FEBS 17267 FEBS Letters 390 (1996) 175–178

An *Arabidopsis thaliana* cDNA encoding PS II-X, a 4.1 kDa component of photosystem II: a bipartite presequence mediates SecA/ΔpH-independent targeting into thylakoids

Soo Jung Kim, David Robinson, Colin Robinson*

Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK

Received 4 June 1996

Abstract Higher plant photosystem II preparations contain a 4.1 kDa polypeptide (subunit X) associated with the oxygen-evolving core complex. We describe the isolation of a cDNA encoding PS II-X from *Arabidopsis thaliana*, in which the C-terminal region is highly homologous to partially sequenced PS II-X from wheat and spinach. The mature protein of 42 residues is preceded by a 74-residue, bipartite presequence similar to those involved in the targeting of nuclear-encoded thylakoid lumen proteins, although hydrophobicity analysis indicates the presence of a single transmembrane span in the mature protein. Moreover, import of pre-PS II-X into the thylakoid membrane of isolated chloroplasts is unaffected by inhibitors of either the Sec- or ΔpH-dependent thylakoidal protein translocases, suggesting a spontaneous insertion mechanism. PS II-X appears to be encoded as a mature protein by the plastid genome in the chlorophyll *a+c*-containing alga, *Odontella sinensis*. We thus propose that the thylakoid transfer signal of *Arabidopsis* pre-PS II-X represents a recent acquisition, in phylogenetic terms, compared with signals of Sec-dependent lumenal proteins.

Key words: Chloroplast; psbX; Photosynthesis; Photosystem II; Protein translocation; Thylakoid

---

1. Introduction

Photosystem II (PS II) is a complex protein-pigment assembly responsible for the transfer of electrons from water to plastoquinone. Over 20 polypeptides are associated with this complex in higher plant chloroplasts, some of which are encoded by the chloroplast genome whereas others are imported after synthesis in the cytosol [1,2]. Some of the proteins have been characterised in considerable detail, but the roles of other PS II components have yet to be rigorously investigated. This applies particularly to a series of low molecular mass polypeptides (in the range of 3–10 kDa) that were first identified through the use of high-resolution polyacrylamide gel electrophoresis techniques [3–5]. N-terminal sequencing of three of these polypeptides, of apparent molecular masses 6.1, 5.0 and 4.1 kDa, suggested that they were not encoded by the chloroplast genome [5], but it is only recently that cDNAs encoding the 5.0 and 6.1 kDa proteins have been isolated. Both proteins are indeed nuclear-encoded. The 5.0 kDa protein is encoded by *PsbT*, and the encoded PS II-T protein has been shown to be located on the lumenal side of the thylakoid membrane [6]. The 6.1 kDa polypeptide, encoded by *PsbW*, is a transmembrane component of the reaction center complex [7].

The cloning of PS II proteins has also led to significant advances in our understanding of the targeting of nuclear-encoded thylakoid proteins, and a surprising variety of targeting pathways has emerged (reviewed in [8]. Most, if not all, of these proteins appear to be imported across the chloroplast envelope membranes by a common mechanism, but multiple pathways have been identified for their subsequent targeting into or across the thylakoid membrane. Lumenal proteins such as the extrinsic 33, 23 and 16 kDa PS II proteins (33K, 23K, 16K) are synthesised with bipartite presequences containing envelope transit and thylakoid transfer signals in tandem. After removal of the envelope transit domain, the transfer signal interacts with translocation machinery in the thylakoid membrane to initiate translocation into the lumenal space; a lumen-facing thylakoidal processing peptidase (TPP) then completes maturation [9]. Translocation of 33K is effected by an ATP-dependent, Sec-type system that is believed to have been inherited from the cyanobacterial-type progenitor of the higher plant chloroplast [10–12], and, in cyanobacteria, 33K is synthesised with a presequence similar in structural terms to both the transfer signals of the higher plant 33Ks and to bacterial signal peptides [13]. Interestingly, 23K and 16K are also synthesised with bipartite presequences, yet the transfer signals specify translocation across the thylakoid membrane by an entirely different mechanism that does not require ATP but which is instead absolutely dependent on the thylakoidal ΔpH [14–16].

The 4.1 kDa PS II protein (PS II-X) is the least well characterised of the low molecular mass PS II proteins. The protein is known to be present in oxygen-evolving core preparations of PS II [5], but has not been cloned from higher plants. However, an open reading frame homologous to the known N-terminal sequence of the 4.1 kDa protein has been identified in the chloroplast genome of the chlorophyll *a+c*-containing alga, *Odontella sinensis* [17], and designated *PsbX*. The predicted amino acid sequence shows no evidence of any presequence. In this report we show that PS II-X is encoded by a nuclear gene in the higher plant, *Arabidopsis thaliana*. We also show that PS II-X is synthesised with a bipartite-type presequence but imported into the thylakoid membrane by a SecA/ΔpH-independent manner.

2. Materials and methods

2.1. Cloning of *PsbX*

A cDNA clone encoding pre-PS II-X was obtained by searching the database of expressed sequence tags (ESTs) using the known N-terminal sequence of the 4.1 kDa PS II protein [5]. Several *A. thaliana* ESTs [18] were identified as potential candidates and obtained from the *Arabidopsis* Biological Resource Center (Columbus, OH). One of these, clone 109H9T7 from a λ zip-lox library (BRL), was found to be full-length and designated pPsbX109. Partial sequence data had been

---

*Corresponding author. Fax: (44) (1203) 523701.

PII S0014-5793(96)00658-8

obtained in the laboratory of T. Newman (Michigan) and we carried out full sequencing of both strands of the cDNA.

2.2. Import and processing studies

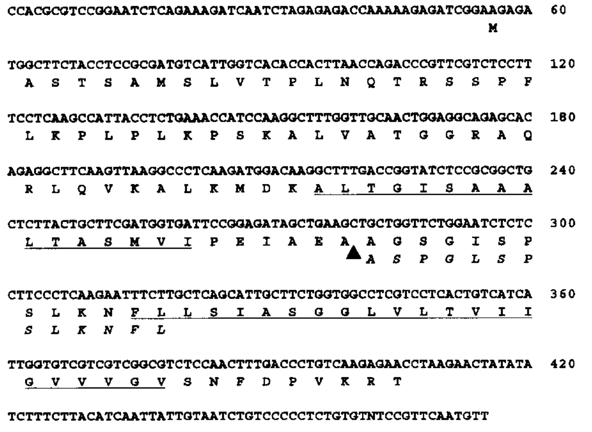
The pSPORT plasmid from clone pPsbX109 was transcribed using T7 RNA polymerase and capped transcripts were translated in a wheatgerm lysate in the presence of [^{3}H]leucine. Pre-PS II-X generated in this way was incubated with intact pea chloroplasts under conditions described elsewhere [14]. Partially purified pea TPP was prepared as detailed in Kirwin et al. [9]; processing of pre-PS II-X was assayed by incubation of 20 μl processing activity with 2 μl pre-PS II-X translation product for 60 min at 27°C.

3. Results and discussion

3.1. Arabidopsis PS II-X is synthesised with a bipartite presequence

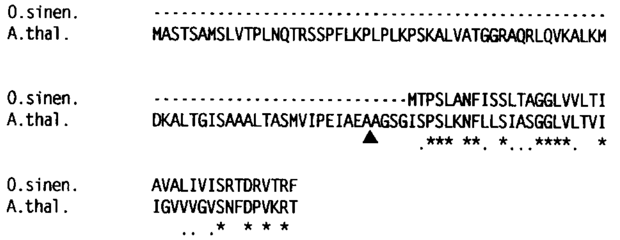
A cDNA clone was obtained by searching the database of expressed sequence tags (ESTs) using the N-terminal sequence data for the 4.1 kDa protein [5]. Several homologous sequences were detected, and one cDNA (designated pPsbX109) was shown to be full-length on the basis of import experiments (see below). The 475 bp nucleotide sequence of this cDNA predicts an open reading frame of 116 residues (Fig. 1). Within this reading frame, residues 75–87 are highly homologous to the 13-residue N-terminal sequence of the wheat 4.1 kDa protein, determined by Ikeuchi et al. [5]. The two sequences share 11 identical residues over this 13 amino acid span, and the C-terminal part of the reading frame is also highly homologous to the *Odontella* open reading frame designated *PsbX* (Fig. 2). PS II-X is therefore encoded by a nuclear gene (*PsbX)* in *Arabidopsis* and synthesised with an N-terminal presequence, as are all known nuclear-encoded thylakoid proteins.

The presequence of PS II-X bears all the hallmarks of a bipartite presequence typical of imported thylakoid lumen proteins. Envelope transit signals are characteristically basic, hydrophilic and enriched in hydroxylated residues [19], and the first domain (approx. 40–50 residues) of the PS II-X presequence shares these characteristics. Thylakoid transfer signals, on the other hand, contain an N-terminal charged region, a hydrophobic core region, and a C-terminal region ending with Ala-X-Ala, the consensus signal for TPP cleavage [19,20]. On this basis, the last approx. 25–30 residues of the presequence correspond to a thylakoid transfer signal that ends with Ala-Glu-Ala (the putative cleavage site is designated by an arrow in Fig. 1). This assignment of the cleavage site is supported by the fact that the known 4.1 kDa polypeptide sequence (from wheat) follows immediately afterwards. Pre-PS II-X therefore comprises a bipartite-type presequence of 74 residues followed by a mature protein of 42 residues.



```
CCACGCGTCCGGAATCTCAGAAAGATCAATCTAGAGAGACCAAAAAGAGATCGGAAGAGA   60
                                                           M

TGGCTTCTACCTCCGCGATGTCATTGGTCACACCACTTAACCAGACCCGTTCGTCTCCTT   120
  A  S  T  S  A  M  S  L  V  T  P  L  N  Q  T  R  S  S  P  F

TCCTCAAGCCATTACCTCTGAAACCATCCAAGGCTTTGGTTGCAACTGGAGGCAGAGCAC   180
  L  K  P  L  P  L  K  P  S  K  A  L  V  A  T  G  G  R  A  Q

AGAGGCTTCAAGTTAAGGCCCTCAAGATGGACAAGGCTTTGACCGGTATCTCCGCGGCTG   240
  R  L  Q  V  K  A  L  K  M  D  K  A  L  T  G  I  S  A  A  A

CTCTTACTGCTTCGATGGTGATTCCGGAGATAGCTGAAGCTGCTGGTTCTGGAATCTCTC   300
  L  T  A  S  M  V  I  P  E  I  A  E  A  A  G  S  G  I  S  P
                                       ▲ A  S  P  G  L  S  P

CTTCCCTCAAGAATTTCTTGCTCAGCATTGCTTCTGGTGGCCTCGTCCTCACTGTCATCA   360
  S  L  K  N  F  L  L  S  I  A  S  G  G  L  V  L  T  V  I  I
  S  L  K  N  F  L

TTGGTGTCGTCGTCGGCGTCTCCAACTTTGACCCTGTCAAGAGAACCTAAGAACTATATA   420
  G  V  V  V  G  V  S  N  F  D  P  V  K  R  T

TCTTTCTTACATCAATTATTGTAATCTGTCCCCCTCTGTGTNTCCGTTCAATGTT
```

Fig. 1. Nucleotide and predicted amino acid sequence of pPsbX109, encoding pre-PS II-X from *Arabidopsis thaliana*. The homologous N-terminal sequence of the wheat 4.1 kDa protein [5] is given in italics underneath amino acid residues 75–76 of the *Arabidopsis* protein sequence; the arrow preceding these sequences denotes the TPP processing site. The hydrophobic core region of the thylakoid transfer signal and the putative transmembrane sequence of the mature protein are underlined.



```
O.sinen.    --------------------------------------------------
A.thal.     MASTSAMSLVTPLNQTRSSPFLKPLPLKPSKALVATGGRAQRLQVKALKM


O.sinen.    ----------------------------MTPSLANFISSLTAGGLVVLTI
A.thal.     DKALTGISAAALTASMVIPEIAEAAGSGISPSLKNFLLSIASGGLVLTVI
                                   ▲     .*** **. *...****.  *


O.sinen.    AVALIVISRTDRVTRF
A.thal.     IGVVVGVSNFDPVKRT
               .. .*  * * *
```

Fig. 2. Alignment of pre-PS II-X with the PsbX open reading frame from *Odontella sinensis* chloroplast DNA. The *A. thaliana* (A. thal.) and *O. sinensis* (O. sinen.) sequences were aligned using the CLUSTAL V programme; identical residues are denoted by asterisks, conserved residues by dots. Arrow denotes TPP cleavage site.

Although bipartite presequences are usually found in the presequences of lumenal proteins, hydrophobicity analysis strongly points to PS II-X being a transmembrane protein (Fig. 3). The hydropathy plot predicts a single transmembrane span with a very high degree of probability (hydrophobic segment II) and also illustrates the presence of the hydrophobic core region in the presequence (segment I).

3.2. Translocation of PS II-X into the thylakoid membrane is azide- and nigericin-insensitive

Although the sequence data alone suggest that PS II-X is synthesised with a bipartite presequence and targeted into the thylakoid membrane, experimental data are required to confirm these points. Pre-PS II-X was therefore synthesised by transcription of pPsbX109 in vitro, followed by cell-free translation in the presence of [^{3}H]leucine, and the precursor protein was then imported into intact chloroplasts. We furthermore tested whether the protein follows the Sec-related or ΔpH-dependent pathways by including diagnostic inhibitors of the two systems: azide has been shown to markedly inhibit the Sec-type pathway (probably by inhibiting the ATPase action of SecA) and nigericin is a proton ionophore that totally blocks translocation by the ΔpH-driven mechanism [21,22,14,15]. The data are shown in Fig. 4A. The control panel (minus inhibitors) shows that PS II-X is synthesised as a precursor of apparent molecular mass 12 kDa which is imported into the chloroplasts and processed to a mature size of apparent molecular mass 5.5 kDa (lane C). The mature protein is protected from protease digestion of the organelles (lane C+). Mature PS II-X is localised in the thylakoid fraction (lane T) with little or no protein apparent in the stroma (lane S). Although the imported protein runs as 5.5 kDa rather than 4.1 kDa, this is a minor difference which probably reflects differences in electrophoresis methods.

Inclusion of nigericin has essentially no effect on the import profile in Fig. 4A, demonstrating that a ΔpH is not required for the insertion of this protein into the thylakoid membrane.

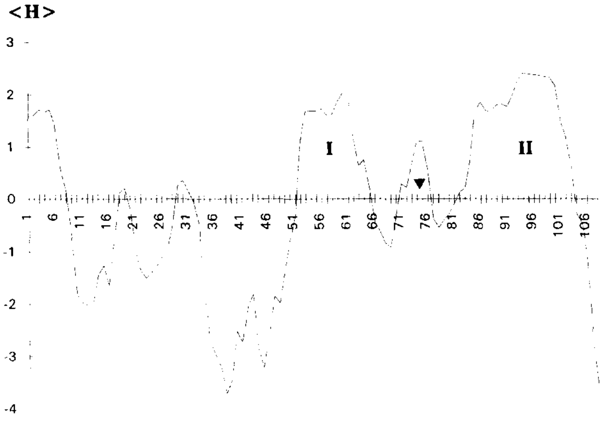


Fig. 3. Hydropathy plot [24] of pre-PS II-X. The pre-PS II-X sequence was analysed using the TopPred II program with a core window of 7 residues and a full window of 9 residues. Mean hydrophobicity is plotted against residue number; the arrow denotes the TPP processing site. Also depicted are the hydrophobic core region of the thylakoid transfer signal (I) and the putative transmembrane span (II).

Neither is PS II-X targeted by the azide-sensitive Sec-type mechanism; azide has little or no effect on the amount of mature-size PS II-X that is present in the thylakoid fraction. However, there is a slight accumulation of pre-PS II-X in the protease-treated chloroplasts, which could reflect a slight inhibition of the insertion process if it is the full precursor protein that is inserted into thylakoids. It is probable that this reflects a non-specific effect of the azide, because azide has a far more dramatic effect on the translocation of known substrates for the Sec machinery (for example, 33K [21,22]). It should also be pointed out that, if this species is inside the chloroplasts, pre-PS II-X represents a much lower proportion of imported protein than is immediately apparent. Mature PS II-X contains only a third of the radiolabeled leucine originally present in the precursor protein (10 of the 15 leucines are in the presequence), hence the mature size protein is in fact by far the predominant imported species in all three incubations.

Most proteins synthesised with bipartite presequences (though not all) are cleaved to intermediate-size forms in the stroma, but we do not yet know whether this applies to PS II-X. An intermediate-size polypeptide is present in some samples from incubations carried out in the presence of azide or nigericin, but this polypeptide is only evident in chloroplasts treated with protease. We have found that both pre-PS II-X and imported PS II-X are very resistant to proteolysis, and this intermediate band is in fact due to proteolysis of

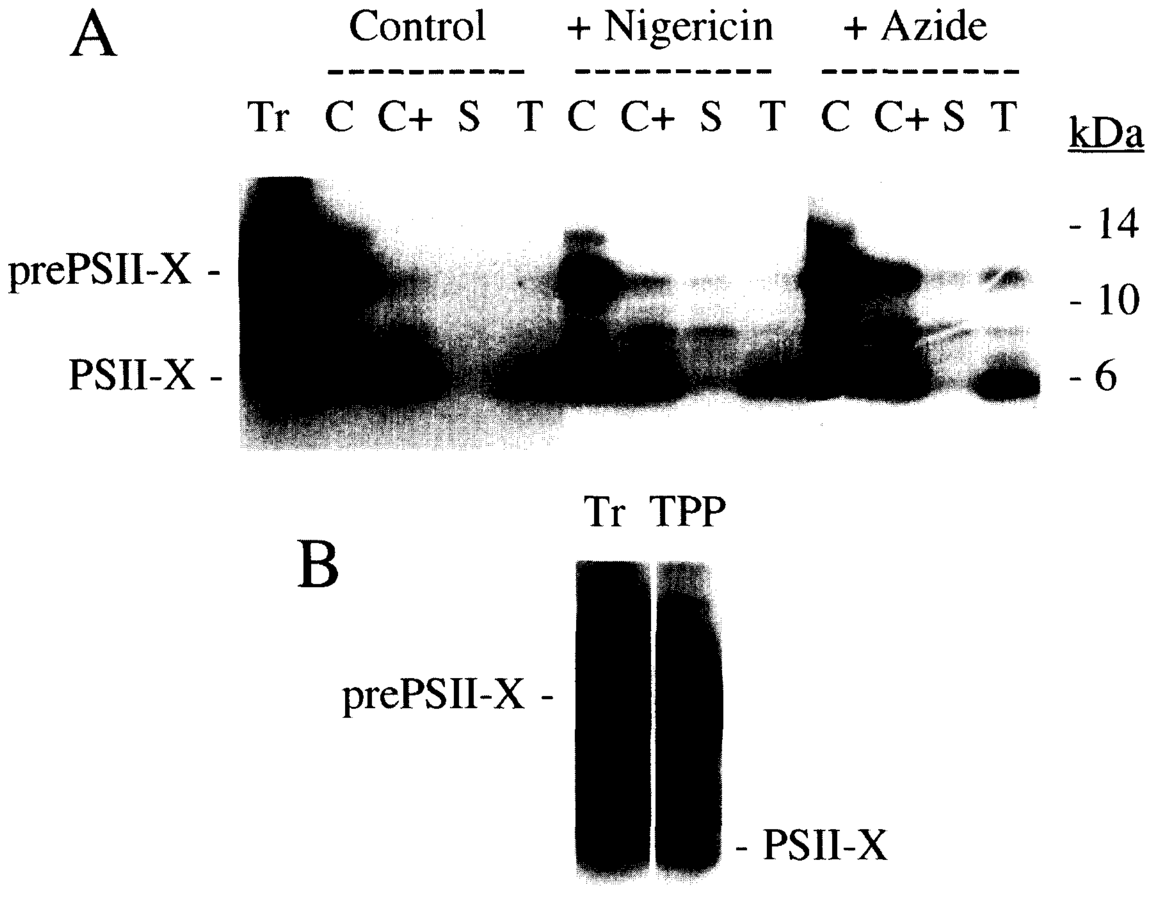


Fig. 4. Import of pre-PS II-X into pea chloroplasts and localisation of the processed product. (A) Pre-PS II-X was synthesized in vitro by transcription-translation of pPsbX109 as detailed in Section 2, and the radiolabeled translation product (lane Tr) was incubated with intact pea chloroplasts. Incubations were carried out in the absence of inhibitors (control) or in the presence of 2 μM nigericin or 10 mM Na azide as indicated. After incubation, samples of the chloroplasts were analysed directly (lane C) or after protease treatment of the organelles (lane C+). Other samples were lysed and centrifuged to generate samples of stromal fraction (S) or thylakoids (lanes T). The mobilites of marker proteins are given on the right-hand side. (B) Pre-PS II-X (lane Tr) was incubated with partially purified pea TPP activity (lane TPP) for 60 min.

bound pre-PS II-X and is hence absent from import incubations that are not protease-treated (lanes 'C'). For the same reason, we have not been able to use protease sensitivity as a test of whether the imported PS II-X is lumenal or transmembrane (the latter type of protein should be sensitive to digestion of thylakoids). Imported, mature-size PS II-X resists digestion by a wide range of proteases (data not shown), possibly because only a small section is predicted to remain exposed on the stromal face of the membrane.

Finally, Fig. 4B shows that incubation of the precursor protein with highly purified pea TPP activity [9] leads to the generation of mature-size PS II-X. Because TPP is so highly specific for thylakoid transfer signals [20], we conclude that that pre-PS II-T does indeed contain a site for this enzyme and, by inference, a thylakoid transfer signal.

3.3. Implications for the evolution of pre-PS II-X

Although pre-PS II-X clearly contains a bipartite presequence, the biogenesis of this protein differs in two intriguing ways from that of the vast majority of other proteins bearing bipartite presequences. First, import of PS II-X is not affected by standard inhibitors of either the Sec-type or ΔpH-driven thylakoidal protein translocases, suggesting that a different import pathway is utilised. The same applies to the import of pre-PsbW [7]. It may be that these proteins follow a variation of the Sec-dependent route, perhaps utilising some elements of the Sec machinery but not others, and further experiments using thylakoid import assays are in progress to test this possibility. However, the discovery of a *PsbX* coding sequence in the chloroplast genome of *O. sinensis* reveals a second difference. Cyanobacterial counterparts of nuclear-encoded, Sec-dependent lumenal proteins of higher plants (for example, 33K, plastocyanin and PS I subunit F) are invariably synthesised with transfer signals/ signal sequences that mediate interaction with translocation apparatus (presumably a Sec system) in the thylakoid membrane. The same applies when these lumenal proteins are found to be encoded by the chloroplast genome; PS I subunit F is chloroplast-encoded in *O. sinensis* and clearly synthesised with a signal sequence [17]. However, PS II-X in the same organism is not synthesised with a presequence and the implication is that the acquisition of the transfer signal in *Arabidopsis* is a recent event in phylogenetic terms. Indeed, higher plant PS II-X appears similar in every respect to a higher plant protein found to follow a novel import pathway: CF_0II. This protein is likewise synthesised with a bipartite presequence but requires neither ATP nor a ΔpH for integration into the thylakoid membrane [23]. Instead, a spontaneous mechanism has been proposed in which the two hydrophobic sections in pre-CF_0II (one in the transfer signal and one in the mature protein) partition into the membrane and allow the intervening sequence (the N-terminus of the mature protein) to flip across. The cyanobacterial homolog of CF_0II is furthermore synthesised without a signal sequence. The biogenesis and phylogeny of CF_0II have tended to be regarded as exceptional, but our data on PS II-X indicate that the transfer of the *PsbX* gene to the nucleus likewise necessitated the acquisition of a transfer signal for targeting into the thylakoid membrane. The inhibitor data suggest that this transfer signal fulfils a similarly unusual role, and it will be interesting to examine in detail the integration mechanisms for PS II-X and PS II-W in order to determine whether a mainstream pathway operates for the targeting of this group of proteins.

Acknowledgements: This work was supported by Biotechnology and Biological Sciences Research Council grant P05148 to C.R. We are grateful to Gunnar von Heijne for advice concerning hydrophobicity profiles, and to the Arabidopsis Biological Resource Center for their support.

References

[1] Andersson, B. and Styring, S. (1991) Curr. Top. Bioenerg. 16, 1–81.
[2] Gruissem, W. (1989) Cell 56, 161–170.
[3] Ljungberg, U., Henrysson, T., Rochester, C.P., Åkerlund, H.-E. and Andersson, B. (1986) Biochim. Biophys. Acta 849, 112–120.
[4] Ikeuchi, M. and Inoue, Y. (1988) Plant Cell Physiol. 29, 1233–1239.
[5] Ikeuchi, M., Takio, K. and Inoue, Y. (1989) FEBS Lett. 242, 263–269.
[6] Kapazoglou, A., Sagliocco, F. and Dure, L. (1995) J. Biol. Chem. 270, 12197–12202.
[7] Lorkovic, Z.J., Schröder, W.P., Pakrasi, H.B., Irrgang, K.-D., Herrmann, R.G. and Oelmüller, R. (1995) Proc. Natl. Acad. Sci. USA 92, 8930–8934.
[8] Robinson, C. and Klösgen, R.B. (1994) Plant Mol. Biol. 26, 15–24.
[9] Kirwin, P.M., Elderfield, P.E., Williams, R.S. and Robinson, C. (1988) J. Biol. Chem. 263, 18128–18132.
[10] Nakai, M., Goto, A., Nohara, T., Sugita, D. and Endo, T. (1994) J. Biol. Chem. 269, 31338–31341.
[11] Yuan, J., Henry, R., McCaffrey, M. and Cline, K. (1994) Science 266, 796–798.
[12] Hulford, A., Hazell, L., Mould, R.M. and Robinson, C. (1994) J. Biol. Chem. 269, 3251–3256.
[13] Kuwabara, T., Reddy, K.J. and Sherman, L.A. (1987) Proc. Natl. Acad. Sci. USA 84, 8230–8235.
[14] Mould, R.M. and Robinson, C. (1991) J. Biol. Chem. 266, 12189–12193.
[15] Cline, K., Ettinger, W. and Theg, S.M. (1992) J. Biol. Chem. 267, 2688–2696.
[16] Klösgen, R.B., Brock, I.W., Herrmann, R.G. and Robinson, C. (1992) Plant Mol. Biol. 18, 1031–1034.
[17] Kowallik, K.V., Stoebe, B., Schaffran, I., Kroth-Pancic, P. and Freier, U. (1995) Plant Mol. Biol. Rep. 13, 336–342.
[18] Newman, T., de Bruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E. and Somerville, C. (1994) Plant Physiol. 106, 1241–1255.
[19] Von Heijne, G., Steppuhn, J. and Herrmann, R.G. (1989) Eur. J. Biochem. 180, 535–545.
[20] Shackleton, J.B. and Robinson, C. (1991) J. Biol. Chem. 266, 12152–12156.
[21] Knott, T.G. and Robinson, C. (1994) J. Biol. Chem. 269, 7843–7846.
[22] Henry, R., Kapazoglou, A., McCaffrey, M. and Cline, K. (1994) J. Biol. Chem. 269, 10189–10192.
[23] Michl, D., Robinson, C., Shackleton, J.B., Herrmann, R.G. and Klösgen, R.B. (1994) EMBO J. 13, 1310–1317.
[24] Claros, M.G. and Von Heijne, G. (1994) CABIOS 10, 685–686.