayer (Inouye and Halegoua, 1980) or for an electrostatic interaction with SecA, which also has a net negative charge (Akita *et al.*, 1990). Whatever the role of the positive charge(s), it is most unlikely that the side chain(s) bearing the charges are recognized in any specific manner, simply because the arginine, lysine and histidine side chains are so different in structural terms. The importance of the arginine side chains in the 23K transfer signal suggests that they are involved in a far more specific, localized, protein–protein interaction, and the bulkier end groups of the arginine side chains (when compared with lysine side chains) are presumably important in this respect.

Materials and methods

Plants

Seedlings of *Pisum sativum*, var Feltham First, were grown for 8–9 days under a 12 h photoperiod. Light intensity was 50 μmol photons/m/s.

Mutagenesis of precursors

A cDNA insert encoding wheat pre-23K (James and Robinson, 1991) and a genomic clone encoding spinach pre-PC (Rother *et al.*, 1986) were cloned into M13 mp18, and codons 51 and/or 52 of pre-23K, and 47 and 48 of pre-PC, were altered to the residues shown in Figure 2. Oligonucleotide-directed mutagenesis was carried out using an Amersham International plc kit, following the manufacturer's instructions. The mutations were identified by DNA sequencing, and clones carrying the desired mutations were completely sequenced to confirm the absence of any additional mutations. The inserts were then cloned into pGEM 4Z (Promega Biotech) and the precursor proteins synthesized by *in vitro* transcription–translation. Wild-type pre-23K and pre-PC were synthesized as previously described (Robinson *et al.*, 1994). In order to generate the H-16K mutant, a cDNA insert encoding spinach pre-16K (Jansen *et al.*, 1987) was cloned into M13 mp9, and codons 68, 70, 71, 73, 74, 76, 77, 78, 80 and 81 were altered using the gap duplex method (Kramer *et al.*, 1984). The oligonucleotide used (5′-TTGGGCTTCGCCGCATCTCTTTTGGGTGAAGGTATTACTGGTAAGGGTGCTCTTGCTGAG-3′) resulted in the amino acid sequence shown in Figure 7. After confirmation of the introduced sequence by nucleotide sequence analysis, the insert was cloned into pBSC M13$^-$ (Stratagene) and the protein synthesized by transcription–translation.

Import assays

Assays for the import of proteins into isolated pea chloroplasts were as described by Mould and Robinson (1991) except that nigericin, when present, was used at 2 μM. Incubations with azide present (10 mM) were carried out as detailed in Knott and Robinson (1994). Thylakoid import assays were as described by Brock *et al.* (1993). Precursor proteins for use in competition studies were prepared as detailed in Michl *et al.* (1994). Quantitation of import efficiencies was achieved using a Molecular Dynamics PhosphorImager.

Acknowledgements

This work was supported by Biotechnology and Biological Sciences Research Council grant PO2474, by a Science and Engineering Council Studentship to A.M., and by an EMBO long-term fellowship to S.B. We also gratefully acknowledge a grant from the British–German Academic Research Collaboration Programme, funded by the British Council and the German Academic Exchange Scheme.

References

Akita,M., Sasaki,S., Matsuyama,S. and Mizushima,S. (1990) *J. Biol. Chem.*, **265**, 8164–8169.
Bassham,D.C., Bartling,D., Mould,R.M., Dunbar,B., Weisbeek,P., Herrmann,R.G. and Robinson,C. (1991) *J. Biol. Chem.*, **266**, 23606–23610.
Bovy,A., de Vrieze,G., Borrias,M. and Weisbeek,P. (1992) *Plant Mol. Biol.*, **19**, 491–492.
Briggs,L.M., Peroraro,V.L. and McIntosh,L. (1990) *Plant Mol. Biol.*, **15**, 633–642.
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.
Chitnis,P.R., Purvis,D. and Nelson,N. (1991) *J. Biol. Chem.*, **266**, 20146–20151.
Cline,K., Ettinger,W. and Theg,S.M. (1992) *J. Biol. Chem.*, **267**, 2688–2696.
Creighton,A.M., Hulford,A., Mant,A., Robinson,D. and Robinson,C. (1995) *J. Biol. Chem.*, **270**, 1663–1669.
Haehnel,W., Jansen,T., Gause,K., Klösgen,R.B., Stahl,B., Michl,D., Huvermann,B., Karas,M. and Herrmann,R.G. (1994) *EMBO J.*, **13**, 1028–1038.
Hageman,J., Robinson,C., Smeekens,S. and Weisbeek,P. (1986) *Nature*, **324**, 567–569.
Henry,R., Kapazoglou,A., McCaffrey,M. and Cline,K. (1994) *J. Biol. Chem.*, **269**, 10189–10192.
Hulford,A., Hazell,L., Mould,R.M. and Robinson,C. (1994) *J. Biol. Chem.*, **269**, 3251–3256.
Inouye,M. and Halegoua,S. (1980) *CRC Crit. Rev. Biochem.*, **7**, 339–371.
Izard,J.W. and Kendall,D.A. (1994) *Mol. Microbiol.*, **13**, 765–773.
James,H.E. and Robinson,C. (1991) *Plant Mol. Biol.*, **17**, 179–182.
James,H.E., Bartling,D., Musgrove,J.E., Kirwin,P.M., Herrmann,R.G. and Robinson,C. (1989) *J. Biol. Chem.*, **264**, 19573–19576.
Jansen,T., Rother,C., Steppuhn,J., Reinke,H., Bayreuther,K., Jansson,C., Andersson,B. and Herrmann,R.G. (1987) *FEBS Lett.*, **216**, 234–240.
Karnauchov,I., Cai,D., Schmidt,I., Herrmann,R.G. and Klösgen,R.B. (1994) *J. Biol. Chem.*, **269**, 32871–32878.
Klösgen,R.B., Brock,I.W., Herrmann,R.G. and Robinson,C. (1992) *Plant Mol. Biol.*, **18**, 1031–1034.
Knoetzel,J. and Simpson,D. (1993) *Plant Mol. Biol.*, **22**, 337–345.
Knott,T.G. and Robinson,C. (1994) *J. Biol. Chem.*, **269**, 7843–7846.
Kramer,W., Drutsa,V., Janse,H.-W., Kramer,B., Pflugfelder,M. and Fritz,H.-J. (1984) *Nucleic Acids Res.*, **12**, 9441–9456.
Kuwabara,T., Reddy,K.J. and Sherman,L.A. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 8230–8235.
Laudenbach,D.E., Herbert,S.K., McDowell,C., Fork,D.C., Grossman, A.R. and Straus,N.A. (1990) *Plant Cell*, **2**, 913–924.
Mant,A., Nielsen,V.S., Knott,T.G., Møller,B.L. and Robinson,C. (1994) *J. Biol. Chem.*, **269**, 27303–27308.
Meadows,J.W. and Robinson,C. (1991) *Plant Mol. Biol.*, **17**, 1241–1243.
Meadows,J.W., Hulford,A., Raines,C.A. and Robinson,C. (1991) *Plant Mol. Biol.*, **16**, 1085–1087.
Michl,D., Robinson,C., Shackleton,J.B., Herrmann,R.G. and Klösgen, R.B. (1994) *EMBO J.*, **13**, 1310–1317.
Mould,R.M. and Robinson,C. (1991) *J. Biol. Chem.*, **266**, 12189–12193.
Mould,R.M., Shackleton,J.B. and Robinson,C. (1991) *J. Biol. Chem.*, **266**, 17286–17290.
Nakai,M., Goto,A., Nohara,T., Sugita,D. and Endo, T. (1994) *J. Biol. Chem.*, **269**, 31338–31341.
Newman,S.M., Eannetta,N.T., Yu,H., Prince,J.P., de Vicente,M.C., Tanksley,S.D. and Steffens,J.C. (1993) *Plant Mol. Biol.*, **21**, 1035–1051.
Nielsen,P.S. and Gausing,K. (1987) *FEBS Lett.*, **225**, 159–162.
Nielsen,V.S., Mant,A., Knoetzel,J., Møller,B.L. and Robinson,C. (1994) *J. Biol. Chem.*, **269**, 3762–3766.
Philbrick,J.B. and Zilinskas,B.A. (1988) *Mol. Gen. Genet.*, **212**, 418–425.
Robinson,C. and Klösgen,R.B. (1994) *Plant Mol. Biol.*, **26**, 15–24.
Robinson,C., Cai,D., Hulford,A., Hazell,L., Michl,D., Brock,I., Schmidt,I., Herrmann,R.G. and Klösgen,R.B. (1994) *EMBO J.*, **13**, 279–285.
Rother,C., Janson,T., Tyagi,A., Tittgen,J. and Herrmann,R.G. (1986) *Curr. Genet.*, **11**, 171–176.
Sasaki,S., Matsuyama,S. and Mizushima,S. (1990) *J. Biol. Chem.*, **265**, 4358–4363.
Seidler,A. and Michel,H. (1990) *EMBO J.*, **9**, 1743–1748.

Shackleton,J.B. and Robinson,C. (1991) *J. Biol. Chem.*, **266**, 12152–12156.
Shestakov,S.V., Anbudurai,P.R., Stanbekova,G.E., Gadzhiev,A., Lind, L.K. and Pakrasi,H.B. (1994) *J. Biol. Chem.*, **269**, 19354–19359.
Sommer,A., Ne'eman,E., Steffens,J.C., Mayer,A.M. and Harel,E. (1994) *Plant Physiol.*, **105**, 1301–1311.
Steppuhn,J., Hermans,J., Nechushtai,R., Ljungberg,U., Thümmler,F., Lottspeich,F. and Herrmann,R.G. (1988) *FEBS Lett.*, **237**, 218–224.
Theg,S.M., Bauerle,C., Olsen,L., Selman,B. and Keegstra,K. (1989) *J. Biol. Chem.*, **264**, 6730–6736.
von Heijne,G., Steppuhn,J. and Herrmann,R.G. (1989) *Eur. J. Biochem.*, **180**, 535–545.
Yuan,J., Henry,R., McCaffrey,M. and Cline,K. (1994) *Science*, **266**, 796–798.
Zhang,L., Pakrasi,H.B. and Whitmarsh,J. (1994) *J. Biol. Chem.*, **269**, 5036–5042.

Received on February 22, 1995; revised on March 22, 1995